UNIVERSITY OF EDINBURGH
Business School

Enhancing Corporate Bankruptcy and Financial Distress Prediction: A Multi-Method Approach with Profiling and Network Analysis

Jinxian Zhao

Submitted for the degree of Doctor of Philosophy

University of Edinburgh

August, 2024

Abstract

This thesis encompasses three interrelated studies on the prediction of corporate bankruptcy and financial distress, a critical area for various stakeholders including businesses, financial institutions, investors, regulatory bodies, auditors, and academics.

The first study presents a comprehensive survey, classification, and critical analysis of the literature on corporate bankruptcy and financial distress prediction. It covers definitions, prediction methodologies, data pre-processing, feature selection, model implementation, performance criteria, and evaluation methodologies, providing a critical analysis to inspire future research directions.

The second study introduces a novel approach to bankruptcy prediction by leveraging company relational information through complex network analysis. Using board of directors' networks, node embeddings and communities, the study devises new corporate governance drivers to enhance prediction accuracy. Empirical results from UK companies listed on the London Stock Exchange demonstrate that these network-based drivers significantly improve prediction performance.

The third study proposes a two-step methodology for bankruptcy prediction utilising companies' financial profiles. Initially, unsupervised cluster analysis generates group financial profiles from accounting information. Subsequently, supervised artificial intelligence models predict bankruptcy based on these profiles. This approach, tested on data from UK companies listed on the London Stock Exchange, shows improved prediction performance and enhanced model interpretability.

Together, these studies contribute to the advancement of bankruptcy prediction methodologies by integrating comprehensive literature analysis, innovative network-based drivers, and a two-step financial profile-based prediction approach.

Lay Summary

This thesis explores new ways to predict corporate bankruptcy and financial distress, which is crucial for business owners, banks, investors, and regulators who need to manage financial risks. Predicting corporate bankruptcy and financial distress is important because it helps prevent financial losses and ensures that companies, especially those in trouble, can get the support they need before it's too late. The research presented in this thesis is divided into three main studies, each offering a different approach to contribute to the current field of research:

The first study provides a detailed overview of existing methods used to predict bankruptcy and financial distress. It looks at how different types of financial data are used, how prediction models are built, and how their effectiveness is measured. This review helps to understand what has been done so far and where future research can make improvements.

The second study introduces a new way to predict bankruptcy by looking at how companies are connected through their board of director members. By analysing these connections as a network, the study identifies new drivers that can be used to predict financial risks. This approach shows that how companies are governed and how their directors are connected can give important clues about their financial health.

The third study focuses on using financial profiles, which are like financial 'fingerprints' of companies, to predict corporate bankruptcy. By grouping companies with similar financial characteristics together, the study creates profiles that make it easier to see which companies are at risk of bankruptcy. This method not only improves prediction performance, but also makes the results easier to understand.

Together, these studies show that combining traditional financial analysis with new techniques, such as complex network analysis and financial profiling, can lead to better ways of predicting when companies might face financial troubles. This research offers valuable insights that can help improve financial decision-making and risk management in the business world.

Acknowledgements

I am sincerely grateful to my supervisors, Professor Jamal Ouenniche and Dr Johannes De Smedt, for their exceptional guidance and whole-hearted support throughout my research journey. Your assistance was particularly invaluable during the hard time of the 2019 COVID pandemic, and, finally, I MADE IT!

I want to express my deepest gratitude to my family, especially my parents, for their unwavering support and encouragement throughout my PhD journey. Your dedication to raising and educating me has been the foundation of all my achievements.

My sincere thanks go to all the staff at the University of Edinburgh for their assistance and support over the years. Thank you!

I would like to extend my heartfelt appreciation to the ‘Dalmahoy Sauna Committee’, particularly Mr Charles Common, and David, Graeme, Paul, Ricky, Lewis, Fiona and James. The enjoyable times we shared and your support during my PhD study have been immensely appreciated.

I am also thankful to the Institute Faculty of Actuaries, the Scottish flying society, particularly Tayside Aviation and ACS Flight Training, as well as the Scottish golf industry, especially Dumbarnie Links, Cabot Highlands, and Kingsbarns Golf Links. Additionally, I appreciate the support from various organisations and individuals, especially, VisitScotland, The R&A, Sina.com, Augusta National Golf Club, Mr David Scott, Mr Andy Tseng, Mr Jinbin Ge and Mr Tony Zhou, for facilitating many of my professional activities.

Finally, I am deeply grateful to the Scottish people for their remarkable friendliness and help. Your hospitality has made my life in Scotland both enjoyable and unforgettable.

To all of my friends, thank you, without the support and encouragement from you, I would not be who I am today.

THANK YOU ALL!

Contents

1 General Introduction

1.1 Survey, Classification, and Critical Analysis of the Literature on Corporate Bankruptcy and Financial Distress Prediction ... 3

1.2 A Complex Network Analysis Approach to Bankruptcy Prediction using Company Relational Information-based Drivers ... 3

1.3 A Two-Step Approach to Bankruptcy Prediction using Companies' Financial Profiles ... 4

2 Survey, Classification and Critical Analysis of the Literature on Corporate Bankruptcy and Financial Distress Prediction

2.1 Introduction ... 5

2.2 Research streams in bankruptcy and financial distress prediction ... 8

2.3 Definitions of bankruptcy and financial distress ... 9

2.4 Design of prediction models or classifiers ... 9

2.4.1 Statistical & probabilistic models ... 9

2.4.2 Stochastic models ... 11

2.4.3 Artificial intelligence/machine learning models ... 12

2.5 Design of new drivers or evaluation of established drivers ... 23

2.6 Feature selection methodologies ... 25

2.7 Performance criteria and their measures, and performance evaluation methodologies ... 26

2.8 Data and markets of research ... 28

2.9 Critical analysis of the literature ... 29

2.9.1 Definition of bankruptcy and financial distress issue ... 29

2.9.2 Data imbalance issue ... 30

2.9.3 Outliers' issue ... 31

2.9.4 Cut-off point issue ... 31

2.10 Conclusions and future research directions ... 32

2.A Appendix ... 35

2.A.1 Classification of the definitions of corporate bankruptcy and financial distress in the BP and FPD surveyed literature ... 35

2.A.2 Classification of methodologies and models or methods in the BP/FDP literature 37
2.A.3 The drivers used in the BP and FPD surveyed literature 38
2.A.4 Classification of feature selection methods in BP and FDP.. 43
2.A.5 Classification of performance criteria & measures in the BP and FDP problem 45
2.A.6 Markets of analysis covered in the surveyed literature on BP & FDP 49
2.A.7 The databases used in current literature in BP and FDP study 50

3 A Complex Network Analysis Approach to Bankruptcy Prediction using Company Relational Information-based Drivers

3.1 Introduction 53
3.2 Literature review 55
3.2.1 Development of bankruptcy prediction models 55
3.2.2 Corporate Governance Indicators in bankruptcy prediction 57
3.2.3 Company relational information-based drivers in bankruptcy prediction..................... 59
3.3 Research questions 63
3.4 Methodology proposal 64
3.4.1 Complex company networks based on board members' connections........................... 64
3.4.2 Node embedding and design of board members' network-based drivers 68
3.5 Data and design of experiments 74
3.5.1 Data and its sources 74
3.5.2 Financial ratios and corporate governance indicators 76
3.5.3 Generic prediction model 76
3.5.4 Performance measures of classification models 78
3.5.5 Sampling strategy and experimental setup 78
3.6 Empirical Analysis and Findings 80
3.6.1 Analysis of results of Experimental Setup I 80
3.6.2 Analysis of results of Experimental Setup II 88
3.7 Conclusion and future research directions 92
3.A Appendix 94
3.A.1 Summary statistics of processed dataset 94
3.A.2 Results of CRIs using machine learning models for prediction with Experimental Setup I 95
3.A.3 An example of prediction results using CRIs generated by different community detection methods with Experimental Setup I 95
3.A.4 An example of statistical properties of successful and failed predictions........................ 96

3.A.5 Full results of Experimental Setup II 96

4 A Two-Step Approach to Bankruptcy Prediction using Companies' Financial Profiles

4.1 Introduction 98

4.2 Literature review 101

4.3 Research questions 106

4.4 Methodological proposal 108

4.5 Design of experiments 118

4.5.1 Data & experimental setup 118

4.5.2 Financial ratios & Corporate governance indicators 120

4.5.3 Implementation decisions of the two-step method 121

4.6 Empirical analysis and findings 126

4.6.1 Analysis of profiles 126

4.6.2 Analysis of results of profile-based prediction models using profile-based drivers 129

4.6.3 Analysis of results of profile-based prediction models based on profile-specific methodologies 130

4.6.4 Analysis of results of profile-based prediction models based on meta-learning structure 131

4.7 Conclusion and future research directions 138

4.A Appendix 141

4.A.1 The centroids of financial profiles 141

5 Discussion and General Conclusion

5.1 The evolution of bankruptcy prediction research 143

5.2 Enhancing bankruptcy prediction with network analysis 144

5.3 Advancements in financial profiling for bankruptcy prediction 145

5.4 General conclusions 146

5.5 Future research directions 148

Bibliography

Chapter 1

General Introduction

Corporate bankruptcy prediction and financial distress prediction have been an important field of research for nearly half a century, attracting significant attention, especially since the Global Financial Crisis of 2008. This field serves as a critical intersection of financial stability, regulatory oversight, and risk management, addressing challenges that affect a broad spectrum of economic activities. The ability to predict corporate bankruptcy and financial distress has crucial implications for various stakeholders, including businesses, financial institutions, investors, regulatory bodies, and academics, as these events carry prohibitive legal and financial costs and increase the expected costs for financial institutions to hedge against such risks.

Historically, the first modern study of bankruptcy prediction can be traced back to 1932, when Fitzpatrick (Fitzpatrick, 1932) successfully distinguished between failed and healthy companies by analysing their accounting ratios. Since the 1960s, a myriad of accounting-based statistical and probabilistic models has been developed to predict corporate bankruptcy and financial distress. Early influential works include those of Beaver (1966), Altman (1968), Ohlson (1980), and Zmijewski (1984). These studies laid the groundwork for future research by identifying key financial ratios and other relevant factors that could predict corporate bankruptcy and financial distress.

The field has evolved significantly since then, with researchers around the world continuously developing and refining methodologies to enhance prediction accuracy. Early studies primarily utilised accounting information-based drivers[1], commonly referred to as financial ratios (FRs). Over time, the scope of information has expanded to include audit information, corporate governance information, corporate social responsibility (CSR) / environment, social and

[1] In this thesis, we use the terms 'drivers', 'features', 'explanatory variables' and 'variables' as synonyms.

governance (ESG) information, market information, macroeconomic information, and more recently, textual information.

In recent years, the development of machine learning and artificial intelligence has revolutionised the field of bankruptcy and financial distress prediction. These technologies enable the analysis of vast amounts of data and the application of advanced computational techniques, leading to significant improvements in prediction accuracy. Machine learning approaches have also enabled the integration of multidimensional data sources, providing a more nuanced understanding of financial distress. This shift underscores the transformative potential of technology in addressing long-standing challenges in financial prediction. Modern research in this area has focused on various aspects, including the design of new prediction models, the design of new drivers, feature selection methodologies, and performance evaluation criteria and methods. Additionally, researchers have addressed challenges such as data sample imbalance and the computational requirements of prediction models, investigating various techniques such as over-sampling, under-sampling, bagging, and boosting.

The importance of this field has been amplified by its relevance to global economic stability and resilience. Recent global events such as Brexit, the escalation of global trade disputes, the COVID-19 pandemic, and the US presidential election have placed businesses in unprecedentedly challenging positions. These events underscore the need for robust prediction methodologies capable of navigating economic uncertainties and guiding stakeholders in mitigating financial risks. As businesses confront these challenges and seek new solutions to evaluate risks, the significance of predicting bankruptcy and financial distress has reached a new level.

This thesis seeks to address these critical challenges by presenting three interrelated studies that collectively advance the field of bankruptcy and financial distress prediction. These studies are unified by their focus on integrating innovative methodologies, leveraging advanced technologies, and addressing existing limitations in prediction, with each contributing uniquely to the field.

1.1 Survey, Classification, and Critical Analysis of the Literature on Corporate Bankruptcy and Financial Distress Prediction[2]

This study provides an up-to-date, comprehensive survey and critical analysis of existing methodologies and models of the literature on bankruptcy and financial distress predictions. It examines definitions of bankruptcy and financial distress, prediction methodologies, data preprocessing techniques, feature selection methods, and performance evaluation criteria and methods. By synthesising and critically analysing the extensive body of work in this domain, this study establishes a foundation for understanding the evolution of methodologies and their application in the face of global economic shocks, such as the 2008 financial crisis and the COVID-19 pandemic. This comprehensive study also highlights gaps and potential future research directions, offering insights into the current landscape of bankruptcy prediction research.

This research has been published in *Machine Learning with Application* (*15 (3):* 100527, DOI: 10.1016/j.mlwa.2024.100527).

1.2 A Complex Network Analysis Approach to Bankruptcy Prediction using Company Relational Information-based Drivers[3]

This study introduces a novel approach to bankruptcy prediction by leveraging company relational information through complex network analysis. It proposes new corporate governance drivers that exploit company relational information using board of directors' networks, the concept of node embeddings obtained by mapping large directors' networks to lower-dimensional spaces, and the concept of communities. This innovative methodology underscores the importance of relational corporate governance information, particularly in capturing complex interdependencies that traditional financial metrics might overlook, and different complex network structures that capture the diversity of connectivity patterns observed in directors' networks have been thoroughly investigated. Empirical results from UK companies listed on the London Stock Exchange demonstrate that these network-based drivers significantly improve prediction performance, highlighting the importance of relational corporate governance information in bankruptcy prediction.

[2] Zhao, J., Ouenniche, J., & De Smedt, J. (2024). Survey, classification and critical analysis of the literature on corporate bankruptcy and financial distress prediction. *Machine Learning with Applications*, *15(3)*, 100527.

[3] Zhao, J., Ouenniche, J., & De Smedt, J. (2024). A Complex Network Analysis Approach to Bankruptcy Prediction using Company Relational Information-based Drivers. *Knowledge-Based Systems*, *300(4)*, 112234.

This research has been published in *Knowledge-based Systems* (*300 (4):* 112234, DOI: 10.1016/j.knosys.2024.112234).

1.3 A Two-Step Approach to Bankruptcy Prediction using Companies' Financial Profiles

This study proposes a two-step methodology for bankruptcy prediction utilising companies' financial profiles to address the common issues of sample heterogeneity. In the first step, company accounting information is used to generate group financial profiles through unsupervised cluster analysis. These profiles encapsulate specific financial patterns that are common among certain subsets of companies, providing general descriptions of their characteristics. In the second step, supervised machine learning models predict bankruptcy based on these profiles, combined with other information such as financial and corporate governance information. This approach provides a novel framework for understanding how nuanced financial behaviours contribute to bankruptcy risks, enabling more tailored and actionable predictions. By combining clustering techniques with predictive modelling, this study demonstrates the potential of profiling techniques to improve both prediction accuracy and model interpretability.

The three studies are connected by a shared goal: to improve bankruptcy prediction through the integration of innovative drivers, advanced methodologies, and rigorous evaluations. The first study provides a critical overview of the field, setting the stage for the methodological advancements explored in the second and third studies. The second study introduces relational drivers that extend beyond traditional financial metrics, whilst the third study addresses data heterogeneity through financial profiling. Together, these studies form a cohesive narrative that advances the state of the art in bankruptcy prediction.

The contributions of this thesis are manifold. By synthesising the existing literature, introducing novel methodologies, and addressing persistent challenges in the field, the thesis provides a comprehensive resource for researchers, financial analysts, and policymakers, offering a detailed analysis of the current landscape of bankruptcy prediction research and paving the way for future innovations in this crucial field. By integrating complex network analysis and financial profiling techniques with traditional and modern prediction models, this thesis advances the understanding and accuracy of corporate bankruptcy and financial distress prediction, addressing the growing complexity and urgency of predicting these critical events.

Chapter 2

Survey, Classification and Critical Analysis of the Literature on Corporate Bankruptcy and Financial Distress Prediction

2.1 Introduction

Corporate bankruptcy prediction (BP) and financial distress prediction (FDP) have been a field of investigation by researchers around the world for nearly a half century. However, increasing attention has been paid to this evergreen subject since the Global Financial Crisis of 2008 (Alaminos et al., 2016), as predictions can have a significant impact on the decisions and returns of various stakeholders (Alam et al., 2021).

Corporate bankruptcy and financial distress events are not desirable (Gordon, 1971), as their legal and financial costs are prohibitive (Weiss, 1990). In addition, these events increase the expected costs for financial institutions, such as banks, to hedge against the risk of these events happening (Alnassar & Chin, 2015). To reduce exposure to risk and catch early warning signs, stakeholders including investors, bankers and governments are proactively looking for solutions to effectively analyse and predict corporate bankruptcy and financial distress events. The earliest study in terms of modern bankruptcy prediction can be traced back to 1932 when Fitzpatrick (1932) presented a successful way to distinguish between failed and healthy companies by analysing 20 pairs of companies' accounting ratios. Since the 1960s, several accounting-based statistical and probabilistic models have been proposed to predict corporate bankruptcy and financial distress (e.g., Beaver, 1966; Altman, 1968; Ohlson, 1980; Zmijewski, 1984; Zavgren, 1985). In addition, several studies focused on the identification of bankruptcy and financial distress drivers (e.g., Altman, 1968; Liang et al., 2016; Tobback et al., 2017; Tinoco et al., 2018; Mai et al., 2019), on one hand, and others focused on the design of methodologies for selecting such drivers (e.g., Tsai, 2009; Lin et al., 2014; Tian & Yu, 2017; Uthayakumar et al., 2020; Kou et al., 2021), on the other hand. Furthermore, a stream of literature investigated the causes of the under-performance of bankruptcy and financial distress prediction models such as data sample imbalance (e.g., Le et al., 2018; Veganzones & Séverin, 2018; Zoričák et al., 2020, Shen et al., 2020). In the era of the rise of machine learning and artificial intelligence technologies, the prediction of bankruptcy and financial distress has been lifted to a whole new level, as more and more new methodologies managed to improve prediction accuracy through the design features of new models (e.g., Ouenniche & Tone, 2017; Ouenniche et al., 2018a, 2018b, 2018c; 2019; Hosaka, 2019; Matin et al., 2019; Yuan et al., 2022), the type of implementation decisions such as over-sampling and under-sampling (e.g., Le et al., 2018; Sun et al., 2020; Shen et al., 2020; Du et al., 2020), and bagging and boosting (e.g., West et al., 2005; Zięba et al., 2016; Jones & Wang, 2019; Chen et al., 2020). On the other hand, another stream of research focused on reducing the computational requirements of prediction models or methods using parallel implementations that take advantage of today's powerful computers, or to be more specific, GPUs (e.g., Huang & Yen, 2019; Le et al., 2019).

Accurate prediction of bankruptcy and financial distress is already a crucial input to important decision-making processes in a normal economic climate, and even more so when the world is being subjected to important shocks. The year 2020 has witnessed drastic changes in the global political and economic environment, and events such as Brexit (Adler-Nissen et al., 2017), the escalation of global trade disputes (Salvatore, 2020), the COVID-19 pandemic (Fernandes, 2020) and the US presidential election (Baccini et al., 2021) have already put businesses either locally or globally into an unprecedented challenging position.

As businesses go through economic hardship and try to find new solutions to evaluate risks, the importance of predicting bankruptcy and financial distress takes a whole new level. In response to the growing complexity and urgency of predicting corporate bankruptcy and financial distress, this chapter offers a unique and timely contribution to the field by dedicating a comprehensive qualitative survey of the literature. The primary motivation behind our comprehensive survey stems from the need to synthesise and critically analyse the vast array of existing methodologies and models that have emerged in the wake of significant global socks, such as the 2008 financial crisis and the COVID-19 pandemic. Unlike previous reviews, our work not only provides several classifications of the literature based on different criteria but also provides an insightful critical analysis that highlights gaps and potential future directions. The significance of our work extends to a variety of stakeholders, including researchers, academics, financial analysts, and policymakers, by offering a comprehensive resource that paints the current landscape of bankruptcy prediction research and paves the way for future innovations in this crucial field.

Our survey uses a state-of-the-art review and critical analysis of the literature to summarise and classify the literature and critically analyse it. Our selection process of research studies on BP & FDP consists of identifying and retrieving papers from Web of Science (WoS), Scopus, Google Scholar, ResearchGate and SSRN databases using combinations of the following keywords: ‘Bankruptcy Prediction’, ‘Financial Distress Prediction’, ‘Default Prediction’, and ‘Business Failure Prediction’. In addition, we used the following filters: year of publication, document types, web of science categories, publications titles, publishers, and journal rankings. The papers were extracted in April 2020, followed by several monthly systematic updates until November 2022. Papers focussing on methodological innovations, new applications, or new applications in BP and FDP of existing methodologies were manually selected. The outcome comprised 293 documents published over the period 1966-2022 consisting of 249 published articles, 27 proceedings papers, 13 early access articles, and 4 book chapters. All documents are published in English. The 293 documents were then carefully analysed, resulting in a set of 149 most relevant papers to our survey. The other 144 documents discarded were out of the scope of our study (e.g., qualitative studies).

The remainder of this chapter unfolds as follows. Section 2.2 provides a classification of research on bankruptcy and financial distress prediction into the main research streams. Section 2.3 investigates the definitions and criteria of bankruptcy and financial distress. Section 2.4 provides a painting of the landscape of research on the design of prediction models or classifiers as applied to BP & FDP. Section 2.5 focuses on research on the design of new drivers or the evaluation of established drivers. Section 2.6 looks into feature selection methodologies. Section 2.7 covers the criteria and their measures for assessing the performance of prediction models as well as the methodologies for evaluating the predictive performance of classifiers. Section 2.8 provides an overview of the data and markets of research in bankruptcy and financial distress prediction studies. Section 2.9 provides a critical analysis of the literature, and finally, Section 2.10 summarises and concludes our study, as well as provides a list of potential future research directions.

2.2 Research streams in bankruptcy and financial distress prediction

Our survey of the literature on BP and FDP research revealed that there are mainly six active research streams: (1) definition and coding of bankruptcy and financial distress; (2) design of new prediction models/classifiers or new application of existing ones; (3) design of new drivers or evaluation of established drivers; (4) design or evaluation of feature selection methods; (5) design of methodologies for the performance evaluation of prediction models; and (6) issues affecting the performance of prediction models. **Table 2.1** provides a classification of the papers surveyed in each of these research streams. To each of these research streams, we dedicate a section where we provide relevant information unless a research stream is under-covered in the BP and FDP literature, in which case such research stream is covered in the critical analysis section.

Table 2.1 Research Streams and sample of papers in BP/FDP literature

Research Streams	Papers/Authors & Year
Definition and coding of bankruptcy and financial distress	All surveyed papers adopt one or more definitions of bankruptcy and/or financial distress
Design of new classifiers or new application of existing ones	Altman et al. (1977), Ohlson (1980), Taffler (1984), Zmijewski (1984), Zavgren (1985), Frydman et al. (1985), Odom & Sharda (1990), Tsukuda & Baba (1994), Bryant (1997), Shumway (2001), Sarkar & Sriram (2001), Paradi et al. (2004), Cielen et al. (2004), Shin et al. (2005), Sun & Shenoy (2007), Bharath & Shumway (2008), Ahn & Kim (2009), Min & Jeong (2009), Chen et al. (2009), Hu (2009), Li et al. (2009), Li & Sun (2009), Li & Ho (2009), Sueyoshi & Goto (2009a), Sueyoshi & Goto (2009b), Nanni & Lumini (2009), Yeh et al. (2010), Tseng & Hu (2010), Gorgani et al. (2010), Li et al. (2011), Hu & Chen (2011), De Andrés et al. (2011), Chen et al. (2011a), Tsai & Hsu (2013), Moradi et al. (2013), Lu et al. (2013), Cheng et al. (2014), Bauer & Agarwal (2014), Zięba et al. (2016), Afik et al. (2016), Cleofas-Sánchez et al. (2016), Sartori et al. (2016), Zhang & Hu (2016), Du Jardin (2016), Barboza et al. (2017), Volkov et al. (2017), Chou et al. (2017), Sun et al. (2017), Tobback et al. (2017), Wang & Wu (2017), Du Jardin (2017), Ouenniche & Tone (2017), Ouenniche et al. (2018a), Ouenniche et al. (2018b), Ouenniche et al. (2018c), Du Jardin (2018), Miao et al. (2018), Tinoco et al. (2018), Alaka et al. (2018), Oz & Simga-Mugan (2018), Choi et al. (2018), Ahmadi et al. (2018), Kim (2018), Le et al. (2019), Hosaka (2019), Mai et al. (2019), Ouenniche et al. (2019), Gupta & Chaudhry (2019), Huang & Yen (2019), Valencia et al. (2019), Qiu et al. (2020), Wang et al. (2020), Shen et al. (2020), Du et al. (2020), Alam et al. (2021), Almaskati et al. (2021), Du Jardin (2021a), Du Jardin (2021b), Kou et al. (2021), Yuan et al. (2022)
Design of new drivers or evaluation of established drivers	Beaver (1966), Altman (1968), Altman et al. (2008), Xu & Zhang (2009), Altman et al. (2010), Tinoco & Wilson (2013), Piñeiro-Sánchez et al. (2013), Tian et al. (2015), Lu et al. (2015), Geng et al. (2015), Liang et al. (2016), Zięba et al. (2016), Appiah & Chizema (2016) , Almamy et al. (2016), Antunes et al. (2017), Zorn et al. (2017), Lian (2017), Lin & Dong (2018), Tinoco et al. (2018), Andrikopoulos & Khorasgani (2018), Wang et al. (2018), Ninh et al. (2018), Serrano-Cinca et al. (2019), Li & Faff (2019), Jones & Wang (2019), Muñoz-Izquierdo et al. (2019), Ahmad (2019), Matin et al. (2019), Liang et al. (2020), Putra et al. (2020), Bai & Tian (2020), Ashraf et al. (2020), Li et al. (2021), Almaskati et al. (2021), Habermann & Fischer (2023), Kou et al. (2021)
Design or evaluation of feature selection methods	Tsai (2009), Lin et al. (2011a), Chen et al. (2011b), Lin et al. (2011b), Wang et al. (2014), Lin et al. (2014), Tian et al. (2015), Liang et al. (2015), Zhou et al. (2015), Pereira et al. (2016), Zelenkov et al. (2017), Tian & Yu (2017), Chou et al. (2017), Lin et al. (2019), Wang et al. (2020), Uthayakumar et al. (2020), Du et al. (2020), Paraschiv et al. (2021), Kou et al. (2021),
Design of methodologies for the performance evaluation of prediction models	Mousavi et al. (2015), Mousavi et al. (2019), Mousavi et al. (2023), Mousavi & Lin (2020)
Issues affecting the performance of models	Zmijewski (1984), Hsieh (1993), Tsai & Cheng (2012), Zhou (2013), Trabelsi et al. (2015), Gupta et al. (2015), Amendola et al. (2015), Kim et al. (2016), Cleofas-Sánchez et al. (2016), Tinoco et al. (2018), Veganzones & Séverin (2018), Le et al. (2018), Son et al. (2019), Huang & Yen (2019), Nyitrai & Virág (2019), Fernández-Gámez et al. (2020), Chen et al. (2020), Shen et al. (2020), Zoričák et al. (2020), Sun et al. (2020), Shen et al. (2020), Du et al. (2020), De Bock et al. (2020)

It is worth noting that several communities research into BP and FDP including business management, accounting, finance, operational research, and informatics. Unfortunately, researchers from different fields are not always aware of the evolving trends and methodologies across disciplines.

2.3 Definitions of bankruptcy and financial distress

In the BP and FDP literature, most research papers are concerned with addressing a classification problem; that is, classifying entities (e.g., companies) into pre-defined and unordered homogenous classes. Typically, the focus has been on two-class or binary classification problems where the main difference between papers lies in the definition of the risk classes, e.g., healthy vs. non-healthy, non-bankrupt vs. bankrupt, financially sound vs. financially distressed. Note, however, that the definitions of the risk classes adopted by researchers vary across papers. In fact, some authors adopt legal definitions of bankruptcy, others focus on financial distress, and others use hybrid definitions that combine both bankruptcy and financial distress criteria. We refer the reader to **Appendix 2.A.1** for our classification of definitions of corporate bankruptcy and financial distress in the bankruptcy and financial distress prediction literature surveyed in this chapter. It is worth mentioning that one of the conscious or unconscious motivations of researchers to opt for broad definitions of the risk classes is to increase the sample size of the risk class so that some prediction models (e.g., discriminant analysis models, logistic regression models) can be properly trained. In the critical analysis section, we shall point to some negative side effects of the choice of broad definitions of risk classes.

2.4 Design of prediction models or classifiers

The interdisciplinary overlap of the field of BP and FDP has seen a shift from univariate analysis to multivariate analysis, and progression from traditional statistical and probabilistic approaches to advanced artificial intelligence/machine learning-based techniques, from pure single classifier methods to hybrid single classifier methods and classifier ensemble methods, and from static classifiers/ stationary modelling to dynamic ones/dynamic modelling considering time process. The remainder of this section shall be organised into subsections, each focussing on a different category of models.

2.4.1 Statistical & probabilistic models

As BP and FDP is a long-standing topic in the academic literature, the actual first research could not be identified. Generally, it is widely recognised that the systematic study by Fitzpatrick (1932) laid the early groundwork. Fitzpatrick's study compared 13 accounting ratios of 20 paired failed and successful companies, marking a significant early effort in ratio analysis within bankruptcy research. This method was further refined by Beaver in 1966, who conducted a more systematic univariate financial ratio analysis. Beaver (1966) demonstrated that 30 financial ratios could effectively distinguish between bankrupt and healthy companies with a 90% accuracy one year prior to bankruptcy.

However, the theoretical landscape of BP underwent a big shift with the introduction of Altman's Z-Score model (Altman, 1968). This model marked a significant departure from the univariate approach of corporate bankruptcy analysis pioneered by Beaver and others. The Z-Score model is revolutionary in its use of *multivariate discriminant analysis* (MDA), fed with five financial ratios (i.e., *working capital/total assets*, *retained earnings/total assets*, *earnings before interest and taxes (EBIT)/total assets*, *the market value of equity/book value of total liabilities*, and *sales/total assets*), and the results indicated a prediction accuracy of 95%. As it is multi-faceted approach, it addressed the limitations and general scepticism that had surrounded traditional ratio analysis. He further enhanced his model with the introduction of the ZETA model in 1977 (Altman, 1977), which is a variant of the Z-score model where two more financial ratios (i.e., *normalised standard error of 10-year EBIT/total asset*s, *EBIT/total interest payments*) were added. This variant of the original Z-Score model not only maintained high accuracy for one-year bankruptcy prediction (i.e., 96%) but also extended its predictive capability to five years before bankruptcy with a 70% accuracy rate. The Z-Score and its variants started a new era in bankruptcy prediction, shifting the focus from traditional ratio analysis to a more comprehensive, multivariate approach. These models have had a profound and lasting impact, and continue to be widely utilised in the financial world today.

Even though there is empirical evidence that statistical models like MDA deliver a reasonably good performance, the underlying assumptions (i.e., linearity of the relationship between the response variable and the drivers of bankruptcy, classes are normally distributed and homoscedastic, residuals are statistically independent with mean zero) are rather restrictive in practice, as financial data can hardly satisfy these requirements. This limitation underscores the potential oversimplification of complex financial relationships when implementing MDA models. Therefore, new methods have been gradually proposed to increase the validity of the model on real-world datasets, which also allow for much relaxed model assumptions. For example, Ohlson (1980) proposed the *conditional logistic regression model*, or *logit* for short, which is a generalised linear model with a logit link function. This model achieved a 93% prediction accuracy within one year of bankruptcy, showing its strength in handling nonlinear relationships. Ohlson claimed that it is possible to identify a company at risk by analysing four factors, namely, the size of the company as measured by *total assets*, the financial structure of the company as measured by *total liabilities/total assets*, the liquidity of the company as measured by *working capital/total assets*, and *current liabilities/current assets*. Later, Zmijewski (1984) proposed the *probit model*, a generalised linear model with a *probit* link function, to predict the bankruptcy of American companies. This model achieved a 97.3% overall classification accuracy on paired samples. The capability of probabilistic modelling was again improved by Zavgren (1985) using the categories of drivers proposed by Pinches & Mingo (1973); namely, *the return on investment (total income/total capital)*, *capital turnover (sales/net plant)*, *inventory turnover (inventories/sales)*, *financial leverage (debt/total capital)*, *receivable turnover (receivables/inventories)*, *short-term liquidity (quick assets/current liabilities)* and *cash position (cash/total assets)*, which improved the model's long-term predictivity by allowing the model to make predictions five years ahead of the bankruptcy event and maintained a high classification accuracy of 95%. This method and its variants are widely used in practice and often benchmarked against (e.g., Almaskati et al., 2021).

2.4.2 Stochastic models

Apart from the static methodologies mentioned above, which do not take into account differences in company performance or risk profiles over time, some dynamic models were also adopted for BP and FDP. For example, Shumway (2001) introduced *discrete-time hazard modelling* for bankruptcy prediction and proved that it effectively captures the temporal dynamics of companies' performance. He compared the performance of the discrete-time hazard model to MDA models and *logit* models and proved its superiority on his dataset. Xu & Zhang (2009) proposed a new *X-score model*, which is a *panel logistic regression model* fed with accounting and stock market data along with the distance-to-default measure estimated with Black & Scholes (1973) & Merton (1974)'s *option pricing model* and proxies for bank dependence and Keiretsu (Flath, 2000) dependence, which are unique features of Japanese institutions, to evaluate the probability of bankruptcy for companies listed on the Tokyo Stock Exchange. The results suggested that their model successfully classified 72.4% of the delisted companies into high-risk categories. The strength of this method lies in its integration of diverse sources of data; however, it might be less generalisable outside the context of Japanese institutions. Miao et al. (2018) used the *multi-period logit model* fed with market data along with the *distance-to-default* measure, as estimated by the *Black-Scholes-Merton model* (Merton, 1974), where historical estimates of volatility and returns are replaced with market-implied measures of volatility and cost of capital. By incorporating both market-implied measures, their model achieved a bankruptcy prediction accuracy of 89%. Gupta & Chaudhry (2019) investigated the role of *value-at-risk* (VaR) and *expected shortfall* (ES) in aggravating companies' likelihood of experiencing financial distress using the *random effect panel logit model* fed with a variety of accounting data as well as *semi-parametric Cornish-Fisher VaR* (VaR_{CF}) (Cornish & Fisher, 1938) and *expected shortfall* measures. The results suggest that longer horizon (3-year and 5-year) tail risk measures contribute positively toward companies' likelihood of experiencing financial distress, and the proposed model achieved an area under the receiver operating characteristic curve (AUC) score of 92% in-sample and an AUC score of 91% out-of-sample in predicting financial distress for US companies.

2.4.3 Artificial intelligence / machine learning models

2.4.3.1 Artificial neural networks (ANNs) models

The 1990s marked a big leap in computational efficiency and artificial intelligence. As more powerful computing devices were developed (Tesler, 1991), new families of BP and FDP models were introduced. Odom & Sharda (1990) first applied *artificial neural networks* (ANNs) (Lippmann, 1987), or to be more specific, the *multi-layer perceptron* (MLP), to the prediction of corporate failure. The comparison with MDA suggested that MLP has good potential in bankruptcy prediction, as it correctly predicts 85.71% of holdout subsamples. Later, Tsukuda & Baba (1994) followed their lead and employed *back-propagation neural networks* (BPNN) with one hidden layer to successfully demonstrate the effectiveness of neural network models in bankruptcy prediction of Japanese companies, and achieved a Type-I error of 16.7% and a Type-II error of 13.0%. Although the theoretical foundation of ANNs was well established in the early 1990s, the massive implementation of network-based algorithms in bankruptcy and financial distress prediction started only in the early 2010s, as a result of the emerging high-performance GPUs and the advancements in deep neural networks – also known as *deep learning* (Schmidhuber, 2015). A number of studies (e.g., Zhou, 2013; Liang et al., 2016; Nyitrai & Virág, 2019) also compared the performance of MLP with other classes of prediction models, e.g., MDA, LR, *support vector machines* (SVM) and *decision trees* (DT), and the results all suggested that, on average, ANNs generally achieve good performance in terms of prediction accuracy. Apart from the multi-layer neural networks, other variants of ANNs-based classifiers, which had already shown good performance in other classification applications, were applied to the problem of bankruptcy and financial distress prediction. For example, Hosaka (2019) applied the '*GoogLeNet*' *convolutional neural network* (CNN) (Szegedy et al., 2015), which is used intensively in image classification, to the binary bankruptcy classification problem and achieved higher performance with an AUC score of 92% compared to DT, MDA, SVM, MLP and Altman's Z-score model. Mai et al. (2019) applied the *natural language* *processing* (NLP) model '*word2vec*' word embedding model (Mikolov et al., 2013) to create a hybrid model with CNN and MLP. This research provided evidence that textual disclosures add value to the accounting-based prediction model, as the AUC score for the model that combines the textual data and numerical data for a 1-year-ahead prediction reaches 85.6%, whilst that of using the numerical data only reaches 80.8%. Although methods based on ANNs generally demonstrate better capability in modelling complex patterns and capturing non-linear relationships, a significant drawback is their computational intensity, as these methods demand substantial data and computational resources, making them less suitable for smaller-scale applications. Additionally, their complexity can lead to issues such as overfitting and lack of interpretability, resulting in challenges in understanding and explaining the decision-making process of the models.

Other major families of statistical, machine learning and artificial intelligence methodologies with application in BP and FDP research include *Naïve Bayes classifiers* (NB), *case-based reasoning* (CBR), *support vector machines* (SVMs), *support vector data descriptions* (SVDDs), *decision trees* (DTs) and *random forests* (RFs), and *ensemble learning*. The latest additions to this list are *data envelopment analysis* (DEA) and *multi-criteria decision analysis* (MCDA) based classifiers.

2.4.3.2 Bayesian theory-based models

Naïve Bayes (NB) and its generalisation; namely, *Bayesian Networks* (BN), are probabilistic classifiers that model a set of variables and their conditional dependencies as a directed acyclic network, where nodes represent the variables, edges represent conditional dependencies between variables, and the edges' weights are conditional probabilities computed using Bayes' theorem. These conditional probabilities are then used to classify observations; to be more specific, the class belonging decision is made using the posterior probabilities of class nodes directly given the values of the vector of features, or indirectly using one or several thresholds of these posterior probabilities to classify observations. NB classifiers assume that all features are conditionally independent, whereas BN classifiers relax this assumption. For example, Sarkar & Sriram (2001), who first introduced *NB* and *BN* classifiers to the field of BP and FDP, relaxed the conditional independence assumption by partitioning the set of features into disjoint subsets of related features for the estimation of posterior probabilities – these subsets are referred to by the authors as *composite attributes* (CAs). These composite attributes are then exploited in computing posterior probabilities by considering such sets of features jointly. Their results suggested that the performance of the proposed BN, referred to as CA-BN, is superior to that of the NB and *C4.5 decision tree* with a classification accuracy of 92.5% for financial distress of banks. Another interesting contribution involving BNs is by Sun & Shenoy (2007) who proposed *cascaded Naïve Bayes* to address the issue of missing values. To be more specific, the authors used a correlation- and partial correlation-based feature selection method to reduce the size of the network to include only drivers referred to as first-order drivers. Then, to compensate for the missing values among first-order variables, second-order variables are identified, i.e., variables that have significant correlations with first-order variables and, therefore, are expected to provide information on the missing values of first-order variables. Their empirical results suggest that their proposal is effective in predicting corporate bankruptcy. Whilst offering flexibility and simplicity, the Bayesian theory-based models, particularly the BNs, can suffer from increased complexity and computational demands. *Naïve Bayes,* although simpler in its implementation, can sometimes oversimplify data relationships due to its assumption of feature independence, potentially limiting its effectiveness on more complex datasets. In the context of bankruptcy prediction, this limitation is particularly relevant as financial data often exhibit strong interdependencies between features, such oversimplification may lead to suboptimal predictions, particularly in scenarios where nuanced relationships between variables play a critical role. For example, a company's solvency risk may depend simultaneously on cash flow and debt levels, making an independence assumption inadequate. These limitations suggests that future applications should focus on further refining these hybrid approaches and incorporating advanced feature selection techniques to mitigate the independence assumption's impact. The reader is referred to **Appendix 2.A.2** for a complete list of references on the NB and BN models in BP and FDP.

2.4.3.3 Case-based reasoning (CBR) models

Case-based reasoning (CBR) was extensively used, as this method is non-parametric and resembles the way humans make decisions; that is, predicting the new case by reusing the historical knowledge in the case base. CBR was first used for bankruptcy prediction by Bryant (1997), however, due to the problems of inefficient feature selection and small sample sizes of the bankrupt companies, the result of this study did not suggest that CBR models perform better than Ohlson (1980)'s logistic regression model. To improve the performance of CBR models, various strategies (e.g., using only a subset of features that prove to be relevant to the target concept obtained, for example, with a feature selection method; using a feature weighting scheme that reflects the relative importance of features; using only relevant instances for reference) were implemented either independently or jointly. For example, Ahn & Kim (2009) embedded CBR into a *genetic algorithm* (GA) which simultaneously optimises feature weighting and instance selection on the training sample; once the stopping criterion of GA is met, the optimal parameters are used to classify instances in the test sample. The proposed *global optimisation of feature weighting and instance selection using GA for CBR* (GOCBR) achieved 86.73% of accuracy in the given holdout data, which improves the prediction accuracy of typical CBR systems by about 6%. Li & Ho (2009) embedded *Fuzzy-CBR* into a *genetic algorithm* which optimises the weight vector of features whose values are expressed in *linguistic terms* (fuzzy terms) by the expert, and the proposed Fuzzy-CBR with GA weighted model achieved 92.36% classification accuracy on a database that contains 746 publicly traded Taiwanese corporations. Chen et al. (2011a) embedded adaptive Fuzzy-CBR into a *particle swarm optimisation* (PSO) algorithm with *time-varying acceleration coefficients* and *time-varying inertia weight* to efficiently control the local and global search ability of PSO, which optimises the neighbourhood size k, the fuzzy strength parameter and the most discriminative subset of features, and claimed that their implementation of CBR, referred to as PTV-PSO-FKNN, produced the best predictions of bankruptcy with a classification accuracy of 92.36% compared to SVM, *basic k-nearest neighbour* (k-NN), BPNN, *probabilistic neural network* (PNN) and *extreme learning machine* (ELM) on a database of Polish companies. Li et al. (2009) exploited the concept of pseudo-criterion and its indifference, preference, and veto thresholds, commonly used by ELECTRE *outranking methods*, to devise a similarity measure based on concordance and discordance indices on each feature to use in implementing k-NN as a retrieval mechanism. They claimed that their implementation of CBR, referred to as OR-CBR, outperforms MDA, Logit, NN, SVM, DT, basic CBR and grey CBR with a leave-one-out accuracy of 91.5% in predicting financial distress of Chinese listed companies. In order to reduce the computational requirements of CBR due to the requirement to compute distances between queries and all cases, on the one hand, and avoid the problem of selecting the parameter k representing the number of nearest neighbours to vote for the class belonging of a query, on the other hand, Li et al. (2011) restricted the relevant instances for reference to the *Ideal Positive and Ideal Negative Cases* in the spirit of TOPSIS and claimed that their implementation of CBR, referred to as SPNIC-based CBR, produced the best predictions of business failure in China with an accuracy of 89.69% compared to MDA, Logit, Probit, CBR with k-NN and CBR with DT methods. The reader is referred to **Appendix 2.A.2** for a complete list of references on CBR models in BP and FDP. Although CBR methods excel in leveraging historical knowledge and adapting to varied data sets, which offer high accuracy in certain implementations, they often face challenges in feature selection and computational complexity. Whilst some variants are able to reduce computational demands, this can sometimes come at the cost of overlooking important case variations, highlighting a trade-off between simplicity and comprehensiveness in these models.

2.4.3.4 Support vector machines (SVMs) methodologies

Another class of non-parametric classifiers, *support vector machines* (SVMs), is also widely adopted in BP and FDP (e.g., Zhou, 2013; Huang & Yen, 2019) – see **Appendix 2.A.2** for a complete list of references on SVM models in BP and FDP. SVMs were first proposed by Vapnik as a class of supervised machine learning models and further developed with his team (Boser et al., 1992), and their robustness in handling high-dimensional data and the capability of handling nonlinear relationships are notable strengths. These classifiers use hyperplanes to separate data instances in different classes, where the parameters of the hyperplane are optimised so as to maximise the margin; i.e., the distance from the hyperplane to the nearest data points on each side of the hyperplane or in each class. These classifiers are commonly referred to as hard-margin linear SVMs and have been extended to soft-margin linear SVMs by relaxing the margin to handle noisy class boundaries in the data. Both hard- and soft-margin linear SVMs assume that data are linearly separable. This assumption has then been relaxed leading to nonlinear SVMs which employ *kernel functions* and have been used in BP and FDP. The introduction of kernel functions allows SVMs to map data from their original feature space into a higher-dimensional space where linear separation becomes possible, effectively enabling SVMs to capture complex, nonlinear relationships in the data. For example, financial ratios or corporate governance measures may not exhibit linear separability in their raw forms but can be better classified when transformed into higher-dimensional spaces using kernel functions such as the *radial basis function* (RBF). However, this flexibility comes at a cost, as kernel-based SVMs are computationally intensive, especially for large datasets or datasets with high dimensionality. This computational demand arises from the need to calculate and store kernel matrices, which grow quadratically with the number of data points. Despite this limitation, the use of kernels is essential in the context of BP and FDP because financial datasets often exhibit nonlinear relationships among features, such as interaction effects between liquidity and leverage, which linear models struggle to address effectively.

The flexibility of SVMs, especially with kernel functions, allows the prediction models to adapt to various nonlinear characteristics of the data, for example, Shin et al. (2005) first introduced nonlinear SVMs to bankruptcy prediction with the *radial basis function* (RBF) as the kernel function, which achieved a better classification accuracy on several datasets ranging from 66.4% to 87.9% compared to *back-propagation neural networks*. Since then, SVMs have gained increasing attention in BP and FDP research, as they demonstrated better prediction performance and generalisation capability. For example, Tobback et al. (2017) proposed a linear kernel SVM fed with a score computed with the relational information learner, the *weighted vote relational neighbour algorithm* (wvRN), to predict corporate bankruptcy for small and medium-sized enterprises (SMEs) in Belgium and the UK. The results suggested that AUC performance increased from 82.86% to 84.71% in the Belgium dataset and from 81.29% to 82.68% in the UK dataset, respectively. Some advantages of SVMs have been pointed out in several papers. For example, using a nonlinear SVM with RBF kernel function, Veganzones & Séverin (2018) found that their SVM model is least prone to the influence of largely imbalanced datasets (i.e., bankrupt companies only represent a small proportion of the overall sample) achieving the highest AUC score of 87.1% on a French bankruptcy research database under the scenario of 80/20 split of data (i.e., 80% non-bankrupt and 20% bankrupt companies) compared to LDA, LR, RFs and back-propagation MLP, and was still able to provide reasonable results even when the splits of 90/10 (AUC of 79.0%) and 95/05 (AUC of 76.4%) were used. A similar comparison has also been made by Zhou (2013), who claimed that a nonlinear *least-square* SVM with the *random under-sampling* method outperforms various prediction techniques including LR, ANN and *C4.5 decision tree*, in US and Japanese corporate bankruptcy prediction with an average AUC score of 84.67% on two largely imbalanced corporate bankruptcy datasets. In addition, Tsai & Cheng (2012) claimed that nonlinear SVM with RFB kernel demonstrates better noise tolerance than three other machine learning models (i.e., LR, C4.5 DT and MLP) in bankruptcy prediction, where outliers are identified using a *k-means*-based outlier detection method and then removed in different proportions. In their study, four bankruptcy databases, three of which were from the UCI machine learning repository (German credit, Australian credit, Japanese credit) and one from the UCSD Competition database, were used, and the results suggest that their SVM model achieves not only the highest average prediction accuracy across the four databases (77%) with 90% outlier removal but is also less affected by different levels of noise within the datasets.

Inspired by Vapnik's SVM model, another class of support vector methodologies, *support vector data description* (SVDD), was proposed by Tax & Duin (2004) as a novel one-class classification model. Instead of separating data points with hyperplanes, SVDD makes the classification by finding the minimum spherically shaped boundary that separates the target class samples from the samples in other classes. SVDD is widely used in outlier detection, however, due to the nature of datasets on bankrupt and financially distressed companies, which usually have a small proportion of bankrupt and financially distressed companies as compared to the much larger proportion of healthy companies, this method shows its strength and is thereby adopted by a handful of research papers in BP and FDP (e.g., Gorgani et al., 2010; Moradi et al., 2013; Yuan et al., 2022). For example, Moradi et al. (2013) applied SVDD with Gaussian kernel function to predict corporate financial distress using an Iranian database and the results suggested that SVDD performs better than an unsupervised machine learning model; namely, *fuzzy c-means clustering*, with a prediction accuracy of 91.9%, 85%, and 78% in the year of the financial distress event occurrence, one year before the event, and two years before the event, respectively.

Although SVMs and SVDD both offer robust solutions in BP and FDP due to their capability to handle high-dimensional and nonlinear data, their performance can be computationally demanding and requires careful fine-tuning, especially in selecting appropriate kernel functions. Attention has increasingly shifted toward optimising kernel computations through techniques such as approximate kernel methods and leveraging distributed computing frameworks, which have been employed to address scalability challenges when applying kernel-based SVMs to large financial datasets.

2.4.3.5 Decision trees (DT)

Decision trees for classification (DTs) are also generally adopted in BP and FDP studies (e.g., Tsai & Cheng, 2012; Zhou, 2013; Nyitrai & Virág, 2019) – see **Appendix 2.A.2** for a complete list of references on DT models in BP and FDP. DTs are non-parametric supervised learning methods that predict the value of a target variable by learning simple decision rules inferred from the data features using a *recursive partitioning algorithm*, their intuitive nature and ease of interpretation make them a popular choice among researchers. Frydman et al. (1985) were the first to use DTs in bankruptcy prediction. Their results suggested that the proposed model performs better than *discriminative analysis* (DA) models in terms of misclassification costs on actual, cross-validated, and bootstrapped validation sets of a customised American bankruptcy dataset. Apart from the standard classification tree algorithm and its successor *C4.5* (e.g., Tsai & Cheng, 2012; Zhou, 2013; Choi et al. 2018), models such as *classification and regression tree* (CART) (e.g., Tsai & Hsu, 2013; Du Jardin, 2018; Nyitrai & Virág, 2019, Liang et al., 2016) and *chi-square automatic interaction detection tree* (CHAID) (e.g., Serrano-Cinca et al., 2019, Nyitrai & Virág, 2019) were also adopted by some other researchers. These variations of DTs offer flexibility in modelling and are good at handling different types of data. Although DTs are usually very competitive in-sample, they tend to overfit and lack generalisation capability out-of-sample (Kamber et al., 1997; Myles et al., 2004; Bramer, 2007). To alleviate this specific issue, many researchers claimed that by incorporating single decision trees into a *random forest* (RF), the classification performance of the single tree model could be significantly improved (e.g., Volkov et al., 2017; Barboza et al., 2017; Veganzones & Séverin, 2018). RFs were amongst the first ensemble models to be introduced in BP and FDP. Alam et al. (2021) compared a group of machine learning models (i.e., SVM, J48 DT, *logistic model tree*, RF and *decision forest*) in bankruptcy prediction and the results suggested that RF as well as its enhanced version of error reduction, *decision forest*, lead to the highest prediction accuracy of 98.9% and 99.0%, respectively, on the validation set of a SMOTE processed Polish companies' dataset, whilst the others only achieve an accuracy around 93%. On the other hand, Shen et al. (2020) proposed a new dynamic financial distress prediction modelling framework named *adaptive neighbour SMOTE-recursive ensemble approach* (ANS-REA), which is capable of handling multiple unbalanced data streams, and implemented such framework using a variety of classifiers including RFs. The authors claimed that the performance of RF-based ANS-REA is significantly better (average AUC of 91.38%) than their counterparts based on single DT and bagging-DT as well as other supervised learning models (i.e., oblique RF, SVM, *Bayesian model*, and *kernel ridge regression*) in predicting financial distress of Chinese companies.

Although DTs and RF demonstrate better model interpretability, which allows for easy understanding of the decision-making process, both can be computationally intensive, requiring careful tuning of parameters for optimal performance.

2.4.3.6 Ensemble learning models

Ensemble learning is an umbrella term for methods that combine multiple predictions through various voting and aggregating schema, where each prediction is devised by a different model or method referred to as a base learner. Since first introduced in the late 1970s by Dasarathy & Sheela (1979), various ensemble learning models have been proposed over the years; e.g., *stacked generalisation* (Wolpert, 1992); *bootstrap aggregation* or *bagging* for short (Breiman, 1996); *Adaptive Boosting* or *AdaBoost* for short (Freund et al., 1999); *gradient boosting* (Friedman, 2001) and *extreme gradient boosting* or *XGBoost* for short (Chen & Guestrin, 2016). Generally, the performance of ensemble learning models is better than that of a single learner due to their better global optimisation capabilities, feature representation, and overfitting avoidance (Sagi & Rokach, 2018). *Ensemble learning* models received wide popularity since their introduction by West et al. (2005) to BP and FDP research, where they investigated three ensemble strategies, namely, *cross-validation*, *bagging* and *boosting*, with MLP as the base learner for credit scoring and bankruptcy prediction on three different datasets. The results suggested that the ensemble strategies reduce the generalisation errors estimated by a single MLP model by 3-5%. Later, Barboza et al. (2017) compared the predictive performance of *RF* (with *classification tree* as the base learner), *AdaBoost* (with *classification tree* as the base learner) and *bagging* (with *classification tree* as the base learner) ensemble models on a North American corporate bankruptcy dataset and benchmarked them against MDA, LR, SVM-linear, SVM-RBF, and MLP. The results suggested that *RF*, *AdaBoost* and *bagging* achieved the highest AUC out-of-sample of 92.92%, 92.97% and 92.48%, respectively, compared to the benchmark single classifier models. However, these strengths also come with increased computational demands and complexity in tuning and implementation. Kim (2018) proposed a *stacking ensemble model* to predict the financial distress of the hospitality sector in the US. In this research, *polynomial kernel SVM*, *back-propagation multi-layer perception neural network* and *J4.8 decision tree* are selected as the level 0 base learners, and the outputs of the base learners are further processed by the level 1 SVM meta learner to discover their best combination for classification. The proposed stacking model achieves overall classification accuracy of 90.97%, 95.57% and 97.82% in three different sectors of sample companies (i.e., restaurants, hotels, and amusement companies), respectively. Le et al. (2018) proposed a *cluster-based boosting* (CBoost) model with the *instance hardness threshold* feature selection algorithm to classify a strongly imbalanced Korean bankruptcy dataset. The CBoost model is largely based on *AdaBoost*, however, as it allows for weight initialisation for each data point with the *k-mean clustering* algorithm, it is able to effectively handle the class imbalance problem by increasing the weight value of data samples in the minority class. The proposed model achieves an AUC of 86.8% and outperforms several other machine learning models, i.e., modified AdaBoost (GMBoost), MLP, DT and RF. Jones & Wang (2019) adopted the *TreeNet gradient-boosting decision tree* for bankruptcy prediction and extended the binary bankruptcy classification into more complex multi-class settings with up to five states of failure (i.e., active firms, active firms in default, firms in bankruptcy or liquidation proceeding, firms dissolved through bankruptcy or liquidation, and firms dissolved for reasons other than bankruptcy). The three-state (i.e., active firms; active firms that are in default or subject to an insolvency proceeding; firms in bankruptcy or liquidation process or firms dissolved through bankruptcy or liquidation) and five-state models achieved the highest out-of-sample AUC scores of 95.1% and 91.2%, respectively, for one year before the bankruptcy event. Huang & Yen (2019) claimed that *extreme gradient boosting* (XGBoost), which is a scalable *gradient boosted decision tree* that allows for penalty regularisation and feature subsampling, achieved high performance in classification on a Taiwanese financial distress dataset with the highest prediction accuracy of 90.6% across seven different combinations of spans of quarterly data prior to the failure event. Chen et al. (2020) developed *boosting-SVM* (*Boosted-pSVM*) and *bagging-SVM* (*Bagged-pSVM*) based on *proportion support vector machine* (*pSVM*) proposed by Yu et al. (2013). They claimed that the proposed models addressed the issue of conducting semi-supervised training with only a proportion of instances being labelled, and the proposed ensemble methods with RBF kernel yield better performance with 91.67% classification accuracy. Sun et al. (2020) proposed a model combining *synthetic minority oversampling technique* (SMOTE), *AdaBoost-SVM*, and *time weighting* (SMOTE-ADASVM-TW). This ensemble model addressed the problem of how to effectively construct dynamic financial distress prediction models that can handle several imbalanced time-related data streams and the results showed that this model achieved good performance with a 91.22% overall prediction accuracy. We refer the reader to **Appendix 2.A.2** for a complete list of references on ensemble learning models.

2.4.3.7 Data envelopment analysis (DEA) models

Classifiers based on *Data Envelopment Analysis* (DEA) are also adopted in bankruptcy and financial distress prediction by a group of researchers (e.g., Premachandra et al., 2009; Yeh et al., 2010; Ouenniche & Tone, 2017) – see **Appendix 2.A.2** for a complete list of references on DEA models in BP and FDP. In the classification context, DEA could be considered as a special kind of nonparametric machine learning methodology that computes multi-criteria scores for observations with an efficiency meaning, which can offer a unique perspective in assessing company performance. The first papers on DEA in bankruptcy were concerned with bankruptcy assessment rather than bankruptcy prediction; in fact, Paradi et al. (2004), Cielen et al. (2004), and Premachandra et al. (2009) were all concerned with assessing the risk profiles of companies using implicitly or explicitly the concept of *worst-practice frontier* introduced by Paradi et al. (2004). Later, Sueyoshi & Goto (2009a,2009b) proposed a *DEA-discriminant analysis classifier*, combining DEA's efficiency scoring with discriminant analysis for enhanced predictive capability. Ouenniche & Tone (2017) proposed a classifier where a new DEA-based classifier is used for in-sample predictions and a CBR-based classifier trained on the class predictions provided by the DEA-based classifier is used for out-of-sample predictions. The classifier proposed by Ouenniche & Tone (2017) was tested on a dataset of companies listed on the London Stock Exchange and achieved a very high performance (close to 100% on Type I error, Type II errors, Sensitivity, and Specificity) similar to the ones achieved by Ouenniche et al. (2018a, 2018b, 2018c; 2019).

Despite the promising performance of DEA-based classifiers, this category of methods requires expert knowledge in DEA, which is not common among researchers in predictive analytics.

2.4.3.8 Multi-criteria decision analysis (MCDA) models

In recent years, a new family of *multi-criteria decision analysis* (MCDA) classifiers was proposed and used in bankruptcy and financial distress prediction (e.g., Hu & Chen, 2011; Li et al., 2011; Ouenniche et al., 2018a) – see **Appendix 2.A.2** for a complete list of references on MCDA classifiers. MCDA classifiers excel at handling multiple criteria simultaneously, providing a comprehensive assessment of bankruptcy risk. Ouenniche et al. proposed a hybrid classifier referred to as an integrated in-sample and out-of-sample prediction framework, where new MCDA-based classifiers (i.e., TOPSIS, DRPM, VIKOR, EDAS) are used for in-sample predictions, and a CBR-based classifier trained on the class predictions provided by the MCDA-based classifier is used for out-of-sample predictions (e.g., Ouenniche et al., 2018a, 2018b, 2018c; 2019); these classifiers were tested on a dataset of companies listed on the London Stock Exchange and achieved a very high performance (close to 100% on Type I error, Type II errors, Sensitivity and Specificity).

Although MCDA-based classifiers offer depth and comprehensiveness in their analysis along with promising performance, the choice of their parameters requires some effort put into it and thus requires expert knowledge in MCDA.

The methodologies for solving BP and FDP problems keep evolving – see **Table 2.2** for the key milestones for corporate bankruptcy and financial distress prediction models. As many advancements are made in other fields of predictive analysis (i.e., business management, operational research, and informatics), more and more novel and sophisticated methodologies, as well as prediction frameworks, are expected to be seen in the field of BP and FDP.

Table 2.2 Key milestones for corporate bankruptcy and financial distress prediction models

Year	Author(s)	Major Contribution
1932	Fitzpatrick	First systematic modern corporate failure prediction research.
1966	Beaver	First systematic univariate discriminant ratio analysis in bankruptcy prediction.
1968	Altman	First multivariate discriminant analysis use in bankruptcy prediction & development of the Z-score model.
1980	Ohlson	First use of logistic regression in bankruptcy prediction.
1984	Zmijewski	First use of Probit in bankruptcy prediction.
1985	Frydman et al.	First use of classification trees (recursive partitioning) in bankruptcy prediction.

1990	Odom & Sharda	First use of artificial neural networks (multi-layer perceptron) in bankruptcy prediction.
1994	Tsukuda & Baba	First use of back-propagation neural networks in bankruptcy prediction.
1997	Bryant	First use case-based reasoning in bankruptcy prediction.
2001	Shumway	First use of hazard modelling / multi-period logit in bankruptcy prediction.
2001	Sarkar & Sriram	First use of Naïve Bayesian networks in financial distress prediction
2005	Shin et al.	First use of support vector machines bankruptcy prediction.
2005	West et al.	First use of ensemble classifiers (MLP ensembles) in bankruptcy prediction.
2009	Ahn & Kim	First use of metaheuristics (genetic algorithm) to simultaneously optimise feature weighting and instance selection in case-based reasoning for bankruptcy prediction.
2009	Li & Ho	First use of fuzzy-CBR bankruptcy prediction.
2011	De Andrés et al.	First use of multivariate adaptive regression splines in bankruptcy prediction.
2013	Tsai & Hsu	First use of stacked generalisation or stacking to design a classification framework in bankruptcy prediction, where the aim is to filter out unrepresentative training data, i.e., less noisy class labels
2016	Zięba et al.	First use of extreme gradient boosting (XGBoost) in bankruptcy prediction.
2017	Ouenniche & Tone	First DEA-based classification framework for in-sample and out-of-sample bankruptcy prediction.
2018	Ouenniche et al.	First MCDA-based classification framework for in-sample and out-of-sample bankruptcy prediction.
2019	Hosaka	First use of convolutional neural network in bankruptcy prediction.
2019	Mai et al.	First use of natural language processing (Word2Vec) in bankruptcy prediction.

2.5 Design of new drivers or evaluation of established drivers

The choice or selection of drivers with which the BP and FPD models are fed plays an important role in their predictive performance. In the academic literature, drivers are either prespecified by the researcher (e.g., Ohlson, 1980; Cielen et al., 2004; Li & Ho, 2009) or selected with a feature selection method (e.g., Tsai, 2009; Kim et al., 2016; Paraschiv et al., 2021). Our survey of the academic literature on drivers used in BP and FDP revealed that such drivers fall within seven main categories: *accounting information-related drivers*, *audit information-related drivers*, *corporate governance-related drivers*, *corporate social responsibility-related drivers*, *market information-related drivers*, *macroeconomic information-related drivers*, and *media information-related drivers* – see **Appendix 2.A.3** for details on the drivers used in previous research within each of these categories. Note however that the conventional information with which prediction models are fed is **accounting information** and remains the dominant category. This category of information focuses on the financial health and performance of a company as reflected in its accounting records, it can also be further refined into six subcategories; namely, *asset utilisation measurements*, *operational performance measurements*, *cashflow measurements*, *liquidity measurements*, *capital structure & solvency measurements*, and *return on investment measurements*. As accounting information could eventually be ‘manipulated’, **audit information**, which is derived from audit reports and focuses on the evaluation of a company’s financial statements, could be used to ease off this issue, as auditors would investigate any lack of compliance with the accounting standards (Lennox, 1999); however, only very few studies supplemented accounting information with audit information (e.g., Altman et al., 2010; Piñeiro-Sánchez et al., 2013; Matin et al., 2019; Muñoz-Izquierdo et al., 2019). Although accounting information, as conveyed by accounting or financial ratios, is an important source of information, it does not take into account external factors related to the market and the economy. Therefore, over time, researchers have adopted additional drivers, such as **market information** (e.g., *stock return*, *market capitalisation*), which are based on market perceptions and reactions to a company's general performance, and **macro-economic information** (e.g., *GDP growth rate*, *inflation rate*, *unemployment rate*), which involves drivers related to the broader economic environment in which a company operates. In addition, as the concept of **corporate social responsibility (CSR)** developed along with its categories of criteria and their rebranding into **economic, social and governance (ESG)** criteria are increasingly being taken into account by investors. This information reflects the commitment and performance of a company in areas such as environmental sustainability, social responsibility, and ethical practices. It often includes measures of environmental impact, community engagement, and labour practices. Some BP and FDP studies complemented accounting information with ESG measures and concluded that companies with a higher prior history of positive CSR engagement are less likely to go bankrupt (e.g., Lin & Dong, 2018; Habermann & Fischer, 2023). Of particular interest to a variety of stakeholders including shareholders, **corporate governance**, which focuses on the management and organisational structure of a company, has been increasingly scrutinised and thus several studies (e.g., Appiah & Chizema, 2016; Ahmad, 2019; Liang et al., 2020) supplemented accounting information with corporate governance information (i.e., *the proportion of outsider directors*, *CEO/Chair duality*, *board size*). As the scope of research is expanding, in recent years, **media information** from some unconventional sources has gradually been adopted by many researchers such as company reports information (e.g., Wang et al., 2018; Ahmadi et al., 2018; Hosaka, 2019; Mai et al., 2019), conventional media information such as financial media reports (e.g., Lu et al., 2013; Lu et al., 2015), and social media information such as Facebook feed (e.g., Putra et al., 2020).

Due to the fact that BP and FDP studies rest heavily on evaluating the ability of a company to meet its maturing obligations (Zavgren, 1985), it is understandable that drivers which can reflect the company's capital structure and liquidity, such as *total liabilities/total assets* (e.g., Ohlson, 1980; Bryant, 1997; Chen et al., 2009; Geng et al., 2015), *working capital/total assets* (e.g., Altman, 1968; Shumway, 2001; De Andrés et al., 2011; Zorn et al., 2017) and *current ratio* (*current assets/current liabilities*) (e.g., Zmijewski, 1984; Tsukuda & Baba, 1994; Tian et al., 2015; Le et al., 2019), are widely adopted in the literature. Furthermore, as companies need to generate sufficient income and produce sustainable profits to survive in the long run, it is also expected that drivers related to asset utilisation and profitability, such as *sales/total* *assets* (e.g., Odom & Sharda, 1990; Min & Jeong, 2009; Almamy et al., 2016; Nyitrai & Virág, 2019) and *net income/total assets* (e.g., Beaver, 1966; Yeh et al., 2010; Tobback et al., 2017), would also be widely adopted. Note however that many authors failed to provide any justification for their choice of drivers, on the one hand, and many others used drivers that look like 'accounting' drivers but do not make sense in the accounting context and failed to justify the rationale behind their choices – see category 'unclassified' in **Appendix 2.A.3**, on the other hand.

2.6 Feature selection methodologies

As there is no generally agreed upon set of drivers to use for BP and FDP, on the one hand, and there is an increasing diversity of information and its sources, on the other hand, automation of the feature selection process is desirable. The feature selection process is concerned with finding an optimal subset of features that can effectively predict the response variable. The dimensionality of the feature space is generally optimised using a multi-stage procedure where the typical stages are concerned with *subset generation*, *subset evaluation*, and *subset validation* along with *stopping criteria* (see for example, Liang et al., 2015). Using a feature selection process has several benefits, including improved interpretation and performance of prediction models, and reduction of model complexity and computational requirements (see, for example, Tsai, 2009). Reflecting the evolutionary trend in bankruptcy prediction methodologies, feature selection has also progressed from simpler single variable analysis to more realistic multivariate models. Early studies in BP and FDP used to conduct feature selection by utilising the domain knowledge of the experts (see, for example, Li & Ho, 2009), which results in most cases in selecting only those ratios that are widely accepted or tested. In recent years, technological progress along with the increased number of data sources and categories led researchers to increasingly focus on automatic feature selection methodologies.

Feature selection methodologies could be divided into two broad categories, namely, supervised and unsupervised methodologies. So far, most BP and FDP studies have focused on supervised methodologies. Supervised feature selection methods are further divided into three categories, namely, *filters*, *wrappers*, and *embedded methods*. *Filters* are concerned with selecting the most relevant features from a set of pre-determined features using common statistical techniques, and thus are sometimes referred to as relevance-based feature selection methods. Examples of these filters in BP and FDP include statistical tests such as *t-tests* (e.g., Min & Jeong, 2009; Liang et al., 2016; Mousavi et al., 2019), *Pearson correlation coefficient* (e.g., Sun & Shenoy, 2007; Appiah & Chizema, 2016; Paraschiv et al., 2021), and *principal component analysis* (e.g., Sueyoshi & Goto, 2009b; Lin et al., 2011b; Mousavi et al., 2019) – see **Appendix 2.A.4** for more examples of filters. Despite some advantages of filters such as their computational simplicity and mathematical tractability, they do not interact with the classification algorithm and thus generally produce weaker overall prediction performance (Lin et al., 2014). *Wrappers* are concerned with searching for the best feature subset that improves a prediction model performance as a whole, and thus are sometimes referred to as predictive accuracy-based feature selection methods. Often, these methods are designed so that the prediction model is embedded within a search method based on specific search strategies (e.g., complete or exhaustive search, heuristic search, nondeterministic search). Examples of wrappers in BP and FDP include *genetic algorithms* (e.g., Liang et al., 2015; Chou et al., 2017; Zelenkov et al., 2017), *particle swarm optimisation* (e.g., Chen et al., 2011a; Liang et al., 2015; Uthayakumar et al., 2020) and *sequential forward selection* (e.g., Zhou et al., 2015; Paraschiv et al., 2021) – see **Appendix 2.A.4** for more examples of wrappers. Note however that the design of wrappers depends on the nature of the prediction model or method and can be computationally expensive. Finally, *embedded methods* are concerned with incorporating feature selection as part of the process of building a specific prediction model for a specific application, and thus can be effective but could be computationally expensive depending on the computational requirements of specific strategies and prediction models. Examples of embedded methods for feature selection in BP and FDP include *decision trees* (e.g., Min & Jeong, 2009; Hosaka, 2019; Du et al., 2020), *least absolute shrinkage and selection operator* (LASSO) (e.g., Amendola et al., 2015; Volkov et al., 2017; Wang et al., 2018) and *ridge regression* (Pereira et al., 2016). We refer the reader to **Appendix 2.A.4** for our classification of feature selection methods in the BP and FDP literature surveyed in this chapter.

2.7 Performance criteria and their measures, and performance evaluation methodologies

Researchers in bankruptcy and financial distress prediction adopted a wide range of performance criteria and their measures – see **Appendix 2.A.5** for a summary of the performance criteria and their measures covered in the surveyed literature. There is no general consensus on which criteria and their measures are best to use for the performance evaluation of classifiers. Note however that, for bankruptcy and financial distress prediction of companies, the discriminatory power criterion and the correctness of categorical prediction criterion are more appropriate, especially for assessing the performance of machine learning two-class classifiers.

In addition to these criteria, it is also important to consider the context in which these classifiers are evaluated. This includes examining their performance across different paired group classifications, such as *training sampling vs. testing sampling*, *single industry vs. cross-industry sampling*, and *balanced vs. imbalanced sampling*. These distinctions are crucial, as they can significantly impact the effectiveness and generalisability of prediction models.

Measures used in the literature on BP and FDP to assess classifier performance can be classified into several categories depending on the classification criterion. In this chapter, we classify such measures into (1) *single objective-based vs. multiple objectives-based measures*, and (2) *local vs. global measures*, to provide some insight into these measures and their use in BP and FDP where actual datasets are always unbalanced. The classification of measures into single objective-based measures (e.g., Type I error, Type II error, Sensitivity, Specificity, Precision or Positive Predictive Value, Negative Predictive Value) vs. multiple objectives-based measures (e.g., Misclassification Rate or Cost, Accuracy) suggests, based on both empirical evidence and conceptual modelling, that the choice of multiple objectives-based measures for assessing the performance of classifiers will in general disadvantage such classifiers, as compared to single objective-based measures in that the optimisation of these measures results in a compromise solution, since the objectives are conflicting in nature, and a classifier would never properly classify both all positive and all negative cases. In fact, under a multiple objectives' performance measure, a classifier can achieve high performance simply by classifying all cases as negative cases since positive cases only represent a very small proportion of the total number of cases. Therefore, to avoid any performance bias due to the unbalanced nature of the datasets and related issues, we recommend the use of single objective-based measures for most application areas – at least at the model evaluation stage. On the other hand, the classification of measures into local vs. global measures suggests that one would ideally use measures from both categories, as the local measures focus on measuring the performance of a classifier with respect to a single cut-off point (e.g., Type I error, Type II error, Accuracy), whilst global measures focus on measuring the performance of a classifier with respect to a range of cut-off points and, most importantly, inform the analyst or researcher on how the performance of the classifier compares to the performance of a random classifier (e.g., AUC, Gini).

It is common that a given classifier performs very well under a certain performance measure but performs very poorly with respect to another (Seliya et al., 2009) – this issue could be resolved by using appropriate multi-criteria ranking frameworks. Unlike the traditional performance evaluation framework, where competing prediction models or classifiers are ranked based on a single measure of a single criterion, multicriteria performance evaluation frameworks were proposed. Following the lead of Xu & Ouenniche (2012a, 2012b) and Ouenniche et al. (2014a, 2014b) who first proposed several multicriteria frameworks for assessing the performance of forecasting models of continuous variables, Mousavi et al. (2015, 2019, 2023) and Mousavi & Lin (2020) proposed several multicriteria frameworks for assessing the performance of prediction models with applications in bankruptcy and financial distress prediction. To be more specific, Mousavi et al. (2015) proposed a super-efficiency DEA-based framework for ranking a variety of bankruptcy prediction models under multiple criteria. Then, Mousavi & Ouenniche (2018) proposed a slacks-based context-dependent DEA framework to evaluate competing distress prediction models as well as a hybrid cross-benchmarking-cross-efficiency framework as an alternative methodology for ranking entities (e.g., prediction models) that are heterogeneous. Finally, Mousavi et al. (2023) proposed a Malmquist-Data Envelopment Analysis multi-period performance evaluation framework for assessing competing static and dynamic statistical models to predict financial distress and using it to address a variety of research questions.

2.8 Data and markets of research

The majority of international publications on BP and FDP have focused on developed countries with developed economies, and only a small proportion of publications are related to developing economies. Some of the reasons behind this phenomenon could be summarised as follows. First, most international journals publishing research on BP and FDP are in English, and therefore publications in other languages might be underestimated (Dimitras et al., 1996; Alaka et al., 2018). Second, since BP and FDP research involves empirical testing and validation, the availability of complete and reliable data is crucial; therefore, researchers are more likely to conduct their experiments using data on companies based in developed markets. Last but not least, researchers tend to use well-maintained, uniformly organised databases developed by well-known data providers to avoid collecting and processing a large amount of data themselves; however, most of the high-quality commercial databases are generally focused on developed markets, and less attention has been given to the developing markets. **Appendix 2.A.6** provides a summary of the market of analysis covered in the surveyed literature on BP and FDP.

In terms of the data adopted by researchers for the BP and FDP studies, the literature surveyed based on the North American markets mainly employed the *COMPUSTAT* database, which is a comprehensive market and corporate finance database maintained by Standard and Poor's (e.g., Bryant, 1997; Bharath & Shumway, 2008; Lian, 2017; Li & Faff, 2019; Qiu et al., 2020). On the other hand, studies focussing on European markets mainly use *LSEG Datastream* (e.g., Tseng & Hu, 2010; Tinoco & Wilson, 2013; Andrikopoulos & Khorasgani, 2018) and *Moody's Bureau van Dijk* (e.g., De Andrés et al., 2011; Chen et al., 2011b; Bauer & Agarwal, 2014; Sartori et al., 2016). Several different sources of data are also mentioned in the surveyed literature, namely, commercial database providers, universities and research institutes, regulators & government agencies, commercial banks, financial journals, and stock exchanges. We provide a detailed summary of the databases used in the surveyed literature on BP and FDP studies in **Appendix 2.A.7**.

2.9 Critical analysis of the literature

In this section, we present a critical analysis of the literature on corporate bankruptcy and financial distress prediction around some major issues.

2.9.1 Definition of bankruptcy and financial distress issue

In practice, all corporations experience, at some point in time, some operating difficulties which may or may not result in financial distress. When these operating difficulties lead to some form of financial distress, the action taken by the relevant stakeholders to address this situation depends on the severity of the financial distress and could be either an internal solution to the problem or an external solution. However, when an internal solution route is chosen and proves unsuccessful, the stakeholders turn to an external solution which, broadly speaking, has two possible outcomes: resolving the issue or discontinuing operations. As this process suggests, distressed corporations do not necessarily end up filing for bankruptcy, and those that file for bankruptcy do not necessarily end up discontinuing their operations and ceasing to exist. The variety of possible trajectories that a corporation might go through in this process and the rather broad definitions adopted by researchers and partitioners alike typically result in heterogeneous classes in that the corporations that fall into each risk class have different risk profiles. Typically, this heterogeneity issue affects the empirical performance of prediction models or methods; attempts have been made by De Andrés et al. (2011) and Du Jardin (2016; 2017; 2021a; and 2021b) who tried to address this issue. Another weakness of some definitions of risk classes lies in the vagueness of the criteria, and in some cases, no definition is provided by the authors. Last, but not least, most authors do not justify their choices of the definitions they adopt.

2.9.2 Data imbalance issue

In the BP and FDP research, unhealthy companies only represent a small proportion of the overall sample resulting in imbalanced datasets of healthy and unhealthy companies, which may significantly degrade the performance of prediction models (Nyitrai & Virág., 2019). In fact, if the number of unhealthy companies is significantly lower than the number of healthy ones, it is worthwhile for the machine to simply classify all companies into the healthy category to achieve a higher rate of classification accuracy (Volkov et al., 2017). This challenge has catalysed an evolution in research methodologies, moving from traditional, simpler sampling techniques to more sophisticated AI-driven methods to balance the datasets.

Early studies in BP and FDP generally adopted *convenience-balanced sampling* to select and pair healthy and unhealthy companies according to their characteristics, e.g., industry and asset size (e.g., Beaver, 1966. Altman, 1968. Zavgren, 1985). However, this sample-match technique introduces bias into the distribution of the training data, which will lead to untrustworthy predictions of the test sample data (Cheng et al., 2014). To address these concerns, a number of systematic sampling approaches have been proposed to solve this problem. Examples include *over-sampling methods* such as *random over-sampling* (Zhou, 2013) and *synthetic minority over-sampling* (SMOT) (Cheng et al., 2014); *under-sampling methods* such as *random under-sampling* (Alam et al., 2021), *instance hardness threshold* (IHT) (Le et al., 2018) and *clustering-based under-sampling* (CUS) (Du et al., 2020). Nowadays, these sampling methods are viewed as viable approaches to alleviate this issue by reproducing artificially synthesised samples of the minority class (or algorithmically reducing oversized samples of the majority class) until the classes are almost equally distributed.

Over-sampling methods, particularly synthetic approaches like SMOTE, are widely adopted due to their ability to retain the full diversity of the original majority class whilst addressing class imbalance by generating synthetic minority samples. This is especially relevant in the context of BP and FDP, where the minority class often contains critical patterns that, if lost, could significantly impair prediction performance. However, over-sampling can affect feature importance by altering the distribution of data, potentially introducing noise or redundancy. Despite this, over-sampling is considered appropriate for BP and FDP because the alternative, under-sampling, risks discarding valuable information from the majority class, which is detrimental in high-dimensional financial datasets. Recent studies (e.g., Sun et al., 2020; Shen et al., 2020) have shown that over-sampling methods such as SMOTE can improve both sensitivity and overall model performance without disproportionately skewing feature importance, particularly when combined with robust feature selection techniques that mitigate potential noise. Thus, the general adoption of over-sampling in BP and FDP research reflects a trade-off that prioritises preserving minority class information whilst leveraging advanced feature engineering to address its potential drawbacks.

2.9.3 Outliers' issue

In general, outliers refer to observations with extreme values that usually lie away from the rest of the data. In the BP and FDP literature, the authors had different views on outliers. In fact, some authors simply ignored this issue. However, other authors used some techniques to address the issue, such as *winsorisation* (e.g., Bharath & Shumway, 2008; Gupta & Chaudhry, 2019; Lian, 2017), *logarithm transformation* (e.g., Altman et al., 1977; Habermann & Fischer, 2023), *tangent hyperbolic transformation* (Tinoco et al., 2018), *pruning* (Cielen et al., 2004), and *clustering* (Tsai & Cheng, 2012). Finally, some authors treated unhealthy companies as outliers that can be detected and classified through the implementation of outlier detection methodologies (e.g., Gorgani et al., 2010; Moradi et al., 2013; Yuan et al., 2022).

2.9.4 Cut-off point issue

For many bankruptcy and financial distress prediction models (e.g., MDA, LR, NB), the classification of companies into healthy and unhealthy classes requires a comparison of each company score and a cut-off score value. In early BP studies, the cut-off score value is set to 0.5. However, in some studies, the cut-off score values are set by the authors based on their experiences due to difficulties in estimating prior probabilities (Trabelsi et al., 2015), and this can inevitably lead to biased results in prediction. In comparative studies, where one compares the performance of the proposed models with the benchmark models, several studies failed to re-estimate the parameters and calibrate the cut-off score values, which may produce unfair comparisons (Almamy et al., 2016). Other studies determined an optimal cut-off score value with respect to a specific performance measure; however, in the scenario of modelling with a largely imbalanced dataset, setting a cut-off score value so as to maximise the overall correct classification rate may result in a high correct classification rate of the majority class whilst sacrificing that of the minority class (Veganzones & Séverin, 2018). Attempts have been made to address this issue by, for example, minimising the *expected misclassification cost* instead of the *overall correct classification rate* to determine the cut-off score value (e.g., Trabelsi et al., 2015; Du Jardin, 2021a), or using AUC measure to compare models to avoid concerns about different cut-off score values across different models (Almaskati et al., 2021).

2.10 Conclusions and future research directions

The topic of corporate bankruptcy and financial distress prediction has attracted the attention of many researchers for several decades and continues to evolve with more and more advanced prediction methodologies and issue fixation solutions being proposed. This study contributes to this domain of research by providing an up-to-date state-of-the-art review, classification and critical analysis of the current BP and FDP literature, where six major research streams are identified and discussed, namely, the definition and coding of bankruptcy and financial distress; design of new prediction models/classifiers or new application of existing ones; design of new drivers or evaluation of established drivers; design or evaluation of feature selection methods; design of methodologies for the performance evaluation of prediction models; and issues affecting the performance of prediction models and related solutions. By painting the landscape of research and analysing each research stream, this study would serve as a guide for researchers who intend to explore this field of study and/or contribute to its development.

Our analysis of the surveyed papers revealed a clear trend in terms of prediction methodologies of bankruptcy and financial distress, where more emphasis is put on advanced machine learning and artificial intelligence models such as *ensemble learning* models. Ensemble learning models and other methodological advances, such as DEA- and MCDA-based prediction methods, have contributed to further pushing the boundaries of research. Overall, it is fair to conclude that there is no single methodology that is better than the others, as each methodology has its strengths and weaknesses, some of which are design-related, and others are implementation decisions related, which have made the design and implementation of new ensemble models more popular.

Importantly, this chapter provides the foundational context for the thesis by critically analysing the current landscape of bankruptcy and financial distress prediction research. By identifying the gaps and trends in the literature, it underscores the need for innovative methodologies, such as those presented in **Chapter 3** and **Chapter 4**. The findings emphasise the necessity of integrating advanced machine learning, network analysis, and financial profiling techniques to address the limitations of traditional approaches. In this way, the chapter establishes the relevance and significance of the novel contributions proposed in the subsequent chapters, linking the reviewed literature to the thesis's aim of enhancing the accuracy, robustness, and interpretability of bankruptcy prediction models.

Our classification of the literature on bankruptcy and financial distress prediction into research streams suggests that the research directions in this field are much broader than focussing only on developing new prediction models or classifiers alone. Future research directions and recommendations to mitigate the risks associated with corporate bankruptcy and financial distress prediction could focus on the following ten key areas:

Enhancing Models' Interpretability: There is a growing need for interpretable models in the prediction of bankruptcy and financial distress. Future research could focus on developing methodologies that not only provide powerful predictions, but also offer insight into the reason behind these predictions. This could involve integrating explainable AI techniques with current methodologies to make models more transparent and trustworthy, especially for stakeholders who rely heavily on these predictions for decision making.

Improving Data Quality and Diversity: The issue of data imbalance and heterogeneity in datasets is always a significant challenge in predictive modelling. Future research could explore more sophisticated data sampling and preprocessing techniques, including developing adaptive algorithms that can dynamically balance datasets and handle a variety of financial indicators for different clusters or homogeneous subsamples, hence improving the robustness and generalisability of the models.

Incorporating Real-Time Data and External Drivers: In an era of rapid economic and political changes, the inclusion of real-time data and economic drivers in predictive models can be crucial. Research in this direction could focus on integrating real-time financial data, global economic indicators, and sentiment analysis from news and social media to enhance the predictive capabilities of bankruptcy and financial distress prediction models.

Developing Advanced Ensemble and Hybrid Models: The use of ensemble learning and hybrid models has shown promise in improving prediction performance. Future research should focus on exploring new combinations of machine learning and statistical methods, potentially integrating novel AI techniques for a more holistic approach.

Exploring the Impact of Regulatory and Compliance Factors: The influence of regulatory changes and compliance requirements on corporate financial health is an underexplored area. Future studies could examine how changes in regulations, both domestically and internationally, impact financial distress and bankruptcy risks, potentially leading to the development of more dynamic and adaptable prediction models.

Cross-Industry and Cross-Country Comparisons: Research spanning various industries and countries can offer broader insights into bankruptcy prediction. This could involve comparative studies that assess the applicability and effectiveness of existing models across different economic sectors and geographical regions, leading to more universally applicable models.

Focussing more on Small and Medium Enterprises: Much of the current research is focused on large corporations. Given the economic importance of SMEs, future research should also address the prediction of bankruptcy and financial distress in this segment, considering their unique financial structures and challenges.

Empirical Application in Risk Mitigation and Financial Impact Estimation: Future research should emphasise the practical application of bankruptcy and financial distress prediction models in risk mitigation and financial impact analysis. This includes developing strategies based on predictive insights to prevent or mitigate the severity of financial distress. Additionally, quantifying the monetary benefits and cost-effectiveness of various risk mitigation measures is also crucial, as it can provide executable information for stakeholders who focus on economic outcomes and the real-world effectiveness of prediction models.

Investigating the Impact of Macroeconomic Events: Future research should focus on understanding how significant macroeconomic events, such as the 2008 global financial crisis and the COVID-19 pandemic, affect the predictions of bankruptcy and financial distress. This includes examining how these events affect financial indicators and other key predictive drivers within machine learning models, as well as the performance of prediction models. Research in this area is vital to develop models that can adapt and accurately reflect the realities of economic turbulence, thus enhancing the robustness and relevance of predictions in varying economic conditions.

Ethical Considerations and Bias Mitigation: Finally, as artificial intelligence and machine learning models become more popular in financial predictions, it is crucial to consider ethical implications and biases in these models. Future research should focus on developing fair and unbiased models, ensuring that they do not inadvertently disadvantage certain groups or companies.

2.A Appendix

2.A.1 Classification of the definitions of corporate bankruptcy and financial distress in the BP and FPD surveyed literature

Category of Definition	Type of Definition	Definition / Criteria	Papers / Authors & Year
Bankruptcy / Legal Definitions	At some stage of the bankruptcy legal process	*A company has filed a bankruptcy petition.*	Zmijewski (1984)
		A company has filed for bankruptcy.	Sarkar & Sriram (2001), Shin et al. (2005), Min & Jeong (2009), Kim et al. (2016), Barboza et al. (2017)
		A company has filed for bankruptcy protection.	Muñoz-Izquierdo et al., (2019)
		A company has filed for any type of bankruptcy within 5 years of delisting.	Shumway (2001)
		A company had entered statutory bankruptcy proceedings.	Serrano-Cinca et al. (2019)
		A company against which a bankruptcy or liquidation procedure was initiated.	Nyitrai & Virág (2019)
		A company has filed for bankruptcy or was liquidated.	Qiu et al. (2020)
		A company has been liquidated due to insolvency.	Tobback et al. (2017)
		A company is liquidated or reorganised.	Veganzones & Séverin (2018),
		A company is liquidated, reorganised, or ruled by court decision as bankrupt.	Du Jardin (2018), Du Jardin (2021b)
		A company's assets are liquidated in order to fulfil as much debt as possible and the company is no longer a going concern, or a company is in reorganisation, which involves the settlement of debt repayment between a company and its creditors whilst the company continues to exist.	Zoričák et al. (2020)
		A company is in either of the following states, i.e., bankruptcy, in compulsory dissolution, or ceased to exist following compulsory dissolution.	Matin et al. (2019)
		A company is either in bankruptcy or in the process of recovery.	Huang & Yen (2019)
		A company has filed a bankruptcy petition under Chapter X of the US National Bankruptcy Act of 1938, or under Chapter XI (reorganisation) or Chapter VII (liquidation) of the US Bankruptcy Reform Act of 1978.	Altman (1968), Altman et al. (1977), Ohlson (1980), Zavgren (1985), Frydman et al. (1985), Odom & Sharda (1990), Bryant (1997), Paradi et al. (2004), Sun & Shenoy (2007), Bharath & Shumway (2008), Chen et al. (2009), Trabelsi et al. (2015), Tian et al. (2015), Zorn et al. (2017), Tian & Yu (2017), Lin & Dong (2018), Mai et al. (2019), Bai & Tian (2020), Habermann & Fischer (2023), Almaskati et al. (2021)
		A company has entered liquidation, administration, or receivership, following the UK Insolvency Act of 1986.	Altman et al. (2008), Gupta et al. (2015), Appiah & Chizema (2016), Andrikopoulos & Khorasgani (2018)
		A company has declared bankruptcy or submitted a restructuring plan to the French court.	Chen et al. (2011b)
		A company triggered a bankruptcy procedure prescribed by Italian law in 2012.	Sartori et al. (2016)
		A company has made a judicial declaration of bankruptcy under the Spanish 2004 bankruptcy act 22/2003.	De Andrés et al. (2011)
		A company has been declared bankrupt by The Federation of Belgian Chambers of Commerce or has obtained a concordat.	Cielen et al. (2004)
Financial Distress Definitions	Abnormal financial conditions	*A company's EBITDA is lower than the financial expenses and has a negative growth in the market value for two consecutive years.*	Tinoco & Wilson (2013), Tinoco et al. (2018), Mousavi et al. (2019)
		A company's interest cover ratio is less than one, or EBIT is less than the interest payments.	Ninh et al. (2018)
		A company having negative EBITDA over interest expenses, receiving negative EBIT and having negative net income before special items, for two consecutive years.	Fernández-Gámez et al. (2020)
		A company meets any of the following three criteria: negative growth in average market value in any two consecutive years, EBITDA is less than financial expenses, and operating cashflow is less than financial expenses.	Gupta & Chaudhry (2019)

		A company has negative net profit in two successive years or its net assets per share is lower than its stock book value.	Sun et al. (2020)
		A company has negative net income for two consecutive years.	Oz & Simga-Mugan (2018)
	Stock exchanges' criteria	*A company is under special treatment by the Teheran Stock Exchange.*	Moradi et al. (2013)
		A company is under special treatment (ST) by Shanghai and Shenzhen stock exchanges under the 'Special Treatment' regulations specified by the China Securities Regulatory Commission (CSRC).	Li et al. (2009), Li & Sun (2009), Zhou et al. (2015), Zhang & Hu (2016), Wang & Wu (2017), Choi et al. (2018), Wang et al. (2018), Sun et al. (2020), Wang et al. (2020), Shen et al. (2020), Mousavi & Lin (2020), Du et al. (2020), Li et al. (2021), Yuan et al. (2022)
		A company faced financial difficulties and is delisted from the US stock exchange.	Kim (2018)
		A company is delisted from the Tokyo Stock Exchange (TSE) due to liquidation, rehabilitation, reorganisation, or failure to meet listing conditions.	Xu & Zhang (2009)
		A company is delisted from the US stock exchange due to bankruptcy, liquidation or poor performance, with CRSP delisting codes of 400 and 550-585.	Cheng et al. (2014)
		A company satisfies any one of the following conditions: reorganisation, bankruptcy, full-value delivery (per-share book value below 5TWD), stock transaction suspension, or withdrawal from the Taiwan stock market.	Lin et al. (2011a), Lin et al. (2011b), Lin et al. (2014)
		A company is delisted from the Japanese stock exchange due to its inadequate financial performance, i.e., failed to pay any dividend in the last two annual periods before its delisting.	Sueyoshi & Goto (2009a), Sueyoshi & Goto (2009b)
		A company is delisted from the Japanese stock exchange due to bankruptcy, rehabilitation, or reorganisation procedures, excessive debt, suspension of bank transactions, and termination of business activities (excluding mergers).	Hosaka (2019)
		A company's stock is announced to be 'terminated' due to the following reasons, i.e., having a credit crisis, having net operating loss, failing to pay debts, or violating regulations set by Taiwan Stock Exchange Corporation (TSE).	Yeh et al. (2010)
		A company has been declared bankrupt according to the operating rules of the Taiwan Stock Exchange Corporation.	Liang et al. (2015), Liang et al. (2016), Chou et al. (2017)
Hybrid Definitions	Databases' listing criteria	*A company that is dropped from a commercial database.*	Beaver (1966), Tseng & Hu (2010)
		A company has stopped reporting financial statements.	Valencia et al. (2019)
		A company meets two criteria set by the COMPUSTAT Database, i.e., the state alert (STALTQ0) and the reason for deletion (DLRSN).	Li & Faff (2019)
		A company whose reason for deletion is marked as 'bankruptcy' or 'liquidation' in the original COMPUSTAT North American Database.	Zhou (2013)
		A company is flagged with codes 16 (in Receivership), 20 (in Administration) or 21 (Cancelled and Assumed valueless) by the London Share Price Database (LSPD).	Bauer & Agarwal (2014), Mousavi et al. (2015), Ouenniche & Tone (2017), Ouenniche et al. (2018a), Ouenniche et al. (2018b), Ouenniche et al. (2018c), Ouenniche et al. (2019)
		A company meets any of the following criteria of Bureau van Dijk ORBIS database, i.e., default of payment; firms subject to insolvency proceedings; firms subject to bankruptcy proceeding; firms which are dissolved (through bankruptcy); firms in liquidation, vi) inactive firms (no precision).	Jones & Wang (2019)
		A company is in either of the following three states of Bureau van Dijk Amadeus database, i.e., has been legally declared not to be able to pay its creditors and is under court supervision; has no longer existed due to ceased activities and liquidation process, or has exited the database with unknown reason.	Amendola et al. (2015)
		A company is recorded as 'delisted', 'managed' or '100% margin' by the Taiwan Economic Journal database (TEJ).	Lu et al. (2013)

	Definitions chosen by authors	*A company is characterised as having a negative net worth, having declared bankruptcy, having gone through restructuring, receiving a bailout, having full-delivery stocks, ceased stock market trading, incurred gigantic losses or delisted from Taiwan Stock Exchange or GreTai Securities Market (GTSM).*	Lu et al. (2015)
		A Company is classified as financially distressed whenever it meets the following two conditions: (1) the firm is inactive, has merged, is suspended, dissolved, or undergone liquidation (either voluntary or by court order), gone bankrupt or equivalent; (2) its net income is negative for three consecutive years.	Ashraf et al. (2020)
Did not provide a definition of bankruptcy or financial distress	Taffler (1984), Hsieh (1993), Tsukuda & Baba (1994), West et al. (2005), Ahn & Kim (2009), Hu (2009), Li & Ho (2009), Premachandra et al. (2009), Nanni & Lumini (2009), Altman et al. (2010), Chen et al. (2011a), Hu & Chen (2011), Li et al. (2011), Tsai (2009), Tsai & Cheng (2012), Tsai & Hsu (2013), Wang et al. (2014), Zięba et al. (2016), Almamy et al. (2016), Afik et al. (2016), Pereira et al. (2016), Du Jardin (2016), Volkov et al. (2017), Antunes et al. (2017), Zelenkov et al. (2017), Du Jardin (2017), Le et al. (2018), Le et al. (2019), Son et al. (2019), Chen et al. (2020), Alam et al. (2021), Cleofas-Sánchez et al. (2016), Lian (2017), Miao et al. (2018), Alaka et al. (2018), Ahmad (2019), Lin et al. (2019), Uthayakumar et al. (2020), Liang et al. (2020), De Bock et al. (2020), Du Jardin (2021a), Kou et al. (2021)		

2.A.2 Classification of methodologies and models or methods in the BP/FDP literature

Methodology / Models		**Papers / Authors & Year**
Univariate Financial Ratio Analysis	-	FitzPatrick (1932)[B][1], Beaver (1966)[B]
Statistical Models	Multivariate Discriminant Analysis Models	Linear Discriminant Analysis (LDA): Altman (1968)[B], Altman et al. (1977)[B], Taffler (1984)[B], Xu & Zhang (2009)[B], Almamy (2016)[B], Qiu et al. (2020)[B], Habermann & Fischer (2023)[B], Alam et al. (2021)[B], Oz & Simga-Mugan (2018)[D], Ninh et al. (2018)[D], Andrikopoulos & Khorasgani (2018)[D] Linear Discriminant Analysis with Regularisation Terms: Volkov et al. (2017)[B] Quadratic Discriminant Analysis (QDA): Altman et al. (1977)[B]
Probabilistic Models	Regression Models	Probit Model: Zmijewski (1984)[D] Logistic Regression (LR): Ohlson (1980)[B], Zavgren (1985)[B], Altman et al. (2008)[B], Altman et al. (2010)[B], Piñeiro-Sánchez et al. (2013)[B], Trabelsi et al. (2015)[B], Appiah & Chizema (2016)[B], Zorn et al. (2017)[B], Lin & Dong (2018)[B], Serrano-Cinca et al. (2019)[B], Li & Faff (2019)[B], Tinoco & Wilson (2013)[D], Lu et al. (2013)[D], Lu et al. (2015)[D], Lian (2017)[D], Tinoco et al. (2018)[D], Ahmad (2019)[D], Fernández-Gámez et al. (2020)[D], Ashraf et al. (2020)[D] Logistic Regression with Regularisation Terms: Pereira et al. (2016)[B], Volkov et al. (2017)[B] Quadratic Interval Logistic Regression: Tseng & Hu (2010)[B] Generalised Additive Model: Valencia et al. (2019)[B] Least-squares Probabilistic Classifier: Zoričák et al. (2020)[B] Multivariate Adaptive Regression Splines (MARS): De Andrés et al. (2011)[B]
	Bayesian Theory-based Models	Naïve Bayes (NB): Hsieh (1993)[B], Trabelsi et al. (2015)[B], Sarkar & Sriram (2001)[D] Bayesian Network (BN): Sun & Shenoy (2007)[B]
Stochastic Models	Option Pricing Theory-Based Models	Black-Scholes/Merton Model: Bharath & Shumway (2008)[B], Miao et al. (2018)[D]
	Stochastic Process-Based Models	Discrete-time Hazard Model: Shumway (2001)[B], Trabelsi et al. (2015)[B], Gupta et al. (2015)[B], Tian & Yu (2017)[B], Du Jardin (2017)[B], Du Jardin (2018)[B], Bai & Tian (2020)[B], Gupta & Chaudhry (2019)[D], Li et al. (2021)[D] Gaussian Process-Based Model: Antunes et al. (2017)[B]
AI/ML Models	Case-based Reasoning Models	Case-based Reasoning with k-Nearest Neighbour (CBR+kNN): Bryant (1997)[B], Ahn & Kim (2009)[B], Min & Jeong (2009)[B], Li et al. (2011)[B], Sartori et al. (2016)[B], Ouenniche et al. (2018a)[B], Ouenniche et al. (2018b)[B], Ouenniche et al. (2018c)[B], Ouenniche et al. (2019)[B], Li & Sun (2009)[D], Li et al. (2009)[D] Case-based Reasoning with Fuzzy k-Nearest Neighbour (CBR+F-kNN): Chen et al. (2011a)[B], Li & Ho (2009)[D]
	Decision Trees	Decision Trees (DT): Cielen et al. (2004)[B], Du Jardin (2017)[B], Muñoz-Izquierdo et al., (2019)[B], Frydman et al. (1985)[D], Sarkar & Sriram (2001)[D] Classification and Regression Tree (CART): Liang et al. (2016)[B], Du Jardin (2018)[B], Nyitrai & Virág (2019)[B], Almaskati et al. (2021)[B] Chi-square Automatic Interaction Detection Decision Tree (CHAID): Serrano-Cinca et al. (2019)[B], Nyitrai & Virág (2019)[B] Logistic Model Tree (LMT): Alam et al. (2021)[B]
	Neural Networks Models	Multi-layer Perceptron (MLP): Odom & Sharda (1990)[B], Tsukuda & Baba (1994)[B], West et al. (2005)[B], Tsai (2009)[B], Kim et al. (2016)[B] Deep Neural Networks: Mai et al. (2019)[B], Paraschiv et al. (2021)[B], Alaka et al. (2018)[D], Matin et al. (2019)[D] Convolutional Neural Network (CNN): Hosaka (2019)[B], Mai et al. (2019)[B], Ahmadi et al. (2018)[D], Matin et al. (2019)[D] Radial Basis Function (RBF) Neural Network: Tseng & Hu (2010)[B], Uthayakumar et al. (2020)[D] Self-Organizing Maps (SOM) Neural Network: Wang & Wu (2017)[D], Du Jardin (2021b)[B] Deep Belief Neural Network (DBN): Huang & Yen (2019)[D]

		Neuro-fuzzy Neural Network: Chen et al. (2009)[B] Single-hidden layer Feedforward Neural Network trained with Extreme Learning Machine (ELM): Du Jardin (2021b)[B] Learning Vector Quantisation (LVQ)/Neural Network trained with Winner-take-all Learning-based Algorithm: Chen et al. (2011b)[B] Hybrid Associative Memory with Translation (HACT) Neural Network: Cleofas-Sánchez et al. (2016)[D], Huang & Yen (2019)[D]
	Support Vector Methodologies	Support Vector Machines (SVM): Shin et al. (2005)[B], Nanni & Lumini (2009)[B], Tobback et al. (2017)[B], Chen et al. (2020)[B], Lin et al. (2011a)[D], Lin et al. (2011b)[D], Zhang & Hu (2016)[D], Wang & Wu (2017)[D], Huang & Yen (2019)[D], Sun et al. (2020)[D] Support Vector Machines based on Rough Set Theory (SVM-RST): Yeh et al. (2010)[B] Support Vector Data Description (SVDD): Gorgani et al. (2010)[B], Moradi et al. (2013)[D], Yuan et al. (2022)[D]
	Ensemble Learning Models	*Non-stacked Homogeneous Ensembles / Ensemble classifiers with single basic model*: Adaptive Boosting (AdaBoost): West et al. (2005)[B], Wang et al. (2014)[B], Barboza et al. (2017)[B], Du Jardin (2017)[B], Le et al. (2018)[B], Chen et al. (2020)[B], Du Jardin (2021a)[B], Sun et al. (2017)[D], Sun et al. (2020)[D] Gradient Boosting (GB): Zelenkov et al. (2017)[B], Jones & Wang (2019)[B], Son et al. (2019)[B], Du et al. (2020)[D] Extreme Gradient Boosting (XGBoost): Zięba et al. (2016)[B], Volkov et al. (2017)[B], Le et al. (2019)[B], Son et al. (2019)[B], Du Jardin (2021a)[B], Du Jardin (2021b)[B], Kou et al. (2021)[B] Bootstrap Aggregating (Bagging): West et al. (2005)[B], Nanni & Lumini (2009)[B], Chen et al. (2020)[B], Shen et al. (2020)[D], Wang et al. (2020)[D] Random Forest (RF): Volkov et al. (2017)[B], Barboza et al. (2017)[B], Du Jardin (2017)[B], Zelenkov et al. (2017)[B], Du Jardin (2021a)[B], Shen et al. (2020)[D] Random Subspace (RS): Nanni & Lumini (2009)[B], Du Jardin (2016)[B], Du Jardin (2018)[B], Du Jardin (2021a)[B], Du Jardin (2021b)[B], Wang et al. (2018)[D], Wang et al. (2020)[D] Rotation Forest (RF): Nanni & Lumini (2009)[B], Du Jardin (2017)[B], Du Jardin (2018)[B], Du Jardin (2021b)[B] Isolation Forest (IF): Zoričák et al. (2020)[B] Class Switching (CS) Ensemble: Nanni & Lumini (2009)[B] *Non-stacked Heterogeneous Ensemble models / Ensemble classifiers with multiple basic models*: Class-belonging based Voting Ensemble Classifiers: Zelenkov et al. (2017)[B] Class-belonging Probability based Ensemble Classifiers: Choi et al. (2018)[D] Outputs Weighting-based Ensemble Classifiers: Du et al. (2020)[D], Shen et al. (2020)[D] *Stacked Heterogeneous Ensemble models* Stacked Generalisation (Stacking): Tsai & Hsu (2013)[B], Kim (2018)[D], Liang et al. (2020)[D]
	Data Envelopment Analysis-Based Models	Paradi et al. (2004)[B], Cielen et al. (2004)[B], Premachandra et al. (2009)[B], Sueyoshi & Goto (2009a)[B], Sueyoshi & Goto (2009b)[B], Yeh et al. (2010)[B], Ouenniche & Tone (2017)[B], Almaskati et al. (2021)[B]
	MCDA-Based Models	Elimination and Choice Translating Reality (ELECTRE): Hu (2009)[B], Hu & Chen (2011)[B] Outranking Relation Theory (OR): Li et al. (2009)[D] Evaluation Based on Distance from Average Solution (EDAS): Ouenniche et al. (2018a)[B] Reference Point Methods (RPMs): Ouenniche et al. (2018b)[B] Technique for Order Preference by the Similarity to Ideal Solution (TOPSIS): Li et al. (2011)[B], Ouenniche et al. (2018c)[B] VIKOR Method: Ouenniche et al. (2019)[B] PROMETHEE Multi-criteria Decision Aid: Hu & Chen (2011)[B]

Remarks:
1, [B] - bankruptcy prediction; [D] – financial distress prediction

2.A.3 The drivers used in the BP and FPD surveyed literature

Category of information	Category of Measurement (s)	Drivers/Authors & Year
Accounting Information	Asset Utilisation Measurements	*Sales/Total Assets* (Altman, 1968); *Sales/Net Plant* (Zavgren, 1985); *Net Sales/Average Fixed Assets (also: Fixed Assets Turnover)* (Cheng et al., 2014); *Revenue/Current Number of Employees (also: Turnover per Employee)* (Chen et al., 2011b); *Value Added[1]/Number of Employees (also: Value Added per Employee)* (Antunes et al., 2017);
	Operating Performance Measurements	*Sales/Cost of Sales* (Ahn & Kim, 2009); *Cost of Sales/Average Payable Accounts* (Geng et al., 2015); *Inventories/Sales (also: Inverse of Inventory Turnover)* (Zavgren, 1985); *Main Business Costs/Average Inventories* (Wang et al., 2020); *Costs of Sales/Average Inventories* (Cheng et al., 2014); *Debtor Days Ratio (also: Account Receivables/Annual Credit Sales * 365 Days)* (Chen et al., 2011b); *Creditor Days Ratio (also: Account Payables/Cost of Sales * 365 Days)* (Chen et al., 2011b); *Payables Turnover (also: Credit Purchases/Average Accounts Payable)* (Ahn & Kim, 2009); *Receivables Turnover (also: Net Credit Sales/Average Accounts Receivable)* (Li & Ho, 2009); *Operating Income[2]/Number of Employee* (Lin et al., 2014); *Gross Income/Sales* (Li et al., 2011); *Operating Income/Sales* (Mai et al., 2019); *Gross Profit/Net Sales* (Lin et al., 2014); *EBIT/Sales* (Tian et al., 2015); *Change in EBIT/Sales* (Zhou et al., 2015); *Depreciation and Amortisation/Sales* (Fernández-Gámez et al., 2020);

		Net Profits/Sales (Sun et al., 2020); *Net Profits/Net Sales* (Nyitrai & Virág, 2019); *Net Income/Sales* (Bryant, 1997); *Net Income/Net Sales* (Andrikopoulos & Khorasgani, 2018); *Operating Income Growth Rate* (Wang & Wu, 2017); *Operating Profits* (Le et al., 2019);
	Other Profitability Measures	*Profit Before Tax/Sales* (Taffler, 1984); *Net Profits* (Le et al., 2019); *Net Income Growth Rate* (Kim, 2018) *Retained Earnings/Net Sales* (Cheng et al., 2014);
	Cashflow measurements	*Cashflow/Net Sale* (Lin et al., 2011b); *Cashflow/Operating Revenue* (Amendola et al., 2015); *Cashflow/Total Assets* (Premachandra et al., 2009); *Cashflow/Equity* (Lin et al., 2014); *Cashflow/Net Worth*[3] (Valencia et al., 2019); *Cashflow/Total Liabilities* (Beaver, 1966); *Cash Re-investment Ratio*[4] (Lin et al., 2014); *Cash/Current Liabilities* (Le et al., 2019); *Cash/Total Liabilities* (Wang & Wu, 2017); *Operating Cashflow/Current Liabilities (also: Operating Cashflow Coverage Ratio)* (Lin & Dong, 2018); *Net Cashflow from Investment Activity per Share* (Sun et al., 2017); *Financial Expenses /Cashflow* (Du Jardin, 2018); *Cashflow/Financial Liabilities* (Du Jardin, 2017);
	Liquidity Measurements	*Sales/Cash* (Bryant, 1997); *Current Assets/Sales* (Chen et al., 2011a); *Current Liabilities/Sales* (Mai et al., 2019); *Quick Assets/Sales* (Bryant, 1997); *Working Capital*[5]*/Sales* (Huang & Yen, 2019); *No Credit Interval*[6] (Ouenniche et al., 2018c); *Current Ratio (also: Current Assets/Current Liabilities)* (Beaver, 1966); log(*Current Ratio*) (Gupta et al., 2015); *Quick Ratio (also: (Current Assets – Inventories)/Current Liabilities)* (Lu et al., 2015); *Quick Assets/Current Liabilities* (Zavgren, 1985); *Working Capital/Total Assets* (Ninh et al., 2018);
	Capital Structure & Solvency Measurements	*Current Assets/Total Assets* (Bryant, 1997); *Current Assets/Total Liabilities* (Ouenniche et al., 2018c); *Current Liabilities/Total Assets* (Min & Jeong, 2009); *Current Liabilities/Total Liabilities* (Tian et al., 2015); *(Current Liabilities-Cash)/Total Assets* (Tian & Yu, 2017); *Quick Assets/Total Assets* (Taffler, 1984); *Quick Assets/Net Capital Employed*[7] (Taffler, 1984); *Short-term Liabilities/Book Value of Equity* (Gupta et al., 2015); *Short-term Liabilities/Net Worth* (Andrikopoulos & Khorasgani (2018)); *Long-term Liabilities/Equity* (Chen et al., 2011a); *Long-term Liabilities/Total Assets* (Du Jardin, 2017); *Fixed Assets/Total Assets* (Geng et al., 2015); *Fixed Assets/Net Worth* (Cheng et al., 2014); *Financial Liabilities/Equity* (Chen et al., 2011b); *Financial Liabilities/Capital Employed* (Antunes et al., 2017); *Total Liabilities/Tangible Assets* (Li et al., 2011); *Total Liabilities/Total Assets* (Ohlson, 1980); *Total Liabilities/Net Capital Employed* (Taffler, 1984); *Total Liabilities/Net Worth* (Bryant, 1997); *Total Liabilities/Equity* (Geng et al., 2015); *Total Liabilities Exceed Total Assets (Dummy Variable)* (Bryant, 1997); *Net Assets/Total Assets* (Shin et al., 2005); *Net Assets per Share* (Li et al., 2011); *Tangible Assets/Total Assets* (Li et al., 2021); *Intangible Assets/Total Assets* (Gupta et al., 2015); *Working Capital/Net Worth* (Taffler, 1984); *Working Capital/Number of Employee* (Jones & Wang, 2019); *Net Worth/Total Assets* (Taffler, 1984); *Capital Employed/Total Liabilities* (Andrikopoulos & Khorasgani, 2018); *Capital Employed/Fixed Assets* (Antunes et al., 2017); *Net Capital Employed/Total Liabilities Excluding Deferred Tax* (Taffler, 1984); *Equity/Total Assets* (Tian et al., 2015); *Equity/Total Liabilities* (Serrano-Cinca et al., 2019); *Equity/Fixed Assets* (Geng et al., 2015); *Parent Company Owner's Equity/Invested Capital* (Du et al., 2020); *Equity per Share* (Geng et al., 2015); *Retained Earnings/Total Assets* (Alaka et al., 2018); *Retained Earnings/Current Liabilities* (Tian et al., 2015); *Interest Expenses/Sales* (Tsukuda & Baba, 1994); *Financial Expenses/Sales* (Gupta & Chaudhry, 2019); *Net Interest Expenses/Sales* (Min & Jeong, 2009); *Earnings Before Interest Expenses/Interest Expenses* (Cheng et al., 2014); *EBIT/Interest Expenses* (Premachandra et al., 2009); *Interest Coverage Ratio (Also: EBITDA/Interest Expense)* (Amendola et al., 2015); *Financial Expenses/EBITDA* (Veganzones & Séverin, 2018);

		Financial Expenses/Value Added (Du Jardin, 2017); *Interest Expenses/Value Added* (Tsukuda & Baba, 1994); *Funds Provided by Operations/Total Liabilities* (Ohlson, 1980); *Non-operational Expenses/Sales* (Tsukuda & Baba, 1994);
	Return on Investment Measures	*Operating Income/Total Assets* (Shin et al., 2005); *Net Profit/Total Assets* (Li et al., 2011) *Net Income Growth Rate* (Bryant, 1997); *Net Income/Current Assets* (Zhou et al., 2015); *Net Income/Total Assets (also: Return on Assets)* (Lian, 2017); *Net Income/Total Liabilities* (Valencia et al., 2019); *Net Income/Net Worth* (Yeh et al., 2010); *Net Income/Equity (also: Return on Equity)* (Lin et al., 2014); *Net Income/Number of Shares (also: Net Income per Share)* (Zhou et al., 2015); *Negative Net Income for Last Two Years (Dummy Variable)* (Bryant, 1997); *Change in Net Income/Total Assets* (Zhou et al., 2015); *Operating Income After Depreciation/Total Assets* (Mousavi et al., 2019); *Operating Income Before Tax/Total Assets* (Lin et al., 2014); *Operating Income After Tax per Share* (Lin et al., 2014); *Operating Profit Margin* (Chen et al., 2011b); *Gross Profit/Total Assets* (Mousavi & Lin, 2020); *Net Profit After Interest and Taxes/Working Capital* (Cheng et al., 2014); *Profit Before Tax/Current Liabilities* (Taffler, 1984); *Net Profit Growth Rate* (Geng et al., 2015); *Net Profit/Current Assets* (Li & Sun, 2009); *Net Profit/Total Assets* (Sun et al., 2017); *Net Profit/Fixed Assets* (Li et al., 2011); *Net Profit/Equity* (Li & Sun, 2009); *Change in Total Profit/EBIT* (Zhou et al., 2015); *EBIT/Total Assets* (Odom & Sharda, 1990); *EBIT/Total Liabilities* (Taffler, 1984); *EBIT/Capital Employed (also: Return on Capital Employed)* (Chen et al., 2011b); *EBIT/Long-term Capital* (Sun et al., 2020); *EBITDA* (Trabelsi et al., 2015); *EBITDA/Total Assets* (Hu, 2009); *EBITDA/Permanent Equity* (Du Jardin, 2018); *Change in Return on Assets* (Zhou et al., 2015); *Change in Return on Equity* (Zhou et al., 2015); *Earnings per Share* (Li & Sun, 2009); *Continuous 4 Quarterly Earnings per Share* (Lin et al., 2011a); *Unallocated Profit per Share* (Sun et al., 2020); *Annual Abnormal Returns*[8] (Tinoco & Wilson, 2013); *Cumulative Abnormal Return* (Trabelsi et al., 2015); *Cumulative Average Abnormal Return* (Ahmad, 2019); *Standard Deviation of the Abnormal Return* (Trabelsi et al., 2015); *Excess return over the S&P 500 index* (Tian et al., 2015);
	Others	*Firm Size*[9] (Amendola et al., 2015); *Firm Size as proxied by Sales* (Li & Ho, 2009); *Firm Size as proxied by* log(*Sales*) (Tian & Yu, 2017); *Firm Region* (Amendola et al., 2015); *Business Sector*[10] (Putra et al., 2020); *Transactional Information*[11] (Kou et al., 2021) *Depreciation Expenses* (Min & Jeong, 2009); *Depreciation of Tangible Assets* (Antunes et al., 2017); *Net Interest Expenses* (Min & Jeong, 2009); *Tax Expenses* (Jones & Wang, 2019); *Tax Rates* (Lin et al., 2011a);
Extra Accounting Information		*Firm Age* (Appiah & Chizema, 2016); *Number of Employees Last Available Year* (Chen et al., 2011b)
Audit Information	Engagement-Level Indicators	*Account Audited (Dummy Variable)* (Altman et al., 2010); *Auditor Type*[12] (Piñeiro-Sánchez et al., 2013); *Proportion of Audited Years* (Piñeiro-Sánchez et al., 2013); *Number of Different Auditors Hired* (Piñeiro-Sánchez et al., 2013); *Average Length of Auditors' Contracts* (Piñeiro-Sánchez et al., 2013); *Temporal Matches between Auditor Changes and Changes in the Opinion* (Piñeiro-Sánchez et al., 2013); *Explicit Obstructionism*[13] (Piñeiro-Sánchez et al., 2013); *Delays in Filing of Annual Financial Statements* (Piñeiro-Sánchez et al., 2013); *Late Filing Days of Company Accounts* (Altman et al., 2008); log(*Number of Days Late in Filing Financial Reports*) (Gupta et al., 2015); *Non-compliance with the Obligation of Auditing Accounts* (Piñeiro-Sánchez et al., 2013);
	Firm-Level Indicators	*Audit Opinion in Audit Report*[14] (Muñoz-Izquierdo et al., 2019); *Emphasis of Matter in Audit Report*[15] (Muñoz-Izquierdo et al., 2019); *Number of Comments in Audit Report* (Muñoz-Izquierdo et al., 2019); *Scope Limitation/GAAP Violation in Audit Report*[16] (Muñoz-Izquierdo et al., 2019); *Ratio between Qualified Reports and Total Number of Reports* (Piñeiro-Sánchez et al., 2013); *Number of Critical Qualified Audit Reports*[17] (Piñeiro-Sánchez et al., 2013); *Audit Qualification as 'Severe'* (Altman et al., 2010); *Audit Qualification as 'Going Concern'* (Altman et al., 2010); *Financial Reporting Quality (Proxies)* [18] (Ashraf et al., 2020)

Corporate Governance Information	Board of Directors Characteristics	*Number of Board Directors* (Liang et al., 2020); *Number of Inside Directors* (Liang et al., 2020); *Number of Independent Directors* (Liang et al., 2020); *Number of Non-Paid Persons in Board of Directors* (Mousavi & Lin, 2020); *Proportion of Outsider Directors* (Appiah & Chizema, 2016); *Proportion of Male Directors on the Board* (Almaskati et al., 2021); *Number of Auditors on the Board* (Amendola et al., 2015); *CEO Age* (Almaskati et al., 2021); *CEO Change*[19] (Almaskati et al., 2021); *CEO-Chairman Duality*[20] (Almaskati et al., 2021); *Director Busyness*[21] (Almaskati et al., 2021); *Average of Education in the Board of Directors* (Mousavi & Lin, 2020); *Director External Experience*[22] (Almaskati et al., 2021); *Director Internal Experience*[23] (Almaskati et al., 2021); *Director Company Experience*[24] (Almaskati et al., 2021); *Nomination Committee Effectiveness*[25] (Appiah & Chizema, 2016). *Board Directors' External Connection with Other Companies* (Tobback et al., 2017)
	Remuneration & Reward	*Number of Compensation Members* (Liang et al., 2020); *Directors' Compensation* (Almaskati et al., 2021); *Average Salaries of Top 3 Executives in Recent 3 Years* (Mousavi & Lin, 2020);
	Shares & Shareholders Information	*Number of Shareholders* (Amendola et al., 2015); *Annual Change in Shareholder Numbers* (Jones & Wang, 2019); *Number of Shares Held by Board of Supervisors* (Mousavi & Lin, 2020); *Average Number of Shares of Board in Recent 3 Years* (Mousavi & Lin, 2020); *Shareholding of Top 10 Shareholders* (Mousavi & Lin, 2020); *Top 5 Institutional Shareholdings* (Ahmad, 2019);
	Relational Information	*Inclination to Membership in a Keiretsu*[26] (Xu & Zhang, 2009); *Political Connections*[27] (Ahmad, 2019).
Corporate Social Responsibility (CSR) / Environment, Social & Governance (ESG) Information	Corporate Social Responsibility	*Corporate Social Responsibility Level*[28] (Lin & Dong, 2018); *Refinitiv ESG Combined Score*[29] (Habermann & Fischer, 2023); *Environmental Pillar Score*[30] (Habermann & Fischer, 2023); *Social Pillar Score*[31] (Habermann & Fischer, 2023); *Governance Pillar Score*[32] (Habermann & Fischer, 2023); *Number of County Court Judgements*[33] *(CCJ)* (Altman et al., 2010); *Number of County Court Judgements Pending* (Gupta et al., 2015); *Outstanding County Court Judgements Amount* (Gupta et al., 2015); *Real Value of County Court Judgements* (Altman et al., 2010);
	Corporate Innovation	*R&D Productivity*[34] (Bai & Tian, 2020); *R&D Expense/Operating Revenue* (Jones & Wang, 2019); *Annual R&D Expenditure/Sales* (Bai & Tian, 2020); *Number of Triadic Patents Granted to the Firm* (Bai & Tian, 2020); *Knowledge Capital*[35]*/Total Assets* (Bai & Tian, 2020); *Organisation Capital*[36]*/Total Assets* (Bai & Tian, 2020);
Textual Information	Company Reports Information	*Management Discussion and Analysis (MD&A) section of 10-K report* (Mai et al., 2019); *Text Information of the management's statement and auditor's report* (Matin et al., 2019); *Text Information in Business Management Reports* (Ahmadi et al., 2018);
	Conventional Media Information	*News-Corpus Variable*[37] (Lu et al., 2013);
	Social Media Information	*Time Since Last Social Media Post* (Putra et al., 2020); *Exhaustive List of Social Media-based Drivers* – see (Putra et al., 2020);

Market Information	Stock Market	*Stock Return* (Taffler, 1984);
		Beta of Stock Return[38] (Afik et al., 2016);
		Sigma of Stock Return[39] (Shumway, 2001);
		Annual Increase in Cumulative Market Return (Ashraf et al., 2020);
		Stock Price (Tinoco & Wilson, 2013);
		log(*Stock Price*) (Mai et al., 2019);
		Share Price/Tangible Assets (Sun et al., 2020);
		
		Market Value of Total Assets (Bharath & Shumway, 2008);
		Market Value of Total Assets/Book Value of Total Assets (Gupta & Chaudhry, 2019);
		Total Liabilities/Market Value of Total Assets (Bauer & Agarwal, 2014);
		Tax/Market Value of Total Assets (Gupta & Chaudhry, 2019);
		Net Income/Market Value of Total Assets (Bauer & Agarwal, 2014);
		Cash/Market Value of Total Assets (Bauer & Agarwal, 2014);
		Expected Return on Total Assets (Bharath & Shumway, 2008);
		
		Market Value of Equity[40] (Afik et al., 2016);
		log(*Market Value of Equity*) (Tinoco et al., 2018);
		log(*Market Value of Equity/Total S&P Market Value*) (Gupta & Chaudhry, 2019);
		Change in Market Capitalisation (Zorn et al., 2017);
		Market Value of Equity/Total Assets (Shumway, 2001);
		Market Value of Equity/Total Liabilities (Qiu et al., 2020);
		Market Value of Equity/Total Capital (Altman et al., 1977);
		Market Value of Equity/Book Value of Equity (Miao et al., 2018);
		Growth of Market Value of Equity per Share/Book Value of Equity per Share (Barboza et al., 2017);
		Firm's Size Relative to the Total Size of the FTSE All-share Market Value (Tinoco & Wilson, 2013);
		Volatility of Equity (Bharath & Shumway, 2008);
		Distance-to-Default Measure[41] *(DD)* (Xu & Zhang, 2009);
		Value-at-Risk[40] *(VaR)* (Gupta & Chaudhry, 2019);

		Expected Shortfall[41] (Gupta & Chaudhry, 2019);
	Bond Market	*10-year Treasury Bond Constant Maturity Rate* (Lin & Dong, 2018); *Long-term Government Yield* (Fernández-Gámez et al., 2020);
	Money Market	*UK Short-Term (3-month) Treasury Bill Rate (Deflated)* (Tinoco et al., 2018); *UK Short-term (3-month) Treasury Bill Rate (Inflation-adjusted)* (Tinoco & Wilson, 2013); *One-year Treasury Bill Rate* (Ninh et al., 2018);
Macroeconomic Information	Leading Indicators	*Money Supply* (Fernández-Gámez et al., 2020); *Government Debt* (Fernández-Gámez et al., 2020); *Public Debt Growth Rate* (Jones & Wang, 2019);
	Coincident Indicators	*Gross Capital Formation (% of GDP)* (Fernández-Gámez et al., 2020); *Lag GDP Growth Rate (1 year)* (Mousavi & Lin, 2020); *Quarterly Change in GDP* (Lin & Dong, 2018); *Monthly Industrial Production Growth Rate* (Lin & Dong, 2018); log(*Total Assets/GNP Price-Level Index*) (Ohlson, 1980);
	Lagging Indicators	*Base Borrowing Rate* (Li et al., 2021); *Real Interest Rate* (Fernández-Gámez et al., 2020); *Short-Term Interest Rate* (Fernández-Gámez et al., 2020); *Risk-Free Interest Rate* (Bharath & Shumway, 2008); *Risk Premium* (Fernández-Gámez et al., 2020); *Inflation Rate* (Lu et al., 2015); *Inflation Growth Rate* (Jones & Wang, 2019); *Consumer Price Index* (Fernández-Gámez et al., 2020); *Retail Price Index (base 100)* (Tinoco & Wilson, 2013); *Exchange Rate* (Fernández-Gámez et al., 2020); *Unemployment Rate* (Lu et al., 2015);
Regulatory Information		*Exhaustive List of Regulatory-based Drivers* – for details, see Fernández-Gámez et al. (2020);
Unclassified		*Depreciation and Amortisation/EBIT* (Fernández-Gámez et al., 2020); *Inventories Growth Rate/Sales* (Ahn & Kim, 2009); *Financial Expenses Growth Rate/Total Assets* (Ahn & Kim, 2009); *Growth of Inventories/Inventories* (Tian et al., 2015); *Stock (Inventories)/Total Assets* (Gupta et al., 2015); *Inventories/Total Assets* (Mousavi et al., 2019); *Quick Assets/Inventories* (Mousavi et al., 2019); *Retained Earnings/Inventories* (Bryant, 1997); *Interest Expenses/Equity* (Lin et al., 2014); *Financial Expenses/Net Income* (Veganzones & Séverin, 2018); *Financial Expenses/Total Assets* (Du Jardin, 2017); *Financial Expenses/Total Liabilities* (Ahn & Kim, 2009); *Trade Debtors/Total Assets* (Gupta et al., 2015); *Trade Creditors/Total Assets* (Gupta et al., 2015); *Trade Creditors/Total Liabilities* (Gupta et al., 2015); *Operational Capital/Number of Employees* (Tsukuda & Baba, 1994); *Capital Expenditures/Total Assets* (Zorn et al., 2017); *Tax/Total Assets* (Tseng & Hu, 2010).

Remarks:

1, Value Added: Value created by the company using its own production factors, namely, physical capital and human capital. Typically, it can be measured by *Total Revenues–All Purchases from Other Firms*;

2, Operating Income: Operating Income can be proxied by earnings before interest and taxes (EBIT) or earnings before interest, taxes, depreciation, and amortisation (EBITDA);

3, Net Worth: The value of a company calculated by deducting the total liabilities from the total assets;

4, Cash Re-investment Ratio: Defined as (Increase in Fixed Assets + Increase in Working Capital)/(Net Income + Non-cash Expenses – Non-cash Sales – Dividends);

5, Working Capital: Defined as Current Asset-Current Liability;

6, No Credit Interval (NCI): The 'no credit' interval is calculated as ((quick assets –current liabilities)/daily operating expenses), which is a measure of short-term liquidity;

7, Capital Employed: The total amount of capital a company uses to operate, which is calculated as (Total Assets – Current Liabilities);

8, Annual Abnormal Returns: Abnormal Returns are estimated as the lagged cumulative abnormal return of individual firms. It is assumed that a low level of a firm's abnormal returns relative to those of the FTSE All Share Index will result in a higher probability of falling into the financial distress/failure category. Each firm's past residual return in year t was calculated as the cumulative monthly return of the twelve months prior to the year where the financial distress event was observed, minus the FTSE All Share Index cumulative monthly return for the same period $(t-1)$;

9, Firm Size: The European classification is used. Micro firms are those having a number of workers less than 10 and sales less than 2million euros; small firms are those having a number of workers between 11 and 50 and sales between 2 and 10millions; medium firms are those having a number of workers between 51 and 250 and sales between 10 and 50millions; large firms are those having a number of workers greater than 350 and sales greater than 50millions;

10, Business Sector: Small and medium-sized enterprises (SMEs) category based on SBI2008;

11, Transactional Information: Including customers' ID, quantity of the payment, reciprocal account ID, transaction remarks, time-stamp of the operation;

12, Auditor Type: Computed as categorical variable (individual, local/national society, *big four*);

13, Explicit Obstructionism: Lack of collaboration by the company, and massive failure of mailing circulation and balance confirmation;

14, Audit Opinion in Audit Report: Dummy variable with a value of 1 if the auditor's report is qualified and 0 if it is unqualified;

15, Emphasis of Matter in Audit Report: Dummy variable with a value of 1 if the auditor's report has an Emphasis of Matter paragraph, 0 otherwise;

16, Scope Limitation/GAAP Violation in Audit Report: Categorical variable with a value of 0 if no qualifications appear in the report, 1 if the audit report has a qualification due to a scope limitation or due a GAAP violation, and 2 if the report shows both;

17, Number of Critical Qualified Audit Reports: Critical uncertainties relevant to the company's survival, or, noncompliance with the GAAP that are empirically related to business failure situation;

18, Financial Reporting Quality (Proxies): Poor financial reporting quality may create difficulties in assessing a firm's position regarding the payment of debts and dividends. Financial Reporting Quality is measured by, 1, Earning quality, if reported income is consistently informative, stock returns should adjust after incorporating this and all other available information, which reflects the high quality of financial reporting; 2, Accruals quality, financial distress and difficult market conditions may lead some companies to adopt fraudulent accounting practices, for example, by understating the cost of goods sold and total R&D expenses, and by relaxing credit terms to increase revenues.

19, CEO Change: Dummy variable, 1 if the CEO was replaced during the last three years, 0 otherwise;

20, CEO-Chairman Duality: Dummy variable, 1 if the chairman is a current or ex-CEO, 0 otherwise;

21, Director Busyness: Dummy variable, 1 if the majority of outside directors hold three directorships or more, 0 otherwise;

22, Director External Experience: The average number of years that a director sits on the boards of other companies;

23, Director Internal Experience: The average number of years that a director sits on the board of the company;

24, Director Company Experience: The average number of years that a director works in the company;

25, Nomination Committee Effectiveness: As measured by a composite index consisting of the nomination committee's presence, independence, chairman independence, size and frequency of meetings;
26, Inclination to Membership in a Keiretsu: A proxy for the dependence of the company on its supporting bank, which is calculated based on various criteria. i.e., the characteristics and historical background of the group or the company; sources and amounts of bank loans; board directors sent by or sent to the nucleus or other group companies; the company's attitude towards the group; and the company's connections with other group or nongroup companies. As such, a company's inclination to be a member of a Keiretsu group is rated on a scale of zero to four asterisks;
27, Political Connections: Dummy variable, for political connected firm and 0 otherwise;
28, Corporate Social Responsibility Level: Five Environmental, Social, Government (ESG) qualitative dimensions. Namely, employee relations, diversity, product, community and environment;
29, Refinitiv ESG Combined Score: An overall company score based on the reported information in the environmental, social and corporate governance pillars (ESG score) with an ESG controversies overlay;
30, Environmental Pillar Score: A company's impact on living and non-living natural systems, including the air, land and water, as well as complete ecosystems;
31, Social Pillar Score: A company's capacity to generate trust and loyalty with its workforce, customers and society, through its use of best management practices;
32, Governance Pillar Score: A company's systems and processes, which ensure that its board members and executives act in the best interests of its long-term shareholders;
33, County Court Judgements: A county court judgment (CCJ) arises from a claim made to the court following the non-payment of unsecured debt (usually trade debts). Where the creditor's claim is upheld by the court, a CCJ is issued. This is an order from the court stating that the debt must be settled. After being issued, either a CCJ is satisfied or it remains outstanding;
34, R&D productivity: Defined as the firm-specific output elasticity of R&D, and is computed by estimating the production function with a random coefficients model that allows for heterogeneity in the output elasticity for R&D. The R&D productivity measure represents the percentage increase in revenue from a 1% increase in R&D expenditure, other things held constant;
35, Knowledge Capital: Accumulated past R&D expenditure;
36, Organisation Capital: Accumulated a fraction of past Selling, General, and Administrative (SG&A) expenditure;
37, News-Corpus Variable: Measured by Distress Intensity of Default-Corpus (DIDC). A higher DIDC indicates a relatively higher intensity of default probability, and vice versa, whilst the level of DIDC also indicates the balance of coverage in the financial media for positive versus negative news;
38, Beta of Stock Return: The beta of stock returns is estimated in a standard technique using the CRSP value weighted return of NYSE/NASDAQ/AMEX index as the market index;
39, Sigma of Stock Return: Measures the history volatility of stocks. Each firm's sigma for year *t* is calculated by regressing each stock's monthly returns in year *t*-1 on the value-weighted NYSE/AMEX index return for the same year;
40, Market Value of Equity: the combined market value of all shares of stock, preferred and common;
41, Distance-to-Default Measure: Measures the distance between the current value of assets and the debt amount in terms of the volatility, that is, the standard deviation of the growth rate, of the assets;
42, Value-at-Risk (VaR): Follows the recommendation of the Basel Accord and considers probability levels of 99% (α=0.01) to estimate downside risk measures. *Semi-parametric Cornish-Fisher VaR* (VaR_{CF}), which considers higher moments in the return distribution, is adopted;
43, Expected Shortfall: Provides the information about the size of the loss in the event when the VaR confidence level is breached. *Cornish-Fisher VaR* is used to estimate *expected shortfall*, denoted as ES_{CF}.

2.A.4 Classification of feature selection methods in BP and FDP

Classification of Feature Selection Methods	Feature Selection Methods	Papers / Authors & Year
Domain Knowledge	Referring to Previous Studies	Beaver (1966), Altman et al. (1977), Ohlson (1980), Zmijewski (1984), Zavgren (1985), Frydman et al. (1985), Odom & Sharda (1990), Hsieh (1993), Bryant (1997), Shumway (2001), Paradi et al. (2004), Cielen et al. (2004), West et al. (2005), Bharath & Shumway (2008), Chen et al. (2009), Hu (2009), Premachandra et al. (2009), Sueyoshi & Goto (2009a), Tseng & Hu (2010), Altman et al. (2010), Gorgani et al. (2010), De Andrés et al. (2011), Hu & Chen (2011), Tinoco & Wilson (2013), Zhou (2013), Piñeiro-Sánchez et al. (2013), Trabelsi et al. (2015), Mousavi et al. (2015), Lu et al. (2015), Zhou et al. (2015), Amendola et al. (2015), Afik et al. (2016), Du Jardin (2016), Barboza et al. (2017), Tobback et al. (2017), Zorn et al. (2017), Ouenniche & Tone (2017), Lian (2017), Ouenniche et al. (2018a), Ouenniche et al. (2018b), Ouenniche et al. (2018c), Lin & Dong (2018), Miao et al. (2018), Andrikopoulos & Khorasgani (2018), Oz & Simga-Mugan (2018), Choi et al. (2018), Ahmadi et al. (2018), Alaka et al. (2018), Li & Faff (2019), Mai et al. (2019), Ouenniche et al. (2019), Muñoz-Izquierdo et al., (2019), Gupta & Chaudhry (2019), Mousavi et al. (2019), Matin et al. (2019), Qiu et al. (2020), Bai & Tian (2020), Ashraf et al. (2020), Almaskati et al. (2021)
	Expert Knowledge	Ahn & Kim (2009), Min & Jeong (2009), Li & Ho (2009), Altman et al. (2010), Bauer & Agarwal (2014)
Filter Methods	Independent Samples *t*-test	Shin et al. (2005), Ahn & Kim (2009), Min & Jeong (2009), Tsai (2009), Li et al. (2011), Liang et al. (2015), Zhou et al. (2015), Liang et al. (2016), Kim et al. (2016), Mousavi et al. (2019), Liang et al. (2020), Shen et al. (2020), Mousavi & Lin (2020), De Bock et al. (2020)
	Fisher *F*-test	Altman (1968), Altman et al. (1977), Li & Sun (2009), Geng et al. (2015), Almamy et al. (2016), Veganzones & Séverin (2018), Zoričák et al. (2020), Du et al. (2020)
	Chi-square Statistics	Chen et al. (2011b), Kou et al. (2021)
	Information Gain (IG)	Chen et al. (2011b), Wang et al. (2014), Alaka et al. (2018), Lin et al. (2019), Wang et al. (2020), Kou et al. (2021)
	Mutual Information (MI) Measure	Sarkar & Sriram (2001)
	Correlation Analysis	Sarkar & Sriram (2001), Paradi et al. (2004), Xu & Zhang (2009), Tsai (2009), Sueyoshi & Goto (2009b), Lin et al. (2011a), Moradi et al. (2013), Barboza et al. (2017), Veganzones & Séverin (2018), Tinoco et al. (2018), Ninh et al. (2018), Serrano-Cinca et al. (2019), Li & Faff (2019), Hosaka (2019), Ahmad (2019), Yuan et al. (2022)
	Pearson Correlation Analysis	Sun & Shenoy (2007), Appiah & Chizema (2016), Bai & Tian (2020), Paraschiv et al. (2021)
	Mann-Whitney Test (Wilcoxon Rank Sum Test)	Tsukuda & Baba (1994), Du Jardin (2016), Du Jardin (2017), Du Jardin (2018), Kim (2018), Shen et al. (2020)
	Variance Inflation Factor (VIF)	Gupta et al. (2015), Zhou et al. (2015), Appiah & Chizema (2016), Sartori et al. (2016), Serrano-Cinca et al. (2019), Sun et al. (2020), Li et al. (2021), Mousavi & Lin (2020)

	Stepwise Multivariate Discriminant Analysis	Altman et al. (1977), Shin et al. (2005), Min & Jeong (2009), Tsai (2009), Li et al. (2009), Li et al. (2011), Cheng et al. (2014), Liang et al. (2016), Kim et al. (2016), Sun et al. (2017), Liang et al. (2020), Sun et al. (2020)
	Stepwise Linear Discriminant Analysis	Taffler (1984), Liang et al. (2015), Nyitrai & Virág (2019)
	Stepwise Logistic Regression	Ahn & Kim (2009), Min & Jeong (2009), Li et al. (2011), Tinoco & Wilson (2013), Lu et al. (2013), Liang et al. (2015), Liang et al. (2016), Sartori et al. (2016), Nyitrai & Virág (2019), Fernández-Gámez et al. (2020), Liang et al. (2020), Ashraf et al. (2020)
	Principal Component Analysis	Taffler (1984), Tsai (2009), Sueyoshi & Goto (2009b), Nanni & Lumini (2009), Lin et al. (2011b), Mousavi et al. (2015), Mousavi et al. (2019), Mousavi & Lin (2020)
	Factor Analysis (Principal axis factoring, maximum likelihood, alpha factoring)	Tsai (2009), Mousavi et al. (2019)
	Tree-Based Filter (Extreme Random Tree)	Zoričák et al. (2020)
	Unsupervised Laplacian Score	Zoričák et al. (2020)
	ReliefF Algorithm	Zoričák et al. (2020), Kou et al. (2021)
	Iterative Relief (I-RELIEF) Algorithm	Lin et al. (2011a)
	Isometric Mapping (ISOMAP)	Lin et al. (2011b), Wang & Wu (2017)
	Laplacian Eigenmaps (LE)	Wang & Wu (2017)
	Locally Linear Embedding (LLE)	Wang & Wu (2017)
	Entropy-based Ranking	Zhou et al. (2015)
	Receiver Operational Curve (ROC)-Based Ranking	Zhou et al. (2015)
Wrapper Methods	Rough Set Theory (RST)-Based Wrapper	Yeh et al. (2010)
	Particle Swarm Optimisation (PSO)-based Wrapper	Chen et al. (2011a), Liang et al. (2015), Uthayakumar et al. (2020)
	Genetic Algorithm (GA)-Based Wrapper	Ahn & Kim (2009), Min & Jeong (2009), Li & Ho (2009), Hu (2009), Chen et al. (2011b), Lin et al. (2014), Liang et al. (2015), Zhou et al. (2015), Liang et al. (2016), Kim et al. (2016), Zelenkov et al. (2017), Chou et al. (2017), Huang & Yen (2019), Lin et al. (2019), Uthayakumar et al. (2020)
	Non-dominated Sorting Genetic Algorithm II (NSGA-II)	De Bock et al. (2020), Kou et al. (2021)
	Ant Colony Optimisation (ACO)-Based Wrapper	Uthayakumar et al. (2020)
	Grey Wolf Optimisation (GWO)-Based Wrapper	Uthayakumar et al. (2020)
	Deep Belief Network (DBN)-Based-Wrapper	Huang & Yen (2019)
	Recursive Feature Elimination (RFE)	Liang et al. (2016), Zoričák et al. (2020)
	Logistic Regression Recursive Feature Elimination Cross-validation (LR-REFCV)	Yuan et al. (2022)
	Sequential Forward Selection	Zhou et al. (2015), Paraschiv et al. (2021)
Embedded Methods	Decision Tree (DT)-Based Embedded Feature Selection	Min & Jeong (2009), Du et al. (2020)
	CART-Based Embedded Feature Selection	Hosaka (2019)
	XGBoost-Based Embedded Feature Selection	Son et al. (2019), Du et al. (2020)
	TreeNet-Based Relative Variable Importance	Jones & Wang (2019)
	Least Absolute Shrinkage and Selection Operator (LASSO)	Tian et al. (2015), Amendola et al. (2015), Pereira et al. (2016), Volkov et al. (2017), Tian & Yu (2017), Wang et al. (2018), Du et al. (2020), Bai & Tian (2020), Paraschiv et al. (2021)
	Sparse Group LASSO	Wang et al. (2020)
	Ridge Regression	Pereira et al. (2016)
	Generalised Additive Model Selection (GAMSEL)	Valencia et al. (2019)

	Non-linear Subspace Multiple-kernel Machine (NS-MKL)	Zhang & Hu (2016)
	Artificially Synthetic Features	Zięba et al. (2016)
Not Mentioned/ Not Used		Tsai & Cheng (2012), Tsai & Hsu (2013), Cleofas-Sánchez et al. (2016), Antunes et al. (2017), Le et al. (2018), Le et al. (2019), Chen et al. (2020), Alam et al. (2021), Habermann & Fischer (2023)

2.A.5 Classification of performance criteria & measures in the BP and FDP problem

Performance Criterion	Measures of Performance Criterion	Definition and Introduction	Papers / Authors & Year
Correctness of Categorical Prediction	Type-I & Type-II Errors	In the context of BP and FDP, a Type-I error occurs when a prediction model incorrectly classifies a financially healthy company as in financial distress or predicts bankruptcy when it is not the case. In other words, it is a false positive result where the model wrongly identifies a healthy company as unhealthy. On the other hand, a Type II error occurs when a prediction model fails to identify a financially distressed company, incorrectly classifies it as healthy, or predicts no bankruptcy when the company is actually in financial distress.	Beaver (1966), Altman (1968), Altman et al. (1977), Ohlson (1980), Taffler (1984), Zavgren (1985), Frydman et al. (1985), Hsieh (1993), Tsukuda & Baba (1994), Bryant (1997), Cielen et al. (2004), Chen et al. (2009), Tsai (2009), Premachandra et al. (2009), Sueyoshi & Goto (2009a), Sueyoshi & Goto (2009b), Nanni & Lumini (2009), Yeh et al. (2010), Chen et al. (2011a), Chen et al. (2011b), Lin et al. (2011a), Lin et al. (2011b), Tsai & Cheng (2012), Tsai & Hsu (2013), Piñeiro-Sánchez et al. (2013), Lu et al. (2013), Lin et al. (2014), Wang et al. (2014), Trabelsi et al. (2015), Liang et al. (2015), Lu et al. (2015), Amendola et al. (2015), Mousavi et al. (2015), Liang et al. (2016), Almamy et al. (2016), Sartori et al. (2016), Pereira et al. (2016), Du Jardin (2016), Barboza et al. (2017), Antunes et al. (2017), Zelenkov et al. (2017), Chou et al. (2017), Du Jardin (2017), Ouenniche & Tone (2017), Ouenniche et al. (2018a), Ouenniche et al. (2018c), Oz & Simga-Mugan (2018), Du Jardin (2018), Wang et al. (2018), Kim (2018), Serrano-Cinca et al. (2019), Huang & Yen (2019), Mousavi et al. (2019), Ouenniche et al. (2019), Lin et al. (2019), Fernández-Gámez et al. (2020), Liang et al. (2020), Mousavi & Lin (2020), Du Jardin (2021b), Almaskati et al. (2021), Yuan et al. (2022)
	Misclassification Rate or Cost	The Misclassification Rate or Misclassification Cost in the context of BP and FDP refers to the rate or cost at which a prediction model or classifier incorrectly classifies companies into different financial health categories. The overall misclassification rate or cost is a combination of both false positive and false negative rates and provides a comprehensive measure of the model's performance in predicting financial distress and bankruptcy. However, it is important to note that the specific costs associated with misclassification can vary depending on the context and the consequences of misjudging a company's financial health.	Altman et al. (1977), Frydman et al. (1985), Hsieh (1993), Chen et al. (2009), Chen et al. (2011b), Lu et al. (2013), Bauer & Agarwal (2014), Trabelsi et al. (2015), Lu et al. (2015), Amendola et al. (2015), Mousavi et al. (2015), Liang et al. (2016), Zięba et al. (2016), Kim et al. (2016), Serrano-Cinca et al. (2019), Liang et al. (2020), De Bock et al. (2020), Almaskati et al. (2021)
	Sensitivity/True Positive Rate/Recall & Specificity/True Negative Rate	Sensitivity (or, True Positive Rate/Recall) measures the ability of a prediction model to correctly identify financially distressed or bankrupt companies (the positive class). It is the ratio of true positives (correctly identified unhealthy companies) to the total number of actual unhealthy companies. $Sensitivity = \frac{True\ Positives}{True\ Positives + False\ Negatives}$ Specificity (or, True Negative Rate) measures the ability of a prediction model to correctly identify financially healthy companies (the negative class). It is the ratio of true negatives (correctly identified healthy companies) to the total number of actual healthy companies. $Specificity = \frac{True\ Negatives}{True\ Negatives + False\ Positives}$	Li & Ho (2009), Zhou (2013), Tinoco & Wilson (2013), Zhou et al. (2015), Geng et al. (2015), Mousavi et al. (2015), Cleofas-Sánchez et al. (2016), Barboza et al. (2017), Antunes et al. (2017), Zelenkov et al. (2017), Ouenniche & Tone (2017), Veganzones & Séverin (2018), Ouenniche et al. (2018a), Ouenniche et al. (2018c), Huang & Yen (2019), Mousavi et al. (2019), Ouenniche et al. (2019), Fernández-Gámez et al. (2020), Sun et al. (2020), Mousavi & Lin (2020), Uthayakumar et al. (2020), Almaskati et al. (2021)
	Precision/Positive Predictive Value	Precision (or, Positive Predictive Value) measures the ability of a prediction model to correctly identify financially distressed or bankrupt companies (the positive class) among the ones it predicts as positive cases. Precision is the ratio of true positives (correctly identified unhealthy companies) to the total number of companies predicted as distressed (both true positives and false positives).	Li & Ho (2009), Geng et al. (2015), Antunes et al. (2017), Zelenkov et al. (2017), Ahmadi et al. (2018), Fernández-Gámez et al. (2020), Sun et al. (2020), Mousavi & Lin (2020), Putra et al. (2020), Du et al. (2020), Almaskati et al. (2021)

		$$Specificity = \frac{True\ Positives}{True\ Positives + False\ Positives}$$	
	Geometric Mean of TPR & TNR	The Geometric Mean of True Positive Rate (TPR) and True Negative Rate (TNR) is a metric used to evaluate the overall performance of a prediction model. It is a balanced measure that takes into account both sensitivity and specificity. $$G-Mean = \sqrt{TPR \times TNR}$$	Cleofas-Sánchez et al. (2016), Kim et al. (2016), Veganzones & Séverin (2018), Le et al. (2019), Zoričák et al. (2020), Sun et al. (2020), Shen et al. (2020), Yuan et al. (2022)
	Harmonic Mean of Precision & Recall (F1 Score, F2 Score)	The Harmonic Mean of Precision and Recall, often referred to as the F1 Score, combines precision and recall into a single metric to provide a balanced evaluation of a prediction model's performance. A higher F1 score indicates a better overall balance between precision and recall. $$F1\ Score = \frac{2 \times Precision \cdot Recall}{Precision + Recall}$$ The F2 score is a variation of the F1 score that places more weight on recall than precision. $$F1\ Score = \frac{(1+\beta^2) \cdot Precision \cdot Recall}{(\beta^2 \cdot Precision) + Recall}$$ β is a parameter that controls the balance between precision and recall. A higher β value gives more weight to recall, making the metric more sensitive to false negatives.	Zhou (2013), Volkov et al. (2017), Antunes et al. (2017), Ahmadi et al. (2018), Hosaka (2019), Uthayakumar et al. (2020), Sun et al. (2020), Shen et al. (2020), Mousavi & Lin (2020), Putra et al. (2020), Du Jardin (2021a)
	Classification Accuracy	Classification Accuracy measures the overall correctness of a prediction model's classifications across all classes. In BP and FDP, it measures the proportion of companies correctly classified as bankruptcy (or, financially distressed) or not. A higher accuracy score indicates that the model is making more correct predictions. $$Accuracy = \frac{Number\ of\ correct\ predictions}{Total\ number\ of\ predictions}$$	Altman (1968), Altman et al. (1977), Odom & Sharda (1990), Bryant (1997), Sarkar & Sriram (2001), Paradi et al. (2004), Cielen et al. (2004), Shin et al. (2005), West et al. (2005), Sun & Shenoy (2007), Ahn & Kim (2009), Min & Jeong (2009), Chen et al. (2009), Hu (2009), Li & Sun (2009), Li et al. (2009), Li & Ho (2009), Premachandra et al. (2009), Xu & Zhang (2009), Tsai (2009), Sueyoshi & Goto (2009a), Sueyoshi & Goto (2009b), Nanni & Lumini (2009), Yeh et al. (2010), Tseng & Hu (2010), Gorgani et al. (2010), Chen et al. (2011a), Lin et al. (2011a), Lin et al. (2011b), Li et al. (2011), Hu & Chen (2011), Tsai & Cheng (2012), Tsai & Hsu (2013), Zhou (2013), Tinoco & Wilson (2013), Moradi et al. (2013), Piñeiro-Sánchez et al. (2013), Lu et al. (2013), Lin et al. (2014), Wang et al. (2014), Liang et al. (2015), Lu et al. (2015), Zhou et al. (2015), Geng et al. (2015), Mousavi et al. (2015), Liang et al. (2016), Cleofas-Sánchez et al. (2016), Almamy et al. (2016), Zhang & Hu. (2016), Sartori et al. (2016), Kim et al. (2016), Pereira et al. (2016), Du Jardin (2016), Barboza et al. (2017), Volkov et al. (2017), Antunes et al. (2017), Zelenkov et al. (2017), Sun et al. (2017), Wang & Wu (2017), Chou et al. (2017), Du Jardin (2017), Tinoco et al. (2018), Oz & Simga-Mugan (2018), Du Jardin (2018), Alaka et al. (2018), Wang et al. (2018), Ahmadi et al. (2018), Kim (2018), Serrano-Cinca et al. (2019), Jones & Wang (2019), Son et al. (2019), Huang & Yen (2019), Ahmad (2019), Mousavi et al. (2019), Muñoz-Izquierdo et al. (2019), Fernández-Gámez et al. (2020), Chen et al. (2020), Uthayakumar et al. (2020), Du et al. (2020), Ashraf et al. (2020), Alam et al. (2021), Putra et al. (2020), Du Jardin (2021a), Du Jardin (2021b)
	Cohen's Kappa Value	Cohen's Kappa Value provides a robust measure of the prediction model's performance that considers both correct predictions and the potential for chance agreement. It is particularly useful for evaluating models in situations where class imbalances or skewed class distributions exist. A higher Kappa value indicates better agreement between the model's predictions and the true labels. $$\kappa = \frac{P_0 - P_e}{1 - P_e}$$ P_0 represents the observed agreement, P_e represents the expected agreement. $$P_e = \frac{(TP_r + TN_r)\cdot(TP_r + FP_r)}{N^2} + \frac{(FN_r + FP_r)\cdot(TN_r + FN_r)}{N^2}$$ where TP_r is the total number of predicted true positives; TN_r is the total number of predicted true negatives; FP_r is the total number of predicted false positives; FN_r is the total number of predicted false negatives; N is the total number of instances.	Ahmadi et al. (2018), Uthayakumar et al. (2020), Shen et al. (2020), Mousavi & Lin (2020)
	Matthews Correlation	Matthews Correlation Coefficient measures the quality of a prediction model's classifications, taking into account both true and false positives and	Uthayakumar et al. (2020), Putra et al. (2020)

	Coefficient (MCC)	negatives. In the context of BP and FDP, MCC assesses the model's ability to correctly identify companies in distress whilst minimising incorrect classifications. It is particularly useful when dealing with imbalanced datasets and is less sensitive to class distribution. $MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$ The MCC value ranges from -1 to 1, where MCC= 1, indicates perfect classification by the model. MCC= 0, indicates random classification. MCC= −1, indicates complete disagreement between the model's predictions and the true labels.	
Discriminatory Power	Area Under Receiver Operating Characteristic Curve (AUC)	The Area Under ROC is a robust metric for evaluating the discriminative power of the BP and FDP models. The ROC curve is a graphical representation of a prediction model's performance (i.e., True Positive Rate vs. False Positive Rate) as the discrimination threshold is varied, and the Area Under ROC (AUC) quantifies the overall ability of the model to distinguish between the two classes. The AUC is calculated by finding the area under the ROC curve and measures the ability of the model to discriminate between the two classes across all possible threshold values. The AUC values range from 0 to 1. A model with AUC = 1 indicates a perfect discrimination, where it can perfectly distinguish between the two classes; A model with AUC = 0.5 indicates that the model's performance is no better than random guess; A model with AUC between 0.5 and 1 indicates varying degrees of discrimination, with higher values indicating better discrimination.	Altman et al. (2008), Nanni & Lumini (2009), Altman et al. (2010), Chen et al. (2011a), Tinoco & Wilson (2013), Zhou (2013), Bauer & Agarwal (2014), Cheng et al. (2014), Gupta et al. (2015), Zhou et al. (2015), Tian et al. (2015), Amendola et al. (2015), Mousavi et al. (2015), Liang et al. (2016), Zięba et al. (2016), Afik et al. (2016), Kim et al. (2016), Du Jardin (2016), Barboza et al. (2017), Tobback et al. (2017), Volkov et al. (2017), Antunes et al. (2017), Tian & Yu (2017), Du Jardin (2017), Veganzones & Séverin (2018), Le et al. (2018), Miao et al. (2018), Andrikopoulos & Khorasgani (2018), Ninh et al. (2018), Lin & Dong (2018), Du Jardin (2018), Choi et al. (2018), Wang et al. (2018), Alaka et al. (2018), Serrano-Cinca et al. (2019), Li & Faff (2019), Le et al. (2019), Hosaka (2019), Mai et al. (2019), Jones & Wang (2019), Gupta & Chaudhry (2019), Son et al. (2019), Nyitrai & Virág (2019), Ahmad (2019), Matin et al. (2019), Mousavi et al. (2019), Valencia et al. (2019), Zoričák et al. (2020), Wang et al. (2020), Shen et al. (2020), Mousavi & Lin (2020), Putra et al. (2020), Du et al. (2020), De Bock et al. (2020), Bai & Tian (2020), Ashraf et al. (2020), Li et al. (2021), Paraschiv et al. (2021), Du Jardin (2021a), Du Jardin (2021b), Almaskati et al. (2021), Kou et al. (2021), Yuan et al. (2022)
	Kolmogorov-Smirnov (K-S) Statistic	The Kolmogorov-Smirnov (K-S) Statistic measures the maximum vertical distance between the cumulative distribution functions (CDFs) of the observed and expected distributions. In the context of BP and FDP model evaluation, the K-S statistic can help assess how well a prediction model's predicted probabilities align with the expected distribution of probabilities for the positive class. $K - S\ Statistic = max\lvert F(x) - G(x) \rvert$ Where $F(x)$ is the CDF of the observed data; $G(x)$ is the CDF of the theoretical or expected distribution.	Tinoco & Wilson (2013), Bauer & Agarwal (2014), Gupta et al. (2015), Mousavi et al. (2015), Mousavi et al. (2019), Jones & Wang (2019), Mousavi & Lin (2020), Li et al. (2021), Almaskati et al. (2021)
	Gini Index	The Gini Index quantifies the inequality or impurity in a set of values. In the context of BP and FDP, the Gini Index measures how well a prediction model separates the positive class (i.e., financially distressed companies) from the negative class (i.e., healthy companies). The Gini Index is closely related to the AUC-ROC, which typically ranges from 0 to 1. $Gini\ = 2 \times AUC - 1$ A higher Gini Index suggests that the model is better at distinguishing between the two classes, resulting in a better separation of positive and negative observations. A Gini Index closer to 1 indicates better model performance.	Tinoco & Wilson (2013), Mousavi et al. (2015), Gupta et al. (2015), Mousavi et al. (2019), Mousavi & Lin (2020), Almaskati et al. (2021)
	Cumulative Accuracy Profile (CAP) Curve & Related Statistic, i.e., Accuracy Ratio	Accuracy Ratio is a single measure derived from the Cumulative Accuracy Profile (CAP) curve. It quantifies the performance of the prediction model in terms of its ability to discriminate between positive and negative instances. In the context of financial distress and bankruptcy prediction, the CAP curve and Accuracy Ratio provide insights into how well the prediction model ranks and identifies financially distressed companies, allowing one to evaluate its discriminatory power more comprehensively. It compares the area between the CAP curve and the random model (a 45-degree diagonal line) to the area between the perfect model (a step curve that goes straight up to 100% accuracy) and the random model.	Amendola et al. (2015), Kim et al. (2016), Li & Faff (2019), Mai et al. (2019), Almaskati et al. (2021)

		$$Accuracy\ Ratio = \frac{Observed\ Area}{Maximum\ Possible\ Area}$$ A higher Accuracy Ratio suggests that the model's predictions are more effective in separating positive and negative instances than random chance.	
Calibration Accuracy	Brier Score (BS)	Brier Score quantifies the accuracy of predicted probabilities generated by the BP and FDP model. It measures the mean squared difference between the predicted probabilities and the actual binary outcomes. $$Brier\ Score = \frac{1}{N}\sum_{i=1}^{N}(P_i - O_i)^2$$ Where P_i is the predicted probability for the positive class for the ith instance; O_i is the actual binary outcome (0 or 1) for the ith instance. A lower Brier Score indicates that the prediction model's predicted probabilities are closer to the actual outcomes and reflects better model calibration.	Lin et al. (2011a), Lin et al. (2014), Mousavi et al. (2015), Tian et al. (2015), Mousavi & Lin (2020)
	Log-Loss Score	Log-Loss Score measures the negative log-likelihood of the true binary outcomes (0 or 1) given the predicted probabilities. It quantifies how well the model's predicted probabilities align with the actual binary outcomes, with lower scores indicating better calibration and confidence in the probability estimates. $$LogLos = -\frac{1}{N}\sum_{i=1}^{N}(y_i \cdot \log(p_i) + (1 - y_i) \cdot \log(1 - p_i))$$ The Log-Loss Score typically ranges from 0 to ∞. A lower Log-Loss Score indicates better model calibration and higher confidence in the probability estimates. In the context of BP and FDP, a lower Log-Loss Score suggests that the prediction model's predicted probabilities provide accurate and well-calibrated estimates of the likelihood of financial distress.	Matin et al. (2019)
	Decile Rankings (Rate of actual positive cases in each decile, where deciles are determined based on the probability of bankruptcy)	Decile Rankings is a technique used to assess the effectiveness of a prediction model in ranking instances based on their predicted probabilities. This measure involves dividing the dataset into ten equal groups, or in another word, deciles, based on the predicted probabilities of a positive outcome. For BP and FDP problems, it helps to evaluate how well the model's predicted probabilities discriminate between positive and negative cases (e.g., healthy and bankrupt companies). The measure includes the following steps: 1, Use the prediction model to generate predicted probabilities for each instance in the dataset for the positive class (e.g., the likelihood of financial distress or bankruptcy). 2, Sort the instances in descending order of their predicted probabilities, which means the instances with the highest predicted probabilities are at the top, and those with the lowest predicted probabilities are at the bottom. 3, Divide the sorted dataset into ten equal groups, with each group containing an approximately equal number of instances. 4, For each decile, calculate the rate or proportion of actual positive cases. This is typically done by counting the number of positive cases (true positives) within each decile and dividing it by the total number of instances in that decile. 5, Analyse and compare the rates of actual positive cases across the deciles.	Shumway (2001), Bharath & Shumway (2008), Tian et al. (2015), Tian & Yu (2017), Miao et al. (2018), Mai et al. (2019), Bai & Tian (2020), Paraschiv et al. (2021)
Informational Efficiency / Information Content	Akaike Information Criterion (AIC)	The Akaike Information Criterion (AIC) is a statistical measure used for model selection and model comparison. The AIC quantifies the quality of a statistical model whilst considering the balance between the model's ability to fit the data well (goodness of fit) and its complexity. $$AIC = -2 \cdot \log \text{likelihood} + 2 \cdot number\ of\ parameters$$ where log-likelihood quantifies the goodness of fit of the model, and the number of parameters represents the complexity of the model. The primary goal of using AIC is to select a model with the lowest AIC value among a set of candidate prediction models. A lower AIC value suggests a better trade-off between model fit and complexity.	Tian et al. (2015), Appiah & Chizema (2016), Tian & Yu (2017), Paraschiv et al. (2021), Almaskati et al. (2021)

	Schwarz Criterion (SC)	The Schwarz Criterion (SC), also known as the Bayesian Information Criterion (BIC), is a statistical measure for model selection and comparison. BIC quantifies the quality of a statistical model whilst considering the balance between the fit of the model and the complexity of the model. It is particularly useful for choosing among a set of competing models. $BIC = -2 \cdot \log \text{likelihood} + \log(n) \cdot number\ of\ parameters$ Where n represents the sample size. As with the AIC, a lower BIC value indicates a better trade-off between model fit and complexity.	Appiah & Chizema (2016)
	Log-likelihood Statistic	Log-likelihood statistic is a measure used in statistical modelling and hypothesis testing to assess the goodness of fit of a model to a given set of data. It is calculated on the basis of the likelihood function, where the likelihood function measures the probability of observing the given data under a specific statistical model. For a given statistic model with parameters θ, the log-likelihood statistic $L(\theta)$ is calculated as the natural logarithm of the likelihood function. A higher log-likelihood value indicates a better fit of the prediction model to the data. It suggests that the model's parameter estimates provide a good explanation of the observed instances.	Appiah & Chizema (2016), Jones & Wang (2019), Miao et al. (2018)
	Friedman Test	The Friedman Test is a non-parametric statistical test used to determine whether there are statistically significant differences among multiple related groups. In the context of BP and FDP, the Friedman test is used to determine whether there are statistically significant differences in the prediction performance among multiple prediction models. $F = \frac{12}{k(k+1)} \sum_{i=1}^{k} R_i^2 - 3k(k+1)$ where k is the number of models, R_i is the rank of ith model.	Son et al. (2019)
	McFadden's Pseudo-R Square	In the context of financial distress and bankruptcy prediction, McFadden's Pseudo-R Square can be used to evaluate the goodness of fit of a logistic regression model or a choice model that predicts the likelihood of financial distress or bankruptcy. It quantifies how well the model explains the variation in the dependent variable relative to a null or baseline model. $R^2_{McFadden} = 1 - \frac{Loglikelihood\ of\ the\ model}{Loglikelihood\ of\ the\ null\ model}$ A higher McFadden's Pseudo-R Square suggests a better-fitting model with greater explanatory power.	Paraschiv et al. (2021)

2.A.6 Markets of analysis covered in the surveyed literature on BP & FDP

Category of Market	Market	Literature
		Panel A: Bankruptcy focused studies
Developed Markets	USA	Beaver (1966), Altman (1968), Altman et al. (1977), Ohlson (1980), Zavgren (1985), Odom & Sharda (1990), Hsieh (1993), Bryant (1997), Shumway (2001), Paradi et al. (2004), West et al. (2005), Sun & Shenoy (2007), Bharath & Shumway (2008), Chen et al. (2009), Hu (2009), Premachandra et al. (2009), Tsai (2009), Tsai & Hsu (2013), Zhou (2013), Cheng et al. (2014), Wang et al. (2014), Trabelsi et al. (2015), Tian et al. (2015), Afik et al. (2016), Barboza et al. (2017), Zorn et al. (2017), Volkov et al. (2017), Lin & Dong (2018), Li & Faff (2019), Mai et al. (2019), Qiu et al. (2020), Bai & Tian (2020), Habermann & Fischer (2023), Almaskati et al. (2021)
	UK	Taffler (1984), Altman et al. (2008), Tseng & Hu (2010), Bauer & Agarwal (2014), Gupta et al. (2015), Mousavi et al. (2015), Appiah & Chizema (2016), Almamy et al. (2016), Tobback et al. (2017), Tian & Yu (2017), Ouenniche & Tone (2017), Ouenniche et al. (2018a), Ouenniche et al. (2018b), Ouenniche et al. (2018c), Ouenniche et al. (2019)
	Japan	Tsukuda & Baba (1994), Xu & Zhang (2009), Tsai (2009), Sueyoshi & Goto (2009a), Sueyoshi & Goto (2009b), Nanni & Lumini (2009), Tsai & Cheng (2012), Zhou (2013), Antunes et al. (2017), Tian & Yu (2017), Hosaka (2019)
	France	Chen et al. (2011b), Du Jardin (2016), Tian & Yu (2017), Du Jardin (2017), Veganzones & Séverin (2018), Du Jardin (2018), De Bock et al. (2020), Du Jardin (2021a), Du Jardin (2021b)
	S. Korea	Shin et al. (2005), Ahn & Kim (2009), Min & Jeong (2009), Kim et al. (2016), Le et al. (2018), Le et al. (2019), Son et al. (2019)
	Taiwan	Yeh et al. (2010), Hu & Chen (2011), Liang et al. (2016), Chou et al. (2017), Lin et al. (2019), Chen et al. (2020)
	Germany	West et al. (2005), Tsai (2009), Nanni & Lumini (2009), Tsai & Cheng (2012), Antunes et al. (2017), Tian & Yu (2017), Lin et al. (2019)

	Spain	De Andrés et al. (2011), Piñeiro-Sánchez et al. (2013), Muñoz-Izquierdo et al., (2019)
	Belgium	Cielen et al. (2004), Tobback et al. (2017), Jones & Wang (2019), De Bock et al. (2020)
	Australia	West et al. (2005), Tsai (2009), Nanni & Lumini (2009), Tsai & Cheng (2012), Antunes et al. (2017), Lin et al. (2019)
	Italy	Sartori et al. (2016), De Bock et al. (2020)
	Netherlands	Putra et al. (2020)
	Norway	Paraschiv et al. (2021)
	European Union	Serrano-Cinca et al. (2019), Fernández-Gámez et al. (2020)
Developing Markets	Poland	Chen et al. (2011a), Zięba et al. (2016), Nyitrai & Virág (2019), Alam et al. (2021)
	Iran	Gorgani et al. (2010), Moradi et al. (2013)
	Hungary	Nyitrai & Virág (2019)
	China	Li et al. (2011), Kou et al. (2021)
	Russia	Zelenkov et al. (2017)
	Colombia	Valencia et al. (2019)
	Slovak	Zoričák et al. (2020)
Panel B: Studies focused on Financial Distress		
Developed Markets	USA	Zmijewski (1984), Frydman et al. (1985), Sarkar & Sriram (2001), Cleofas-Sánchez et al. (2016), Lian (2017), Miao et al. (2018), Kim (2018), Gupta & Chaudhry (2019), Liang et al. (2020)
	UK	Altman et al. (2010), Tinoco & Wilson (2013), Tinoco et al. (2018), Andrikopoulos & Khorasgani (2018), Alaka et al. (2018), Mousavi et al. (2019), Ashraf et al. (2020)
	Taiwan	Li & Ho (2009), Lin et al. (2011a), Lu et al. (2013), Lin et al. (2014), Liang et al. (2015), Lu et al. (2015), Huang & Yen (2019)
	Germany	Liang et al. (2015), Cleofas-Sánchez et al. (2016), Ahmadi et al. (2018), Uthayakumar et al. (2020)
	Australia	Liang et al. (2015), Cleofas-Sánchez et al. (2016), Uthayakumar et al. (2020)
	Italy	Amendola et al. (2015)
	Japan	Cleofas-Sánchez et al. (2016)
	S. Korea	Choi et al. (2018)
	Denmark	Matin et al. (2019)
Developing Markets	China	Li & Sun (2009), Li et al. (2009), Liang et al. (2015), Zhou et al. (2015), Geng et al. (2015), Zhang & Hu (2016), Sun et al. (2017), Wang & Wu (2017), Wang et al. (2018), Sun et al. (2020), Wang et al. (2020), Shen et al. (2020), Li et al. (2021), Mousavi & Lin (2020), Du et al. (2020), Yuan et al. (2022)
	Poland	Uthayakumar et al. (2020)
	Iran	Cleofas-Sánchez et al. (2016)
	Vietnam	Ninh et al. (2018)
	Pakistan	Ashraf et al. (2020)
	Malaysia	Ahmad (2019)

2.A.7 The databases used in current literature in BP and FDP study

Type of Provider or Database	Provider	Database Name	Literature
Commercial Databases Providers	Moody's	Moody's Industrial Manual	Beaver (1966), Altman (1968), Altman et al. (1977), Ohlson (1980), Odom & Sharda (1990), Hu (2009), Wang et al. (2014)
	Standard & Poor's	COMPUSTAT	Zmijewski (1984), Zavgren (1985), Frydman et al. (1985), Hsieh (1993), Bryant (1997), Shumway (2001), West et al. (2005), Sun & Shenoy (2007), Bharath & Shumway (2008), Chen et al. (2009), Tinoco & Wilson (2013), Zhou (2013), Wang et al. (2014), Tian et al. (2015), Afik et al. (2016), Barboza et al. (2017), Lian (2017), Tian & Yu (2017), Miao et al. (2018), Lin & Dong (2018), Kim (2018), Li & Faff (2019), Mai et al. (2019), Gupta & Chaudhry (2019), Liang et al. (2020), Qiu et al. (2020), Almaskati et al. (2021)
	LSEG	Datastream	Xu & Zhang (2009), Tseng & Hu (2010), Bauer & Agarwal (2014), Ouenniche & Tone (2017), Ouenniche et al. (2018a), Ouenniche et al. (2018b), Ouenniche et al. (2018c), Ouenniche et al. (2019), Tinoco et al. (2018), Andrikopoulos & Khorasgani (2018), Oz & Simga-Mugan (2018), Ahmad (2019), Mousavi et al. (2019), Ashraf et al. (2020), Habermann & Fischer (2023)
		Thomson One	Tinoco & Wilson (2013), Appiah & Chizema (2016), Tinoco et al. (2018)
		WorldScope Fundamentals	Appiah & Chizema (2016), Tinoco et al. (2018)
	Bloomberg		Almamy et al. (2016), Oz & Simga-Mugan (2018), Ninh et al. (2018)
	Dun & Bradstreet	Altares	Veganzones & Séverin (2018)
		Dun & Bradstreet Database (List of Bankruptcy Firms)	Beaver (1966)
	Wolters Kluwer	Capital Changes (Capital Changes Reporter)	Zmijewski (1984)
	The Nikkei	Nikkei Annual Corporation Reports	Tsukuda & Baba (1994)
		Nikkei Economic Electronic Databank System (NEEDS)	Sueyoshi & Goto (2009b), Hosaka (2019)

	Bureau van Dijk (BvD)	Aida	Sartori et al. (2016)
		Amadeus	Amendola et al. (2015), Ahmadi et al. (2018), Serrano-Cinca et al. (2019), Fernández-Gámez et al. (2020), Ashraf et al. (2020),
		Bel-first	Tobback et al. (2017), Volkov et al. (2017)
		Diane	Chen et al. (2011b), Du Jardin (2016), Antunes et al. (2017), Du Jardin (2017), Du Jardin (2018), Du Jardin (2021a), Du Jardin (2021b)
		Fame	Bauer & Agarwal (2014), Tobback et al. (2017), Alaka et al. (2018), Andrikopoulos & Khorasgani (2018)
		Orbis	Jones & Wang (2019)
		Sabi	De Andrés et al. (2011), Piñeiro-Sánchez et al. (2013), Muñoz-Izquierdo et al., (2019)
	GuoTaiAn (GTA)	China Stock Market and Accounting Research Database (CSMAR)	Zhou et al. (2015), Geng et al. (2015), Lian (2017), Sun et al. (2017), Wang & Wu (2017), Wang et al. (2018), Sun et al. (2020), Wang et al. (2020), Shen et al. (2020), Li et al. (2021), Mousavi & Lin (2020), Du et al. (2020), Yuan et al. (2022)
	Wind Data Service	Wind-Economic Database	Li et al. (2021), Yuan et al. (2022)
	NICE Investors Service		Choi et al. (2018), Son et al. (2019)
	Exact		Putra et al. (2020)
	New Generation Research Inc.		Paradi et al. (2004)
	Korea Credit Guarantee Fund (KODIT)		Shin et al. (2005)
	Informa Plc	Informa D&B	De Andrés et al. (2011)
	CreditScorer Ltd		Altman et al. (2008), Altman et al. (2010)
Universities & Research Institutes	London Business School (LBS)	London Share Price Database (LSPD)	Tinoco & Wilson (2013), Bauer & Agarwal (2014), Mousavi et al. (2015), Ouenniche & Tone (2017), Ouenniche et al. (2018a), Ouenniche et al. (2018b), Ouenniche et al. (2018c), Tinoco et al. (2018), Ouenniche et al. (2019)
	University of Leeds	The Credit Management Research Centre	Altman et al. (2008), Altman et al. (2010), Gupta et al. (2015)
	University of California, Irvine (UCI)	UCI Machine Learning Repository	Tsai (2009), Nanni & Lumini (2009), Chen et al. (2011a), Tsai & Hsu (2013), Liang et al. (2015), Cleofas-Sánchez et al. (2016), Antunes et al. (2017), Lin et al. (2019), Nyitrai & Virág (2019), Uthayakumar et al. (2020), Chen et al. (2020)
	University of California, Los Angeles (UCLA)	UCLA-LoPucki Bankruptcy Research Database (BRD)	Trabelsi et al. (2015), Zorn et al. (2017), Miao et al. (2018), Liang et al. (2020), Bai & Tian (2020)
	University of California, San Diego (UCSD)	UCSD competition dataset	Tsai (2009), Tsai & Cheng (2012), Tsai & Hsu (2013), Cleofas-Sánchez et al. (2016)
	New York University (NYU)	Altman's Bankruptcy Database	Premachandra et al. (2009), Sueyoshi & Goto (2009a)
	Centre for Research in Security Prices (CRSP)		Shumway (2001), Tian et al. (2015), Afik et al. (2016), Li & Faff (2019), Mai et al. (2019), Bai & Tian (2020), Almaskati et al. (2021),
	Emerging Markets Information Service (EMIS)		Zięba et al. (2016), Alam et al. (2021)
	Wharton Research Data Services		Cheng et al. (2014)
	Institute for Systems and Computer Engineering, Technology and Science (INESCTEC)	Laboratory of Artificial Intelligence and Decision Support (LIAAK)	Tsai & Hsu (2013)
	The PACAP Research Centre		Xu & Zhang (2009)
	Laboratory of Artificial Intelligence and Computer Science at the University of Porto (LIACC)		Tsai (2009)
	Iran Stock Market and Accounting Research Database		Gorgani et al. (2010), Moradi et al. (2013)
Regulators & Government Agencies and Industry Associations	U.S. Securities and Exchange Commission (SEC)	U.S. SEC 10K Report	Zmijewski (1984), Mai et al. (2019)
		U.S. SEC 8K Report	Sun & Shenoy (2007), Almaskati et al. (2021)
	UK Company House		Altman et al. (2008), Altman et al. (2010), Appiah & Chizema (2016)
	Bundesanzeiger (Also: Federal Legal Gazette, Germany)		Ahmadi et al. (2018)
	National Bank of Belgium		Cielen et al. (2004)

	Danish Business Authority	Central Business Register	Matin et al. (2019)
	Dutch Chamber of Commerce		Putra et al. (2020)
	Spanish Bankruptcy Public Registry		Muñoz-Izquierdo et al., (2019)
	Brønnøysund Register Centre (Norway)		Paraschiv et al. (2021)
	National Bureau of Statistics of China		Yuan et al. (2022)
	Superintendence of Companies of Colombia		Valencia et al. (2019)
	World Bank		Ninh et al. (2018), Ashraf et al. (2020),
Commercial Banks/Financial Institutions	A Korean Commercial Bank (no name is provided)		Ahn & Kim (2009)
	A Korean Financial Company (no name is provided)		Le et al. (2018), Le et al. (2019)
	Shandong City Commercial Bank Alliance		Kou et al. (2021)
	An Iranian Private Commercial Bank (no name is provided)		Cleofas-Sánchez et al. (2016)
	KPMG Peat Marwick		Sarkar & Sriram (2001)
Financial Journals	The Wall Street Journal (WSJ)	Wall Street Journal Index	Ohlson (1980), Zmijewski (1984), Shumway (2001)
	Taiwan Economic Journal (TEJ)	TEJ Databank	Yeh et al. (2010), Lin et al. (2011b), Hu & Chen (2011), Lu et al. (2013), Lin et al. (2014), Lu et al. (2015), Liang et al. (2015), Liang et al. (2016), Chou et al. (2017), Lin et al. (2019), Huang & Yen (2019)
	Financial Times (FT)	Extel Database	Tseng & Hu (2010)
Stock Exchanges	Shanghai & Shenzhen Stock Exchange		Li & Sun (2009), Li et al. (2009), Li et al. (2011), Zhang & Hu (2016)
	Karachi Stock Exchange (KSE)		Ashraf et al. (2020)

Chapter 3

A Complex Network Analysis Approach to Bankruptcy Prediction using Company Relational Information-based Drivers

3.1 Introduction

Corporate bankruptcy prediction has become a widely investigated field of research, with a large number of methodologies and techniques that have been put forward to help a variety of stakeholders make sound financial decisions and prevent unnecessary losses. The aftermath of the 2008 global financial crisis saw increasing attention to accurate prediction of corporate bankruptcy, as banks and other financial institutions face a much more volatile economic environment, stricter supervisory regulations, and tougher public scrutiny. Basel III, which was agreed upon in 2009 and came into effect in 2023, has introduced much more refined capital requirements that are more sensitive to the risks of the underlying assets (Basel Committee, 2010). It is an indication that banks and other financial institutions may impose stricter requirements for credit granting, which, undoubtedly, would result in higher financial costs for riskier businesses in the foreseeable future (Bauer et al., 2014). Under this scenario, developing more efficient and effective methodologies for corporate bankruptcy prediction would not only benefit creditors such as banks, by significantly reducing the risk in capital investment and credit management, but also can help ease the systemic risk of the broad economy.

Researchers around the world are constantly developing new methodologies and improving existing ones to achieve more accurate bankruptcy predictions. Dating back to the 1960s, since Beaver (1966) and Altman (1968) first scientifically investigated the field of modern bankruptcy prediction, the methodologies in this field have evolved for more than a half century, with a great number of statistical, probabilistic, and artificial intelligence models being proposed and validated in empirical settings; see §3.2 for an overview of prediction models and methodologies. As discussed in **Chapter 2**, these methodologies encompass a wide array of machine learning and artificial intelligence techniques, such as ensemble models, which have significantly advanced the field. However, the literature also highlights a crucial gap —— whilst the focus has largely been on designing new prediction models or refining existing ones, the role of company relational information, which could reflect interconnectedness between businesses, remains largely underexplored. **Chapter 2** underscores that understanding these connections is key to improving bankruptcy prediction performance, thereby motivating the focus of this chapter.

As to the information with which these prediction models are fed, early studies mainly used accounting information-based drivers, commonly referred to as financial ratios (FRs); then increasingly such information has been supplemented with one or several categories of information such as audit information, corporate governance information, corporate social responsibility (CSR) / environment, social and governance (ESG) information, market information, macroeconomic information, and lately textual information. However, although businesses can be viewed as connected cells of the economy, so far, most research papers have not considered connectivity relationships between businesses. As identified in the **Chapter 2** review, this omission leaves a gap in understanding the role of inter-business relationships, particularly those formed through corporate governance structures such as board interconnections. Addressing this limitation is critical for advancing both theoretical and practical approaches to bankruptcy prediction. We argue that in real-life business settings, businesses are connected in many ways, and these connections should be taken into account in the modelling and prediction of the risk profiles of companies. Therefore, in this research, we shall model and analyse these relationships using the concepts and methods of *complex networks analysis* to devise new drivers.

There are many ways that companies can connect to each other, such as supply chain belonging, transactions, ownership, etc. It is well recognised that the quality of the management team can significantly affect the survival of a business, and it is also very common that different companies may 'share' some of their managers for various reasons (Almaskati et al., 2021). Therefore, it is reasonable to make the following two *assumptions*: (1) the personal experience of managers may drive the risk of the company going bankrupt and (2) companies that are connected to other companies through the board of directors (BoD) may experience contagion effects when a connecting company goes bankrupt. In this study, we intend to investigate the contribution of company relational information data, with an emphasis on connections through the board of directors, to improving bankruptcy prediction. In summary, we intend to investigate whether relational information data from the company can improve the performance of prediction models.

Therefore, the contributions of our work to the current literature are threefold. First, we propose new drivers for bankruptcy prediction by learning algorithms on companies' board of directors' networks using the concept of node embeddings obtained by mapping a large board of directors' network to a lower-dimensional space. The main features of the learning algorithms are their independence from the downstream prediction task, and the representations or node embeddings can be learnt in a purely unsupervised way. Second, we investigate different complex network structures that capture the diversity of connectivity patterns at the board level observed in directors' networks for bankruptcy prediction, namely, *company-company direct connection networks*, *company-director-company bipartite connection networks* and *stacked time-weighted company connection networks*. Our numerical results suggest that all network structures capture valuable information; however, drivers based on a stacked time-weighted company connection network structure significantly improve the performance of prediction models.

The remainder of this chapter unfolds as follows. Section 3.2 provides an overview of the landscape of research in the field of bankruptcy prediction, as well as other relevant progress in company network analysis. Section 3.3 presents our research questions along with their justifications. Section 3.4 presents our methodological proposal along with details on the design of our experiments and the types and sources of data for the empirical evaluation of the new drivers. Section 3.5 reports the empirical findings and provides a detailed analysis of the performance of the prediction models. Finally, Section 6 presents the concluding remarks and possible directions for future research.

3.2 Literature review

3.2.1 Development of bankruptcy prediction models

Corporate bankruptcy prediction is a long-standing topic in the academic literature. In 1966, Beaver (1966) first implemented *financial ratio analysis* in corporate bankruptcy prediction. Later, Altman (1968) followed the idea of financial ratio analysis and proposed the well-known *Z-score* model, which extended *univariate ratio analysis* to *multivariate discriminant analysis (MDA)*. The performance of MDA models was then improved by Ohlson (1980) and Zavgren

(1985) in the 1980s with *logistic regression analysis* (*logit* models), which relax the strict assumptions of MDA. This family of models is still widely used today due to its simplicity and interpretability. The 1990s and early 2000s saw a huge leap in computational efficiency and the emergence of the concept of machine learning and artificial intelligence (Lu, 2019). As a result, new families of bankruptcy prediction methodologies, namely, *artificial neural networks* (Odom & Sharda, 1990), *case-based reasoning* (Bryant, 1997), and *support vector machines* (Shin, 2005), were successfully implemented in the field of bankruptcy prediction and achieved significant performance improvement. Later, the idea of *ensemble learning* gradually dominated the research direction in developing new bankruptcy prediction models in the 2010s (Du Jardin, 2016) with new models, including but not limited to *random forest* (e.g., Barboza et al., 2017; Volkov et al., 2017; Veganzones & Séverin, 2018;), *AdaBoost* (e.g., West et al., 2005; Du Jardin, 2016; Sun et al., 2017), *bootstrap aggregating* (*bagging*) (e.g., Nanni & Lumini, 2009; Du Jardin, 2016; Chen et al., 2020), and *extreme gradient boosting* (e.g., Zięba et al., 2016; Son et al., 2019; Huang & Yen, 2019), which generally beat the prediction accuracy of single-classifier models and significantly improved the robustness of model performance for bankruptcy prediction. Compared to statistical and probabilistic models, machine learning models (e.g., neural networks) can include non-linear relationships between input variables and are able to detect complex patterns of the features other than simple causal links (Geng et al., 2015). However, they face interpretability problems, as most of them are 'black-boxed', and therefore it is much more difficult for researchers to analyse the contribution of input variables to overall model performance (Wang et al., 2018). In addition to the above-mentioned mainstream prediction models, a number of researchers have developed other types of prediction methodologies such as *discrete-time hazard models* (e.g., Shumway, 2001; Gupta et al., 2015), which take the multiperiod time series data into account to make consistent estimations over the years; *data envelopment analysis* (DEA) models (e.g., Paradi et al., 2004; Sueyoshi & Goto, 2009a,b; Ouenniche & Tone, 2017), which are nonparametric machine learning models that classify the samples by computing multi-criteria scores with an efficiency meaning; and *multi-criteria decision analysis* (MCDA) models (e.g., Hu & Chen, 2011; Li et al., 2011; Ouenniche et al., 2018a), which account for multiple conflicting criteria to solve risk class classification problems.

In addition to the development of prediction models, another main research stream, which focusses on the design of new bankruptcy prediction drivers, has been extensively investigated by many researchers. The drivers with which bankruptcy prediction models are fed play an important role in their performance. In the academic literature, drivers used in bankruptcy prediction fall within seven main categories (see Zhao et al., 2024), namely, *accounting information-related drivers* (e.g., Altman, 1968; Taffler, 1984; Bryant, 1997), *audit information-related drivers* (e.g., Piñeiro-Sánchez et al., 2013; Matin et al., 2019; Muñoz-Izquierdo et al., 2019), *corporate governance-related drivers* (e.g., Lin et al., 2010; Appiah et al., 2016; Liang et al., 2020), *corporate social responsibility-related drivers* (e.g., Gupta et al., 2015; Lin & Dong, 2018; Habermann & Fischer, 2021), *market information-related drivers* (e.g., Taffler, 1984; Bharath & Shumway, 2008; Afik et al., 2016), *macroeconomic information-related drivers* (e.g., Lu et al., 2015; Lin & Dong, 2018; Fernández-Gámez et al., 2020), and *media information-related drivers* (e.g., Lu et al., 2013; Ahmadi et al., 2018; Putra et al., 2020). Early studies mainly used accounting information-based drivers, and then increasingly such information has been supplemented with one or several of the above-mentioned categories. However, despite connections between businesses – whether direct or indirect – observed in real-life business settings, so far, most papers do not consider connectivity relationships between businesses. We argue that these connections should be taken into account in the modelling and prediction of the risk profiles of companies. Therefore, in this research, we shall model and analyse these relationships using the concepts and methods of *complex network analysis* to devise new corporate governance drivers.

3.2.2 Corporate Governance Indicators in bankruptcy prediction

Most of the early bankruptcy prediction studies fail to take non-financial information, such as *corporate governance indicators* (CGIs), into account. However, in recent years, there is an increasing trend of using non-financial factors in bankruptcy prediction, as more information and related data are being provided by commercial databases. Corporate governance plays an important role in affecting the probability of bankruptcy of a company, because agency risk and moral hazard can increase if managers sacrifice the company's long-term interests and instead focus on their short-term personal fulfilment (Jensen et al., 1986). The issue of information asymmetry and hidden information, as exemplified by cases such as Enron scandal (Arnold & De Lange, 2004) and the 2008 financial crisis (Liao et al., 2013), highlights the limitations of traditional financial indicators in capturing the full scope of corporate risks. Instances where managers deliberately withhold or distort information to present a more favourable financial outlook underscore the importance of robust corporate governance mechanisms. In such scenarios, non-financial indicators, such as CGIs, play a crucial role in addressing risks associated with moral hazard and agency conflicts. By implementing proper corporate governance measures, a company can reduce managerial information asymmetry (Cormier et al., 2010), prevent the manipulation of company financial statements (Beasley, 1996), and consequently reduce the probability of going bankrupt (Fich & Slezak, 2008).

Therefore, many researchers adopted CGIs in their bankruptcy prediction research and investigated their ability to reflect fundamental risks (e.g., Lu et al., 2015; Appiah et al., 2016; Liang et al., 2020; Mousavi & Lin, 2020). In the remainder of this chapter, we shall refer to these drivers as the conventional CGIs. In this section, we limit our reporting to articles that focus on the value added of CGIs. For example, Lin et al. (2010) investigated the role of non-financial features related to corporate governance in bankruptcy prediction. A total of 42 CGIs were considered and 6 of them were selected by stepwise logistic regression and proved effective in predicting bankruptcy for listed companies in Taiwan. The authors claimed that CGIs provide a promising solution for assessing bankruptcy risk, as the CGI-based model alone achieves similar performance compared to the FR-based model alone. Furthermore, the proposed hybrid model that combines FRs and CGIs provides much higher prediction accuracy than that using FRs alone or CGIs alone. Li et al. (2020) analysed the effectiveness of 33 CGIs in predicting financial distress of Chinese listed companies and claimed that board independence, state ownership, management compensation, and personal characteristics of board members are significantly associated with the risk of financial distress. They also claimed that CGIs alone have limited capacity to detect financial distress, whereas FRs alone can do relatively better. However, this issue can be solved by combining CGIs and FRs together, which provides significantly improved prediction performance. Almaskati et al. (2021) conducted a horse race of bankruptcy prediction models for US listed companies to examine the role of CGIs in improving model performance. They re-estimated seven well-received bankruptcy prediction models (e.g., *discriminant analysis*, Altman, 1968; *logit model*, Ohlson, 1980; *probit model*, Zmijewski, 1984; *hazard analysis*, Shumway, 2001, Bharath & Shumway, 2008 and Campbell et al., 2008; and *DEA additive model*, Premachandra et al., 2009) and compared them with their counterparts augmented with 12 CGIs, which reflect the board characteristics. Empirical results suggest that the addition of CGIs to the different models significantly improves their prediction performance, and the additional explanatory power provided by CGIs improves as the time to bankruptcy increases. This suggests that CGIs can provide earlier and more accurate warnings of the potential for bankruptcy of a company. However, note that the value added of CGIs seems to be market dependent, as suggested by Liang et al. (2016), who looked into the performance of an exhaustive list of CGIs from five categories, namely board structure, ownership structure, cash flow rights, key person retained, and others, in predicting bankruptcy for companies listed on both Chinese and Taiwanese markets. The results suggest that the predictability of the combination of FRs and CGIs is no better than that obtained by using FRs alone for companies in the Chinese market. The authors claim that attention should be paid to the definition of bankruptcy, as well as to the extent to which CGIs relate to the characteristics of companies in the chosen market, which can significantly influence the usefulness of CGIs for empirical prediction of bankruptcy.

3.2.3 Company relational information-based drivers in bankruptcy prediction

Since all companies are embedded in a giant network of broad economies through all kinds of interconnections (e.g., supply chain belonging, transactions, board of directors), it is necessary to evaluate bankruptcy cases comprehensively rather than individually. *Company Relational Information-based drivers* (CRIs) provide information on the relationship and connectivity between business entities. One of the key advantages of CRIs is their potential to uncover collusive behaviour within interconnected networks (Howard, 1954; Asch & Seneca, 1975). Whilst collusive behaviour — such as preferential treatment, information sharing, or unethical alliances — among closely connected entities can amplify systemic risks, network-based approaches are capable of detecting such hidden vulnerabilities. By analysing the structure, strength, and evolution of inter-company relationships, CRIs can identify anomalous patterns, such as overly centralised power dynamics or unusually strong ties between entities, which may indicate collusion.

Several researchers have investigated the effectiveness of CRIs in predicting bankruptcy and other applications using both informal and formal network analyses. In terms of informal network analysis, examples include Fujiwara (2006; 2008), Xu & Zhang (2009), and Lian (2017). To be more specific, early studies by Fujiwara (2006; 2008) suggest that companies can be subject to the 'link effect' in terms of bankruptcy risks in that if a company goes into financial insolvency state or bankruptcy, other companies in its upstream or downstream of the production network may suffer a secondary effect that drives them too into bankruptcy. In fact, despite not using a formal network analysis, Fujiwara (2008) analysed a dataset of bankruptcy cases in Japan and claimed that about 10% of bankruptcy cases are due to the link effect in the supplier-customer network. Also, Xu & Zhang (2009) used the proportion of a company's stock held by its banks as a proxy for the dependence of the company on its banks, and used a 4-level index that reflects the company's characteristics, historical background, source and amounts of bank loans, board composition, company's attitude toward the group, and other external connections as a proxy for the *company's inclination to be in a Keiretsu* or Keiretsu dependence. These two relational variables, along with the accounting and market-based variables used by Altman (1968) and Ohlson (1980), are then fed to a hazard model to test how they contribute to bankruptcy prediction. Their empirical results suggest that these relational variables further improve the ability to predict the bankruptcy of Japanese listed companies. Furthermore, Lian (2017) provides additional evidence that the link effect is important. To be more specific, using a regression framework, the paper shows that financial distress risk transfers from customer firms to supplier firms and that the effect of financially distressed customers on suppliers is persistent up to two years.

Regarding formal network analyses, examples include Tobback et al. (2017), Constantin et al. (2018), Letizia & Lillo (2019) and Kou et al. (2021). To be more specific, Tobback et al. (2017) used information on past and current directors and managers of companies to build a network of SMEs where two companies are linked if they share or have shared a member of the board of directors and/or the management board. Regarding the weights of the links, seven weighting schemas (Stankova et al., 2015) have been tested, and the hyperbolic tangent seems to be the best weighting scheme for the two datasets used, namely the Belgian and UK datasets of SMEs. Then, a smoothed version of the weighted-vote relational neighbour (s-wvRN) classifier (Macskassy & Provost, 2007) is used to transform the relationships between companies in the network thus constructed into bankruptcy prediction scores calculated as the weighted sum of the bankruptcy probabilities of a company's neighbours, where a neighbour's bankruptcy probability is set to either 0 or 1 depending on whether it went bankrupt or not. The authors compared the performance of three separate models, namely, s-wvRN, a linear SVM model fed with financial ratios that have been shown to be good predictors of bankruptcy, a linear SVM model that adds the relational bankruptcy prediction score as an additional variable next to the financial data, and reached the conclusion that relational scores represent valuable information to incorporate in bankruptcy prediction models. Constantin et al. (2018) estimate tail-dependence networks based on equity prices, where network nodes represent banks and a network link between any two banks exists if the standard z-score of the extremal dependence measure (Poon et al., 2004) between any two banks' return series is one. Note that in this application, the measure of extremal dependence is based on the probability of having extreme negative values that occur simultaneously for the series of returns of two banks. Then, the following variables representing network effects and contagion-related information are chosen as drivers alongside bank-specific balance sheet variables, country-specific banking sector variables, and country-specific macro-financial variables: (1) 'Network Dummy': a dummy variable that indicates for each bank whether there are any vulnerable (i.e., distressed) banks to which it is estimated to be connected ('neighbours') through tail dependence; (2) 'Network Sum': a variable that counts how many vulnerable neighbouring banks the bank has in its estimated tail-dependence network; (3) 'Country Dummy': a dummy variable that indicates for each bank whether there are other banks being signalled as vulnerable in the same country; and (4) 'Country Share': the share of vulnerable banks of total banks in the respective country. Constantin et al. (2018) claim that the variables representing network effects and contagion-related information enhance the predictive power of the classifier used, namely, logistic regression. Letizia & Lillo (2019) used transactional data from 2.4 million Italian companies to build monthly payment networks. For each month of 2014, the nodes of the monthly network represent active firms in that month, and a link between any two firms exists if there is a payment between the source and the recipient firms with weight equal to the payment amount. If multiple transactions occur between the same (ordered) pair of nodes, the weight of the link is the sum of the amounts of the payments. Therefore, the monthly payment networks thus constructed are directed and weighted networks. The authors showed that network metrics representing its characteristics (i.e., in- and out-degree; weighted fraction of in- and out-neighbours with a given rating – Hight Medium, Large, or NA; rank of the class in the hierarchy inferred by agony minimisation; membership in community inferred by modularity maximisation, and sum of in- and out-strength) can be informative of the risk of a firm and used them as predictors of the missing ratings fed into several classifiers (i.e., multinomial logistic, classification trees, neural networks). Empirical results suggest that network metrics and community structure can be used successfully to predict missing ratings with machine learning models. Regarding Kou et al. (2021), they constructed three Chinese SME networks, namely, the manager network, shareholder network, and payment network, where nodes represent SMEs, two nodes in the manager network are connected by an edge if they have common managers and the edge weight is the number of common managers, two nodes in the shareholder network are connected by an edge if they have common shareholders and the edge weight is the number of common shareholders, two nodes in the payment network – also referred to as the transaction network – are connected by an arc if they have at least one direct payment interaction or payment from one SME to the other in the last 180 days, and the edge weight is the sum of all payments from one SME to the other in the last 180 days. Thus, the manager and shareholder networks are undirected, and the payment network is directed. The constructed networks were then used for feature generation based on node centrality measures, community measures, and risk propagation measures – see Kou et al. (2021) for details on the list of 98 measures used. Finally, the predictive power of these network-based features (excluding the usual accounting-based features) is evaluated by comparing the classification performance of seven different models, namely, linear discriminant analysis, logistic regression, support vector machine, decision tree, random forest, XGBoost, and neural network, fed with five combinations of features. Their empirical results suggest that network-based information improves the performance of SMEs bankruptcy prediction and show that payment network-based variables have the highest importance weights. Long et al. (2022) proposed a novel framework for evaluating the credit risk of SMEs by quantifying relational risk from publicly available risk events (i.e., administrative penalty and loan dispute) within firm networks. They developed a method to assign a *relational risk score* that incorporates the time-dependent impact of risk events reported by government agencies, such as legal judgments or regulatory penalties. Each firm in the network is assigned a risk score based on its direct and indirect connections to other firms that have experienced risk events. This quantification takes into account the type of network connection (e.g., direct ownership vs. shared management), the nature of the risk event, and its recency, with consideration of the impact of risk events diminishing over time. The relational risk scores are then integrated with traditional credit assessment factors such as financial ratios and demographic characteristics in machine learning models (i.e., Logit, Random Forests, and eXtreme Gradient Boosting). Their empirical findings suggest positive value added of incorporating relational risk into companies' credit evaluations for SMEs.

Using company relational information in bankruptcy prediction can be challenging, as there are many issues that can impede the performance of the models/proposed drivers. For example, using large-scale datasets that contain thousands of companies would introduce high dimensionality in network representation, which could make a graph overly large, hence resulting in high sparsity in feature representation matrices; the complex nature of network structures and diversities in local network patterns may contain valuable hidden information that could be helpful for prediction, however, it is typically difficult to excavate them; processing large-scale networks of company relational information would require high data storage and high computational power; furthermore, it would also set a higher standard on the efficiency and compatibility for the learning models. To address these issues, in this study, we implement graph embedding techniques to reduce the dimensionality and complexity in the representation of the network-based company relational information, and we further exploit and examine that hidden information by proposing new network-based drivers for the downstream prediction tasks.

3.3 Research questions

In this chapter, we intend to address seven research questions related to the implementation of *Company Relational Information-based Drivers* (CRIs) in bankruptcy prediction, on one hand, and the tools used for devising such CRIs, on the other hand. CRIs based on board of directors' connections, which account for the interconnections of business entities through their boards of directors and extracted from complex network structures, provide structural managerial information rather than financial information. These CRIs are another class of CGIs.

Our first set of research questions is concerned with whether CRIs add value to the prediction exercise and can be stated as follows.

(1) Are CRIs based on board of directors' connections a good alternative to conventional CGIs? Or more specifically, do CRIs based on the board of directors' connections provide more predictive value than the conventional CGIs?

(2) Does complementing Financial Ratios (FRs) with CRIs provide improved bankruptcy prediction performance than using FRs alone?

(3) Does complementing Financial Ratios (FRs) with CRIs provide an improved bankruptcy prediction performance than complementing FRs with CGIs?

Our second set of research questions is concerned with whether specific network designs and related CRIs drivers are better at enhancing bankruptcy prediction or not and can be stated as follows:

(4) How do CRIs based on company-company direct connection networks and CRIs based on company-director-company bipartite connection networks compare in enhancing the predictive performance of models?

(5) How CRIs based on company-company direct connection networks (resp. company-director-company bipartite connection networks) and CRIs based on stacked time-weighted company connection networks compare in enhancing the predictive performance of models?

Our third and last set of research questions is concerned with finding out about the ability of specific network designs and related CRIs drivers to predict short- and long-term corporate bankruptcy, and can be stated as follows.

(6) How do CRIs based on different network structures compare in predicting short- and long-term corporate bankruptcy?

(7) Do CRIs based on time-stacked networks predict long-term corporate bankruptcy better than static network structures?

3.4 Methodology proposal

In this section, we present our methodological proposal for engineering features using learning algorithms over the board of directors' networks of companies. In the remainder of this section, we shall first present the four structures of complex company networks based on board members' connections (§4.1). Then we shall present the concept of node embeddings and how to devise them, as well as how to design our board members' network-based drivers (§4.2). **Fig. 3.1** below provides a flowchart depicting the main steps of our methodology from data collection to the use of network-based drivers for bankruptcy prediction including how embeddings are utilised in the downstream task of prediction.

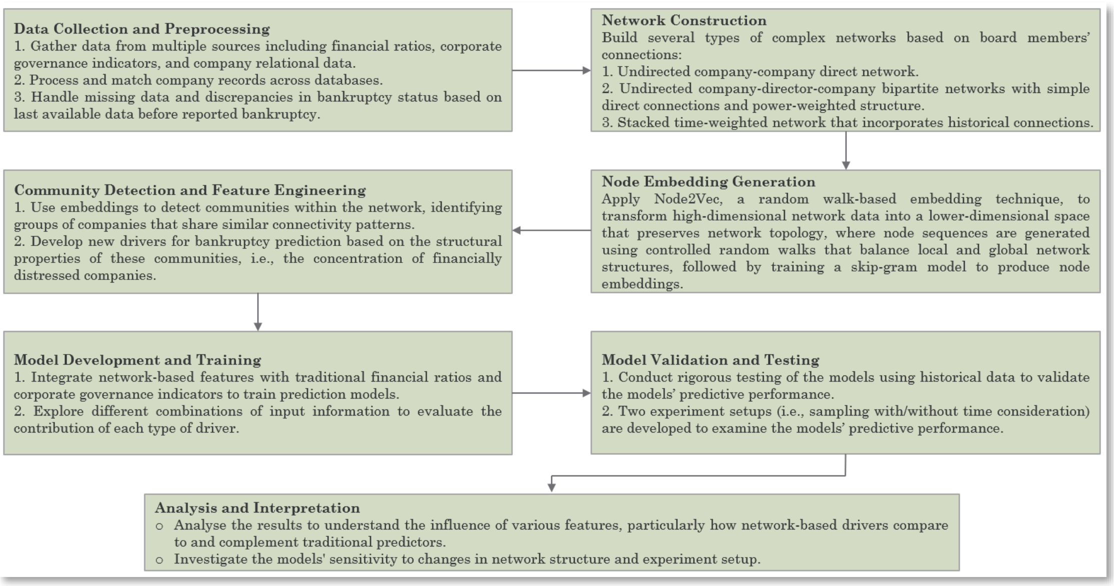

Fig. 3.1 Flowchart illustrating the pipeline of the proposed methodology for bankruptcy prediction

3.4.1 Complex company networks based on board members' connections

In this research, we focus on the design of complex network-based drivers of bankruptcy and financial distress to feed into BP models. In particular, the focus shall be on the connections of board members to other companies / the external world. The rationale is that board members of a given company who sit on the boards of financially distressed or bankrupt firms carry a failure risk through their input to the board. Therefore, we assume that the characteristics and personal experience of managers may drive the risk of a company going bankrupt (Platt & Platt, 2012; Almaskati et al., 2021), and that companies that are connected to other companies through the board of directors may experience contagion effects when a connecting company goes bankrupt (Tobback et al., 2017). To reflect the relationships between board members and companies, three static complex networks, namely, an undirected company-company direct network, an undirected company-director-company bipartite network with simple direct connections, and an undirected company-director-company bipartite network weighted by directors' power, are built based on board members' connections. The undirected company-company direct network and the undirected company-director-company bipartite network with simple direct connections assume that companies sharing board members are directly influenced by each other, as evidenced by Mizruchi (1996) and Kramarz & Thesmar (2013), whilst the weighted bipartite network accounts for director dominance through the board size, as shown by Raheja (2005) and Liang et al. (2020). Additionally, a stacked time-weighted network structure based on company-company direct connections is developed to examine whether considering the time dynamics in company connections has any impact on the performance of prediction models. The stacked time-weighted network structure assumes that historical connections have lasting effects, as evidenced by Tobback et al. (2017). Hereafter, we shall provide the details of each network structure along with its graphical illustration.

a. *Undirected company-company direct network:* Vertices represent companies, edges connecting the vertices represent the board of directors' connections, and the weights of the edges are chosen as the number of common board of directors' members (see **Fig. 3.2** and **Table 3.1**).

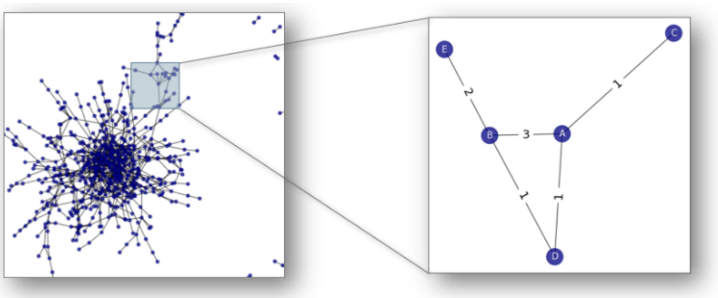

Fig. 3.2 An illustration of undirected company-company direct network

b. *Undirected company-director-company bipartite network with simple direct connections:* Vertices represent both companies (blue vertices) and directors (red vertices) in a bipartite network structure, edges connect companies to directors sitting on their board, and the weights of the edges are all set to be 1 which reflects the simple direct connections between companies and directors (see **Fig. 3.3** and **Table 3.1**).

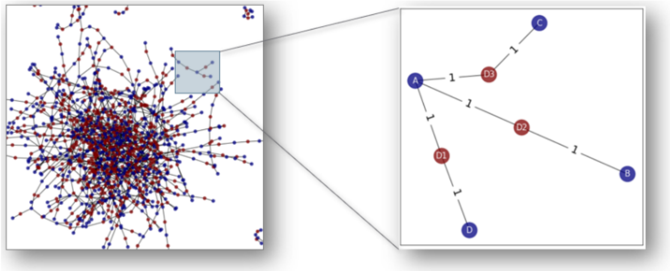

Fig. 3.3 An illustration of undirected company-company direct network

c. *Undirected company-director-company bipartite network weighted by directors' power:* Vertices represent both companies (blue vertices) and directors (red vertices) in a bipartite network structure, edges connect companies to directors sitting on their board, and the weights of the edges represent the power of directors as measured by the inverse of the size of the board of directors on which they sit (see **Fig. 3.4** and **Table 3.1**).

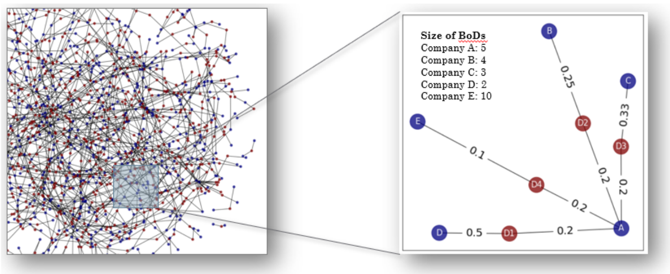

Fig. 3.4 An illustration of undirected company-director-company bipartite network

d. *Undirected stacked time-weighted company-company network:* Vertices represent companies, edges connecting the vertices represent the board of directors' connections, and the weights of the edges of each period network are chosen as the number of joint board of directors' members. A stacked time-weighted network is created by combining the *s* previous years' network connections with the current ones (see **Fig. 3.5**). The weights of previous years' connections are adjusted to reflect the decaying effect as time goes by, based on the concept of 'half-life' (Burton & Kebler, 1960; Knudsen & Lien, 2023). Specifically, we apply a half-life of one year to the influence of these connections, meaning the weight is reduced by half each subsequent year. This decay rate is chosen to realistically model the diminishing influence of past board relationships on current company operations and risk, acknowledging that whilst the impact reduces over time, it remains significant for several years. Details of the specific weighting calculations are provided in **Table 3.2**. This stacked time-weighted structure allows us to explore how historical relationships between companies evolve and their prolonged impact on corporate governance and financial health.

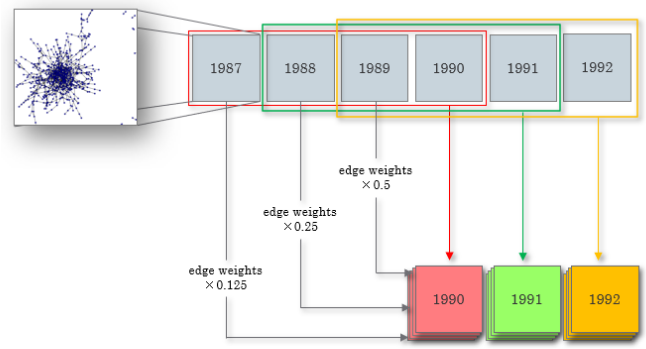

Fig. 3.5 An illustration of undirected stacked time-weighted company-company network

To be more specific, edge weights of the stacked time-weighted network are determined using two different weighting schemas. The *first (respectively, second) weighting scheme* uses the sum (respectively, maximum) of the weights of each pair of company connections across the previous s years to determine the weights in the stacked network. The first approach assumes that a longer duration of connection between companies indicates a stronger relationship; on the contrary, the second approach assumes that the most recent connections indicate more substantial relationships between companies. **Fig. 3.6** provides an illustration of the calculation of edge weights under both weighting schemas along with formulae, and **Table 3.2** provides the pseudocode for computing these weights along with the definition of notation.

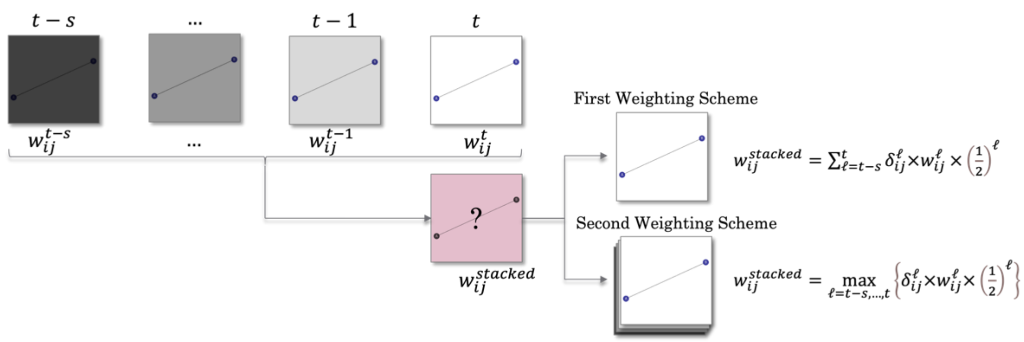

Fig. 3.6 Methods of calculation of edge weights of the stacked time-weighted networks

Table 3.1 Calculation of edge weights for network structures a, b, and c

a, Company-Company Direct Network	
Notation:	
	i, j: Companies
	k: Director
	$\delta_{ij} = 0,1$ depending on whether a link exists between company i and j
	w_{ij} = weight of the edge between company i and j
Weight Calculation Formula:	
	$w_{ij} = \sum_k \delta_{ik} \cdot \delta_{kj}$, where $\delta_{ik} = 1$ if director k sits on the board of company i, otherwise $\delta_{ik} = 0$
b, Bipartite Network with Simple Direct Connections	
Notation:	
	i: Company
	k: Director
	w_{ik} = weight of the edge between company i and director k
Weight Calculation Formula:	
	$w_{ik} = 1 \ \forall i, k$ (indicating direct connection)

c, Bipartite Network Weighted by Directors' Power
Notation:
i: Company
k: Director
d_{ik} = the size of the board of directors of company i on which director k sits
w_{ik} = weight of the edge between company i and director k
Weight Calculation Formula:
$w_{ik} = 1/d_{ik}\ \ \forall i,k$ (indicating direct connection)

Table 3.2 Calculation of edge weights of the stacked time-weighted networks (network structure d)

d, Stacked time-weighted network structure
Parameters: Number of years to stack (s)
Notation:
$\delta_{ij}^{\ell} = 0{,}1$ depending on whether a link exists between company i and j in period ℓ
w_{ij}^{ℓ} = original weights of link (i,j) in period ℓ
First weighting scheme:
For t =1990 to 2019 {
Stack the $s+1$ networks of year $t-s$ to year t;
For each pair of nodes (e.g., companies), say i and j, compute the stacked time-weighted network structure weight of their connection as follows:
$w_{ij}^{stacked} = \sum_{\ell=t-s}^{t} \delta_{ij}^{\ell} \cdot w_{ij}^{\ell} \cdot \left(\frac{1}{2}\right)^{\ell}$ }
Second weighting scheme:
For t =1990 to 2019 {
Stack the $s+1$ networks of year $t-s$ to year t;
For each pair of nodes (e.g., companies), say i and j, compute the stacked time-weighted network structure weight of their connection as follows:
$w_{ij}^{stacked} = \max_{\ell=t-s,\dots,t} \left\{\delta_{ij}^{\ell} \cdot w_{ij}^{\ell} \cdot \left(\frac{1}{2}\right)^{\ell}\right\}$ }

3.4.2 Node embedding and design of board members' network-based drivers

The company relational information data typically contains a large volume of entities and relationships. In most cases, business entities are treated as 'nodes' of a topological graph and relationships between these entities are considered as 'edges'. In our application context, this type of network structure inevitably has a highly sparse adjacency matrix, where 99.9% of its entries are zeros, which would result in significant memory and computational requirements and eventually reduce the performance of prediction models. To overcome these issues, we intend to investigate the effectiveness of implementing graph embeddings in complex network analysis for bankruptcy prediction, where the network of connected companies can be better represented, and the performance of the prediction model would benefit from the information provided by the newly proposed company relational information-based drivers, which are learnt through complex network analysis.

The networks presented in the previous section can be overly large (because of the large number of companies and directors) and sparse (because not all companies are connected through their board members). Please refer to **Table 3.3** for empirical evidence. To reduce the dimensionality of the graph representation and improve the knowledge and pattern discovery of the board members' connection networks, we use the concept of *node embedding* to map such large networks into lower-dimensional spaces whilst maximally preserving the structure of the network (i.e., connectivity patterns) as well as reducing noises and biases caused eventually by the sparse nature of these networks. Recall that creating node embeddings is the process of converting nodes in a graph into numerical vectors, AKA node embeddings, or node vectors, derived from the graph topology and the relationships between nodes and capturing structural and semantic information about nodes, making them useful for various tasks. In summary, this process allows one to (1) reduce the dimensionality of graphs, making their processing tractable, (2) help uncover patterns, clusters, and communities within a graph, making them a good source of information for devising drivers, and (3) enable the use of graph relational data in machine learning models.

In this research, we use a three-step process to design company relational information drivers to predict bankruptcy and financial distress. This process is summarised as follows.

Step 1 – Node embeddings generation:

Node embedding methods are techniques used to transform nodes in a graph into numerical vector representations commonly referred to as node embeddings. These methods capture structural and semantic information about nodes in a way that is useful for various graph-analysis tasks. Node embedding methods can be divided into four broad categories, namely, (1) *Node similarity-based methods* (e.g., Jaccard Index, Katz Index, Hub Promoted Index), (2) *Role-based methods* (e.g., Role2Vec, Role Discovery in Networks or RDN for short, Role Aware Random Walk for Network Embedding or RARE for short), (3) *Structural methods*, which can be divided into *matrix factorisation methods* (e.g., Singular Value Decomposition or SVD for short, High Order Proximity Embedding or HOPE for short, and Large-scale Information Network Embedding or LINE for short), *graph factorisation methods* (e.g., Graph Sample and Aggregated Embeddings or GraphSAGE for short), *random walk-based methods* (e.g., Node2Vec, Metapath2Vec, DeepWalk), *graph sketching methods* (e.g., Random Projection, Node Sketching, Edge Sketching, and Graph Sampling), and *hybrid methods* (e.g., Structural Deep Network Embedding or SDNE for short, and Triadic Deep Network Representation or TriDNR for short), and (4) *Deep Learning methods* (e.g., DeepWalk, Graph Convolutional Neural Networks, Graph Autoencoders, and Graph Transformers). Each of these categories of methods has its own strengths and weaknesses, and the choice of method depends on the specific characteristics of the graph data and the task at hand.

In this research, we adopt a random walk-based method, which is a structural method, namely, *Node2Vec* proposed by Grover & Leskovec (2016), to generate numerical representations (i.e., embeddings) of nodes (i.e., companies) in the board members' network. *Node2Vec* is a popular method for generating node embeddings in complex network analysis, with several benefits over other methods. For example, its advantage over *Node Similarity-based methods* (e.g., Jaccard Index or Katz Index) lies in its capability of capturing higher-order proximity and structural information beyond direct node similarities. This capability is crucial for uncovering the intricate interconnections within networks (e.g., board of director networks in our case), where indirect connections often play crucial roles. Its advantage over *Role-based methods* (e.g., Role2Vec) is that it does not require a priori knowledge of structural roles and is capable of learning representations adaptively. This makes *Node2Vec* exceptionally versatile and suitable for application across diverse network structures (e.g., company-director-company bipartite networks, stacked time-weighted company-company networks in our case) where the roles and connections may not be immediately apparent or are subject to changes. Furthermore, *Node2Vec* employs a flexible random-walk strategy that balances between breadth-first and depth-first exploration during the walks. This allows it to capture both local and global structural information in the graph. *Node2Vec* treats random walks on the graph as sequences – it effectively captures the network neighbourhoods around nodes, and learns representations that preserve important structural information. It also allows users to control the balance between *homophily* (focussing on similar nodes) and *structural diversity* (exploring diverse parts of the graph) through adjustable parameters, adding to its versatility. This level of control adds to the method's versatility, allowing for customised embeddings that are reflective of specific network dynamics. Furthermore, when compared to *Structural methods* such as matrix factorisation techniques (e.g., HOPE, SVD) and graph factorisation methods (e.g., GraphSAGE), *Node2Vec* stands out for its computational efficiency and its focus on neighbourhood structures, which are often crucial for predictive accuracy in applications like financial distress and bankruptcy prediction, and when compared to *Deep Learning methods* (e.g., graph convolutional neural networks or generative adversarial networks), which are powerful, however, require extensive datasets and significant training time and may not be practical in all scenarios. *Node2Vec*, on the other hand, with its efficient handling of large datasets and less demanding computational requirements, offers a practical solution that is both effective and scalable.

Inspired by the well-known natural language processing algorithm *Word2Vec* (Mikolov et al., 2013), *Node2Vec* is a flexible and scalable method that leverages the *skip-gram model* from *Word2Vec* to generate embeddings for nodes in a graph – Similarly to the way *Word2Vec* generates a corpus of sentences by randomly sampling the words in texts, *Node2Vec* generates a number of random walks to obtain a series of node sequences. It learns continuous feature representations for nodes in networks by adopting a second-order biased random walks sampling strategy, which allows it to efficiently explore the neighbourhoods and make trade-offs between learning the characters of local communities (i.e., *breadth-first sampling*) and global structural connections (i.e., *depth-first sampling*). The *transition probabilities* of each step of the random walks are calculated by considering both the current and the previous states. The likelihood of immediately revisiting a node in a random walk is controlled by the return hyperparameter p, whilst the In-Out hyperparameter q biases the random walk to move towards nodes closer to the node in the previous step or move further away. Once the random walks (sequences of nodes) are generated, they are then fed into a *skip-gram single hidden layer neural network with negative sampling* for the training of the node representations. Basically, the *skip-gram* model was originally used to train word embeddings for *Word2Vec*, where the input is a one-hot encoded vector that represents the input word, and the output is a vector of probabilities that represents the likelihood of each context word (i.e., the words within the context window of each sampled sentence). The goal of the learning process is to learn the 'weights' of the hidden layer, which are, in fact, the embeddings that represent each sample word. The number of neurons in the hidden layer would also determine the dimension of the word embeddings and should be specified in advance. The *skip-gram* model significantly optimises the training process with the help of the *gradient descent* optimisation strategy, and by implementing *negative sampling*, that is, sampling frequent random walks to decrease the size of training samples and only updating a small percentage of the model's weights with each training sample, this algorithm can largely avoid the problem of overfitting. Finally, nodes (or their embeddings) that appear in a similar context (i.e., sampled walks) are similar, in our case, companies that share great similarities in the board of directors' networks would be assigned similar representations. **Fig. 3.7** demonstrates the process of generating embeddings for companies in complex networks.

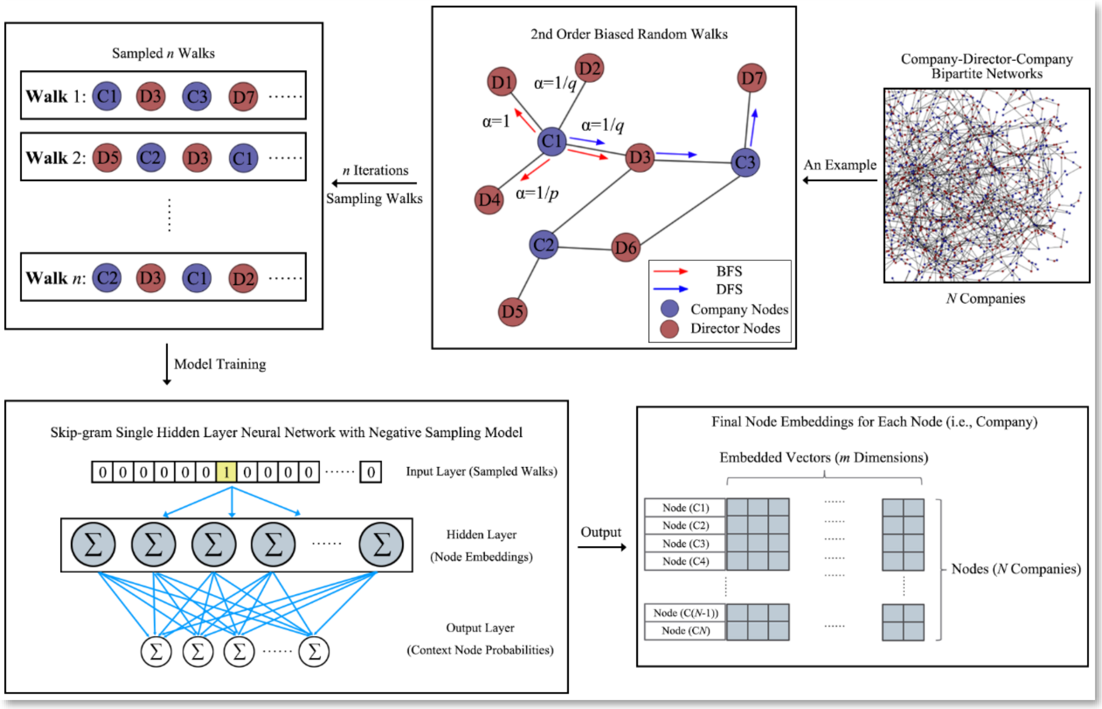

Fig. 3.7 An illustration of undirected company-company direct network

Step 2 – Node embeddings exploitation for community detection:

Methods for identifying communities in complex networks can be divided into seven broad categories, namely, (1) *partitioning methods*, which can be divided into *modularity-based methods* (e.g., Louvain method, Infomap, Fastgreedy), *density-based methods* (e.g., DBSCAN, OPTICS), and *centroid-based methods* (e.g., k-means, k-median, k-medoids); (2) *graph-based methods*, which can be divided into *spectral clustering* (e.g., normalised cut, ratio cut, spectral modularity) and *hierarchical clustering* (e.g., Ward's method, single linkage, complete linkage); (3) *propagation-based methods*, which can be divided into *label propagation methods* (e.g., Label Propagation Algorithm or LPA for short) and *random walk-based methods* (e.g., Community Detection via Random Walks or CRAWL for short, Walktrap); (4) *probabilistic methods*, which can be divided into *Bayesian methods* (e.g., Stochastic Block models) and *Fuzzy clustering methods* (e.g., Fuzzy c-Means or FCM for short); (5) *optimisation methods*, which can be divided into *mathematical programming models* (e.g., integer programmes), *heuristic methods* (e.g., Kernighan-Lin Algorithm) and *metaheuristics* (e.g., Simulated Annealing, Genetic Algorithms, Ant Colony Optimisation, Particle Swarm Optimisation); (6) *game-theoretic methods* (e.g., Community Detection via Game Dynamics or CDGD for short); and (7) *neural network-based methods* (e.g., Graph Convolutional Neural Networks).

In this chapter, we opt for partitioning methods as they often provide straightforward and intuitive interpretations, are often simpler and easier to implement compared to more complex algorithms, and can be more scalable for large networks, particularly density-based ones (Xu & Tian, 2015). In addition, density-based methods can adapt to networks with varying density patterns, allowing for more flexibility in community shape, are well-suited for networks with clusters of different density scales, often excel in identifying outliers or isolated nodes, and usually have fewer parameters to tune compared to some other techniques. Last but not least, we have chosen to implement OPTICS, first proposed by Ankerst et al. (1999), as it offers several advantages. First, it is particularly adept at handling variable density clusters. Unlike traditional density-based methods that rely on a fixed density threshold, OPTICS uses a reachability plot, allowing it to capture clusters with varying densities more effectively. Additionally, in a dataset with clusters of different densities or where the density varies between regions, OPTICS can provide more accurate and flexible clustering results compared to methods with fixed density parameters. Second, it produces a hierarchical representation of clusters in the form of a reachability plot, which allows one to explore and understand the clustering structure at multiple levels of granularity. Third, it includes a mechanism to handle noise using the concept of reachability distance. This allows it to identify and differentiate noise points from actual clusters. Fourth, it is relatively parameter-free compared to some other density-based methods. The absence of a fixed density threshold makes OPTICS more adaptable to datasets with varying cluster sizes and densities. Fifth, it is scalable to large datasets. Its efficiency is attributed to the use of an indexing structure that allows for faster processing of neighbourhood queries. Finally, it provides a reachability plot that offers a more nuanced and interpretable representation of the clustering structure. This helps in understanding the connectivity and hierarchy of clusters.

Step 3 – Design of Community-based Drivers:

A new driver, which we refer to as the *failure proximity score*, is then calculated based on the percentage of bankrupt companies in each cluster:

$$failure\ proximity\ score\ of\ cluster\ k\ =\frac{number\ of\ bankrupt\ companies\ in\ cluster\ k}{number\ of\ companies\ in\ cluster\ k}$$

For example, if there are 4 companies in a cluster, of which one is bankrupt and the other three are healthy, the failure proximity score of such cluster and therefore of each of the companies in the cluster would be ¼=0.25. The rationale of this step is that board members of a given company who are sitting on the boards of financially distressed or bankrupt firms carry a failure risk through their input to the board.

This approach is informed by the *weighted-vote relational neighbour* (wvRN) classifier adopted by Tobback et al. (2017), however, with a critical adaptation – instead of using a weighted average of neighbours' bankruptcy probabilities, we focus on creating a direct, cluster-based bankruptcy proximity measure as a community-based bankruptcy prediction driver. This shift acknowledges the contagion effect within financial networks, which is also similar to the variable '*Network Sum*' adopted by Constantin et al. (2018), but within a defined cluster context (i.e., developing a scaled measure rather than using the absolute value). Furthermore, our proposed driver echoes the principles outlined by Letizia & Lillo (2019), where network characteristics such as the '*weighted fraction of neighbours with given ratings*' are crucial. In our case, the 'failure proximity score' serves a similar purpose, but is streamlined for ease of calculation and interpretation within networks with clusters. This measure not only helps evaluate the immediate risk posed by neighbouring bankrupt companies within a cluster, but also provides good interpretability and ensures the practical application of our model in understanding the dynamics of corporate networks and their propagation of risks.

3.5 Data and design of experiments

In this section, we provide information on the types and sources of data for the empirical evaluation of the new drivers (§5.1) as well as details of how our experiments are designed (§5.2).

3.5.1 Data and its sources

This study focusses on UK companies listed on the *London Stock Exchange* (LSE), where three categories of drivers, namely, *financial ratios* (FRs), *corporate governance indicators* (CGIs), and *company relational information-based drivers* (CRIs), are considered in the analysis. FRs are extracted from *LSEG Datastream* (Datastream, 2022), and CGIs and company relational information (i.e., company-director relationships) data are extracted from *BoardEx* (BoardEx, 2022). Regarding the definition of the response variable, information on company status is extracted from the *London Share Price Database* (LSPD, 2022). In this study, corporate bankruptcy is defined in accordance with bankruptcy legislation in the UK (*the UK Insolvency Act of 1986*); to be more specific, we consider a company as bankrupt whenever it is assigned codes *16* (*Receivership*), *20* (*in Administration*) or *21* (*Cancelled and Assumed valueless*) by *LSPD*.

Since the data are extracted from several databases, we carried out the following steps to organise the data for our analysis. First, we selected the companies listed on LSE from *LSEG Datastream* over the period 1990 to 2019, providing us with company general information and financial ratios of 7165 companies. The selected companies are then matched with the corresponding companies in the *LSPD* and *BoardEx* databases using the *Stock Exchange Daily Official List* (*SEDOL*) code and *fuzzy name matching.* The matched companies are then manually checked to ensure that there is no mismatch. The data matching and merging process produced a total number of 1,922 companies with a total number of 21,631 firm-year observations for our analysis, as we selected only companies that have information across all these databases. The final dataset consists of 21,403 non-bankrupt records and 228 bankrupt records. The details and summary statistics of the final processed dataset are provided in **Appendix 3.A.1**. With respect to company relational information, we consider common board members of companies. By using data on board members' appointments and their overlap periods among different companies, 3,335 board of directors with external connections are identified. Information on company connections through their board of directors is then decomposed into 7,773 company-to-company connecting links and 8,588 company-to-director connecting links. **Table 3.3** provides a detailed information on the connection links for each year of observation. All the above information is extracted from the relevant databases without any further processing except for the bankruptcy status. To be more specific, several companies are assigned a bankruptcy status by LSPD in a given year, but they stopped disclosing data at least one year before the reported year of bankruptcy. To deal with this issue of missing data due to a lack of reporting, which is often an early sign of severe financial difficulties, we changed the bankruptcy year to the first year the company ceased to report data. As reflected in **Table 3.3**, the number of bankrupt companies per year for the original dataset and the pre-processed dataset is different. Additionally, notice that in the pre-processed dataset, for example, there is no bankrupt record for the year 2019 because the companies reported as bankrupt in 2019 were moved to 2018 and 2017 as a result of modifying the reported bankruptcy year.

Table 3.3 Number of companies and company connections in absolute and percentage terms by year

Year	No. Companies According to LSPD	No. Bankrupt Companies According to LSPD (in %)	No. Companies According to Our Modification	No. Bankrupt Companies According to Our Modification (in %)	No. BoD Members with External Connections	No. Company-Director Connections (%)*	No. Company-Company Connections (%)*
1990	407	0 (0.00%)	284	0 (0.00%)	58	122 (0.74%)	67 (0.17%)
1991	415	0 (0.00%)	309	1 (0.32%)	85	181 (0.69%)	102 (0.21%)
1992	425	0 (0.00%)	331	0 (0.00%)	120	263 (0.66%)	152 (0.28%)
1993	447	0 (0.00%)	355	0 (0.00%)	167	368 (0.62%)	221 (0.35%)
1994	487	1 (0.21%)	379	0 (0.00%)	209	464 (0.59%)	288 (0.40%)
1995	533	0 (0.00%)	397	0 (0.00%)	251	557 (0.56%)	351 (0.45%)
1996	598	0 (0.00%)	503	0 (0.00%)	313	706 (0.45%)	453 (0.36%)
1997	668	0 (0.00%)	589	0 (0.00%)	393	885 (0.38%)	559 (0.32%)
1998	708	0 (0.00%)	638	0 (0.00%)	454	1033 (0.36%)	658 (0.32%)
1999	754	0 (0.00%)	677	0 (0.00%)	537	1223 (0.34%)	774 (0.34%)

2000	894	0 (0.00%)	777	0 (0.00%)	665	1525 (0.30%)	980 (0.33%)
2001	972	0 (0.00%)	859	0 (0.00%)	724	1698 (0.27%)	1148 (0.31%)
2002	1028	0 (0.00%)	926	0 (0.00%)	802	1921 (0.26%)	1343 (0.31%)
2003	1071	0 (0.00%)	948	4 (0.42%)	905	2167 (0.25%)	1498 (0.33%)
2004	1233	6 (0.49%)	1044	8 (0.77%)	1101	2656 (0.23%)	1839 (0.34%)
2005	1371	4 (0.29%)	1122	16 (1.43%)	1246	3020 (0.22%)	2095 (0.33%)
2006	1428	16 (1.12%)	1176	21 (1.75%)	1287	3148 (0.21%)	2203 (0.32%)
2007	1406	22 (1.56%)	1139	52 (4.57%)	1245	2997 (0.21%)	2043 (0.32%)
2008	1307	54 (4.13%)	1002	26 (2.59%)	1118	2639 (0.24%)	1735 (0.35%)
2009	1154	31 (2.68%)	906	24 (2.65%)	964	2246 (0.26%)	1444 (0.35%)
2010	1052	25 (2.38%)	834	8 (0.96%)	897	2086 (0.28%)	1351 (0.39%)
2011	973	18 (1.85%)	797	19 (2.38%)	823	1897 (0.29%)	1207 (0.38%)
2012	926	16 (1.73%)	740	10 (1.35%)	758	1756 (0.31%)	1143 (0.42%)
2013	898	11 (1.22%)	729	7 (0.96%)	755	1752 (0.32%)	1140 (0.43%)
2014	923	7 (0.76%)	750	10 (1.33%)	795	1871 (31%)	1247 (0.44%)
2015	916	13 (1.42%)	723	3 (0.41%)	829	1947 (0.32%)	1326 (0.51%)
2016	881	9 (1.02%)	711	7 (0.98%)	792	1844 (0.33%)	1238 (0.49%)
2017	841	7 (0.83%)	676	8 (1.18%)	753	1737 (0.34%)	1143 (0.50%)
2018	825	10 (1.21%)	674	4 (0.59%)	753	1743 (0.34%)	1128 (0.50%)
2019	776	10 (1.29%)	636	0 (0.00%)	734	1659 (0.36%)	1043 (0.52%)

* Connecting links as a percentage of links in the complete graph

3.5.2 Financial ratios and corporate governance indicators

In terms of financial ratios, this research adopts the conventional accounting ratios proposed by Altman (1968), i.e., *working capital / total assets*, *retained earnings / total assets*, *EBITDA / total assets*, *market value of equity / total liabilities*, and *sales / total assets*. The rationale is that Altman's discriminant analysis model is widely adopted as the benchmark in the bankruptcy and financial distress prediction literature, and the aforementioned drivers are still well received in today's financial world (Altman et al., 2014). As this study focusses on the influence of director network-based drivers on bankruptcy prediction, we would also incorporate several board-level CGIs which reflect the board's internal composition and structure as mentioned by other studies (e.g., Li et al., 2020; Liang et al., 2020; Almaskati et al., 2021; Mousavi & Lin, 2020) to investigate their value added and answer some of our research questions. More specifically, we include the level of board exogenous information asymmetry as measured by the size of the board (*Number of Directors*) (Darrat et al., 2016), the board independence as measured by the proportion of board members who are considered as independent (*Percentage of Non-executive Directors*) (Platt & Platt, 2012), and the dominance of the board of directors as measured by whether the company's CEO is simultaneously serving as the company's Chairman/Chairwoman (*Chair / CEO Duality*) (Krause et al., 2014). The detailed description and summary statistics of FRs and CGIs are provided in **Table 3.4** and **Appendix 3.A.1**, respectively.

3.5.3 Generic prediction model

To perform our empirical analysis, we have chosen to implement several *logistic regression* models, or *logit* models for short, *with a regularisation term* fed with a variety of combinations of categories of drivers to determine the value added of each category of drivers and to answer our research questions. Please refer to **Table 3.4** for details on the combinations of categories of drivers considered.

Table 3.4 The combinations of drivers used to feed the model

a) Accounting-based drivers only (Benchmark model)
Financial Ratios (FRs): (1) ***Working Capital/Total Assets***: Compares the net liquid assets to the total assets of the firm. Working capital is defined as the difference between current assets and current liabilities. This ratio can be used to evaluate the short-term solvency of a company. (2) ***Returned Earnings/Total Assets***: Measures the amount of reinvested earnings or losses generated from the total assets. This ratio can be used to evaluate the profitability of a company and its ability to stand up to a bad year of losses. (3) ***Earnings Before Interest, Tax, Depreciation and Amortisation (EBITDA)/Total Assets***: Measures a company's ability to generate profits from its total assets before deducting financial expenses, depreciation, amortisation and taxes. (4) ***Market Value of Equity/Total Liabilities***: Measures how much the company's market value would decline before liabilities exceed the total assets if a company were to become insolvent. This ratio reflects the market's confidence in a company's financial position. (5) ***Sales/Total Assets***: Measures the efficiency of a company that uses its assets to generate sales. This ratio reflects how well the managers handle competition.
b) Corporate Governance-based drivers only
Corporate Governance Indicators (CGIs): (1) ***Number of Directors***: Measures the level of exogenous information asymmetry of the board. (2) ***Percentage of Non-executive Directors***: Compares the number of non-executive directors to the total number of board of directors. This indicator measures the proportion of board members who are considered independent. (3) ***Chair/CEO Duality***: Dummy variable that measures the dominance of the board of directors. If the chair and the CEO are the same person, this indicator gives 1, otherwise, it gives 0.
c) Company Relational Information-based drivers (CRIs) only
(1) ***Company-Company Direct Connections-based Driver (DC)*****:** The *failure proximity score* calculated with the node embeddings generated from the undirected company-company direct network. (2) ***Company-Director-Company Bipartite Connections-based Driver with Simple Direct Connections (SBC)*****:** The *failure proximity score* calculated with the node embeddings generated from the undirected company-director-company bipartite network with simple direct connections. (3) ***Company-Director-Company Bipartite Connections-based Driver with Consideration of Directors' Power (PBC)*****:** The *failure proximity score* calculated with the node embeddings generated from the undirected company-director-company bipartite network with the consideration of the power of the directors. (4) ***Stacked Time-weighted Company-Company Direct Connections-based Driver with the Maximum Weights (MWDC)*****:** The *failure proximity score* calculated with the node embeddings generated from undirected stacked time-weighted company-company network with weights calculated by using the maximum weights of each pair of company connections. (5) ***Stacked Time-weighted Company-Company Direct Connections-based Driver with the Sum of Weights (SWDC)*****:** The *failure proximity score* calculated with the node embeddings generated from the undirected stacked time-weighted company-company network with weights calculated by using the sums of the weights of each pair of company connections across the years.
d) Accounting-based & Corporate Governance-based drivers
A combination of **a)** and **b)**, i.e., ***FRs + CGIs***
e) Accounting-based & Company Relational Information-based drivers
A combination of **a)** and **c)**, including: (1) ***FRs + DC***; (2) ***FRs + SBC***; (3) ***FRs + PBC***; (4) ***FRs + MWDC***; (4) ***FRs + SWDC***
f) Accounting-based, Corporate Governance-based, and Company Relational Information-based drivers
A combination of **a)**, **b)** and **c)**, including: (1) ***FRs +CGIs+ DC***; (2) ***FRs + CGIs+ SBC***; (3) ***FRs + CGIs+ PBC***; (4) ***FRs + CGIs+ MWDC***; (4) ***FRs + CGIs+ SWDC***

We selected the logistic regression model for its direct interpretability and its established benchmark status in the industry (Almaskati et al., 2021). It also provides a more straightforward and transparent framework for quantifying the direct influence of specific drivers on the prediction outcome. This clarity is essential to our objective of demonstrating the efficacy of the proposed driver. In addition, the *logit* framework is popular in the BP and FDP literature. As for the regularisation term, we consider *L1*, *L2* and *ElasticNet* regularisations. For each regularisation strategy, there is a corresponding loss function that should be minimised to learn the parameters of the model. Recall that the implementation decisions are concerned with the choice of the regularisation term (i.e., L1, L2, or ElasticNet) and the value of the regularisation strength, say λ. *Grid-search* with *5-fold cross-validation* is adopted to search for the best combination of these hyperparameters.

3.5.4 Performance measures of classification models

A total of six performance measures are adopted for the evaluation of the performance of the models, where four of them are single criterion-based measures, i.e., *sensitivity* (*true positive rate*) for evaluating the models' ability to correctly predict bankrupt companies; *specificity* (*true negative rate*) for evaluating the models' ability to correctly predict healthy companies; *Type-I error* (*false positive rate*) and *Type-II error* (*false negative rate*) for evaluating the models' lack of correctly predicting the risk class. The reason for choosing the single-criterion-based performance measures is to avoid the biases associated with the weighted / multi-criteria evaluation measures. Two multi-criteria measures are also included, namely, *prediction accuracy,* which evaluates the general predictability of the models measured by the number of correct predictions over the total number of predictions; and the *area under receiver operating characteristic curve* (AUC), which provides an aggregate measure of performance across all possible cut-off points that reflects how the model performs compared to a random classifier. Details of these performance measures, including their definitions and mathematical formulations, are provided in **Table 3.5**.

Table 3.5 Performance measures for classification models

Performance Metric	Definitions	Mathematical Formula
Type-I Error (false positive rate)	Occurs when a healthy company is incorrectly classified as bankrupt (false positive).	$Type - I\ Error = \frac{False\ Positives}{False\ Positives + True\ Positives}$
Type-II Error (false negative rate)	Occurs when a bankrupt company is incorrectly classified as healthy (false negative).	$Type - II\ Error = \frac{False\ Negatives}{True\ Positives + False\ Negatives}$
Sensitivity (true positive rate/recall)	Measures the model's ability to correctly identify bankrupt companies (positive class).	$Sensitivity = \frac{True\ Positives}{True\ Positives + False\ Negatives}$
Specificity (true negative rate)	Measures the model's ability to correctly identify healthy companies (negative class).	$Specificity = \frac{True\ Negatives}{True\ Negatives + False\ Positives}$
Classification Accuracy	Measures the overall correctness of the model's predictions across all classes.	$Accuracy = \frac{Number\ of\ Correct\ Predictions}{Total\ Number\ of\ Predictions}$
Area Under ROC Curve (AUC)	Evaluates the discriminative power of the model by measuring the area under the *receiver operating characteristic* (ROC) curve, which plots *true positive rate* (Sensitivity) against *false positive rate* (1-Specificity) at various threshold values. Higher AUC indicates better performance.	AUC is the area under the ROC curve, typically calculated numerically.

3.5.5 Sampling strategy and experimental setup

In this study, two main experimental setups are adopted for sampling and split of data between training set and test set to address our research questions.

The first type of experimental setup adopts *sampling without time considerations*, referred to as *Experiment Setup I.* It uses a 70%-30% split of data across the years without time considerations; that is, the dataset (firm-year observations over 1990-2019) is randomly sampled 30 times, regardless of time, so that the 30 training samples would each consist of

70% of observations, and the 30 test samples would each consist of 30% of observations, and *random oversampling* is used to address the issue of severe data imbalance.

The second type of experimental setup adopts *sampling with time considerations*, referred to as *Experiment Setup II*, and uses time periods (i.e., years) to split the original dataset into training-test periods. In our experiments, we use the year 2014 to split the overall period 1990-2019 into training and test periods; to be more specific, the training period is chosen as 1990-2014 and the test period is chosen from within 2016 to 2019 based on different time lags to assess the ability of the models to predict 2 years ahead, 3 years ahead, 4 years ahead, and 5 years ahead. In addition, in *Experiment Setup II*, we choose 30 training samples randomly using *random over-sampling* to address the issue of severe data imbalance, where each consists of 70% of the total observations in the training period 1990-2014, and we choose four test samples where each consists of all observations in the individual year of 2016 (2-year lag), 2017 (3-year lag), 2018 (4-year lag), and 2019 (5-year lag), respectively.

Finally, note that regardless of the type of experimental setup, the observations in the sampled training sets are *winsorised* at the 99% percentile to trim extreme values and hence minimise the influence of the outliers. A summary of our methodological choices and experimental steps is provided in the flow chart presented in **Fig. 3.8**.

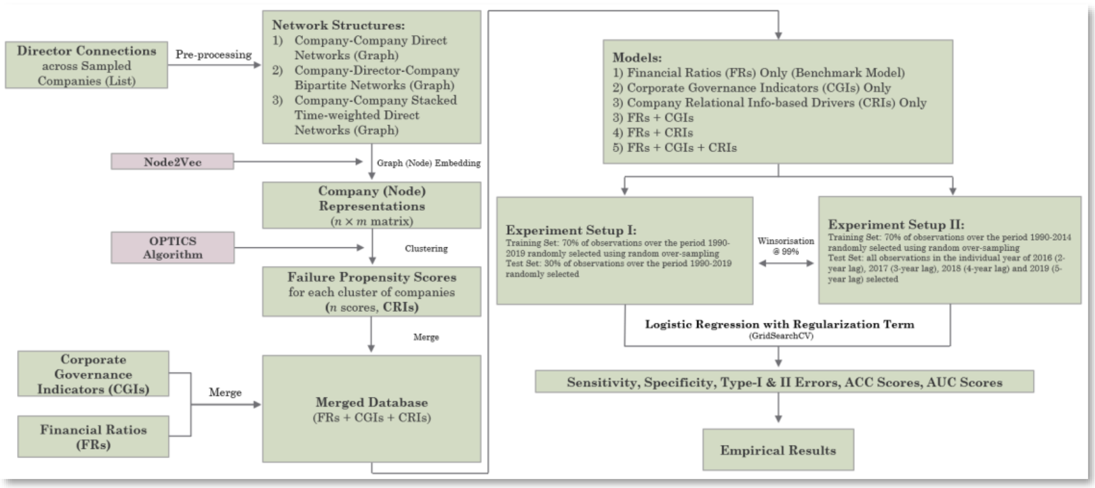

Fig. 3.8 Graphical Summary of Methodological & Experimental Steps

In sum, under the first experimental setup, a total of *18* models with different combinations of drivers are tested to evaluate the value added of CRIs. Among them, the model with FRs as the only drivers is set as the benchmark model against which the performance of models with network-based drivers would be compared; *Two* models with CGIs as input variables (i.e., *CGIs* only and *FRs+CGIs*) are built to examine the contribution of corporate governance information in bankruptcy prediction; another *nine* models which include static complex networks-based drivers (i.e., *DC*, *PBC* and *SBC*) are proposed to determine the best structure of the board of directors' networks as well as examine the best combination of drivers; and, finally, *six* additional models with stacked time-weighted complex networks-based drivers (i.e., *MWDC* and *SWDC*) are proposed to evaluate the value added of taking the dynamic nature of directors' connections into account in designing the network-based bankruptcy prediction drivers. In the second experimental setup, we focus on evaluating the performance of the proposed network-based drivers for short- and long-term predictions. Under this scenario, the best performing models in the first experiment are selected to conduct further tests, where data from 1990 to 2014 are used for in-sample training and the individual years of data between 2016 and 2019 are used for out-of-sample testing. In this experimental setup, the best performing models (i.e., combinations of categories of drivers) identified in Experiment Setup I are chosen to perform comparative analysis, where the long-term predictive ability of the models would be evaluated.

3.6 Empirical Analysis and Findings

The results of the two experimental setups described in the previous section are provided in **Table 3.6** and **Table 3.11**, respectively. In the following two subsections, we shall discuss these results and provide answers to our research questions.

3.6.1 Analysis of results of Experimental Setup I

In this section, we shall analyse the results of Experimental Setup I to answer the first five research questions.

Table 3.6 suggests that the benchmark model, which only uses Altman's *financial ratios* (FRs) as drivers, misclassifies on average more than 30% of bankrupt companies and more than 15% of non-bankrupt companies in the case of our UK dataset of companies listed on the *London Stock Exchange*. However, the out-of-sample performance does not deteriorate much compared to the in-sample performance, and the model still performs much better than a random classifier, as suggested by an AUC of 0.8156. On the other hand, using only *corporate governance indicators* (CGIs) as drivers of bankruptcy results in a model that misclassifies on average about 50% of bankrupt firms and about 45% of non-bankrupt firms. In sum, this model is only slightly superior to a random classifier, as suggested by an AUC of 0.5736. Note that when Altman's FRs are supplemented with CGIs, the predictive performance improves slightly (i.e., by 0.58% in out-of-sample sensitivity and 0.003 in AUC score), suggesting that CGIs only slightly add value, on average, to the detection of bankrupt companies, but only for TPR or sensitivity and AUC. However, statistical tests demonstrate that such improvement is not statistically significant (see **Table 3.7**). Notice that supplementing FRs and CRIs with CGIs (i.e., *FRs+CGIs+DC*, *FRs+CGIs+SBC*, *FRs+CGIs+PBC*) did not enhance model effectiveness in terms of prediction accuracy and AUC, which can be attributed to an overfitting issue, where the model becomes too closely fitted to the training data and unable to generalise effectively on test data.

These results echo the conclusion of Li et al. (2020) that the use of CGIs alone has a limited ability to detect unhealthy firms, however, when combined with other information, such as financial ratios, they would improve the model's overall prediction performance.

Table 3.6 Results of **Experimental Setup I** *(Training: 1990-2019 @70%; Test: 1990-2019 @30%)*

Results - Mean	**Sensitivity (TPR)**		**Specificity (TNR)**		**Type-I Error Rate (FPR)**		**Type-II Error Rate (FNR)**		**Classification Accuracy (ACC)**		**Area Under ROC Curve (AUC)**
	TPR Train	***TPR Test***	***TNR Train***	***TNR Test***	***Type-I Train***	***Type-I Test***	***Type-II Train***	***Type-II Test***	***Acc Train***	***Acc Test***	***AUC Test***
Financial Ratios (FRs) Model / Benchmark Model											
FRs	66.80%	65.85%	83.48%	83.48%	16.52%	16.52%	33.20%	34.15%	83.31%	83.30%	0.8156
Corporate Governance Indicators (CGIs) Models											
CGIs	49.68%	49.12%	59.49%	59.45%	40.51%	40.55%	50.32%	50.88%	59.38%	59.35%	0.5736
FRs+CGIs	68.40%	66.43%	81.11%	81.16%	18.89%	18.84%	31.60%	33.57%	80.98%	81.01%	0.8186
Static Complex Network Models (incl. DC, PBC & SBC)											
DC	47.51%	47.10%	98.09%	98.06%	1.91%	1.94%	52.49%	52.90%	97.55%	97.54%	0.7273
SBC	36.88%	36.70%	99.02%	99.05%	0.98%	0.95%	63.12%	63.30%	98.36%	98.40%	0.6793
PBC	34.84%	34.09%	99.06%	99.07%	0. 94%	0. 93%	65.16%	65.91%	98.38%	98.40%	0.6662
FRs+DC	72.12%	71.36%	91.42%	91.34%	8.58%	8.66%	27.88%	28.64%	91.22%	91.13%	0.9011
FRs+SBC	70.67%	70.70%	90.70%	90.69%	9.30%	9.31%	29.33%	29.30%	90.48%	90.48%	0.8832
FRs+PBC	69.21%	68.71%	90.63%	90.57%	9.37%	9.43%	30.79%	31.29%	90.40%	90.34%	0.8845
FRs+CGIs+DC	73.78%	72.98%	90.75%	90.68%	9.25%	9.32%	26.22%	27.02%	90.57%	90.50%	0.9015
FRs+CGIs+SBC	71.51%	71.08%	89.56%	89.54%	10.44%	10.46%	28.49%	28.92%	89.36%	89.34%	0.8836
FRs+CGIs+PBC	69.78%	68.85%	89.64%	89.60%	10.36%	10.40%	30.22%	31.15%	89.43%	89.39%	0.8834
Stacked Time-weighted Complex Network Models (incl. MWDC & SWDC)											
MWDC	58.16%	58.57%	98.47%	98.48%	1.53%	1.52%	41.84%	41.43%	98.04%	98.07%	0.7866
SWDC	55.87%	55.25%	98.51%	98.53%	1.49%	1.47%	44.13%	44.75%	98.06%	98.08%	0.7701
FRs+MWDC	74.97%	75.31%	92.91%	92.83%	7.09%	7.17%	25.03%	24.69%	92.71%	92.65%	0.9181
FRs+SWDC	74.55%	72.94%	92.48%	92.41%	7.52%	7.59%	25.45%	27.06%	92.29%	92.20%	0.9091
FRs+CGIs+MWDC	76.36%	75.78%	92.74%	92.70%	7.26%	7.30%	23.64%	24.22%	92.56%	92.53%	0.9183
FRs+CGIs+SWDC	76.21%	74.72%	92.26%	92.24%	7.74%	7.76%	23.79%	25.28%	92.09%	92.06%	0.9114

Note: *FRs* – Financial ratios; *CGIs* – Corporate governance indicators; *DC* – Company-Company direct connections-based driver; *SBC* – Company-Director-Company bipartite connections-based driver with simple direct connections; *PBC* – Company-Director-Company bipartite connections-based driver with consideration of directors' power; *MWDC* – Stacked time-weighted Company-Company direct connections-based driver with the maximum weights; *SWDC* – Stacked time-weighted Company-Company direct connections-based driver with the sum of weights.

Table 3.7 Mann-Whitney U Test for **Experimental Setup I** (Comparison against Benchmark Model FRs)

Metric	**Group**	**Model**	**Mean**	**p_value (*Significance)**
Comparison against Benchmark Model (FRs)				
Test Sensitivity (TPR)	Control	*FRs*	65.85%	-
	1	*FRs+CGIs*	66.43%	.4917
	2	*FRs+DC*	71.36%	.0006***
	3	*FRs+SBC*	70.70%	.0020**
	4	*FRs+PBC*	68.71%	.0977
	5	*FRs+CGIs+DC*	72.98%	.0000***
	6	*FRs+CGIs+SBC*	71.08%	.0006***
	7	*FRs+CGIs+PBC*	68.85%	.0919
	8	*FRs+MWDC*	75.31%	.0000***
	9	*FRs+SWDC*	72.94%	.0000***
	10	*FRs+CGIs+MWDC*	75.78%	.0000***
	11	*FRs+CGIs+SWDC*	74.72%	.0000***
Test Specificity (TRN)	Control	*FRs*	83.48%	-
	1	*FRs+CGIs*	66.43%	.0000***
	2	*FRs+DC*	91.34%	.0000***
	3	*FRs+SBC*	90.69%	.0000***
	4	*FRs+PBC*	90.57%	.0000***
	5	*FRs+CGIs+DC*	90.68%	.0000***
	6	*FRs+CGIs+SBC*	89.54%	.0000***
	7	*FRs+CGIs+PBC*	89.60%	.0000***
	8	*FRs+MWDC*	92.83%	.0000***
	9	*FRs+SWDC*	92.41%	.0000***
	10	*FRs+CGIs+MWDC*	92.70%	.0000***
	11	*FRs+CGIs+SWDC*	92.24%	.0000***
Test Accuracy (ACC)	Control	*FRs*	83.30%	-
	1	*FRs+CGIs*	81.01%	.0000***

	2	*FRs+DC*	91.13%	.0000***
	3	*FRs+SBC*	90.48%	.0000***
	4	*FRs+PBC*	90.34%	.0000***
	5	*FRs+CGIs+DC*	90.50%	.0000***
	6	*FRs+CGIs+SBC*	89.34%	.0000***
	7	*FRs+CGIs+PBC*	89.39%	.0000***
	8	*FRs+MWDC*	92.65%	.0000***
	9	*FRs+SWDC*	92.20%	.0000***
	10	*FRs+CGIs+MWDC*	92.53%	.0000***
	11	*FRs+CGIs+SWDC*	92.06%	.0000***
Test AUC	Control	*FRs*	0.8156	-
	1	*FRs+CGIs*	0.8186	.5106
	2	*FRs+DC*	0.9011	.0000***
	3	*FRs+SBC*	0.8832	.0000***
	4	*FRs+PBC*	0.8845	.0000***
	5	*FRs+CGIs+DC*	0.9015	.0000***
	6	*FRs+CGIs+SBC*	0.8836	.0000***
	7	*FRs+CGIs+PBC*	0.8834	.0000***
	8	*FRs+MWDC*	0.9181	.0000***
	9	*FRs+SWDC*	0.9091	.0000***
	10	*FRs+CGIs+MWDC*	0.9183	.0000***
	11	*FRs+CGIs+SWDC*	0.9114	.0000***

*p<.05, **p<.01, ***p<.001

Regarding the first research question, namely, ‘Are CRIs based on the board of directors’ connections a good alternative to conventional CGIs? Or, more specifically, do CRIs based on the board of directors’ connections provide more predictive value than the conventional CGIs?’, numerical results suggest that, for models that use CRIs (i.e., *DC, SBC, PBC*) as sole drivers, they misclassify between 52.49% and 65.91% of bankrupt firms and misclassify between 0. 93% and 1.94% of non-bankrupt firms. Therefore, as compared to the models using only CGIs or only FRs, our CRIs significantly improve the ability of the models to correctly predict non-bankrupt firms, but worsen for bankrupt firms. Further investigation of this result suggests that it is mainly due to the fact that meaningful CRIs, as proxied by the failure proximity score, are generated for companies with external connections; however, there are also many companies without external connections, which means that CRIs could not provide meaningful information for those companies, and thus the performance of CRIs dropped in detecting bankrupt companies.

Regarding the second research question, namely ‘Does complementing FRs with CRIs provide an improved bankruptcy prediction performance than using FRs alone?’, numerical results suggest that when CRIs are used to supplement FRs, regardless of which CRI is used (i.e., *DC, SBC,* or *PBC*), the ability of the models to correctly predict both bankrupt and non-bankrupt companies improves significantly (i.e., by 7.83%, 7.18%, and 7.04% in out-of-sample prediction accuracy, respectively).

Furthermore, with respect to the third research question, namely, ‘Does complementing FRs with CRIs provide an improved bankruptcy prediction performance than complementing FRs with CGIs?’, the numerical results also suggest that the proposed new corporate governance drivers that exploit company relational information, using board of directors’ networks of companies, add much more value to the prediction exercise than the conventional CGIs (i.e., improvements of 10.12%, 9.47% and 9.33% in out-of-sample prediction accuracy, for *FRs+DC*, *FRs+SBC* and *FRs+PBC*, respectively). Notice also that supplementing FRs with CRIs and CGIs at the same time slightly improves the ability of the models to correctly predict bankrupt companies but slightly worsens the ability of the models to correctly predict non-bankrupt companies, suggesting that CGIs are not discriminatory enough for healthy companies. However, due to their better ability to detect bankrupt companies, these models actually provide much more useful information, since financial service providers generally focus on reducing Type-II errors that can cause greater damage to credit granting. In summary, empirical evidence, as summarised in **Table 3.6**, suggests that the answers to our first three research questions are all positive. In fact, CRIs based on the board of directors' connections provide more predictive value than the conventional CGIs; complementing FRs with CRIs provides an improved bankruptcy prediction performance than using FRs alone; and complementing FRs with CRIs provides improved bankruptcy prediction performance than complementing FRs with CGIs.

Regarding the fourth research question, namely, 'How do CRIs based on company-company direct connection networks and CRIs based on company-director-company bipartite connection networks compare to improve the predictive performance of models?', when examining the influence of the choice of network structure on the performance of network-based bankruptcy prediction drivers, the results suggest that incorporation of the bipartite network structure does not necessarily improve the prediction performance of models. To be more specific, among the static models that use CRIs as drivers, models with CRIs based on company-company direct connection networks (i.e., *FRs+DC* and *FRs+CGIs+DC*) provide the best prediction performance in all aspects, as they achieve the highest sensitivity, specificity, classification accuracy, and AUC scores for both in-sample and out-of-sample predictions. Although not performing as well as the aforementioned models, static models with CRIs based on company-director-company bipartite networks (i.e., *FRs+SBC*, *FRs+PBC, FRs+CGIs+SBC* and *FRs+CGIs+PBC*) still significantly outperform the benchmark model and the model based on FRs and CGIs (see **Table 3.7**). Among them, models with CRIs based on simple direct connection bipartite networks (i.e., *FRs+SBC* and *FRs+CGIs+SBC*) prove superior to those based on bipartite networks taking into account the power of directors (i.e., *FRs+PBC* and *FRs+CGIs+PBC*) in terms of detecting bankrupt companies, as they provide an average improvement of 2.11% improvement in out-of-sample prediction sensitivity, however, it is not statistically significant. The overall performance of these two types of models in terms of specificity, classification accuracy, and AUC is very similar (see **Table 3.8** for results of statistical tests).

With respect to the fifth research question, namely, 'How CRIs based on company-company direct connection networks (resp. company-director-company bipartite connection networks) and CRIs based on stacked time-weighted company connection networks compare to enhance the predictive performance of models?', the best performing static network structure, in this case, the company-company direct connection network (*DC*), is adopted to create a stacked time-weighted network for analysis. The results of CRIs based on stacked time-weighted complex networks (i.e., *MWDC* and *SWDC*) suggest that the implementation of the time-weighted network structure can significantly improve the performance of bankruptcy prediction models compared to static models. To be more specific, the results of models that combine stacked time-weighted CRIs with FRs and CGIs echo performance improvement, as the models based on stacked time-weighted CRIs (i.e., *FRs+MWDC*, *FRs+SWDC*, *FRs+CGIs+MWDC*, and *FRs+CGIs+SWDC*) achieve an average improvement of 2.52% in out-of-sample sensitivity scores over their non-stacked static counterparts (i.e., *FRs+DC*, *FRs+CGIs+DC*) (see **Table 3.8** for the statistical tests of model performance). They also significantly improve the performance of the benchmark mode based on FRs by around 10% in all performance measures in the out-of-sample tests, again showing the superiority of the stacked time-weighted complex network structure for the design of bankruptcy prediction drivers (see **Table 3.7** for the statistical tests of model performance).

Table 3.8 Mann-Whitney U Test for **Experimental Setup I** (Time-weighted vs. Static Models)

Metric	Group	Model	Mean	p_value (*Significance)
Test Sensitivity (TPR)	12	*FRs+MWDC*	75.31%	.0241*
		FRs+DC	71.36%	
	13	*FRs+SWDC*	72.94%	.2804
		FRs+DC	71.36%	
	14	*FRs+CGIs+MWDC*	75.78%	.0574
		FRs+CGIs+DC	72.98%	
	15	*FRs+CGIs+SWDC*	74.72%	.1623
		FRs+CGIs+DC	72.98%	
Test Specificity (TRN)	12	*FRs+MWDC*	92.83%	.0000***
		FRs+DC	91.34%	
	13	*FRs+SWDC*	92.41%	.0000***
		FRs+DC	91.34%	
	14	*FRs+CGIs+MWDC*	92.70%	.0000***
		FRs+CGIs+DC	90.68%	
	15	*FRs+CGIs+SWDC*	92.24%	.0000***
		FRs+CGIs+DC	90.68%	
Test Accuracy (ACC)	12	*FRs+MWDC*	92.65%	.0000***
		FRs+DC	91.13%	
	13	*FRs+SWDC*	92.20%	.0000***
		FRs+DC	91.13%	
	14	*FRs+CGIs+MWDC*	92.53%	.0000***
		FRs+CGIs+DC	90.50%	
	15	*FRs+CGIs+SWDC*	92.06%	.0000***
		FRs+CGIs+DC	90.50%	
Test AUC	12	*FRs+MWDC*	0.9181	.0008***
		FRs+DC	0.9011	
	13	*FRs+SWDC*	0.9091	.0905
		FRs+DC	0.9011	
	14	*FRs+CGIs+MWDC*	0.9183	.0004***
		FRs+CGIs+DC	0.9015	
	15	*FRs+CGIs+SWDC*	0.9114	.0211*
		FRs+CGIs+DC	0.9015	

*p<.05, **p<.01, ***p<.001

In terms of the performance of the two weighting schemas for the calculation of stacked time-weighted CRIs, the results suggest that the weighting scheme based on the maximum weights (i.e., the second weighting scheme) of each pair of company connections (i.e., *MWDC*) generally outperforms the weighting scheme that uses the sums of the weights of each pair of company connections across the years (i.e., *SWDC*) in out-of-sample specificity, prediction accuracy and AUC (see **Table 3.9** for results of statistical tests). Although models based on MWDC (i.e., *MWDC* only, *FRs+MWDC* and *FRs+CGIs+MWDC*) demonstrate better capability in detecting bankrupt companies against models based on SWDC (i.e., *SWDC* only, *FRs+SWDC* and *FRs+CGIs+SWDC*) in out-of-sample sensitivity (around 2.25% improvement), the improvement is not statistically significant. These results indicate that the closeness of the board of directors' connections between companies in the past may not necessarily provide the best insights into the risk profile of a company, instead, the stacked time-weighted network-based drivers would perform better with the appropriate weighting scheme of the directors' connections across the stacked years.

Table 3.9 Mann-Whitney U Test for **Experimental Setup I** (*SBC vs. PBC, MWDC vs. SWDC*)

Metric	Group	Model	Mean	p_value (*Significance)
Test Sensitivity (TPR)	12	*FRs+SBC*	70.70%	.1536
		FRs+PBC	69.71%	
	13	*FRs+CGIs+SBC*	71.08%	.1296
		FRs+CGIs+PBC	68.85%	
	14	*FRs+MWDC*	75.31%	.0891
		FRs+SWDC	72.94%	
	15	*FRs+CGIs+MWDC*	75.78%	.3182
		FRs+CGIs+SWDC	74.72%	
Test Specificity (TRN)	12	*FRs+SBC*	90.69%	.4119
		FRs+PBC	90.57%	
	13	*FRs+CGIs+SBC*	89.54%	.8418
		FRs+CGIs+PBC	89.60%	
	14	*FRs+MWDC*	92.83%	.0016**
		FRs+SWDC	92.41%	
	15	*FRs+CGIs+MWDC*	92.70%	.0012**
		FRs+CGIs+SWDC	92.24%	
Test Accuracy (ACC)	12	*FRs+SBC*	90.48%	.3516
		FRs+PBC	90.34%	
	13	*FRs+CGIs+SBC*	89.34%	.8941
		FRs+CGIs+PBC	89.39%	
	14	*FRs+MWDC*	92.65%	.0008***
		FRs+SWDC	92.20%	
	15	*FRs+CGIs+MWDC*	92.53%	.0009***
		FRs+CGIs+SWDC	92.06%	
Test AUC	12	*FRs+SBC*	0.8832	.8360
		FRs+PBC	0.8845	
	13	*FRs+CGIs+SBC*	0.8836	.9823
		FRs+CGIs+PBC	0.8834	
	14	*FRs+MWDC*	0.9181	.0266*
		FRs+SWDC	0.9091	
	15	*FRs+CGIs+MWDC*	0.9183	.1646*
		FRs+CGIs+SWDC	0.9114	

*p<.05, **p<.01, ***p<.001

The proposed network structures and models are further compared with *Role2Vec* (Ahmed et al., 2018), *DeepWalk* (Perozzi et al., 2014), and *Role Aware Random Walk for Network Embedding* or RARE for short (Zhang et al., 2024). The experimental results, summarised in **Table 3.10**, indicate that *Node2Vec* consistently outperforms *Role2Vec*, *DeepWalk*, and *RARE* across various network structures and metrics in the context of bankruptcy prediction. When using the undirected company-company direct network combined with FRs and CGIs, *Node2Vec* demonstrates superior performance with an out-of-sample AUC score of 0.9015, highlighting its ability to balance local and global network structures effectively. In the company-director-company bipartite networks with both simple direct connections and connections weighted by directors' power, *Node2Vec* still achieves the highest out-of-sample AUC scores (i.e., 0.8836 and 0.8845, respectively), outperforming *Role2Vec* (i.e., 0.8563 and 0.8584, respectively), *DeepWalk* (i.e., 0.8582 and 0.8584, respectively), and *RARE* (i.e., 0.8255 and 0.8272, respectively). This underscores *Node2Vec*'s flexibility in capturing diverse neighbourhood contexts, making it robust for various network structures. In more complex undirected stacked time-weighted company-company networks, *Node2Vec* maintains its lead with out-of-sample AUC scores of 0.9183 and 0.9114, demonstrating its capability to effectively integrate temporal information and capture evolving network dynamics. Whilst *Role2Vec* and *RARE* are role-based methods with *Role2Vec* being popular and *RARE* is the latest promising methodology, their performance in these contexts is lower, with AUC scores of 0.8611 and 0.8571 for *Role2Vec* and 0.8237 and 0.8223 for *RARE*. *DeepWalk*, as a deep learning method incorporating the random walk concept, shows competitive performance with out-of-sample AUC scores of 0.8627 and 0.8793, but still falls short of *Node2Vec*'s comprehensive embedding capabilities. These findings suggest that *Node2Vec* is a robust choice for node embedding in bankruptcy prediction.

Table 3.10 Comparison of Performance of Different Node Embedding Methods with **Experimental Setup I**

Results - Mean	**Sensitivity (TPR)**		**Specificity (TNR)**		**Type-I Error Rates**		**Type-II Error Rates**		**Classification Accuracy (ACC)**		**Area Under ROC Curve (AUC)**
	TPR Train	***TPR Test***	***TNR Train***	***TNR Test***	***Type-I Train***	***Type-I Test***	***Type-II Train***	***Type-II Test***	***Acc Train***	***Acc Test***	***AUC Test***
Undirected Company-Company Direct Network (with FRs and CGIs)											
Node2Vec	73.78%	72.98%	90.75%	90.68%	9.25%	9.32%	26.22%	27.02%	90.57%	90.50%	0.9015
Role2Vec	69.51%	68.70%	87.22%	87.23%	12.78%	12.77%	30.49%	31.30%	87.04%	87.04%	0.8566
DeepWalk	68.30%	66.85%	88.15%	88.15%	11.85%	11.85%	31.70%	33.15%	87.93%	87.93%	0.8604
RARE	68.48%	67.10%	83.10%	83.08%	16.90%	16.92%	31.52%	32.90%	82.95%	82.92%	0.8313
Undirected Company-Director-Company Bipartite Network with Simple Direct Connections (with FRs and CGIs)											
Node2Vec	71.51%	71.08%	89.56%	89.54%	10.44%	10.46%	28.49%	28.92%	89.36%	89.34%	0.8836
Role2Vec	67.52%	67.70%	87.71%	87.64%	12.29%	12.36%	32.48%	32.30%	87.49%	87.43%	0.8563
DeepWalk	69.07%	67.25%	88.81%	88.78%	11.19%	11.22%	30.93%	32.75%	88.60%	88.56%	0.8582
RARE	68.44%	67.12%	81.90%	81.92%	18.10%	18.08%	31.56%	32.88%	81.76%	81.76%	0.8255
Undirected Company-Director-Company Bipartite Network Weighted by Directors' Power (with FRs and CGIs)											
Node2Vec	69.21%	68.71%	90.63%	90.57%	9.37%	9.43%	30.79%	31.29%	90.40%	90.34%	0.8845
Role2Vec	67.52%	67.98%	87.43%	87.41%	12.57%	12.59%	32.48%	32.02%	87.21%	87.21%	0.8584
DeepWalk	66.78%	65.09%	89.11%	89.08%	10.89%	10.92%	33.22%	34.91%	88.87%	88.84%	0.8584
RARE	68.45%	66.41%	82.24%	82.23%	17.76%	17.77%	31.55%	33.59%	82.09%	82.06%	0.8272
Undirected Stacked Time-Weighted Company-Company Network with the Maximum Weights (with FRs and CGIs)											
Node2Vec	76.36%	75.78%	92.74%	92.70%	7.26%	7.30%	23.64%	24.22%	92.56%	92.53%	0.9183
Role2Vec	70.71%	67.74%	88.51%	88.55%	11.49%	11.45%	29.29%	32.26%	88.32%	88.33%	0.8611
DeepWalk	68.69%	65.96%	88.25%	88.32%	11.75%	11.68%	31.31%	34.04%	88.05%	88.09%	0.8627
RARE	67.28%	65.63%	82.76%	82.76%	17.24%	17.24%	32.72%	34.37%	82.59%	82.59%	0.8237
Undirected Stacked Time-Weighted Company-Company Network with the Sum of Weights (with FRs and CGIs)											
Node2Vec	76.21%	74.72%	92.26%	92.24%	7.74%	7.76%	23.79%	25.28%	92.09%	92.06%	0.9114
Role2Vec	68.95%	67.14%	88.22%	88.26%	11.78%	11.74%	31.05%	32.86%	88.02%	88.05%	0.8571
DeepWalk	70.95%	69.75%	89.38%	89.30%	10.62%	10.70%	29.05%	30.25%	89.18%	89.10%	0.8793
RARE	68.53%	67.04%	81.75%	81.82%	18.25%	18.18%	31.47%	32.96%	81.61%	81.66%	0.8223

In order to further explore the potential of CRIs, three additional tests, which incorporate different prevalent machine learning prediction models (i.e., decision tree, multi-layer perceptron, and AdaBoost) fed with the proposed CRIs have been analysed. **Appendix 3.A.2** summarises the results of these tests. The results demonstrate varying levels of predictive performance across different models. Whilst the LR models exhibit robust performance, particularly when combined with *MWDC* and *SWDC*, achieving high AUC scores and balanced sensitivity and specificity, decision tree models, on the other hand, show strong sensitivity in training but exhibit a decline in out-of-sample sensitivity, suggesting their potential of overfitting. The multi-layer perceptron models, although demonstrating high training sensitivity, however, generally suffer from significant drops in out-of-sample sensitivity, which indicates that the complex nature of these models may result in instability in generalisation. The AdaBoost models generally provide a balanced performance similar to logistic regression. Overall, logistic regression remains a good choice due to its consistent performance and balanced trade-offs between sensitivity, specificity, and overall classification accuracy. In addition, the consistency and interpretability of logistic regression, coupled with its competitive performance metrics, indicates its suitability as a reliable model for predicting bankruptcy using CRIs.

Furthermore, we examine other community detection methods such as DBSCAN, *k-means*, and BIRCH, along with OPTICS, using the *FRs+CGIs+DC* model to predict bankruptcy. The results, presented in **Appendix 3.A.3**, indicate that whilst all methods exhibit reasonable predictive performance, there are notable differences. The *k-means* method yielded a test sensitivity of 72.50% and an AUC of 0.8549, demonstrating good overall performance, but slightly lower specificity compared to OPTICS. DBSCAN showed lower out-of-sample sensitivity at 66.64% and an AUC of 0.8186, indicating that it might not capture the clusters as effectively as other methods. BIRCH, with an out-of-sample sensitivity of 72.65% and an AUC of 0.8546, performed similarly to *k-means*. OPTICS, however, achieved the highest AUC of 0.9015 and the best specificity, highlighting its superior ability to handle variable density clusters and noise. These comparisons suggest that whilst partitioning methods like *k-means* and BIRCH are effective, OPTICS remains the preferred choice in our context due to its adaptability and nuanced clustering capability, especially in datasets with varying densities and noise levels.

We also investigated the statistical properties of successful and failed predictions, with an in-depth analysis conducted on the out-of-sample test dataset using the *FRs+CGIs+DC* model. The statistical properties of correctly predicted bankrupt companies (i.e., labelled as 1) and incorrectly predicted bankrupt companies (i.e., labelled as 0) are compared in **Appendix 3.A.4**. For successful predictions, companies tend to have lower *working capital/total assets* and *retained earnings/total assets* ratios, indicating relatively weaker liquidity and profitability performance. Additionally, they also exhibit higher *market value/total liabilities* ratios, suggesting a relatively better market valuation despite they may be in financial distress. Incorrect predictions, on the other hand, were characterised by more extreme values in the above-mentioned financial ratios, particularly with higher variances, which may indicate unstable financial conditions that are hard to capture. The presence of *CEO duality* and a greater *number of directors* was also more prominent in successful predictions, which may reflect the possibly stronger governance structures. These findings suggest that the proposed method works effectively by leveraging financial and governance indicators to identify distress patterns, although extreme or highly volatile indicators can challenge prediction performance.

3.6.2 Analysis of results of Experimental Setup II

In this section, we shall analyse the results of Experimental Setup II to answer the last two research questions.

In this experiment, the best-performing models of each combination of drivers are selected to examine their performance both for short-term and long-term bankruptcy prediction. The selection of these models followed a performance-based approach, where only those demonstrating superior performance across various metrics (AUC, accuracy, sensitivity, and specificity) in Experimental Setup I were further evaluated. However, we provide detailed results for all models in **Appendix 3.A.5**. For models based on static complex network structures, the model that combines FRs, CGIs, and the *Company-Company Direct Connections-based Driver* (i.e., *FRs+CGIs+DC*) is selected; for models based on stacked time-weighted complex network structures, the model that combines FRs, CGIs, and the *Stacked Time-weighted Company-Company Direct Connections-based Driver with the Maximum Weights* (i.e., *FRs+CGIs+MWDC*) is selected. Two aspects were considered in choosing this approach, i.e., efficiency and relevance. Efficiency was ensured by focussing resources on models with proven effectiveness in a controlled environment, thereby making subsequent testing phases practical and manageable. Relevance was achieved by testing only the best models, aligning our testing procedures with typical industry practices where only promising solutions are scaled or deployed.

In addition, the model based on *FRs+CGIs* is also selected to examine the value added of CGIs in long-term prediction, and the model using FRs alone is selected for comparison purposes. The performance results are summarised in **Table 3.11**.

Table 3.11 Results of **Experimental Setup II** *(Training: 1990-2014 @70%; Test: 2016,2017,2018,2019)*

Results - Mean	**Sensitivity (TPR)**		**Specificity (TNR)**		**Type-I Error Rates**		**Type-II Error Rates**		**Classification Accuracy (ACC)**		**Area Under ROC Curve (AUC)**
	TPR Train	***TPR Test***	***TNR Train***	***TNR Test***	***Type-I Train***	***Type-I Test***	***Type-II Train***	***Type-II Test***	***Acc Train***	***Acc Test***	***AUC Test***
Training Period: 1990-2014 @70%; Test Period: 2016; 2-year Lag)											
FRs	69.69%	89.05%	81.43%	79.40%	18.57%	20.60%	30.31%	10.95%	81.29%	79.50%	0.9319
FRs+CGIs	70.11%	84.76%	80.76%	84.68%	19.24%	15.32%	29.89%	15.24%	80.64%	84.68%	0.9335
FRs+CGIs+DC	74.17%	85.71%	90.30%	92.53%	9.70%	7.47%	25.83%	14.29%	90.11%	92.47%	0.9670
FRs+CGIs+MWDC	76.95%	95.24%	92.82%	94.46%	7.18%	5.54%	23.05%	4.76%	92.64%	94.46%	0.9806
Training Period: 1990-2014 @70%; Test Period: 2017; 3-year Lag)											
FRs	69.69%	64.58%	81.43%	82.93%	18.57%	17.07%	30.31%	35.42%	81.29%	82.72%	0.7973
FRs+CGIs	70.11%	65.83%	80.76%	86.48%	19.24%	13.52%	29.89%	34.17%	80.64%	86.23%	0.8047
FRs+CGIs+DC	74.17%	90.00%	90.30%	92.43%	9.70%	7.57%	25.83%	10.00%	90.11%	92.40%	0.9447
FRs+CGIs+MWDC	76.95%	65.00%	92.82%	94.34%	7.18%	5.66%	23.05%	35.00%	92.64%	93.99%	0.8852
Training Period: 1990-2014 @70%; Test Period: 2018; 4-year Lag)											
FRs	69.69%	75.00%	81.43%	79.54%	18.57%	20.46%	30.31%	25.00%	81.29%	79.51%	0.7327
FRs+CGIs	70.11%	54.17%	80.76%	85.15%	19.24%	14.85%	29.89%	45.83%	80.64%	84.97%	0.7079
FRs+CGIs+DC	74.17%	50.00%	90.30%	93.72%	9.70%	6.28%	25.83%	50.00%	90.11%	93.46%	0.7992
FRs+CGIs+MWDC	76.95%	25.00%	92.82%	94.69%	7.18%	5.31%	23.05%	75.00%	92.64%	94.28%	0.6735
Training Period: 1990-2014 @70%; Test Period: 2019; 5-year Lag)											
FRs	69.69%	N/A	81.43%	77.13%	18.57%	22.87%	30.31%	N/A	81.29%	77.13%	N/A
FRs+CGIs	70.11%	N/A	80.76%	83.61%	19.24%	16.39%	29.89%	N/A	80.64%	83.61%	N/A
FRs+CGIs+DC	74.17%	N/A	90.30%	93.86%	9.70%	6.14%	25.83%	N/A	90.11%	93.86%	N/A
FRs+CGIs+MWDC	76.95%	N/A	92.82%	95.08%	7.18%	4.92%	23.05%	N/A	92.64%	95.08%	N/A

Note: *FRs* – Financial ratios; *CGIs* – Corporate governance indicators; *DC* – Company-Company direct connections-based driver; *MWDC* – Stacked time-weighted Company-Company direct connections-based driver with the maximum weights.

Regarding the sixth research question, namely 'How do CRIs based on different network structures compare to predict short- and long-term corporate bankruptcy?', regarding the overall performance of the prediction models, the results of experiments based on different prediction terms for all the models suggest a decreasing trend in prediction performance as the prediction time window increases. Attention must be paid to the results for 2019 in that the out-of-sample sensitivity score, Type-II error, and AUC score are not available due to the fact that there is no bankrupt company in the processed database. Generally, there is a decreasing trend in AUC scores for all prediction models as the prediction term increases from a 2-year lag to a 4-year lag, suggesting that the separability of these models decreases when predicting bankruptcy events very long from now in the future. The results make sense, as the fundamentals of companies change over time, which would introduce many more uncertainties into the future financial stability of companies. Furthermore, it has generally been proven that bankruptcy prediction models struggle to make reliable long-term predictions (e.g., greater than a 3-year lag) as compared to short-term predictions (e.g., less than a 3-year lag) (e.g., Beaver, 1966; Du Jardin & Severin, 2011; Du Jardin, 2015). However, since the AUC scores of the prediction models are all above 0.5, it suggests that the models based on three types of drivers (i.e., FRs, CGIs, and CRIs) are still capable of producing reasonable predictions other than making random guesses.

Regarding the performance of drivers based on accounting information and corporate governance indicators, the results of experiments based on different terms of prediction suggest that models based on FRs alone and FRs with CGIs still show good performance in terms of detecting bankrupt companies, as well as overall prediction accuracy. For the model based solely on FRs, the model achieves a good overall out-of-sample classification accuracy at around 80% from a 2- to 5-year lag. However, when comparing the sensitivity scores and AUC scores for out-of-sample predictions, the results suggest a decrease in the performance of detecting bankrupt companies as the predicted time period increases. The model based on the combination of FRs and CGIs also demonstrates similar behaviour in terms of overall out-of-sample classification accuracy at around 84%; however, the sensitivity scores on out-of-sample predictions drop significantly as the lag increases, from 84.76% for a 2-year lag to 54.17% for 4-year lag. It suggests that adding CGIs may not be helpful in long-term bankruptcy detection; however, this model achieves significantly higher specificity scores (i.e., 5% improvement on average) and overall classification accuracy (i.e., 5% improvement on average) than the model based on FRs alone for out-of-sample predictions. This indicates that CGIs still provide valuable information to predict the financial stability of a company.

Regarding the performance of drivers based on company relational information, the performance of models that are based on CRIs (i.e., *FRs+CGIs+DC* and *FRs+CGIs+MWDC*) suggests that the proposed models perform well in-sample, with both models achieving sensitivity scores at around 75%, which are on average 5% higher than the models without CRIs. Higher in-sample specificity scores (at around 91%) and overall classification accuracy (at around 91%) also prove that CRIs are powerful in distinguishing healthy and bankrupt companies. The out-of-sample results also suggest that models with company relational information-based drivers still perform well and beat benchmark models (i.e., *FRs* and *FRs+CGIs*) within a 3-year lag in sensitivity scores, specificity scores, classification accuracy, and AUC scores. However, the results for 2018 do not support the idea that models with CRIs are performing well, as the sensitivities scores drop significantly (i.e., by 40% for *FRs+CGIs+MWDC* model) and the further resulting drop in AUC scores. After investigating the results, we found that it is due to the lack of bankrupt observations, as there are only four bankrupt companies in the test set for that specific year. In addition, among them, three of those bankrupt companies do not have any external connection.

With respect to the seventh research question, namely, 'Do CRIs based on time-stacked networks predict long-term corporate bankruptcy better than static network structures?', numerical results suggest that, among the two models based on CRIs (i.e., *FRs+CGIs+DC* and *FRs+CGIs+MWDC*), the model based on time-weighted network structure-based CRIs (i.e., *FRs+CGIs+MWDC*) generally performs better than the model with the combination of FRs, CGIs and DC within 2-year lag of prediction, as it achieves higher out-of-sample sensitivity, specificity scores, and classification accuracy. However, since its performance drops in 2017 (3-year lag) and 2018 (4-year lag) both in sensitivity scores and AUC scores, it shows that CRIs based on the time-weighted network structure can be weak when some of the connection information prior to the year of prediction is not available. Furthermore, it is crucial to acknowledge that the structure of the network itself may evolve over time due to changes in corporate governance, board composition, and inter-company relationships. These temporal dynamics could influence the effectiveness of CRIs, as historical connections may no longer represent current realities, potentially impacting both the performance and generalisability of the predictions. This evolution highlights the potential need to consider temporal variations when interpreting the results and conclusions. We conclude that the proposed models based on CRIs are still able to make reliable predictions within a 3-year lag, however, the performance may drop as it is largely dependent on the quality of data in terms of the presence or absence of connections.

In sum, the results of Experimental Setup II suggest that CRIs, when incorporated into models based on different network structures, show varying degrees of effectiveness in predicting short- and long-term corporate bankruptcy. Our results indicate a general trend of decreasing prediction accuracy as the forecast period extends, highlighting the challenges in long-term forecasting due to evolving company fundamentals. Specifically, models that combine FRs, CGIs, and the *Company Direct Connections Driver* (DC) or the *Stacked Time Weighted Company-Company Direct Connections-based Driver with Maximum Weights* (MWDC) demonstrate the best predictive abilities. However, the temporal evolution of the network structure adds complexity to long-term forecasts, as the dynamics of relationships between companies may weaken or strengthen over time. Incorporating methods to dynamically update the network representation could mitigate these challenges and improve model robustness for long-term predictions. Our investigation also reveals that CRIs based on time-weighted network structures tend to outperform those based on static network structures in the context of long-term bankruptcy prediction. However, this superiority is dependent on the quality and completeness of the data, as demonstrated by performance fluctuations in specific years. These findings suggest the potential of CRIs in improving bankruptcy prediction models, especially for medium- to long-term forecasts, whilst also highlighting the limitations and challenges in predictions over longer durations.

In addition, the full results of Experimental Setup II, as presented in **Appendix 3.A.5**, provide detailed insights into the performance of all the proposed models over different time lags. For the 2-year lag, models combining FRs with CGIs and CRIs, especially those using MWDC and SWDC (i.e., *FRs+CGIs+MWDC*, i.e., *FRs+CGIs+SWDC*), show significant predictive ability with high AUC scores and classification accuracy. As the lag increases to 3 and 4 years, there is a notable decline in sensitivity, indicating a reduced ability to predict bankruptcies accurately, although specificity remains relatively high. This trend is evident in most models, underscoring the inherent difficulty in making long-term predictions. In particular, the performance of models solely based on CGIs is considerably lower compared to those incorporating CRIs, which reinforces the added value of relational information in bankruptcy prediction. Overall, the comprehensive results underscore the importance of CRIs in enhancing predictive performance, particularly in short- to medium-term forecasts.

3.7 Conclusion and future research directions

This chapter proposes a significant advance in bankruptcy prediction by integrating a new method that uses complex network analysis focused on company relationships. Our research, focused on UK companies listed on the London Stock Exchange, demonstrates that incorporating company relational information-based drivers, along with traditional financial ratios and corporate governance indicators, enhances the accuracy of bankruptcy predictions.

The key findings of our analysis reveal that, whilst FRs provide a solid foundation for predicting bankruptcy, their effectiveness is notably enhanced when combined with CRIs. These CRIs, derived from board-of-director connections networks, offer valuable information on the interconnectedness of companies, which is often overlooked in conventional models.

Our results indicate that models using a combination of FRs and CRIs are particularly effective, outperforming those that rely solely on FRs or CGIs. Among the various network structures that we examined, company-company direct connection-based CRIs emerged as the most effective in improving prediction accuracy. This highlights the importance of considering direct company relationships in assessing bankruptcy risk.

The study also explores the impact of different network structures on the accuracy of the prediction. It finds that time-weighted network structures, which consider the evolving nature of company relationships over time, further improve the performance of the model. This suggests that dynamic changes in company networks are crucial for an accurate assessment of the risk of bankruptcy.

In summary, our research confirms that incorporating network analysis of company relationships into bankruptcy prediction models provides a more comprehensive and accurate approach. This method has significant potential for financial institutions and analysts, offering a nuanced understanding of the risks associated with corporate bankruptcy.

In summary, this chapter contributes to the overarching narrative of the thesis by demonstrating how network analysis can address critical gaps in traditional bankruptcy prediction methodologies. By integrating company relational information-based drivers into prediction models, this research underscores the importance of capturing the interconnected nature of businesses in an increasingly connected economy. The findings highlight the broader relevance of this approach, offering financial institutions and analysts a more holistic and dynamic framework for assessing bankruptcy risk. Furthermore, this chapter complements the insights from **Chapters 2** by demonstrating how advanced data-driven techniques, such as network analysis, can be used in conjunction with established financial and governance indicators to improve prediction accuracy and interpretability. Collectively, this research advances the thesis's aim of creating robust and adaptable tools for corporate bankruptcy prediction.

Whilst our findings are promising, they also open several avenues for future exploration.

Automatic Feature Selection Techniques: Although our current model leverages proven financial ratios and corporate governance indicators, future studies could employ automatic feature selection techniques, such as *Relief Algorithm*, *Recursive Feature Elimination* (RFE) and *Decision Tree -Based Embedded Feature Selection* (Htun et al., 2023) to choose the final set of drivers from an initial exhaustive list (see for example Zhao et al., 2024) complemented with our proposed CRI drivers.

Extending Across Different Markets: Although our study offers comprehensive insights into the UK market, its applicability to other markets remains untested because the goal is not to provide a global bankruptcy prediction model. Future research could aim to test this methodology in different geographical and regulatory environments to validate the external robustness of the model using eventually different sets of drivers for different markets.

Exploration of Time-Weighted Networks in Other Contexts: The use of time-weighted networks based on the 'half-life' concept has proven effective in our analyses. However, the optimal decay rate might vary by industry or economic condition. Future research could explore dynamic weighting schemes where decay rates are adjusted based on contextual factors, such as economic cycles. This would allow the model to adapt to changing environments, potentially increasing its predictive accuracy.

Integration with Other Data Sources: Integrating additional data sources such as macroeconomic indicators or industry-specific trends could enhance the predictive power. For instance, including macroeconomic factors might help capture external influences on a company's bankruptcy risk that are not evident from financial ratios or governance structures alone.

Application of Alternative State-of-the-Art Models: Whilst our study primarily uses Node2Vec for generating embeddings, the exploration of other neural network architectures, such as Convolutional Neural Networks (CNNs) or Recurrent Neural Networks (RNNs), could offer additional benefits. These architectures are particularly adept at handling sequential and graph-structured data and might provide new insights into the temporal and relational dynamics of corporate bankruptcy.

Leadership and Optimal Board Composition in Network-Based Models: Leadership plays a critical role in corporate governance and directly influences the effectiveness of decision-making within companies. Future research could explore the relationship between board leadership and the predictive power of network-based approaches to bankruptcy. For example, investigating the optimal number of directors on a board is essential, as overly large boards may lead to diluted accountability and technostructure-based governance. Conversely, smaller boards may lack the diversity needed for robust decision-making. Additionally, the influence of strong leadership within networks could be studied to assess its impact on corporate risk and resilience. For example, incorporating metrics that evaluate the presence and quality of leadership within board networks could refine bankruptcy prediction models and offer deeper insights into how governance structures impact financial stability.

By addressing these areas, future research could not only reinforce the findings of this study but also expand the scope of network analysis in financial prediction models. Each direction offers a promising path toward developing more nuanced and adaptable tools for financial analysis and risk assessment.

3.A Appendix

3.A.1 Summary statistics of processed dataset (FRs & CGIs)

	Sales/Total Assets							
	count	**mean**	**std**	**min**	**25%**	**50%**	**75%**	**max**
Healthy	21403	1.1200	0.8746	0.0069	0.4827	0.9373	1.5166	4.4357
Bankrupt	228	1.1273	1.0381	0.0069	0.4176	0.8215	1.4586	4.4357
	Working Capital/Total Assets							
	count	**mean**	**std**	**min**	**25%**	**50%**	**75%**	**max**
Healthy	21403	0.1502	0.2729	-0.7266	-0.0155	0.1206	0.2981	0.8529
Bankrupt	228	-0.0776	0.3580	-0.7266	-0.2910	-0.0468	0.1641	0.8368
	Retained Earnings/Total Assets							
	count	**mean**	**std**	**min**	**25%**	**50%**	**75%**	**max**

Healthy	21403	-0.3538	1.5687	-9.5319	-0.2290	0.1150	0.3016	0.7459
Bankrupt	228	-1.6510	2.7347	-9.5319	-1.9113	-0.5744	-0.0302	0.6815
				EBITDA/Total Assets				
	count	**mean**	**std**	**min**	**25%**	**50%**	**75%**	**max**
Healthy	21403	0.0214	0.3008	-1.6261	0.0015	0.0955	0.1613	0.4421
Bankrupt	228	-0.3088	0.5142	-1.6261	-0.4671	-0.1087	0.0391	0.4421
				Market Value/Total Liabilities				
	count	**mean**	**std**	**min**	**25%**	**50%**	**75%**	**max**
Healthy	21403	5.6401	11.8308	0.05862	0.9091	1.9654	4.6331	79.8472
Bankrupt	228	2.2003	6.4729	0.05862	0.1421	0.5821	1.5502	57.5949
				CEO Chairman Duality				
	count	**mean**	**std**	**min**	**25%**	**50%**	**75%**	**max**
Healthy	21403	0.2647	0.4411	0	0	0	1	1
Bankrupt	228	0.3816	0.4868	0	0	0	1	1
				Number of Directors				
	count	**mean**	**std**	**min**	**25%**	**50%**	**75%**	**max**
Healthy	21403	11.5377	9.5409	1	6	9	14	113
Bankrupt	228	9.6798	4.9123	2	6	8.5	11	39
				Percentage of Non-Executive Directors				
	count	**mean**	**std**	**min**	**25%**	**50%**	**75%**	**max**
Healthy	21403	0.4036	0.1904	0	0.2727	0.3846	0.5	1
Bankrupt	228	0.4093	0.1685	0	0.2920	0.4	0.5	1

3.A.2 Results of CRIs using machine learning models for prediction with Experimental Setup I

Results - Mean	**Sensitivity (TPR)**		**Specificity (TNR)**		**Type-I Error Rates**		**Type-II Error Rates**		**Classification Accuracy (ACC)**		**Area Under ROC Curve (AUC)**
	TPR Train	***TPR Test***	***TNR Train***	***TNR Test***	***Type-I Train***	***Type-I Test***	***Type-II Train***	***Type-II Test***	***Acc Train***	***Acc Test***	***AUC Test***
					Prediction Model: Logistic Regression						
FRs	66.80%	65.85%	83.48%	83.48%	16.52%	16.52%	33.20%	34.15%	83.31%	83.30%	0.8156
FRs+CGIs	68.40%	66.43%	81.11%	81.16%	18.89%	18.84%	31.60%	33.57%	80.98%	81.01%	0.8186
FRs+CGIs+DC	73.78%	72.98%	90.75%	90.68%	9.25%	9.32%	26.22%	27.02%	90.57%	90.50%	0.9015
FRs+CGIs+SBC	71.51%	71.08%	89.56%	89.54%	10.44%	10.46%	28.49%	28.92%	89.36%	89.34%	0.8836
FRs+CGIs+PBC	69.78%	68.85%	89.64%	89.60%	10.36%	10.40%	30.22%	31.15%	89.43%	89.39%	0.8834
FRs+CGIs+MWDC	76.36%	75.78%	92.74%	92.70%	7.26%	7.30%	23.64%	24.22%	92.56%	92.53%	0.9183
FRs+CGIs+SWDC	76.21%	74.72%	92.26%	92.24%	7.74%	7.76%	23.79%	25.28%	92.09%	92.06%	0.9114
					Prediction Model: Decision Tree						
FRs	87.02%	64.77%	78.63%	78.51%	21.37%	21.49%	12.98%	35.23%	78.72%	78.36%	0.7673
FRs+CGIs	89.40%	67.16%	76.80%	76.60%	23.20%	23.40%	10.60%	32.84%	76.93%	76.50%	0.7649
FRs+CGIs+DC	86.40%	69.57%	89.84%	89.70%	10.16%	10.30%	13.60%	30.43%	89.80%	89.49%	0.8317
FRs+CGIs+SBC	86.89%	70.57%	85.86%	85.73%	14.14%	14.27%	13.11%	29.43%	85.87%	85.57%	0.8146
FRs+CGIs+PBC	87.46%	70.57%	86.35%	86.20%	13.65%	13.80%	12.54%	29.43%	86.37%	86.04%	0.8199
FRs+CGIs+MWDC	89.06%	73.92%	91.48%	91.40%	8.52%	8.60%	10.94%	26.08%	91.45%	91.22%	0.8469
FRs+CGIs+SWDC	90.25%	71.36%	88.81%	88.69%	11.19%	11.31%	9.75%	28.64%	88.83%	88.52%	0.8312
					Prediction Model: Multi-Layer Perceptron						
FRs	90.87%	56.55%	84.92%	84.87%	15.08%	15.13%	9.13%	43.45%	84.98%	84.57%	0.7859
FRs+CGIs	96.35%	45.19%	89.61%	89.25%	10.39%	10.75%	3.65%	54.81%	89.68%	88.79%	0.7657
FRs+CGIs+DC	98.85%	54.48%	94.90%	94.58%	5.10%	5.42%	1.15%	45.52%	94.94%	94.16%	0.8300
FRs+CGIs+SBC	97.97%	55.46%	93.98%	93.69%	6.02%	6.31%	2.03%	44.54%	94.02%	93.29%	0.8196
FRs+CGIs+PBC	98.28%	54.07%	93.79%	93.44%	6.21%	6.56%	1.72%	45.93%	93.84%	93.03%	0.8238
FRs+CGIs+MWDC	99.17%	61.89%	96.44%	96.11%	3.56%	3.89%	0.83%	38.11%	96.46%	95.76%	0.8587
FRs+CGIs+SWDC	98.66%	60.94%	95.73%	95.49%	4.27%	4.51%	1.34%	39.06%	95.76%	95.13%	0.8584
					Prediction Model: AdaBoost						
FRs	81.38%	74.26%	74.96%	74.99%	25.04%	25.01%	18.62%	25.74%	75.02%	74.98%	0.8127
FRs+CGIs	81.52%	72.44%	77.45%	77.39%	22.55%	22.61%	18.48%	27.56%	77.50%	77.34%	0.8204
FRs+CGIs+DC	75.43%	68.48%	89.97%	89.97%	10.03%	10.03%	24.57%	31.52%	89.82%	89.76%	0.8714
FRs+CGIs+SBC	82.43%	76.93%	77.62%	77.59%	22.38%	22.41%	17.57%	23.07%	77.67%	77.59%	0.8169
FRs+CGIs+PBC	82.73%	76.92%	80.03%	79.98%	19.97%	20.02%	17.27%	23.08%	80.06%	79.95%	0.8415
FRs+CGIs+MWDC	82.07%	77.44%	91.16%	91.19%	8.84%	8.81%	17.93%	22.56%	91.06%	91.05%	0.9089
FRs+CGIs+SWDC	78.87%	72.95%	88.61%	88.75%	11.39%	11.25%	21.13%	27.05%	88.51%	88.59%	0.8581

Note: *FRs* – Financial ratios; *CGIs* – Corporate governance indicators; *DC* – Company-Company direct connections-based driver; *SBC* – Company-Director-Company bipartite connections-based driver with simple direct connections; *PBC* – Company-Director-Company bipartite connections-based driver with consideration of directors' power; *MWDC* – Stacked time-weighted Company-Company direct connections-based driver with the maximum weights; *SWDC* – Stacked time-weighted Company-Company direct connections-based driver with the sum of weights.

3.A.3 An example of prediction results using CRIs generated by different community detection methods with Experimental Setup I (Model: FRs+CGIs+DC)

Results - Mean	**Sensitivity (TPR)**		**Specificity (TNR)**		**Type-I Error Rates**		**Type-II Error Rates**		**Classification Accuracy (ACC)**		**Area Under ROC Curve (AUC)**
	TPR Train	***TPR Test***	***TNR Train***	***TNR Test***	***Type-I Train***	***Type-I Test***	***Type-II Train***	***Type-II Test***	***Acc Train***	***Acc Test***	***AUC Test***
					Model Example: FRs+CGIs+DC						
OPTICS	73.78%	72.98%	90.75%	90.68%	9.25%	9.32%	26.22%	27.02%	90.57%	90.50%	0.9015
k-means	73.80%	72.50%	82.74%	82.72%	17.26%	17.28%	26.20%	27.50%	82.64%	82.62%	0.8549
DBSCAN	68.56%	66.64%	81.07%	81.12%	18.93%	18.88%	31.44%	33.36%	80.94%	80.97%	0.8186
BIRCH	72.74%	72.65%	83.17%	83.18%	16.83%	16.82%	27.26%	27.35%	83.06%	83.07%	0.8546

3.A.4 An example of statistical properties of successful and failed predictions (Model: FRs+CGIs+DC)

				Working Capital/Total Assets				
Predict	**count**	**mean**	**std**	**min**	**25%**	**50%**	**75%**	**max**
0	12	0.0419	0.2361	-0.3269	-0.1929	0.0365	0.2218	0.3542
1	57	-0.1514	0.3445	-0.7266	-0.3797	-0.1016	0.0733	0.7429
				Retained Earnings/Total Assets				
Predict	**count**	**mean**	**std**	**min**	**25%**	**50%**	**75%**	**max**
0	12	-0.7609	1.7198	-5.8946	-0.8161	-0.1635	0.0033	0.5092
1	57	-1.7127	2.437	-9.5319	-2.1779	-0.7811	-0.1843	0.4068
				EBITDA/Total Assets				
Predict	**count**	**mean**	**std**	**min**	**25%**	**50%**	**75%**	**max**
0	12	-0.1343	0.4709	-1.4950	-0.1350	0.0159	0.0683	0.3380
1	57	-0.3588	0.5219	-1.6261	-0.5592	-0.1493	-0.0139	0.2910
				Market Value/Total Liabilities				
Predict	**count**	**mean**	**std**	**min**	**25%**	**50%**	**75%**	**max**
0	12	5.9700	16.3565	0.0767	0.1558	0.6193	1.8914	57.5949
1	57	1.0625	2.8678	0.0586	0.1340	0.5035	1.0579	21.5423
				Sales/Total Assets				
Predict	**count**	**mean**	**std**	**min**	**25%**	**50%**	**75%**	**max**
0	12	0.8839	0.7855	0.0069	0.2973	0.7184	1.2593	2.7696
1	57	1.1378	0.9692	0.0126	0.5143	0.8198	1.6388	4.4357
				CEO/Chairman Duality				
Predict	**count**	**mean**	**std**	**min**	**25%**	**50%**	**75%**	**max**
0	12	0.3333	0.4924	0	0	0	1	1
1	57	0.3333	0.4756	0	0	0	1	1
				Number of Directors				
Predict	**count**	**mean**	**std**	**min**	**25%**	**50%**	**75%**	**max**
0	12	11.4167	9.6338	4	8	8	9.25	39
1	57	9.614	4.3456	2	7	9	11	29
				Percentage of Non-executive Directors				
Predict	**count**	**mean**	**std**	**min**	**25%**	**50%**	**75%**	**max**
0	12	0.4371	0.2154	0.0769	0.2806	0.5000	0.6062	0.7778
1	57	0.3791	0.1783	0.0909	0.2667	0.3750	0.5	0.8571
				Failure Proximity Score				
Predict	**count**	**mean**	**std**	**min**	**25%**	**50%**	**75%**	**max**
0	12	0	0	0	0	0	0	0
1	57	0.1657	0.1346	0	0	0.1667	0.3333	0.3333

Note: *0* indicate a bankrupt company failed to predict as bankrupt*; 1* indicate a bankrupt company successfully predicted as bankrupt*; FRs+CGIs+DC* model is chosen for analysis.

3.A.5 Full results of Experimental Setup II (Training: 1990-2014 @70%; Test: 2016,2017,2018,2019)

Results - Mean	**Sensitivity (TPR)**		**Specificity (TNR)**		**Type-I Error Rates**		**Type-II Error Rates**		**Classification Accuracy (ACC)**		**Area Under ROC Curve (AUC)**
	TPR Train	***TPR Test***	***TNR Train***	***TNR Test***	***Type-I Train***	***Type-I Test***	***Type-II Train***	***Type-II Test***	***Acc Train***	***Acc Test***	***AUC Test***
				Training Period: 1990-2014 @70%; Test Period: 2016; 2-year Lag)							
FRs	69.69%	89.05%	81.43%	79.40%	18.57%	20.60%	30.31%	10.95%	81.29%	79.50%	0.9319
CGIs	44.38%	44.76%	63.95%	80.52%	36.05%	19.48%	55.62%	55.24%	63.73%	80.16%	0.7150
DC	46.41%	57.14%	97.99%	97.44%	2.01%	02.56%	53.59%	42.86%	97.41%	97.05%	0.7752
SBC	36.08%	14.29%	98.94%	99.57%	1.06%	0.43%	61.92%	85.71%	98.25%	98.73%	0.5693
PBC	35.32%	28.57%	99.00%	99.43%	1.00%	0.57%	64.68%	71.43%	98.28%	98.73%	0.6402
MWDC	59.53%	42.86%	98.43%	98.15%	1.57%	1.85%	40.47%	57.14%	97.99%	97.61%	0.7057
SWDC	55.32%	57.14%	98.51%	97.59%	1.49%	2.41%	44.68%	42.86%	98.02%	97.19%	0.7745
FRs+CGIs	70.11%	84.76%	80.76%	84.68%	19.24%	15.32%	29.89%	15.24%	80.64%	84.68%	0.9335
FRs+DC	72.10%	87.62%	90.99%	92.30%	9.01%	7.70%	27.90%	12.38%	90.78%	92.25%	0.9763
FRs+SBC	71.02%	92.86%	90.47%	91.53%	9.53%	8.47%	28.98%	7.14%	90.25%	91.55%	0.9715
FRs+PBC	71.67%	92.86%	90.29%	91.81%	9.71%	8.19%	28.33%	7.14%	90.08%	91.82%	0.9747
FRs+MWDC	75.57%	99.52%	93.05%	93.74%	6.95%	6.26%	24.43%	0.48%	92.85%	93.80%	0.9814
FRs+SWDC	73.79%	100.00%	92.33%	92.58%	7.67%	7.42%	26.21%	0.00%	92.12%	92.65%	0.9857
FRs+CGIs+DC	74.17%	85.71%	90.30%	92.53%	9.70%	7.47%	25.83%	14.29%	90.11%	92.47%	0.9670
FRs+CGIs+SBC	72.58%	81.43%	89.31%	92.19%	10.69%	7.81%	27.42%	18.57%	89.12%	92.08%	0.9576
FRs+CGIs+PBC	72.52%	85.24%	89.46%	92.67%	10.54%	7.33%	27.48%	14.76%	89.27%	92.59%	0.9667
FRs+CGIs+MWDC	76.95%	95.24%	92.82%	94.46%	7.18%	5.54%	23.05%	4.76%	92.64%	94.46%	0.9806
FRs+CGIs+SWDC	75.42%	100.00%	92.11%	93.27%	7.89%	6.73%	24.58%	0.00%	91.92%	93.34%	0.9869
				Training Period: 1990-2014 @70%; Test Period: 2017; 3-year Lag)							
FRs	69.69%	64.58%	81.43%	82.93%	18.57%	17.07%	30.31%	35.42%	81.29%	82.72%	0.7973
CGIs	44.38%	25.42%	63.95%	80.70%	36.05%	19.30%	55.62%	74.58%	63.73%	80.05%	0.6147
DC	46.41%	62.50%	97.99%	97.31%	2.01%	02.69%	53.59%	37.50%	97.41%	96.89%	0.8013
SBC	38.08%	37.50%	98.94%	99.10%	1.06%	0.90%	61.92%	62.50%	98.25%	98.37%	0.6830
PBC	35.32%	25.00%	99.00%	98.95%	1.00%	1.05%	64.68%	75.00%	98.28%	98.08%	0.6200
MWDC	59.53%	62.50%	98.43%	97.60%	1.57%	2.40%	40.47%	37.50%	97.99%	97.19%	0.8014
SWDC	55.32%	87.50%	98.51%	97.16%	1.49%	2.84%	44.68%	12.50%	98.02%	97.04%	0.9248
FRs+CGIs	70.11%	65.83%	80.76%	86.48%	19.24%	13.52%	29.89%	34.17%	80.64%	86.23%	0.8047
FRs+DC	72.10%	92.92%	90.99%	92.29%	9.01%	7.71%	27.90%	7.08%	90.78%	92.29%	0.9384
FRs+SBC	71.02%	87.50%	90.47%	92.00%	9.53%	8.00%	28.98%	12.50%	90.25%	91.94%	0.8558
FRs+PBC	71.67%	75.00%	90.29%	92.46%	9.71%	7.54%	28.33%	25.00%	90.08%	92.25%	0.8384
FRs+MWDC	75.57%	77.92%	93.05%	93.58%	6.95%	6.42%	24.43%	22.08%	92.85%	93.40%	0.8637
FRs+SWDC	73.79%	96.25%	92.33%	92.66%	7.67%	7.34%	26.21%	3.75%	92.12%	92.71%	0.9554
FRs+CGIs+DC	74.17%	90.00%	90.30%	92.43%	9.70%	7.57%	25.83%	10.00%	90.11%	92.40%	0.9447
FRs+CGIs+SBC	72.58%	75.42%	89.31%	92.09%	10.69%	7.91%	27.42%	24.58%	89.12%	91.89%	0.8506
FRs+CGIs+PBC	72.52%	70.83%	89.46%	92.84%	10.54%	7.16%	27.48%	29.17%	89.27%	92.58%	0.8387

FRs+CGIs+MWDC	76.95%	65.00%	92.82%	94.34%	7.18%	5.66%	23.05%	35.00%	92.64%	93.99%	0.8852
FRs+CGIs+SWDC	75.42%	97.50%	92.11%	93.31%	7.89%	6.69%	24.58%	2.50%	91.92%	93.36%	0.959
Training Period: 1990-2014 @70%; Test Period: 2018; 4-year Lag)											
FRs	69.69%	75.00%	81.43%	79.54%	18.57%	20.46%	30.31%	25.00%	81.29%	79.51%	0.7327
CGIs	44.38%	25.00%	63.95%	81.63%	36.05%	18.37%	55.62%	75.00%	63.73%	81.30%	0.4533
DC	46.41%	25.00%	97.99%	99.70%	2.01%	0.30%	53.59%	75.00%	97.41%	99.26%	0.6235
SBC	38.08%	0.00%	98.94%	100.00%	1.06%	0.00%	61.92%	100.00%	98.25%	99.41%	0.5000
PBC	35.32%	0.00%	99.00%	100.00%	1.00%	0.00%	64.68%	100.00%	98.28%	99.41%	0.5000
MWDC	59.53%	25.00%	98.43%	98.81%	1.57%	1.19%	40.47%	75.00%	97.99%	98.37%	0.6190
SWDC	55.32%	50.00%	98.51%	98.51%	1.49%	1.49%	44.68%	50.00%	98.02%	98.22%	0.7425
FRs+CGIs	70.11%	54.17%	80.76%	85.15%	19.24%	14.85%	29.89%	45.83%	80.64%	84.97%	0.7079
FRs+DC	72.10%	58.33%	90.99%	93.21%	9.01%	6.79%	27.90%	41.67%	90.78%	93.01%	0.8087
FRs+SBC	71.02%	58.33%	90.47%	91.43%	9.53%	8.57%	28.98%	41.67%	90.25%	91.24%	0.7635
FRs+PBC	71.67%	71.67%	90.29%	92.02%	9.71%	7.98%	28.33%	28.33%	90.08%	91.90%	0.7408
FRs+MWDC	75.57%	28.33%	93.05%	93.88%	6.95%	6.12%	24.43%	71.67%	92.85%	93.49%	0.7278
FRs+SWDC	73.79%	51.67%	92.33%	92.97%	7.67%	7.03%	26.21%	48.33%	92.12%	92.73%	0.7832
FRs+CGIs+DC	74.17%	50.00%	90.30%	93.72%	9.70%	6.28%	25.83%	50.00%	90.11%	93.46%	0.7992
FRs+CGIs+SBC	72.58%	25.00%	89.31%	92.35%	10.69%	7.65%	27.42%	75.00%	89.12%	91.95%	0.7230
FRs+CGIs+PBC	72.52%	35.83%	89.46%	92.62%	10.54%	7.38%	27.48%	64.17%	89.27%	92.28%	0.7117
FRs+CGIs+MWDC	76.95%	25.00%	92.82%	94.69%	7.18%	5.31%	23.05%	75.00%	92.64%	94.28%	0.6735
FRs+CGIs+SWDC	75.42%	50.00%	92.11%	93.64%	7.89%	6.36%	24.58%	50.00%	91.92%	93.38%	0.8057
Training Period: 1990-2014 @70%; Test Period: 2019; 5-year Lag)											
FRs	69.69%	N/A	81.43%	77.13%	18.57%	22.87%	30.31%	N/A	81.29%	77.13%	N/A
CGIs	44.38%	N/A	63.95%	82.27%	36.05%	17.73%	55.62%	N/A	63.73%	82.27%	N/A
DC*	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
SBC*	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
PBC*	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
MWDC*	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
SWDC*	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
FRs+CGIs	70.11%	N/A	80.76%	83.61%	19.24%	16.39%	29.89%	N/A	80.64%	83.61%	N/A
FRs+DC	72.10%	N/A	90.99%	93.52%	9.01%	6.48%	27.90%	N/A	90.78%	93.52%	N/A
FRs+SBC	71.02%	N/A	90.47%	91.04%	9.53%	8.96%	28.98%	N/A	90.25%	91.04%	N/A
FRs+PBC	71.67%	N/A	90.29%	91.28%	9.71%	8.72%	28.33%	N/A	90.08%	91.28%	N/A
FRs+MWDC	75.57%	N/A	93.05%	94.79%	6.95%	5.21%	24.43%	N/A	92.85%	94.79%	N/A
FRs+SWDC	73.79%	N/A	92.33%	94.20%	7.67%	5.80%	26.21%	N/A	92.12%	94.20%	N/A
FRs+CGIs+DC	74.17%	N/A	90.30%	93.86%	9.70%	6.14%	25.83%	N/A	90.11%	93.86%	N/A
FRs+CGIs+SBC	72.58%	N/A	89.31%	92.03%	10.69%	7.97%	27.42%	N/A	89.12%	92.03%	N/A
FRs+CGIs+PBC	72.52%	N/A	89.46%	92.32%	10.54%	7.68%	27.48%	N/A	89.27%	92.32%	N/A
FRs+CGIs+MWDC	76.95%	N/A	92.82%	95.08%	7.18%	4.92%	23.05%	N/A	92.64%	95.08%	N/A
FRs+CGIs+SWDC	75.42%	N/A	92.11%	94.68%	7.89%	5.32%	24.58%	N/A	91.92%	94.68%	N/A

Note: *FRs* – Financial ratios; *CGIs* – Corporate governance indicators; *DC* – Company-Company direct connections-based driver; *SBC* – Company-Director-Company bipartite connections-based driver with simple direct connections; *PBC* – Company-Director-Company bipartite connections-based driver with consideration of directors' power; *MWDC* – Stacked time-weighted Company-Company direct connections-based driver with the maximum weights; *SWDC* – Stacked time-weighted Company-Company direct connections-based driver with the sum of weights.

* No performance measures were computed for these entries with N/A due to the lack of bankrupt cases in 2019.

Chapter 4

A Two-Step Approach to Bankruptcy Prediction using Companies' Financial Profiles

4.1 Introduction

Corporate bankruptcy prediction has become a topic that has been extensively investigated by academic researchers, business practitioners, investors, and regulators for more than a half century (Beaver, 1966). Several corporate bankruptcy prediction methodologies have been proposed and implemented to help stakeholders in financial decision making and loss prevention (Fejér-Király, 2015). The last decade has witnessed drastic changes in attitude toward corporate bankruptcy risks, which significantly influences financial institutions on risk assessment and credit management, due to the aftermath of the *global financial crisis* of 2008. In January 2023, *Basel III*, which is a comprehensive set of reformed international standards developed by the *Basel Committee on Banking Supervision*, has come into effect to strengthen the regulation, supervision and risk management of the banking sector (Basel Committee, 2010). The new measures require the banking sectors to improve risk management and governance, as well as increase the ability to absorb shocks arising from financial and economic stress, regardless of the source (European Banking Authority, 2018). Under this scenario, banks and other financial institutions must review and strengthen their risk management procedures in credit granting and develop more efficient and powerful methodologies to detect potential risks of corporate bankruptcy.

Since being first introduced by Beaver (1966) and Altman (1968), modern bankruptcy prediction methodologies have been constantly investigated and improved by researchers. Apart from traditional statistical and probabilistic models (e.g., *multivariate discriminant* *analysis*, *logistic regression*) which generally dominated corporate bankruptcy prediction research in the 1970s and 1980s, new families of artificial intelligence models, such as *decision trees* (e.g., Frydman et al., 1985; Cielen et al., 2004), *support vector machines* (e.g., Shin et al., 2005; Nanni & Lumini, 2009), *artificial neural networks* (e.g., Odom & Sharda, 1990; Tsukuda & Baba, 1994) and *ensemble learning* (e.g., West et al., 2005; Zięba et al., 2016), are gradually introduced into this field of research from other research communities since 1990s, thanks to the rise of information technology and the ever-stop evolution of computing power – see §4.2 for a sample of prediction models and methodologies. As highlighted in **Chapter 2**, a significant trend in the literature is the emphasis on advanced machine learning and artificial intelligence models, with a notable gap in addressing sample heterogeneity within datasets. This chapter seeks to address this gap by introducing a novel profiling-based approach to enhance bankruptcy prediction models' ability to generalise across diverse companies.

Although a large number of investigations have focused on proposing new prediction models or improving the existence ones to achieve better prediction performance, only a few papers have looked into the 'financial profiles' of companies and investigated their contribution to improving classifiers' performance. Beaver (1966) first introduced the concept of *profile analysis* to corporate bankruptcy prediction, where a profile is defined as an outline or a concise sketch of an entity. He performed a profile analysis by comparing the mean values of the financial ratios of bankrupt and non-bankrupt companies and provided evidence that the financial profiles of companies are a good source of information. Nowadays, research progress in profile analysis allows researchers to devise 'group profiles' for companies that share certain similarities, which not only represent the characteristics of individual companies, but also reflect the 'hidden knowledge' that consists of correlated patterns and can further enhance the performance of bankruptcy prediction – see §4.2 for detailed discussion of the definition of profile analysis in bankruptcy prediction. **Chapter 3** builds upon this foundation by exploring company relationships through *company relational information-based drivers* (CRIs), revealing the potential of interconnectedness in enhancing prediction performance. However, whilst **Chapter 3** focuses on the external relational aspects of companies, this chapter takes a complementary perspective by investigating into the internal financial structures and behaviours through financial profiling, thus bridging internal and external dimensions of corporate health.

Therefore, in this chapter, we intend to excavate patterns that can influence the survival of companies and develop relevant financial profiles that can help improve the performance of bankruptcy prediction models. In this research, we make several assumptions that are pivotal to the methodology and framework of the study, informed by research principles borrowed from other domains (e.g., criminology research), including the *consistency assumption*, *homology assumption*, and *distinct categorisation assumption*. To be more specific, two fundamental assumptions are made based on Warikoo (2014) and Kirwan & Power (2012) – The consistency assumption, which assumes that corporations exhibit consistent financial behaviours and patterns that lead up to bankruptcy events; the homology assumption, which hypothesises that corporations with similar financial behaviours or patterns will have resemblances in their financial backgrounds and structures. Furthermore, drawing insights from Huysmans et al. (2006), another vital assumption for this study is that, when engaged in the classification task of determining bankruptcy risk, observations from one financial category are anticipated to be distinctly different from another category.

In accordance with the assumptions mentioned above, in this study, we propose a two-step approach to the prediction of bankruptcy using the financial profiles of companies. In the first step, we use profiling techniques to identify distinct financial profiles for companies. These profiles show specific financial patterns that are common among certain subsets of companies and provide general descriptions of companies' financial positions and activities, which resonate with the assumption of consistency and the assumption of homology. In the second step, we combine the financial profiles identified in the first step with other information, such as financial information and corporate governance information; then, supervised machine learning models are adopted to make bankruptcy predictions based on the characteristics of the financial profiles of companies. Therefore, the main contribution of this study to the current literature is a novel two-step approach for bankruptcy prediction based on financial profiling. By integrating financial profiles into prediction models, this chapter complements the relational approach proposed in **Chapter 3**, illustrating how internal financial structures and external relational dynamics can together create more comprehensive and robust bankruptcy prediction frameworks. This approach integrates traditional financial metrics with financial profiles information to improve the predictions of machine learning models to address the issue of heterogeneity among companies and enhance predictions. In sum, we provide a fresh perspective and further push the boundary of profile-based bankruptcy prediction. Furthermore, by revisiting and improving the idea of 'profile-based' analysis initially introduced by Beaver (1966), we offer a better understanding of how shared financial behaviours between companies can be leveraged for better bankruptcy predictions.

The remainder of the chapter unfolds as follows. Section 4.2 provides an overview of the landscape of research on conventional bankruptcy prediction, as well as the relevant profiling analyses. Section 4.3 presents the research questions along with their justifications. Section 4.4 presents the methodological proposal along with implementation decisions and their justifications. Section 4.5 presents the details of the design of the experiments and the types and sources of data for the empirical evaluation of our proposal. Section 4.6 reports the empirical findings and provides a detailed analysis of the performance of our proposal and its implementation decisions. Finally, section 4.7 presents the concluding remarks and possible directions for future research.

4.2 Literature review

'Profiling' (or 'profiling analysis') is a highly evocative term with multiple meanings. Most of the time, profiling is likely to stand for the technique used to partition individuals in order to identify, for example, criminals (Ferraris et al., 2013). However, the development of other fields of academic research empowers the term 'profiling' with much more meaning. Hildebrandt & Backhouse (2005) claimed that profiling can be described as 'the process of constructing or applying a profile of an individual or a group', where a 'profile' can be considered a 'knowledge construct', which represents a subject, whether it is a person, a thing, or an organisation. They also pointed out that the profiling process involves *techniques* (i.e., profiling methods) and *technologies* (i.e., the combination of tangible instruments and techniques; hardware and software), which are two important pillars to ensure a successful profiling practice. There is always a thin line between 'profile' and 'the description of the original', since, in some definitions, a profile can be recognised as an outline or a concise description of the data (Beaver, 1966). However, they are not the same thing. Fuster et al. (2010) explained that profiles do not describe reality; instead, they should be detected by the aggregation, mining, and cleansing of data. They are based on correlations that cannot be equated with causes or reasons without further inquiry and are probabilistic knowledge. Generally, it is fair to say that profiling can be recognised as a systematic process of discovering and constructing new knowledge from existing data. As mentioned by Ferraris et al. (2013), the definitions adopted by researchers from various backgrounds somehow underline the following standard features of profiling analysis: (1) the central role of data and quantitative techniques, (2) the categorisation as one of the main characteristics, (3) the deduction of new information from something already known (e.g., a behaviour, a specific characteristic), and (4) the use of this information for some purposes, i.e. the importance of domains of application.

Profiling analysis can be classified into three main categories in terms of its form, i.e., organic profiling, which is carried out by non-human organisms (e.g., plants) that allows them to know the surrounding environment and constantly adapt for their own survival; human profiling, which is carried out by humans that allows for the intentional action and conscious reflection; and machine profiling, which is carried out by machines to enable automatic processing and discovery of information (Ferraris et al., 2013). In recent years, there is growing trend of using sophisticated machine learning-based methodologies to perform profiling analysis (Van Otterlo, 2013) which helps to shift the profiling methodologies from non-automatic profiling to automatic and autonomic profiling, where human intervention is minimised and the decision making process is entirely driven by the machine (Ferraris et al., 2013).

Several key domains saw the general adoption of profiling analysis in normal practice, including criminal investigation (e.g., Devery, 2010; Pramanik et al., 2017), financial security (e.g., Edge & Sampaio, 2009; Politou et al., 2019), healthcare (e.g., Lin et al., 2017; Li et al., 2022), human resources (e.g., Vatousios & Happonen, 2022; Rafae & Erritali, 2023), and marketing (e.g., Somervuori, 2014; Wedel & Kannan, 2016). Some studies from other new or cross-subject domains also implemented profiling analysis to tackle specific problems (e.g., Shiers et al., 2006; Godsell et al., 2011; Darcan & Badur, 2012).

Corporate bankruptcy prediction is a hot topic that draws wide attention from the business analytics community, in general, and predictive analytics community, in particular. It is generally accepted that the study by Fitzpatrick (1932) is the first systematic corporate bankruptcy prediction research based on financial ratio analysis, however, Beaver (1966) took the concept of ratio analysis to a new level by proposing a systematic univariate ratio analysis, which is effective in distinguishing between bankrupt and healthy companies one year before the bankruptcy event. Later, Altman (1968) further improved the idea of univariate financial ratio analysis by proposing the famous Altman *Z-score* model, which uses *multivariate discriminant analysis* (MDA) to make effective bankruptcy predictions and is still widely received in the financial world today. The works of Beaver and Altman laid a strong foundation for modern corporate bankruptcy prediction research, with later studies gradually expanding the scope of research streams and successfully branched out various families of prediction models: statistical models, such as *logistic regression* (LR, or *logit* for short) (e.g., Ohlson, 1980; Zavgren, 1985); probabilistic models, such as *probit model* (Zmijewski, 1984); and machine learning models, such as *artificial neural networks* (ANNs) (Odom & Sharda, 1990), *case-based reasoning* (CBR) (Bryant, 1997), *support vector machines* (SVMs) (Shin et al., 2005); ensemble models, such as *random forest* (RF) (e.g., Barboza et al., 2017), *boosting* (e.g., Wang et al., 2014), *bootstrap aggregating* (*bagging*) (e.g., West et al., 2005), *random subspace* (RS) (e.g., Nanni & Lumini, 2009) and *stacked generalisation* (*stacking*) (e.g., Tsai & Hsu, 2013).

Despite the extensive amount of research surrounding the topic of corporate bankruptcy prediction, the integration of profiling analysis into this domain remains a relatively new and less investigated field of research. Generally, there is no consensus on the definition of financial profile due to the fact that its definition is largely influenced by the purpose and methodology of research and the application area. Beaver (1966) is the first study in corporate bankruptcy prediction to mention the term 'profile analysis'. In his study, a company profile is defined as an outline or a concise sketch of the distribution of company accounting ratios, that is, descriptive statistics such as mean value. He also claimed that the comparison of mean values of the accounting ratios between different companies can be called profile analysis. The information provided by individual financial ratios can be recognised as 'individual profile' and can certainly describe a company's financial performance without introducing unnecessary subjectiveness or distortions; however, it does not extract new knowledge from existing ones and also does not outline the differences in general characteristics between failed and healthy companies.

Nowadays, the financial profile in bankruptcy prediction research can generally be recognised as a 'group profile', which focusses more on the representation of certain types of companies that share one or more common financial characteristics. For example, Du Jardin (2016) claimed that a profile reflects a particular financial structure shared by a subset of firms and can be expressed as a set of positions on different scales that characterise their activity, financial structure, profitability, turnover, liquidity, and solvency. Generally, there are two approaches to creating financial profiles for corporate bankruptcy prediction, i.e., the *supervised approach*, which uses a supervised model to detect the shared characteristics of observed companies based on the results of classification, and the *unsupervised approach*, which, on the other hand, uses unsupervised methodologies to discover the common patterns of observed companies based on the results of clustering. For example, Ozgulbas & Koyuncugil (2006) conducted a study to describe the financial profiles of companies through data mining and to obtain vital characteristics of the performance level of the companies that have the best financial profile in the Turkish health sector. The authors took a supervised approach by building a non-binary *Chi-square automatic interaction detector* (CHAID) tree to partition hospitals into 12 different profiles (i.e., the 12 leaves of the CHAID tree) in terms of the level of financial performance, based on cascade splits and merges of sampled hospitals according to different split values determined by Chi-square tests. They claimed that those profiles illustrate what financial indicators hospitals should focus on for good financial performance. The same method has also been used by Koyuncugil & Ozgulbas (2008), where the data of the research are expanded to 6,185 companies in Turkey and a total number of 35 financial profiles are built based on operational risk factors. The proposed supervised CHAID tree model provides highly visualised outputs that are easy for humans to interpret; however, the model would only produce superficial profiling based on the differences in the value of financial ratios rather than identifying the potential strong or weak similarities or relations among those firms within the same profile. Du Jardin (2016) argued that traditional supervised single models fail to capture the patterns of decline of a company, where those patterns can be recognised as prototype situations that a company may experience before going bankrupt. Although ensemble techniques, such as *bagging* and *boosting*, intend to encompass the diversity of situations a company may face shortly before its bankruptcy, they still fail to take into account any explicit characteristics of bankruptcy itself. They proposed a two-step profile-based model, where the first step utilises the *Kohonen map* (i.e., *self-organising map*, or SOM for short) to cluster observed companies into eight financial profiles depending on six performance dimensions (i.e., activity, financial structure, profitability, turnover, liquidity, and solvency), and the second step builds as many supervised classification models as there are profiles to make the prediction. They claimed that a company's financial profile can be used as a priori knowledge to estimate the decision boundary between failed and healthy companies, and the results suggest that profile-based models perform much more accurately and more stable than ensemble models. Du Jardin (2017) later expanded the idea of profiling analysis into financial evolution prototype profiling with a similar methodology, where the *Kohonen map* is again used to partition sampled companies into as many categories of companies as there are different prototypes of variation of their financial situation over time. The profiles created describe the financial stability of the companies, and, later, by assessing the evolution of the financial health of the companies with those profiles, the model is able to make an effective prediction of bankruptcy. When using profiling analysis in bankruptcy prediction, it is worth noticing that the financial profiles generated by the model may not provide good descriptions for a subset of sampled companies, due to the fact that the choice of drivers can largely influence the structure of the profiles. Du Jardin claimed that it is often the case that a part of a company's financial profile measured using some variables may resemble that the company is in good condition; however, another part of its profile suggests the opposite when measured with some other variables (Du Jardin, 2021b). To solve this problem, Du Jardin (2021b) proposed an ensemble methodology that incorporates the '*biclustering*' technique, which simultaneously clusters both sampled companies and a set of pre-selected drivers. Unlike the clusters generated by traditional clustering methods, instead of depending on the same set of drivers, a bicluster depends solely on a subset of drivers chosen from the initial set, and therefore each bicluster is the result of double partitioning of the dataspace (Du Jardin, 2021a). The biclustering technique is based on the *Croeuc algorithm* (Jollois & Nadif, 2004), where the clustering is performed both on rows (i.e., a set of companies) and columns (i.e., a set of drivers) of the data matrix simultaneously by iteratively optimising the partitioning with the *k-means* algorithm. With each bicluster, an ensemble of SOMs with neurons (i.e., nodes) ranging from 2 to 1,000 is calculated; then those ensembles are pruned, and the one made up of the smallest number of networks that achieves the highest performance is selected for prediction. The results suggest that biclustering-based models systematically achieve better results than single models (i.e., *discriminant analysis*, LR, DT, SVM, and *extreme learning machine)* and significantly outperform the models without clustering and simple clustering-based models.

Some other studies on corporate bankruptcy prediction also adopted similar ideas by using unsupervised machine learning models; however, they generally did not mention the concept of financial profile, which itself can provide very much insight in the forecast of future bankruptcy events. For example, Huysmans et al. (2006) proposed an integrated methodology by combining a *self-organising map* with supervised techniques to make predictions. They examined two different integration approaches, where the first approach first trains SOM and then uses the supervised classifier (i.e., MLP) to improve the decision made by the individual neurones of the SOM, and the second approach, on the other hand, creates a 'stacking model', which uses the predictions of the supervised classifier (i.e., MLP) as input to the SOM. The results suggest that using SOM alone only receives lower prediction accuracy, and the integration of SOM with supervised classifiers improves the performance of the SOM model. The results also show that the stacking method yields low prediction performance, as the degree of complexity of this model is high; however, the advantage of the explanatory power of SOM is not appreciated. Although the concept of financial profile is not explicitly mentioned, the author also pointed out that SOM demonstrates strong capability in visualisation, which enables a clear understanding of the decisions made by the machine. Tsai (2014) also developed a hybrid financial distress prediction model based on the combination of clustering techniques (i.e., SOM and *k-means*) and supervised classifiers (i.e., LR, MLP and DT). In this study, clustering is used as a preprocessing method to identify pattern classes for subsequent supervised classifiers. In addition, as clustering techniques can filter out nonrepresentative data, they can also be used to reduce noise and improve performance for the later stage of classification. The results indicate that using SOM to build the hybrid model can provide better prediction performance than using *k-means*, and the best model is based on the combination of SOM and MLP ensembles with the weighted voting approach.

This research aims to push the frontiers of bankruptcy predication research by proposing a novel two-step approach that not only creates financial profiles through advanced unsupervised machine learning techniques but also integrates these profiles into supervised machine learning models for bankruptcy prediction. Unlike previous studies that often treat profiling as a standalone process, this research emphasises the interaction between profiling and prediction, allowing for more accurate assessments of bankruptcy risk. This study contributes to the literature by providing a more holistic and nuanced understanding of how financial profiles can be leveraged to improve bankruptcy predictions, with application in the context of publicly listed companies in the UK.

4.3 Research questions

Given the evolving bankruptcy prediction research and the promising future of implementing financial profiling techniques, this chapter intends to investigate seven research questions related to the application of financial profiles in bankruptcy prediction.

As introduced by Beaver (1966) and further developed in subsequent studies (e.g., Ozgulbas & Koyuncugil, 2006; Du Jardin, 2016; Du Jardin, 2021b), profile analysis provides better understanding of companies' financial behaviours. By examining this methodology's efficacy, we aim to validate its role in the modern landscape of bankruptcy prediction, specifically for listed companies in the UK. Our first research question is about whether profiling analysis adds value to the prediction exercise and can be stated as follows.

(1) Is profiling analysis beneficial in bankruptcy prediction of publicly listed companies in the UK?

Different profiling techniques come with their own strengths and weaknesses. By contrasting their performances, we intend to identify the most robust technique(s) that can consistently and effectively partition or cluster companies based on shared financial characteristics or behaviours. In light of this, we also put forward our next research question:

(2) Which profiling technique yields superior bankruptcy prediction performance?

The identification and analysis of distinct financial profiles, where each is associated with varying levels of financial risk, highlight the importance of understanding the diverse financial behaviours and sample heterogeneity present within the dataset. This leads to our next research question:

(3) How do distinct financial profiles, as identified through profiling techniques, correlate with varying levels of bankruptcy risk? Moreover, how can these profiles be leveraged to enhance the accuracy and interpretability of prediction models?

The way in which financial profiles are integrated into supervised prediction models can significantly influence prediction performance. Whether using a uniform supervised model across all financial profiles or tailoring specific models to specific profiles, the choice of the modelling strategy is crucial in shaping the future direction of profile-based bankruptcy prediction. So, the next research question is:

(4) How does the different potential use of financial profiles affect the performance of supervised prediction models?

The trade-offs between profile-specific prediction models and meta-learner prediction models (i.e., global models) must be scrutinised to determine the best approach for balancing model complexity and predictive power. This leads to our next research question:

(5) To what extent do meta-learner prediction models (global models) improve prediction performance compared to profile-specific prediction models, and what are the trade-offs involved?

The role of feature engineering and feature selection within the profile-based bankruptcy prediction methodology also needs to be investigated. Understanding how these processes influence the model's performance can provide deeper insights into the optimisation of financial profiling methodologies. Therefore, our next research question is:

(6) What is the role of feature engineering and feature selection in enhancing the performance of profile-based prediction models for bankruptcy prediction?

Lastly, As suggested by Zhao et al. (2024), ensemble learning methodologies have gained wide popularity in bankruptcy and financial distress prediction. Integrating profiles into ensemble methods presents an additional layer of analysis, potentially leading to improved prediction accuracy and robustness. To explore this further, we put forward our final research question:

(7) How do profiles-based ensemble methods compare to profiles-based traditional models in terms of prediction accuracy and robustness?

4.4 Methodological proposal

In this section, we present and discuss our methodological proposal for integrating profile analysis in bankruptcy prediction, where our primary focus is on three crucial problems. The first problem is about the determination of financial profiles, where various clustering algorithms such as *k-means* and *hierarchical clustering* are used to cluster companies effectively. The second problem is how to accurately describe each profile. This will involve identifying a set of potential drivers, establishing thresholds for the statistics of these drivers, and discovering patterns. The last problem involves how to develop models for bankruptcy prediction using financial profiles, which involves effectively exploiting financial profiles in prediction models to enhance their performance.

Financial profiling involves creating a comprehensive depiction of a company's financial health and behaviour patterns, and this concept extends beyond the traditional financial ratios and metrics, incorporating a more systematic representation of a company's financial status. The objective of financial profiling is to capture the essence of a company's unique financial patterns, trends, and indicators that could signal its financial stability or provide an early warning of its potential bankruptcy.

In the process of developing financial profiles for companies, profiling methods are applied to group companies with similar financial behaviours. This step aims to discern inherent groupings or patterns within the data, thereby signifying distinct financial profiles. Next, a thorough analysis and characterisation of each group is performed to identify its defining features. This analysis interprets the groups in terms of financial health, risk factors, and potential future trajectories.

A Two-Step Approach for Profile-based Bankruptcy Prediction

Hereafter, we describe the proposed two-step approach for predicting corporate bankruptcy using companies' financial profiles to address the issue of sample heterogeneity. In the first step, company accounting information is used to generate financial profiles using unsupervised cluster analyses. In the second step, supervised artificial intelligence models are used to make bankruptcy predictions based on the properties of the financial profiles of companies either directly or indirectly. The proposed framework and its implementation decisions and their justifications are summarised hereafter.

Step 1: Generating financial profiles

The first step in our proposed methodology is to generate financial profiles using company accounting information through unsupervised cluster analyses. This approach is designed to address the issue of sample heterogeneity by grouping companies with similar financial behaviours, thereby creating distinct profiles that can be leveraged in the subsequent bankruptcy prediction models used in step 2.

In this study, various methodologies were tested to determine the most effective approach for profiling companies in the context of bankruptcy and financial distress prediction. The primary objective is to stratify the data into distinct groups, where each group encapsulates a profile that signifies certain characteristics pertinent to bankruptcy risk. There are two main approaches to devise company profiles, which are the *supervised profiling* approach and the *unsupervised profiling* approach. Supervised profiling employs supervised models to classify companies into distinct profiles based on pre-labelled data, with a focus on identifying shared characteristics pertinent to bankruptcy risk. Techniques such as the *Chi-square Automatic Interaction Detector* (CHAID), as exemplified in studies by Ozgulbas & Koyuncugil (2006) and Koyuncugil & Ozgulbas (2008), demonstrate that the supervised approach has the advantage of producing highly interpretable, visually articulate outputs, facilitating better bankruptcy and financial distress predictions based on the company profiles. However, this method has limitations, such as having a propensity for overfitting and a lack of flexibility in adapting to unusual patterns. Unsupervised profiling, on the other hand, uses unsupervised machine learning algorithms like *k-means clustering* and *self-organising maps* (SOMs) to determine inherent groupings within data without the constraints of pre-set classifications (i.e., labels). This approach is good at discovering new insights and offers a broader understanding of a company’s financial health across various dimensions. Despite these strengths, the complexity in interpreting the resulting profiles and the challenge of integrating those profiles with prediction models remains. Each approach has its own strengths and weaknesses. The choice between these approaches depends on the research goals, the availability of data, and the complexity of the profiling analysis.

Given the nature of bankruptcy and financial distress prediction, which is inherently nonlinear and complex, unsupervised clustering methodologies were chosen for this study due to their advantages in handling the patterns often exhibited in financial data. The unsupervised approach provides an adaptive framework that is critical in financial data analysis, where the underlying structures and relationships within the data are not always apparent. The ability of unsupervised learning techniques to detect subtle, yet significant patterns, and associations among financial variables offers a robust tool to reveal insights that might not be captured through supervised learning techniques. Moreover, unsupervised profiling offers the flexibility required to adapt to new data as it becomes available. In a rapidly changing economic environment, the ability to iteratively refine and evolve company profiles without the need for relabelling or restructuring a model is crucial. This flexibility is key to maintaining the relevance and accuracy of the predictive model over time.

Additionally, unsupervised profiling can be beneficial for exploratory data analysis, often serving as a precursor to more detailed investigations. By not relying on predetermined labels, this approach provides an unbiased overview of the data landscape, essential when dealing with complex financial datasets that could be influenced by groups of interdependent factors. This can lead to the discovery of unconventional indicators of financial distress that might be overlooked when confined to labelled data. By selecting the unsupervised approach, this study is positioned to utilise the full spectrum of information present within the data, offering a comprehensive and nuanced understanding of the factors that contribute to bankruptcy and financial distress.

To better profile company data for bankruptcy and financial distress prediction, this research carefully selected six clustering techniques: *k-means*, *Density-based spatial clustering of applications with noise* (DBSCAN), *hierarchical clustering*, *Gaussian mixture models* (GMM), *spectral clustering*, and *Balanced iterative reducing and clustering using hierarchies* (BIRCH). These methods were chosen based on their distinct strengths in handling various data characteristics inherent in financial datasets. *K-means* was selected for its simplicity and efficiency in handling large datasets, particularly when the number of clusters is known and the data tends to form spherical clusters (Macqueen, 1967). DBSCAN was included for its ability to identify clusters of arbitrary shapes and sizes without requiring a pre-specified number of clusters, making it particularly useful in detecting outliers or noise in financial data (Ester, Martin et al., 1996). Hierarchical clustering was chosen for its capacity to provide a nested clustering structure, which is valuable for understanding the hierarchical relationships between companies (Ward, 1963). GMMs were considered due to their probabilistic approach, which allows for the modelling of overlapping clusters and provides uncertainty estimates for each data point's cluster assignment (He et al., 2010). Spectral clustering was selected for its ability to handle non-linearly separable datasets, leveraging the global structure of the data to identify complex patterns within the financial information (Ng et al., 2001). Finally, BIRCH was chosen for its scalability and efficiency in handling large datasets, particularly when memory resources are limited (Zhang et al., 1996). Together, these methods offer a comprehensive toolkit for capturing the diverse patterns present in the datasets, thereby enabling a robust and nuanced profiling of companies for the purpose of bankruptcy prediction. A summary of these methods is provided in **Table 4.1**.

Table 4.1 Summary of The Chosen Profiling / Clustering Methods

Name	Description	Pros	Cons
K-means	K-means clustering is an unsupervised learning algorithm used to group similar data points together based on their similarity or distance to each other. The algorithm works by first randomly selecting k initial cluster centres (centroids), where k is the number of desired clusters. Each data point is then assigned to the nearest cluster based on its distance to its centroid, usually calculated using the *Euclidean distance*. The mean of each cluster is then calculated, and the centroids are updated to the new mean. The algorithm repeats until the centroids no longer change significantly or a maximum number of iterations is reached.	1. Simple and easy to implement; 2. Fast and efficient for large datasets; 3. Performs well on circular or spherical clusters.	1. Requires the number of clusters to be specified beforehand; 2. Sensitive to the initial random assignment of cluster centroids; 3. Assumes clusters have a spherical shape and equal size.
DBSCAN	DBSCAN (*Density-Based Spatial Clustering of Applications with Noise*) is an unsupervised learning algorithm used to group data points that are close to each other based on their density. The algorithm works by first selecting a random point in the dataset and finding all the neighbouring points within a specified distance (i.e., the *epsilon neighbourhood*). A core point is defined as a data point that has at least a specified number of neighbouring points within its epsilon neighbourhood (i.e., the minimum number of points). Any non-core point that falls within the epsilon neighbourhood of a core point is considered part of the same cluster. Points that are not part of any cluster are labelled as *noise*.	1. Can identify clusters of arbitrary shape and size; 2. Does not require the number of clusters to be specified beforehand; 3. Can handle noise/outliers.	1. Sensitive to the choice of distance metric and parameters (e.g., eps, minimum samples); 2. Computationally expensive for large datasets; 3. May not work well for datasets with varying density.
Hierarchical clustering	Hierarchical clustering is a clustering algorithm that seeks to group together similar data points based on their distance or similarity to each other. The resulting groups are arranged in a hierarchical structure, often represented as a *dendrogram*. The algorithm starts with each data point in its own cluster, and then iteratively merges the two closest clusters together until all data points are in a single cluster. The distance between clusters can be measured using various metrics such as the *Euclidean distance* or the *Manhattan distance*. There are two main types of hierarchical clustering: agglomerative and divisive. Agglomerative clustering starts with each data point in its own cluster and then merges the closest clusters together, whereas divisive clustering starts with all data points in a single cluster and then recursively splits it into smaller clusters.	1. Produces a hierarchy of nested clusters that can be visualised as a dendrogram; 2. No need to specify the number of clusters beforehand; 3. Can handle various types of data (e.g., categorical, binary, numerical).	1. Computationally expensive, especially for large datasets; 2. Can produce a large number of clusters that may be difficult to interpret; 3. Sensitive to the choice of distance metric and linkage criteria.
Gaussian Mixture Models	Gaussian mixture models (GMM) clustering is a probabilistic clustering algorithm for unsupervised learning. It assumes that the data points are generated from a mixture of several Gaussian distributions with unknown parameters. The algorithm works by first randomly initialising the means and covariances of the Gaussian distributions. The *Expectation Maximisation* (EM) algorithm is then used to estimate the parameters of the Gaussian distributions, which represent the centres and spreads of the clusters. The EM algorithm alternates between the E-step, where the probabilities of each data point belonging to each cluster are calculated, and the M-step, where the parameters of the Gaussian distributions are updated based on the probabilities. GMM clustering is a flexible algorithm that can capture complex patterns in the data and can model overlapping clusters. It is also able to estimate the uncertainty associated with each data point's assignment to a particular cluster. However, it can be computationally expensive and sensitive to the initialisation of the parameters.	1. Can identify clusters with different shapes and sizes; 2. Can estimate the probability of a data point belonging to a cluster; 3. Allows for overlapping clusters.	1. Requires the number of clusters to be specified beforehand; 2. Sensitive to the choice of initial parameters and may converge to local optima; 3. Computationally expensive for large datasets.
Spectral Clustering	Spectral clustering is an unsupervised clustering algorithm that uses the spectrum of the data's similarity matrix to group together similar data points. It works by transforming the data into a lower-dimensional space using eigenvalues and eigenvectors. The algorithm starts by constructing a *similarity matrix* that measures the similarity between each pair of data points. The similarity matrix is then transformed into a graph, where each data point is a node, and the similarity between two points is the weight of the edge connecting them. The graph *Laplacian matrix* is then calculated, and its eigenvectors are used to perform a spectral clustering of the data.	1. Can identify clusters of arbitrary shape and size; 2. Can handle non-linearly separable datasets; 3. Can preserve the global structure of the data.	1. Requires the number of clusters to be specified beforehand; 2. Sensitive to the choice of distance metric and number of eigenvectors to use; 3. Computationally expensive for large datasets.
BIRCH (Balanced Iterative Reducing and Clustering using Hierarchies)	BIRCH (*Balanced Iterative Reducing and Clustering using Hierarchies*) is an unsupervised clustering algorithm used to group together data points into clusters based on their similarity or distance to each other. The algorithm is particularly effective at handling large datasets and reducing the memory and time consumption of clustering. The algorithm works by first constructing a tree-like data structure called the *Clustering Feature Tree* (CFT). The CFT is built using a two-phase clustering process, where in the first phase, the algorithm uses a clustering method called	1. Can efficiently process large datasets and reduce the memory requirements; 2. It is scalable to handle large datasets with high-dimensional features; 3. Can provide fast clustering results;	1. Sensitive to the initial selection of parameters; 2. May not perform well for datasets with irregularly shaped clusters or on sparse datasets; 3. May produce suboptimal clustering results compared to other clustering

	CF (i.e., Clustering Feature) to construct subclusters from the input data points. In the second phase, the algorithm constructs a hierarchical tree structure by recursively merging the subclusters until they form a single cluster. The CF method partitions the input data points into small sub-clusters using a clustering criterion that considers the proximity of the data points to a cluster centre and the size of the sub-cluster. The cluster centre and size are updated incrementally as new data points are added to the subcluster.	4, Can perform hierarchical clustering and extract hierarchical representation of the data.	algorithms such as k-means or hierarchical clustering; 4. Require careful tuning of the parameters.

Step 2: Exploiting financial profiles in bankruptcy prediction

We propose two approaches to exploit the financial profiles devised in the previous step. The first approach exploits financial profiles by devising profile-based drivers to feed bankruptcy prediction models with, whereas the second approach exploits the financial profiles or clusters by fitting bankruptcy prediction models to each of these clusters without using any additional drivers. The implementation details of these approaches are summarised hereafter.

First Approach to Profiles Exploitation: Profile-based Drivers

To exploit the profiles learnt in Step 1, three distinct drivers are devised and used to feed bankruptcy prediction models:

(1) **Profile Direct Utilisation**: This type of profile-based driver makes direct use of the learnt profiles as categorical information, which means that the profiles/clusters of the training dataset, say P_k^{train} $(k = 1, \dots, K^{train})$, are used to define a new driver, say x, as a categorical variable with as many categories as the number of clusters of the training dataset, K^{train}, where for each observation i, $x_i = k$ if $x_i \in P_k^{train}$. Since this type of profile-based driver is also required to predict the risk class of the test dataset observations, similarly, the test dataset is profiled into, say K^{test}, profiles or clusters P_k^{test} $(k = 1, \dots, K^{test})$. Then, for each observation i in the test dataset, $x_i = k$ if $x_i \in P_k^{test}$. x thus is defined as a driver. Please refer to **Fig. 4.1** for a graphical illustration of the direct utilisation of profiles as categorical information.

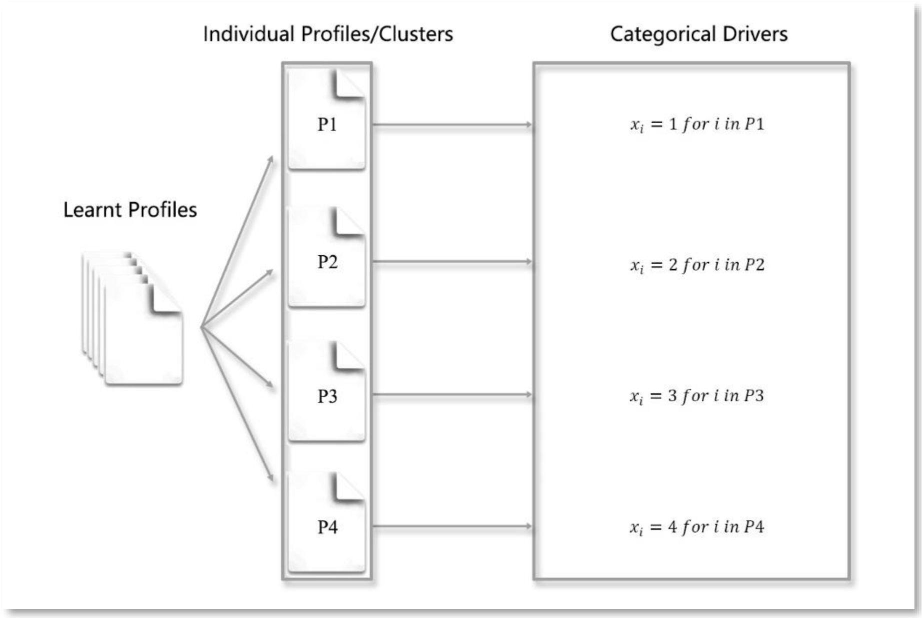

Fig. 4.1 Graphical illustration of the direct utilisation of profiles as categorical information (4 Profiles Example)

(2) **One-Hot Encoding**: A *one-hot encoding* is devised for each profile or cluster, e.g., if the total number of profiles is 5, the encoding (1,0,0,0,0) is for Profile 1, and (0,0,1,0,0) is for Profile 3. Encodings are used as additional drivers. In sum, for each observation i in the training dataset, $e_i = \left(e_k, k = 1, \dots, K^{train}; e_k = 1 \; if \; i \in P_k^{train} \; and \; 0 \; otherwise\right)$, and for each observation or company i – denoted hereafter as C_i , in test dataset, $e_i = (e_k, k = 1, \dots, K^{test}; e_k = 1 \; if \; i \in P_k^{test} \; and \; 0 \; otherwise)$; e thus defined is used as a driver. Please refer to **Fig. 4.2** for a graphical illustration of the one-hot encoding of profiles as categorical information.

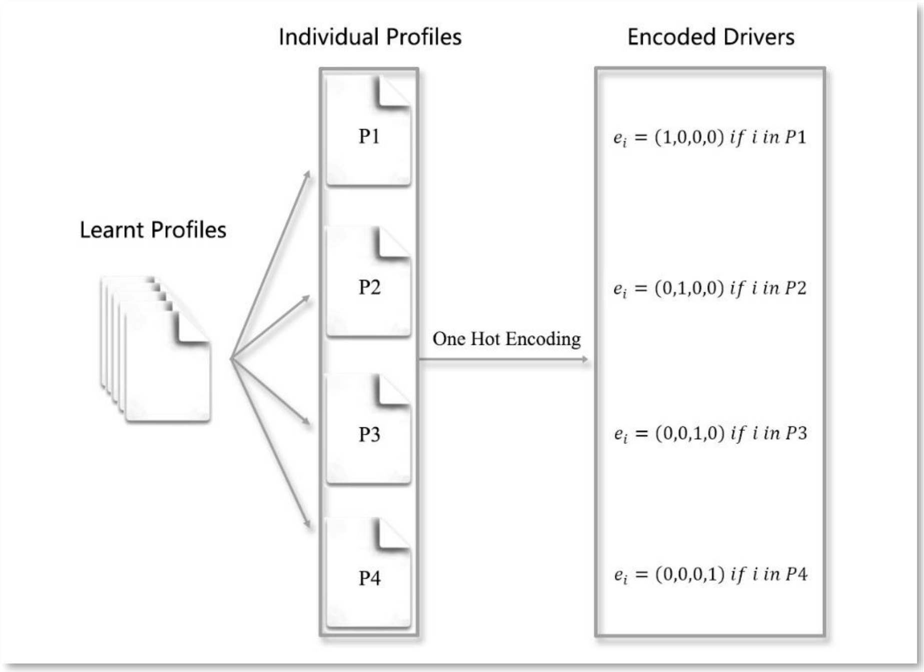

Fig. 4.2 Graphical illustration of the one-hot encoding of profiles as categorical information (4 profiles example)

(3) **Profile Ranking**: Profiles or clusters, say P_k ($k = 1, \dots, K$), are ranked (i.e., renumbered) in descending order of the percentage of bankrupt companies in each cluster so that clusters with higher rankings indicate a higher risk of bankruptcy (e.g., the profiles are ranked from 1 to 4 with profile 1 being the riskiest and profile 4 being the least risky). Such ranking is used to devise a new driver as follows: for each observation i in the training dataset, $r_i = k$ if observation i belongs to renumbered training cluster k ($k = 1, \dots, K^{train}$), and for each observation i in the test dataset, $r_i = k$ if observation i belongs to renumbered test cluster k ($k = 1, \dots, K^{test}$); r thus defined is used as a driver. Please refer to **Fig. 4.3** for a graphical illustration of the ranked profiles as categorical information.

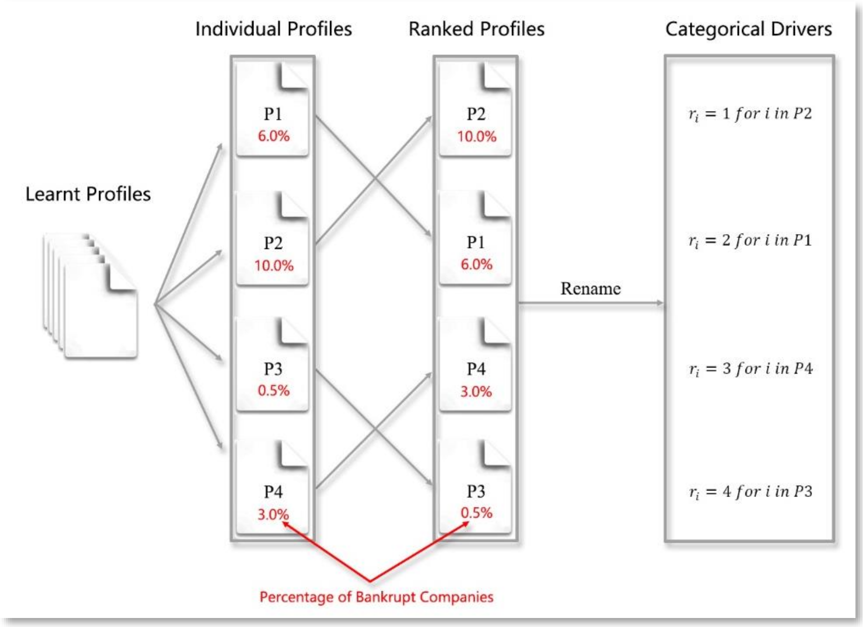

Fig. 4.3 Graphical illustration of the ranked profiles as categorical information (4 profiles example)

Each of these drivers provides a different method for processing and utilising the learnt financial profiles, offering a range of complexity and insight for subsequent analysis and decision-making.

Second Approach to Profiles Exploitation: Training of Bankruptcy Prediction Models on Profiles

In the stage of model fitting and bankruptcy prediction, two approaches are adopted to leverage the profiles obtained in Step 1:

Approach 1 – Building Classifiers for each Profile: This approach is more granular, where a different classifier is built for each individual profile. This means that samples belonging to a specific profile are used to train a dedicated profile-specific classifier that is tailored to the characteristics of that profile. Two separate experiments are conducted to explore the efficacy of this approach:

(1) In the first experiment, *logistic regression* (LR) is used as the exclusive classifier for each profile. LR is a statistical model that, in this context, is used to predict a binary outcome, such as bankrupt or healthy, based on the attributes of the samples within a profile. Please refer to **Fig. 4.4** for a graphical illustration of the process of building each profile-specific LR classifier.

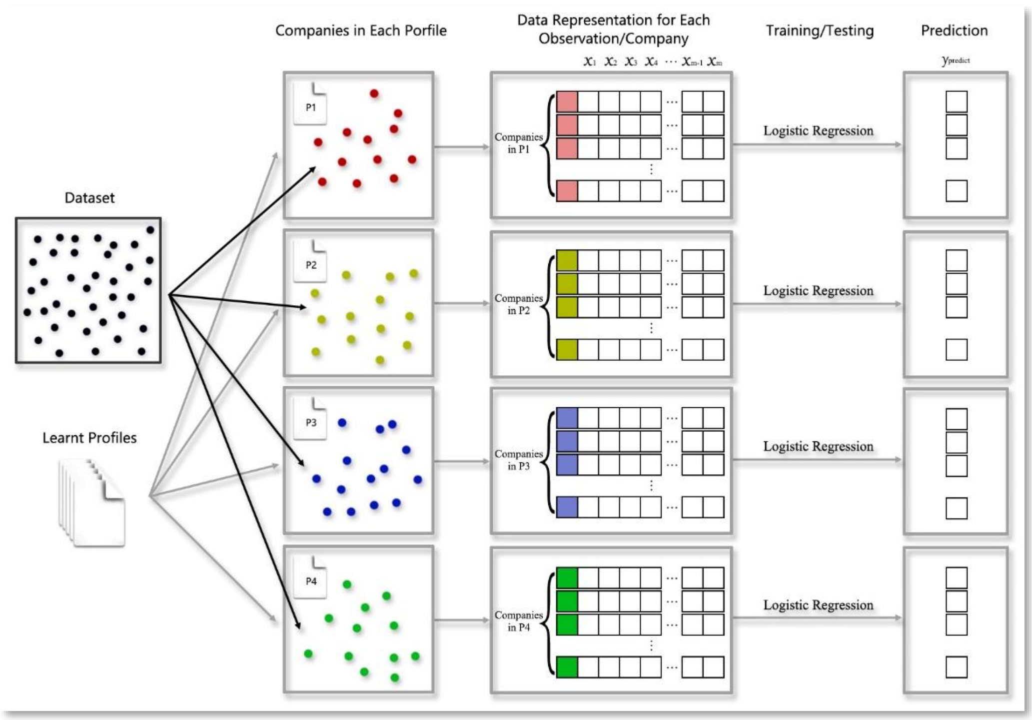

Fig. 4.4 Graphical illustration of the process of building each profile-specific LR classifier (4 profiles example)

(2) In the second experiment, different classes of classifiers are used for different profiles, chosen based on the size and characteristics of the samples within each profile. Such classifiers include LR, *support vector machines* (SVM), *multi-layer perceptrons* (MLP), and *AdaBoost*. Each classifier has its strengths and is selected to match the profile's characteristics and complexity. For example, SVM is used for profiles with clear margin separation, MLP is used for profiles that require deep learning for pattern recognition, and *AdaBoost* is used for profiles that may benefit from an ensemble approach to improve prediction performance. Please refer to **Fig. 4.5** for a graphical illustration of the process of building a different classifier for each profile.

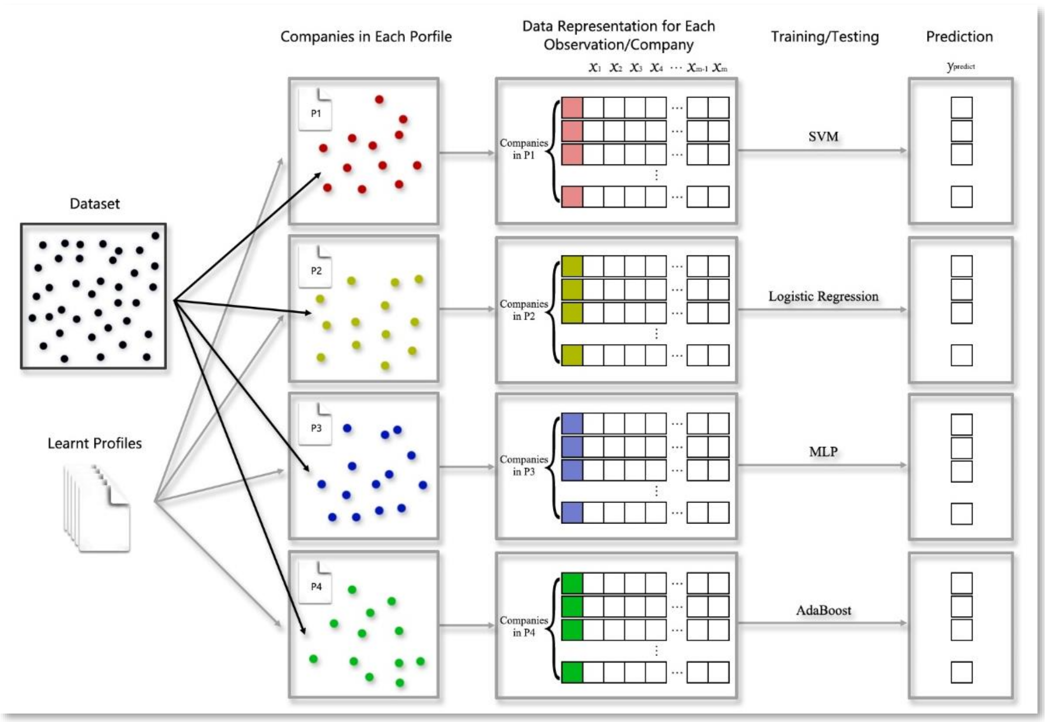

Fig. 4.5 Graphical illustration of the process of building a different classifier for each profile (4 profiles example)

Approach 2 – Developing Meta-learning Structure based on Profile-specific Models: This approach enhances bankruptcy prediction by utilising a '*meta learner*' that integrates predictions from supervised models trained on individual profiles. For each profile or cluster, a specific supervised model is trained to capture the intricate patterns and risk factors unique to that profile. The meta-learner then aggregates the predictions of these profile-specific models, thereby leveraging the diverse financial characteristics and potentially improving the overall predictive performance. Please refer to **Fig. 4.6** for a graphical illustration of using a Meta-learning structure based on profile-specific models for bankruptcy prediction.

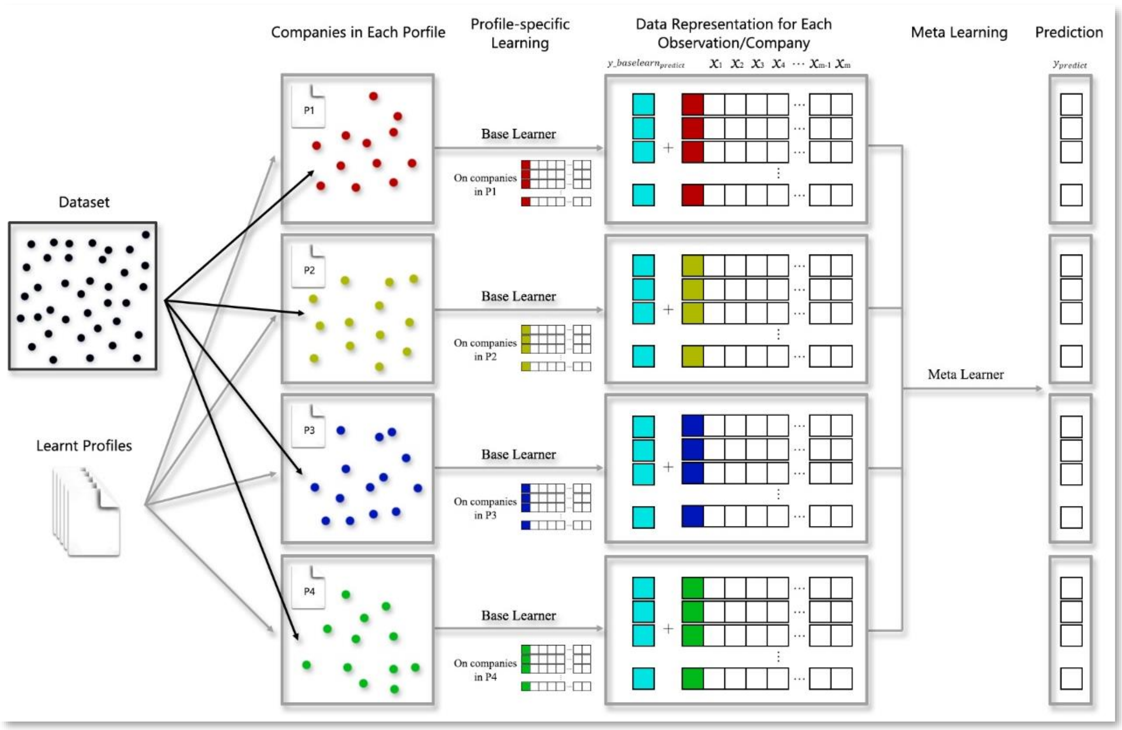

Fig. 4.6 Graphical illustration of using a Meta-learning structure based on profile-specific models for bankruptcy prediction (4 profiles example)

By implementing these two approaches, the learning phase is enhanced not only to incorporate the clustering information into the classification process, but also to optimise the selection of classifier based on the unique properties of each cluster, potentially improving the accuracy and robustness of the predictive model.

4.5 Design of experiments

This section outlines the experimental framework for the two-step bankruptcy prediction approach. We detail the data sources and their compilation as well as the experimental setup in §4.5.1, the criteria for selecting financial ratios and corporate governance indicators in §4.5.2, the development of prediction models combining financial profiles and machine learning techniques in §4.5.3, and the performance measures in §4.5.4.

4.5.1 Data & experimental setup

This study focusses on UK public companies listed on the *London Stock Exchange* (LSE). Two categories of drivers are adopted for analysis, namely *financial ratios* (FRs) and *corporate governance indicators* (CGIs). FRs are extracted from the *LSEG Datastream*, CGIs are extracted from *BoardEx*, and information on company status is extracted from the *London Share Price Database* (LSPD). In this study, corporate bankruptcy is defined in accordance with bankruptcy legislation in the UK (*the UK Insolvency Act of 1986*). In this paper, a company is considered as bankrupt whenever it is assigned by *LSPD* codes *16* (*Receivership*), *20* (*in Administration),* or *21* (*Cancelled and Assumed valueless*). Since data was collected from multiple sources, we performed a sequence of steps to streamline it for this study. First, companies listed on the LSE between 1990 and 2019 are picked from the *LSEG Datastream*. This gave us general information and financial ratios for 7,165 companies. The chosen companies are then aligned with their counterparts in the *LSPD* and *BoardEx* databases by employing the *Stock Exchange Daily Official List* (*SEDOL*) code *hard match* and *fuzzy name match*. A manual inspection is also performed to ensure that all companies are paired correctly, resulting in a refined list of 1,855 companies with a total of 18,540 firm-year observations. The 18,540 firm-year observations consist of 18,335 non-bankrupt records (99.89%) and 205 bankrupt records (0.11%). All these data are taken directly from the respective sources without any further processing except for the bankruptcy status. The only modification made is in the bankruptcy timeline - some companies, whilst labelled as bankrupt by *LSPD* in a particular year, had actually stopped publishing their data at least one year before their bankruptcy was officially recorded. To address this gap, which often indicates an early sign of severe financial difficulties, we changed the bankruptcy year to the first year the company ceased to report data. As reflected in **Table 4.2**, the annual count of bankruptcy cases differs between the original and adjusted datasets. Moreover, it is worth noting that in the adjusted set, 2019 lacks bankruptcy data. This is because companies marked as bankrupt in 2019 had their bankruptcy dates retroactively moved to 2018 or 2017 following our revision.

Table 4.2 Number of companies in absolute and percentage terms by year

Year	No. Companies According to LSPD	No. Bankrupt Companies According to LSPD (in %)	No. Companies According to Our Modification	No. Bankrupt Companies According to Our Modification (in %)
1990	401	0 (0.00%)	1	0 (0.00%)
1991	409	0 (0.00%)	2	0 (0.00%)
1992	419	0 (0.00%)	4	0 (0.00%)
1993	441	0 (0.00%)	17	0 (0.00%)
1994	480	0 (0.00%)	177	0 (0.00%)
1995	526	0 (0.00%)	369	0 (0.00%)
1996	592	0 (0.00%)	387	0 (0.00%)
1997	661	0 (0.00%)	496	0 (0.00%)
1998	698	0 (0.00%)	581	0 (0.00%)
1999	744	0 (0.00%)	613	0 (0.00%)
2000	879	0 (0.00%)	659	0 (0.00%)
2001	955	0 (0.00%)	754	0 (0.00%)
2002	1007	0 (0.00%)	848	0 (0.00%)
2003	1048	0 (0.00%)	885	4 (0.45%)
2004	1204	6 (0.50%)	895	7 (0.78%)
2005	1337	4 (0.30%)	945	11 (1.16%)
2006	1385	19 (1.14%)	1029	18 (1.75%)
2007	1359	31 (2.28%)	1052	48 (4.56%)
2008	1270	74 (5.83%)	974	26 (2.67%)
2009	1127	78 (6.92%)	883	24 (2.72%)
2010	1031	45 (4.36%)	801	8 (0.99%)
2011	952	36 (3.78%)	770	15 (1.95%)
2012	905	27 (2.98%)	715	7 (1.35%)
2013	879	23 (2.61%)	693	6 (0.98%)
2014	904	9 (1.00%)	694	9 (1.30%)
2015	898	25 (2.78%)	695	3 (0.43%)
2016	864	18 (2.09%)	675	7 (1.04%)
2017	826	13 (1.57%)	643	8 (1.24%)
2018	811	21 (3.33%)	652	4 (0.61%)
2019	761	22 (2.89%)	631	0 (0.00%)

The foundation of the experimental methodology is a robust data preparation process, followed by an intricate profiling analysis. For the evaluation of prediction models, a standard train/test split methodology is implemented. The dataset is divided so that 70% is allocated for model training and the remaining 30% for testing. It is crucial to maintain an identical proportion of bankrupt companies within both the training and the test sets to ensure consistency in the training and evaluation; therefore, the percentage of bankrupt companies is kept constant across the training and test datasets. To reinforce the reliability and robustness of the models, 30 distinct training and test sets are generated, allowing us to conduct extensive validation and reduce the risk of overfitting. This rigorous cross-validation strategy facilitates a more comprehensive assessment of the models' predictive performance and generalisation ability for the unseen data. All the datasets have been standardised to have a mean of 0 and a standard deviation of 1. This standardisation ensures that each variable contributes equally to the analysis and prevents any one variable with a larger scale from dominating the model's feature space. In addition, the observations in the sampled training sets are winsorised at the 99% percentile in order to trim extreme values, and hence minimise the influence of the outliers.

4.5.2 Financial ratios & Corporate governance indicators

In terms of financial ratios, this research is based on conventional accounting ratios that are selected based on a meticulous set of criteria. Initially, the dataset has a pool of 85 potential financial ratios, also referred to as drivers. However, as the data come from multiple sources, there are inconsistencies with some records missing certain values. To address this problem, the completeness of the data for each ratio is determined by the fraction of observations that have data for that specific financial ratio out of the total observations. For the purpose of this study, we focused only on the ratios that had a data completeness of more than 70% and further reduced the number of drivers to keep only those that have a very good discrimination between healthy and bankrupt companies based on standard summary statistics. Based on these criteria, the final set of drivers selected is: *Working Capital/Total Assets*, *Current Assets/Current Liabilities*, *Retained Earnings/Total Assets*, *Cash/Total Assets*, *Cash/Current Liabilities*, *Cash Flow/Total Liabilities*, *Total Liabilities/Total Assets*, *Current Assets/Total Liabilities*, *Current Liabilities/Total Assets*, *Sales/Total Assets*, *Net Income/Sales*, *Sales/Current Assets*, *Net Income/Total Assets*, *EBITDA/Total Assets*, *Cost of Sales/Sales*, *Market Value of Equity/Total Liabilities*, and *Earnings per share*.

Apart from the financial ratios, we also include several *corporate governance indicators* (CGIs) which are proven to be effective in the current literature (e.g., Li et al., 2021, Zhao et al., 2024), namely, *Total Number of Directors*, *CEO/Chairman Duality* and *Percentage of Non-executive Directors*, to improve the predictability of the model. Including these CGIs is essential because they encapsulate aspects of a company's governance that are not captured by financial ratios. By integrating these governance measures, this study leverages a more holistic approach in merging traditional financial analysis with corporate governance evaluation to refine the predictive capabilities of prediction models.

Although our method for driver selection is specific and rigorous, it is noteworthy that the drivers we have selected align well with those commonly adopted in the existing literature (Zhao et al., 2024). This result suggests that the selection method adopted is sound and reasonable.

4.5.3 Implementation decisions of the two-step method

4.5.3.1 Selection of profiling methods

In examining various clustering methodologies to effectively profile companies for predicting bankruptcy and financial distress, in this study, six clustering techniques, namely, *k-means*, *DBSCAN*, *Agglomerative hierarchical clustering* (AHC), *GMM*, *Spectral clustering*, and *BIRCH*, are tested and analysed. This analysis directly addresses the research question 2 on identifying the most robust profiling technique that consistently and effectively categorises companies based on shared financial behaviours, thereby enhancing bankruptcy prediction performance. Please refer to **Fig. 4.7** for a graphical illustration of the profiling results obtained with *t-SNE*, a 2-D plotting method.

These methods were selected for their ability to effectively identify patterns within financial datasets. Each method was used to create distinct groups of companies, and their performance was evaluated based on how well they can separate bankrupt companies from healthy ones. The key to answering our research question lies in determining which clustering method can create profiles that exhibit a clear distinction between companies at high risk of bankruptcy and those at lower risk. An essential criterion for evaluating the effectiveness of these clustering methods is ensuring that the intra-group variance (i.e., variance within clusters) is smaller than the inter-group variance (i.e., variance between clusters). This is critical for validating that the created profiles are compact and distinct, reflecting meaningful groupings within the data. Hence, in our case, the most effective model is one that concentrate a high percentage of bankrupt companies into specific clusters whilst maintaining lower percentages in others, indicating a clear distinction between groups. Through this evaluation, we found that *k-means* emerged as the most effective clustering method for this purpose, as it achieves the largest difference between the highest and lowest percentages of bankrupt companies within its clusters, and also provides meaningful company clusters and groups bankrupt companies in a consistent manner (**Table 4.3**). This indicates that *k-means* has superiority in discerning the heterogeneity within the data, which is vital for the identification of bankruptcy signals.

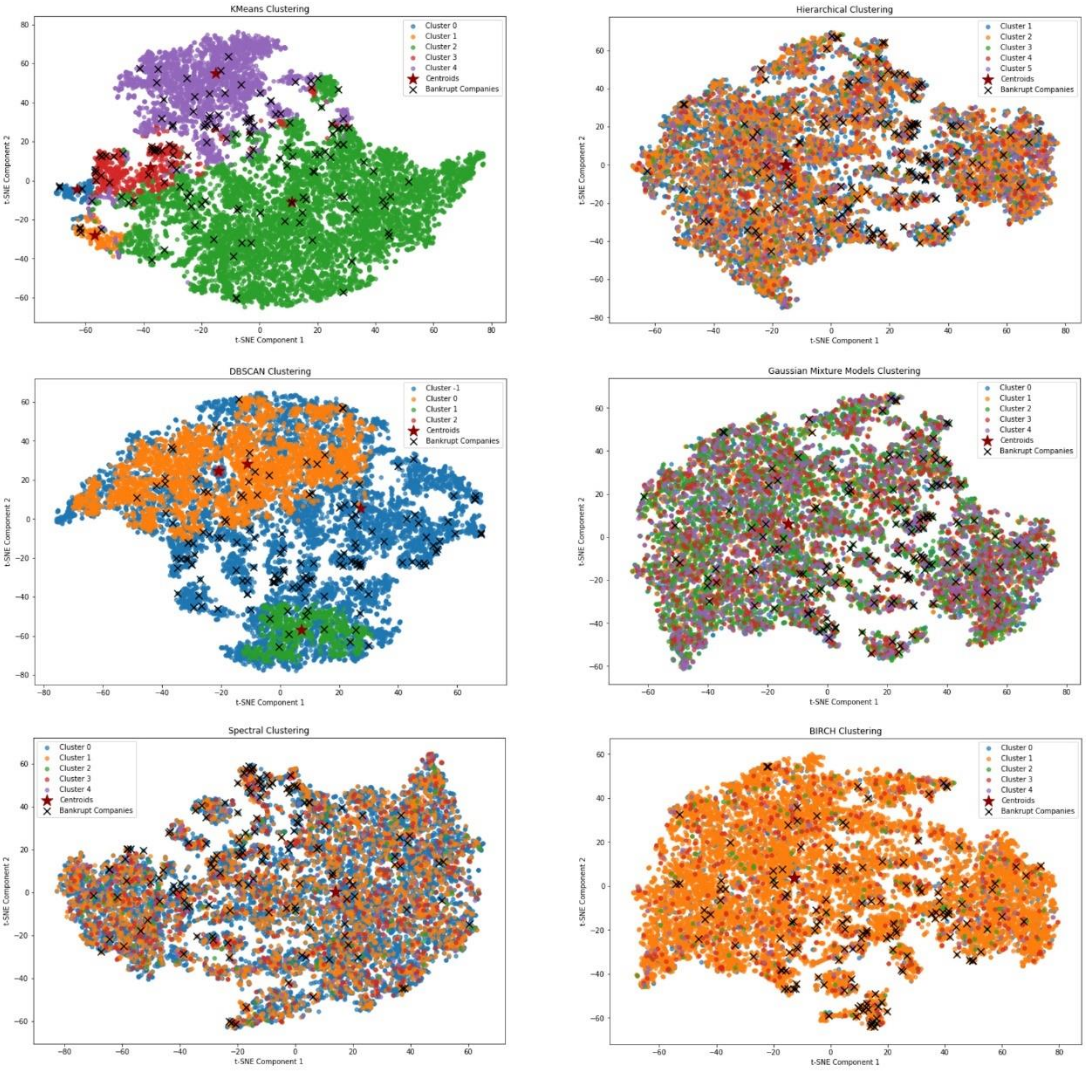

Fig. 4.7 Graphic illustration of profiling results for t-SNE (2-D plots)

Table 4.3 Highest/Lowest Percentage of Bankrupt Companies by Profiling Method

Methods	Highest Percentage of Bankrupt Companies in Clusters	Lowest Percentage of Bankrupt Companies in Clusters	Difference (Highest - Lowest)
K-means	3.82%	0.76%	3.06%
AHC	1.20%	0.47%	0.73%
DBSCAN	1.43%	0.00%	1.43%
GMM	1.23%	0.67%	0.56%
Spectral Clustering	2.72%	0.63%	2.09%
BIRCH	2.38%	0.61%	1.77%

The best/optimal number of profiles for further analysis was determined through *Silhouette analysis*, a method that evaluates the separation distance between clusters, and examination of the 'quality' of clusters, as indicated by the mix of bankrupt and healthy companies in the same cluster. Silhouette analysis calculates the *average silhouette score* ($S_{Dataset}$) for different number of clusters, which ranges from -1 to +1:

$$S_{Dataset} = \frac{1}{N}\sum_{i=1}^{N} s(i)$$

$$s(i) = \frac{b(i) - a(i)}{max\{a(i), b(i)\}}; -1 \leq s(i) \leq 1$$

where N is the total number of companies in the dataset, $s(i)$ is the individual Silhouette score for the i-th company, $a(i)$ is the average distance between the i-th company and all other companies in the same cluster (i.e., intra-cluster distance), and $b(i)$ is the minimum average distance from the i-th company to all companies in the nearest neighbouring cluster (i.e., nearest-cluster distance).

Scores close to +1 suggest the sample is well-separated from neighbouring clusters, whilst scores near 0 indicate proximity to a cluster boundary. Negative scores may imply misclassified samples. This score is critical because it quantifies how well each observation lies within its cluster, providing an essential measure of compactness and separation, which are key indicators of a robust profiling method. A higher average silhouette score corresponds to a model with better-defined clusters. Silhouette analysis indicated fluctuations in scores across different cluster counts (**Fig. 4.8**). However, counts that resulted in low scores were indicative of overlapping or poorly separated clusters, thus being unsuitable.

In our analysis, whilst the two-profile scenario showed the highest average silhouette score of 0.5804, suggesting tight clusters with a clear distinction from each other, it however lacks the granularity required for our specific research goals. Similarly, for the three-profile model, despite a reasonably high score of 0.5145, it does not meet the need for a nuanced categorisation to capture the complex patterns within the data. As progressing to higher profile counts, a dip in the silhouette scores is observed, with four profiles' scores reducing to 0.1208, reflecting a significant decrease in cluster definition. Upon reaching five profiles, the silhouette score grows back to 0.1682, which marks an improvement over the four-profile scenario and indicates a better balance between cluster cohesion and separation. After careful consideration, although this score is lower than that for two- or three-profile scenarios, it provides a more meaningful structure for profiling within the experimental framework. It allows for the discernment of more intricate relationships and patterns that are essential for a comprehensive analysis of bankruptcy risk and provides reasonable separation of observations and good interpretability of the financial profiles.

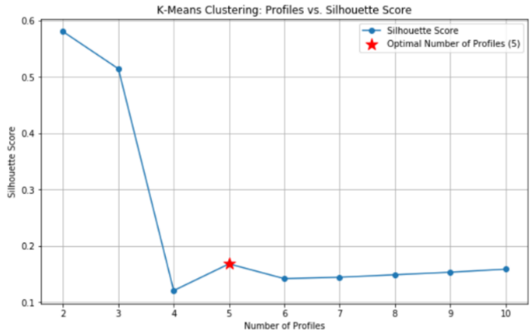

Fig. 4.8 Number of profiles vs. silhouette scores for *k-means* clustering

4.5.3.2 Development of prediction models

For the empirical analysis, 20 supervised models are implemented, and their performance is compared and analysed – see **Table 4.4** for details on these profile-based supervised prediction models.

Table 4.4 The combinations of financial profiles and supervised prediction models

a) Benchmark model without any profiling information
Logistic Regression Model (LR): Logistic regression model with FRs and CGIs as input variables.
b) Prediction Models fed with Profiles-based Drivers
(1) ***Profile Direct Utilisation (PDU)***: LR model fed with FRs and CGIs as input variables as well as the new driver devised with the direct utilisation method that exploits the financial profiles generated by the chosen clustering method. (2) ***Profiling Model with One-Hot Encoding (POH)***: LR model fed with FRs and CGIs as input variables as well as the new driver devised with the one-hot encoding method that exploits the financial profiles generated by the chosen clustering method. (3) ***Profile Ranking Model (PR)***: LR model fed with FRs and CGIs as input variables as well as the new driver devised with the profile ranking method that exploits the financial profiles generated by the chosen clustering method. The financial profiles are ranked by the linear weighting scheme.
c) Profile-specific Prediction Models
(4) ***Profile-Specific Logistic Regression (PSLR)***: For each financial profile or cluster, a dedicated LR model is trained and then used to make predictions for companies within that profile. (5) ***Profile-Specific Prediction Models (PSPM):*** For each financial profile or cluster, a dedicated machine learning model is used to make predictions for companies within that profile. The selection of the machine learning models for each profile is based on the total number of companies within that profile. For each financial profile, if the total number of companies within that profile is between 0-999, SVM is selected; 1,000-2,999, LR is selected; 3,000-4,999, MLP is selected, and above 5,000, AdaBoost is selected.
c) Profiling Models Based on Meta-Learning Structure
(6) ***Meta-Profile Learning Model Using Logistic Regression***: For each financial profile, a dedicated LR model is trained. The predictions from all profile-specific LR models are weighted (linearly or exponentially) by the ranking of each profile, and then combined with the FRs and CGIs as the input variables of the meta-learner (meta LR model). In this method, two weighting schema are compared, resulting in 2 variants with different model settings. (7) ***Meta-Profile Learning Model Using Logistic Regression with PCA or RFE (MPLRPCA, MPLRRFE)***: For each financial profile, a dedicated LR model is trained using features processed by *principal component analysis* (PCA) or *recursive feature elimination* (RFE). The predictions from all profile-specific LR models are weighted by the ranking of each profile, and then combined with the FRs and CGIs as the input variables of the meta-learner (meta LR model). For each method, 3- /5-selected feature scenarios are compared, resulting in 4 variants with different model settings. (8) ***Meta-Profile Learning Model Using Logistic Regression with RFE & Global Feature Reduction (MPLRRFE-Reduce, MPLRRFE-ReSelect)*:** For each financial profile, a dedicated LR model is trained using the selected features from RFE. The predictions from all profile-specific LR models are weighted by the ranking of each profile, and then combined with the reduced features (or, re-selected features) as the input variables of the meta-learner (meta LR model). In the *MPLRRFE-Reduce* variant, the global model reduces the feature set to retain only the most impactful features across profiles. In contrast, the *MPLRRFE-ReSelect* variant reselects features after the initial profile-specific RFE, further refining the feature set for the meta-learner. For each method, 3- /5-selected feature scenarios are compared, resulting in 4 variants with different model settings. (9) ***Meta-Profile Learning Model Using Logistic Regression with PCA/RFE & Global Feature Engineering (MPLRPCA/RFE-Eng)*:** For each financial profile, a dedicated LR model is trained using the combination of features processed by both PCA (feature engineering) and RFE (feature selection). Specifically, for each profile, features are first reduced using RFE to identify key predictors. In addition, PCA is applied to further reduce the dimensionality of the data, extracting principal components that capture the maximum variance within the profile. The profile-specific LR models are then trained on a combined feature set that includes both the top features selected by RFE and the principal components derived from PCA. The predictions from all profile-specific LR models are weighted by the ranking of each profile, and then combined with the reduced features (by RFE) as the input variables of the meta-learner (meta LR model). For each method, 3- /5-engineered feature scenarios are compared, resulting in 2 variants with different model settings. (10) ***Meta-Profile Learning Model Using Voting Ensemble (MPVOTRFE-Reduce, MPVOTRFE-ReSelect, MPVOTPCA/RFE-Eng)*:** This approach uses a similar idea as shown in (8) and (9); however, instead of using the LR model, for each financial profile, a dedicated voting ensemble (i.e., LR, decision trees) is adopted. For each method, only 3-selected/engineered feature scenario is considered, resulting in 3 variants with different model settings.

In this study, *logistic regression* is chosen as the prediction model, due to its straightforward interpretability and its well-established use as a benchmark in the financial sector (Almaskati et al., 2021). It provides a transparent method for assessing the impact of specific variables on the outcomes of predictions, which is crucial for our goal of validating the effectiveness of the proposed drivers. Concerning the model's regularisation term, *L1*, *L2,* and *ElasticNet* penalties are employed, where each of these regularisation options has an associated loss function that needs to be optimised to determine the model's parameters effectively. Furthermore, *grid-search* with *5-fold cross-validation* is applied to identify the optimal mix of hyperparameters. The details of the model settings and experimental setups are provided in **Table 4.5**.

Table 4.5 The model settings and experimental setups of the profiling methodologies adopted in the prediction models

Model	Profile Utilisation	Feature Engineering	Feature Selection	Profile Ranking Weighting Scheme	Profile-specific Learning Model	Global Prediction Model
LR	-	-	-	-	-	LR
PDU	Direct Use (Categorical)	-	-	-	-	LR
POH	One-Hot Encoded (Categorical)	-	-	-	-	LR
PR	Profile-Ranking (Categorical)	-	-	Linear	-	LR
PSLR	Profile-specific Prediction	-	-	-	LR	-
APSM	Profile-specific Prediction	-	-	-	Multi-ML Models	-
MPLR-Linear	Meta-Learning Structure	-	-	Linear	LR	LR (Meta)
MPLR-Exponential	Meta-Learning Structure	-	-	Exponential	LR	LR (Meta)
MPLRPCA-3/5	Meta-Learning Structure	Profile-specific PCA	-	Exponential	LR	LR (Meta)
MPLRRFE-3/5	Meta-Learning Structure	-	Profile-specific RFE	Exponential	LR	LR (Meta)
MPLRRFE-3/5-Reduce	Meta-Learning Structure	-	Profile-specific RFE; Global RFE	Exponential	LR	LR (Meta)
MPLRRFE-3/5-ReSelect	Meta-Learning Structure	-	Profile-specific RFE; Global RFE and Feature Re-select	Exponential	LR	LR (Meta)
MPLRPCA/RFE-3/5-Eng	Meta-Learning Structure	Profile-specific PCA	Profile-specific RFE; Global RFE + Profile-specific PCA	Exponential	LR	LR (Meta)
MPVOTRFE-3-Reduce	Meta-Learning Structure	-	Profile-specific RFE; Global RFE	Exponential	Voting Ensemble	LR (Meta)
MPVOTRFE-3-ReSelect	Meta-Learning Structure	-	Profile-specific RFE; Global RFE and Feature Re-select	Exponential	Voting Ensemble	LR (Meta)
MPVOTPCA/RFE-3-Eng	Meta-Learning Structure	Profile-specific PCA	Profile-specific RFE; Global RFE + Profile-specific PCA	Exponential	Voting Ensemble	LR (Meta)

Each configuration of the model was systematically evaluated to determine its efficacy in leveraging financial profiles for the prediction of bankruptcy. The settings reflect a progression from straightforward logistic regression model to more complex ensemble and meta-learning structures, testing a broad spectrum of approaches from basic to advanced techniques.

4.5.3.3 Performance measures of classification models

Six performance measures are used to evaluate the prediction models. These include four single-criterion-based measures, that is, *sensitivity* (*true positive rate*) to evaluate the ability of the models to correctly predict bankrupt companies; *specificity* (*true negative rate*) to evaluate the ability of the models to correctly predict healthy companies; *Type-I error* (*false positive rate),* and *Type-II error* (*false negative rate*) to evaluate the model's lack of correctly predicting the risk class. Note that the single-criterion-based performance measures are chosen in order to minimise the biases that can occur with weighted or multi-criteria evaluation metrics. Also, two multi-criteria measures are included to evaluate the overall performance of the models, that is, precision that evaluates the overall predictability of the models as measured by the number of correct predictions on the total number of predictions; and the *area under the receiver operating characteristic curve* (AUC), which provides an aggregate measure of performance at all possible cut-off points that reflects how the model performs compared to a random classifier. We recommend readers refer to **Table 4.6** for details of these performance measures, including their definitions and mathematical formula.

Table 4.6 Performance measures for classification models

Performance Metric	Definitions	Mathematical Formula
Type-I Error (false positive rate)	Occurs when a healthy company is incorrectly classified as bankrupt (false positive).	$Type-I\ Error = \frac{False\ Positives}{False\ Positives + True\ Positives}$
Type-II Error (false negative rate)	Occurs when a bankrupt company is incorrectly classified as healthy (false negative).	$Type-II\ Error = \frac{False\ Negatives}{True\ Positives + False\ Negatives}$
Sensitivity (true positive rate/recall)	Measures the model's ability to correctly identify bankrupt companies (positive class).	$Sensitivity = \frac{True\ Positives}{True\ Positives + False\ Negatives}$
Specificity (true negative rate)	Measures the model's ability to correctly identify healthy companies (negative class).	$Specificity = \frac{True\ Negatives}{True\ Negatives + False\ Positives}$
Classification Accuracy	Measures the overall correctness of the model's predictions across all classes.	$Accuracy = \frac{Number\ of\ Correct\ Predictions}{Total\ Number\ of\ Predictions}$
Area Under ROC Curve (AUC)	Evaluates the discriminative power of the model by measuring the area under the *receiver operating characteristic* (ROC) curve, which plots *true positive rate* (Sensitivity) against *false positive rate* (1-Specificity) at various threshold values. Higher AUC indicates better performance.	AUC is the area under the ROC curve, typically calculated numerically.

4.6 Empirical analysis and findings

In the empirical analysis, various combinations of methods are implemented to evaluate the predictive performance of the proposed bankruptcy prediction models. In the following four subsections, we shall discuss these results in detail and provide answers to the research questions.

4.6.1 Analysis of profiles

This section of the analysis directly addresses our research question 3 on how distinct financial profiles correlate with varying levels of bankruptcy risk and how these profiles can enhance the accuracy and interpretability of predictive models. The analysis begins by examining the characteristics of financial profiles identified in the dataset. These profiles were generated through *k-means* clustering, which grouped companies based on similar financial characteristics. The identification of these profiles serves as the foundation for understanding the diverse financial behaviours and risk levels present within the dataset, thus providing crucial insights for addressing the issue of sample heterogeneity.

Each profile represents a distinct financial pattern and is associated with specific risk levels, as outlined in **Table 4.7**. The table provides an example of the key financial characteristics of each profile along with their associated risk levels. The numerical values of the centroids of the profiles of the selected example are provided in **Appendix 4.A.1**.

Table 4.7 The description and key financial characteristics of the profiles (an example from the 30 experiments)

Profile	Overview	Key Financial Characteristics	Risk Level
Profile 1	*The Stable Performer Group*	- Slightly negative *Working Capital/Total Assets*, indicating minimal excess working capital; - Positive *Retained Earnings/Total Assets*, reflecting a good history of profitability; - Balanced *Current Liabilities/Current Assets* ratio, indicating sufficient short-term liquidity; - Positive *Cashflow* metrics, showing operational stability; - Positive *EBIT/Sales* and *Net Income/Sales* ratios, indicating sustainable profitability; - Negative *CEO/Chairman Duality*, suggesting separation of powers in governance.	**Low Risk** Lowest bankruptcy rate of 0.76%.
Profile 2	*The Distressed Group*	- Substantial negative *Working Capital/Total Assets*, suggesting short-term liquidity issues; - Highly negative *Retained Earnings/Total Assets*, indicating long-term accumulated losses; - High *Current Liabilities/Current Assets* ratio, exacerbating short-term liquidity concerns; - Negative *Cashflow* metrics, reflecting operational inefficiencies; - Negative *EBIT/Sales* and *Net Income/Sales* ratios, suggesting operational losses; - Slightly positive *CEO/Chairman Duality*, indicating potential consolidated power.	**Very High Risk** Highest bankruptcy rate of 3.82%.
Profile 3	*The Conservative and Cautious Group*	- Slightly negative *Working Capital/Total Assets*, indicating cautious asset management; - Positive but lower *Retained Earnings/Total Assets*, reflecting conservative profitability; - Balanced *Current Liabilities/Current Assets* ratio, indicating sufficient short-term liquidity; - Marginally positive *Cashflow* metrics, showing careful operational management; - Modest positive *EBIT/Sales* and *Net Income/Sales* ratios, indicating reasonably good profitability; - Strong positive *CEO/Chairman Duality*, suggesting strong consolidated decision-making.	**Moderate Risk** Bankruptcy rate of 1.30%.
Profile 4	*The High-Risk Group*	- Significantly positive *Working Capital/Total Assets*, but overshadowed by severe losses; - Deeply negative *Retained Earnings/Total Assets*, indicating substantial historical losses; - Negative *Current Liabilities/Current Assets* ratio, suggesting strong liquidity; - Extremely negative *Cashflow* metrics, indicating potential insolvency; - Profoundly negative *EBIT/Sales* and *Net Income/Sales* ratios, reflecting strong profitability concerns; - Slightly positive *CEO/Chairman Duality*, with governance overshadowed by financial struggles.	**High Risk** High bankruptcy rate of 3.06%.
Profile 5	*The Outlier Group*	- Extremely high *Working Capital/Total Assets*, suggesting strong liquidity or over-capitalisation; - Negative *Retained Earnings/Total Assets*, indicating substantial historical losses; - Negative *Current Liabilities/Current Assets* ratio, reflecting a strong short-term liquidity; - Extremely positive *Cashflow* metrics, however, with negative profitability ratios, indicating unusual financial behaviour; - CEO/Chairman Duality and other governance indicators varies, suggesting special governance structures/procedures.	**Mixed Risk** Bankruptcy rate of 1.60%.

These financial profiles reveal distinct patterns that relate to the bankruptcy risk. Profile 1 (*The Stable Performer Group*), representing stable performers, shows positive profitability and strong liquidity, resulting in the lowest bankruptcy risk. In contrast, Profile 2 (*The Distressed Group*), the distressed group, exhibits severe liquidity and profitability issues, leading to the highest bankruptcy risk. The other profiles highlight varying degrees of financial health, with Profile 3 (*The Conservative and Cautious Group*) showing cautious but stable performance, Profile 4 (*The High-Risk Group*) marked by high risk despite relatively good liquidity, and Profile 5 (*The Outlier Group*) containing outliers with unusual financial behaviours. These profiles underscore the importance of considering multiple financial dimensions in bankruptcy prediction, as they offer nuanced insights into the factors that drive financial performance and risk, enhancing the accuracy and reliability of predictive models.

These financial profiles can also offer actionable insights for managers. For companies in Profile 1, managers in this group should focus on maintaining the strengths by ensuring consistent cash flow management and avoiding excessive risk-taking behaviour that could destabilise the financial position. In contrast, as profile 2 highlights companies with severe financial vulnerabilities, for companies in this group, immediate corrective actions are critical, including restructuring short-term liabilities, improving operational efficiency, and addressing governance weaknesses to regain stability. Profile 3 represents companies with cautious but stable management, balanced liquidity, and moderate profitability. Managers of these companies should aim to improve profitability whilst ensuring governance remains adaptable to changing market conditions to enhance resilience. Meanwhile, Profile 4 encompasses companies with relatively good liquidity but significant profitability challenges. Managers in this group should prioritise addressing core operational inefficiencies and reassessing strategic investments to reverse declining profitability. Finally, for companies in Profile 5 which demonstrate unique financial behaviours and governance structures, managers must conduct deeper diagnostics to identify the underlying causes of their unusual financial patterns and implement tailored actions to stabilise the companies' performance.

Integrating these profiles into predictive models allows for more targeted and effective identification of companies at risk of bankruptcy. The details of centroids (i.e., sample average) for each financial profile of the given example is provided in **Fig. 4.9**, which demonstrates the clear differences between the financial profiles in terms of each selected feature.

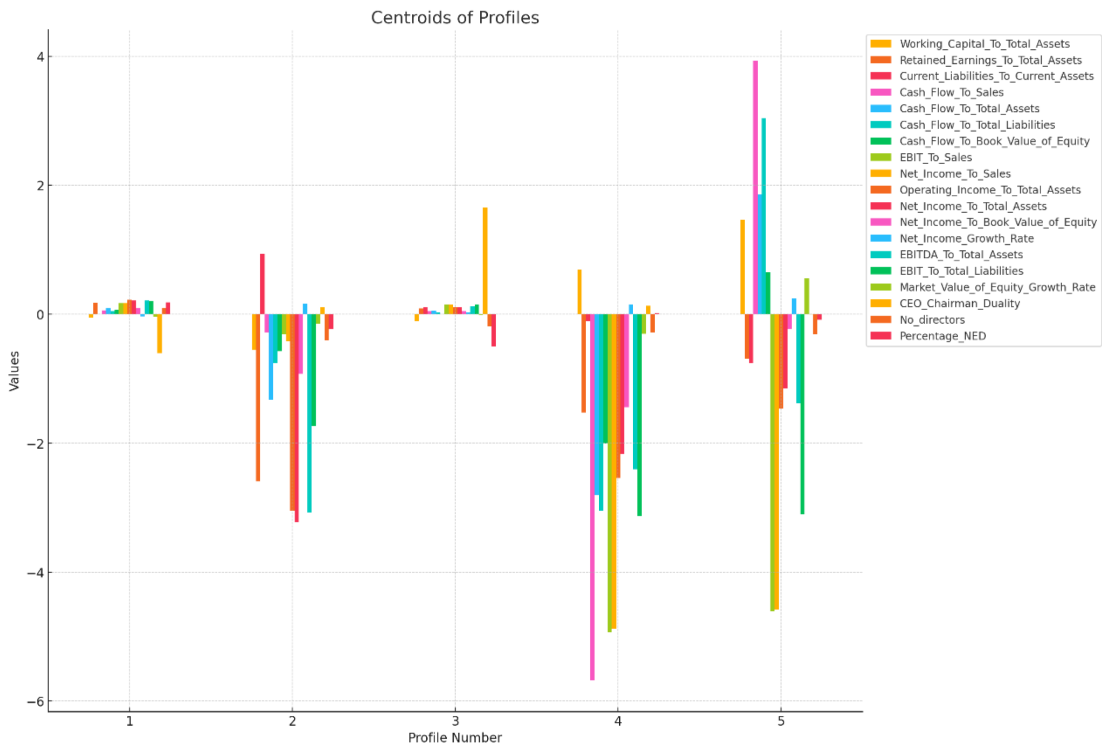

Fig. 4.9 A demonstration of centroids of profiles (an example from experiment)

The identification and analysis of these profiles highlight the diverse financial behaviours and risk levels present within the dataset. Understanding these profiles is crucial as they directly inform the performance of the prediction models. These differences offer not only a foundation for prediction but also a blueprint for managers to implement tailored strategies based on their profile's risk level and characteristics.

4.6.2 Analysis of results of profile-based prediction models using profile-based drivers

In this section, we specifically address our research questions 3 and 4 regarding the impact of integrating financial profiles into supervised prediction models, focussing on how different modelling strategies influence prediction performance.

The first approach involved profile-based prediction models based on profile-based drivers. The three key methodologies in this category, namely, *profile direct utilisation* (PDU), *profiling with one-hot encoding* (POH), and *profile ranking model* (PR) are tested and analysed. The results of these experiments are presented in **Table 4.7**.

The model with direct utilisation of financial profiles (PDU) integrates financial profiles directly into the *logistic regression* model. With sensitivity of 73.78% in out-of-sample test sets, it demonstrates an improvement over the benchmark logistic regression model, which stands at 64.72%, by 9.06%. The improvement suggests that direct utilisation of financial profiles can enhance the detection of bankrupt companies (see **Table 4.8** for the statistical tests of model performance). In terms of *profiling with one-hot encoding model* (POH) and *profile ranking model* (PR), both models show an enhanced sensitivity in out-of-sample test sets (i.e., 73.57% and 73.81%, respectively) over the benchmark model, and the highest AUC score is achieved by *POH* model at 0.8047, which is a 3.85% improvement.

The subtle, yet notable differences in the performance of *PDU*, *POH*, and *PR* illustrate the impact of how financial profiles are integrated into predictive models. The improvement in the sensitivities of the out-of-sample of these three models over the benchmark model suggested that profile analysis adds value to the detection of bankrupt companies. Both *POH* and *PR* provide more sophisticated methodologies that enhance the LR model's ability to generalise and focus on meaningful data through effective feature transformation and prioritisation, and these insights suggest that utilising financial profiles is not only necessary but critical in refining predictive models and improving the models' robustness.

Table 4.7 Results of experiments of profiling models based on *profile operation* and *profile-specific methodologies*

Results - Mean	Sensitivity (TPR)		Specificity (TNR)		Type-I Error Rate (FPR)		Type-II Error Rate (FNR)		Classification Accuracy (ACC)		Area Under ROC Curve (AUC)
	TPR Train	*TPR Test*	*TNR Train*	*TNR Test*	*Type-I Train*	*Type-I Test*	*Type-II Train*	*Type-II Test*	*Acc Train*	*Acc Test*	*AUC Test*
Logistic Regression / Benchmark Model											
LR	67.01%	64.72%	71.47%	71.40%	28.53%	28.60%	32.99%	35.28%	71.42%	71.33%	0.7662
Profiling Models Based on Profile Operation											
PDU	79.12%	73.78%	74.98%	74.88%	25.01%	25.11%	20.88%	26.21%	75.03%	74.88%	0.8024
POH	79.05%	73.57%	74.94%	74.86%	25.06%	25.14%	20.95%	26.42%	74.98%	74.84%	0.8047
PR	79.05%	73.81%	74.95%	74.82%	25.04%	25.18%	20.95%	26.18%	75.00%	74.81%	0.8029
Profiling Models Based on Profile-Specific Methodologies											
PSLR	75.31%	69.23%	73.16%	72.95%	26.84%	27.05%	24.69%	30.77%	73.19%	72.90%	0.7799

APSM	84.18%	64.47%	74.47%	73.94%	25.53%	26.06%	15.82%	35.53%	74.58%	73.84%	0.7318

Note: *LR* – Logistic Regression; *PDU* – Profile Direct Utilisation of Profiles; *POH* – Profiling Model with One-Hot Encoding; *PR* – Profile Ranking Model; *PSLR* – Profile-Specific Logistic Regression Model; *APSM* – Profile-Specific Dedicated Machine Larning Model.

4.6.3 Analysis of results of profile-based prediction models based on profile-specific methodologies

The application of dedicated *logistic regression* models to individual financial profiles (i.e., *profile-specific logistic regression*, or PSLR) and adaptive model selection based on profile characteristics (i.e., *adaptive profile-specific models*, or APSM) offers some other insights of profile analysis in bankruptcy prediction (**Table 4.7**).

For the *PSLR* model, when the LR models are tailored to each financial profile, the model exhibits a lower overall effectiveness compared with the direct utilisation of profiles (i.e., *PDU*, *PHO* and *PR*), with lower out-of-sample sensitivity (i.e., 69.23%), specificity (i.e., 72.95%) and AUC (i.e., 0.7799) (see **Table 4.8** for the statistical tests of model performance). This suggests potential challenges in model generalisation when overly customised to specific data subsets. These findings highlight the inherent trade-offs between customisation and generalisation, a crucial consideration for determining the most effective use of financial profiles in supervised models. Similarly to the *PSLR* model, *adaptive profile-specific models* (APSM) experienced the same issue — This approach leverages distinct machine learning models chosen based on the specific characteristics of each profile, which, ideally, is an excellent strategy as it allows for the adaptation of models to the unique characteristics of each company profile, however, whilst *APSM* shows a remarkable in-sample sensitivity (i.e., 84.18%), its out-of-sample sensitivity plummets to 64.47%, underscoring a severe overfitting issue. The lowest AUC score of 0.7318 among all models confirms that, whilst adaptive methodologies may promise high customisation, they risk failing in unpredicted environments.

Segmentation into different financial profiles, although providing a tailored approach, often results in a significantly smaller sample size for each supervised model, which can severely hinder the models' ability to learn effectively and generalise beyond training data. The complexity introduced by customising models for different profiles can also contribute to difficulties in model training and generalisation, as complex models require more data to validate their predictions and prevent overfitting (Rajput et al., 2023). Variability across profiles can exacerbate these challenges if the profiles are not sufficiently distinct or if their defining characteristics do not have strong predictive value, potentially leading to models that are too specialised to their training data, which again, contribute to the overall overfitting of the model. Moreover, both *PSLR* (based on multiple LR models) and *APSM* (based on various machine learning models) are highly sensitive to the features selected during the modelling process. Suboptimal features can lead to inadequate or misleading signals being used to train the models, further impacting the model's performance and generalisability. These challenges underscore the importance of robust dataset sizes, careful feature selection, and model simplicity to ensure that the models are predictive and generalisable.

Moving forward, it may be beneficial to explore models that balance customisation with the ability to generalise across broader datasets, in particular, integrating frameworks which can not only address the limitations of overfitting seen in profile-specific methodologies, but also enhance the model's capacity to learn from complex, multi-dimensional data from the financial profiles, offering a more nuanced and scalable solution to bankruptcy prediction.

Table 4.8 Mann-Whitney U test for profiling models based on *profile operation* and *profile-specific* methodologies (Comparison against Benchmark Model LR)

Metric	Group	Model	Mean	p_value (*Significance)
		Comparison against Benchmark Model (LR)		
Test Sensitivity (TPR)	Control	*LR*	64.72%	-
	1	*PDU*	73.78%	.0000***
	2	*POH*	73.57%	.0000***
	3	*PR*	73.81%	.0000***
	4	*PSLR*	69.23%	.0101*
	5	*APSM*	64.47%	.9528
Test Specificity (TRN)	Control	*LR*	71.40%	-
	1	*PDU*	74.88%	.0000***
	2	*POH*	74.86%	.0000***
	3	*PR*	74.82%	.0000***
	4	*PSLR*	72.95%	.0002***
	5	*APSM*	73.94%	.0000***
Test Accuracy (ACC)	Control	*LR*	71.33%	-
	1	*PDU*	74.88%	.0000***
	2	*POH*	74.84%	.0000***
	3	*PR*	74.81%	.0000***
	4	*PSLR*	72.90%	.0001***
	5	*APSM*	73.84%	.0000***
Test AUC	Control	*LR*	0.7662	-
	1	*PDU*	0.8024	.0000***
	2	*POH*	0.8047	.0000***
	3	*PR*	0.8029	.0000***
	4	*PSLR*	0.7799	.0339*
	5	*APSM*	0.7318	.0000***

*p<.05, **p<.01, ***p<.001

4.6.4 Analysis of results of profile-based prediction models based on meta-learning structure

This subsection provides the empirical results of profiling models structured around 'meta-learning'. It addresses the research question 5 on exploring the trade-offs between profile-specific and global models, specifically within a meta-learning context. It also addresses the research question 6 on the role of feature engineering and feature selection in enhancing the performance of profile-based prediction models for bankruptcy prediction. The performance of models focussing on different weights for the profile ranking (i.e., *MPLR-Linear* and *MPLR-Exponential*), and models focussing on using different profile-specific feature engineering/selection methods (i.e., *MPLRPCA* and *MPLRRFE*), are presented in **Table 4.9**. These models are designed to enhance the predictive capabilities of the financial profiles by incorporating meta-learning structures that employ different strategies for weighting profile riskiness and conducting feature engineering/selection in bankruptcy prediction.

The *MPLR-Linear* model, which uses a linear weighting of profiles, shows an in-sample training sensitivity of 79.86% and an out-of-sample test sensitivity of 66.94%. Compared to the benchmark LR model, which has an out-of-sample test sensitivity of 64.72%, the *MPLR-Linear* model suggests a slight improvement by 2.22%. However, it also records an AUC of 0.7615, which is lower than the benchmark LR's AUC of 0.7662. The subtle improvement in test sensitivity and decrease in AUC, together with a significant gap between the training and test sensitivity, suggest that the linear weighting approach is suffered from the overfitting issue, and may not optimise the balance between sensitivity and specificity as effectively as the benchmark model (see **Table 4.10** for the statistical tests of model performance). In contrast, the *MPLR-Exponential* model applies an exponential weighting scheme to the ranking of profiles, aiming to enhance sensitivity to profiles likely associated with higher bankruptcy risk. It achieves a higher out-of-sample test sensitivity of 71.40% and AUC of 0.7897, improving upon the benchmark by 6.68% and 2.35%, respectively, however, the out-of-sample test specificity for this model is 76.02%, showing a slight reduction compared to the *MPLR-Linear* model (77.94%) (see **Table 4.10** for the statistical tests of model performance). Despite this, the overall performance of *MPLR-Exponential* model indicates a better handling of the weights of the profile ranking that could potentially discover more nuanced bankruptcy signals than the linear method, albeit with a modest trade-off in specificity.

Comparing *MPLR-Linear* and *MPLR-Exponential* models highlights the trade-offs involved in integrating meta-learning structures with logistic regression for bankruptcy prediction. The *MPLR-Exponential* model, with its focus on exponentially scaling the riskiness of profiles, shows a promising direction by significantly improving out-of-sample test sensitivity without a substantial drop in other performance metrics. However, the relatively lower improvement in other performance criteria underscores the complexity and challenges of effectively incorporating meta-profile learning into bankruptcy prediction models, which indicates that further adjustments to the models are needed.

Table 4.9 Results of experiments of profiling models based on *meta-learning structure*

Results - Mean	Sensitivity (TPR)		Specificity (TNR)		Type-I Error Rate (FPR)		Type-II Error Rate (FNR)		Classification Accuracy (ACC)		Area Under ROC Curve (AUC)
	TPR Train	*TPR Test*	*TNR Train*	*TNR Test*	*Type-I Train*	*Type-I Test*	*Type-II Train*	*Type-II Test*	*Acc Train*	*Acc Test*	*AUC Test*
Logistic Regression / Benchmark Model											
LR	67.01%	64.72%	71.47%	71.40%	28.53%	28.60%	32.99%	35.28%	71.42%	71.33%	0.7662
Profiling Models Based on Meta-Learning Structure											
MPLR-Linear	79.86%	66.94%	78.04%	77.94%	21.96%	22.07%	20.14%	33.06%	78.06%	77.82%	0.7615
MPLR-Exponential	79.21%	71.40%	76.16%	76.02%	23.82%	23.98%	20.79%	28.60%	76.21%	75.97%	0.7897
MPLR Models with Profile-specific Feature Engineering/Selection											
MPLRPCA-3	79.00%	73.63%	75.01%	74.91%	24.99%	25.09%	21.00%	26.36%	75.06%	74.90%	0.8029
MPLRPCA-5	79.17%	73.92%	74.97%	74.87%	25.03%	25.13%	20.82%	26.07%	75.02%	74.86%	0.8025

MPLRRFE-3	79.02%	73.67%	75.19%	75.13%	24.80%	24.87%	20.98%	26.33%	75.24%	75.12%	0.8009
MPLRRFE-5	79.08%	73.37%	75.41%	75.30%	24.59%	24.70%	20.92%	26.63%	75.45%	75.28%	0.7994

Note: *LR* – Logistic Regression; *MPLR-Linear* – Meta-Profile Learning Model Using Logistic Regression with Linearly Weighted Profile Ranking; *MPLR-Exponential* – Meta-Profile Learning Model Using Logistic Regression with Exponentially Weighted Profile Ranking; *MPLRPCA-3/5* – Meta-Profile Learning Model Using Logistic Regression with Profile-specific PCA Feature Engineering (3 or 5 principal components); *MPLRRFE-3/5* – Meta-Profile Learning Model Using Logistic Regression with Profile-specific RFE Feature Selection (3 or 5 selected features).

The results of *MPLRPCA* and *MPLRRFE* models (i.e., *MPLRPCA-3*, *MPLRPCA-5*, *MPLRRFE-3* and *MPLRRFE-5*), which represent further developments within the meta-learning framework introduced earlier, suggest that the implementation of profile-specific feature engineering/selection for the profiles can be helpful for the meta-learning framework. These models incorporate feature engineering technique—*principal component analysis* (PCA), and feature-selection technique — *recursive feature elimination* (RFE) to refine the feature sets used in profile-specific logistic regression model for bankruptcy prediction.

The *MPLRPCA-3* and *MPLRPCA-5* models employ PCA to reduce profile-specific feature dimensions to 3 and 5 principal components, respectively. The *MPLRPCA-3* model shows an out-of-sample test sensitivity of 73.63% and specificity of 74.91%, achieving an AUC of 0.8029. Slightly enhancing these metrics, the *MPLRPCA-5* model achieves an out-of-sample test sensitivity of 73.92% and specificity of 74.87% and an AUC of 0.8025. Both models demonstrate an improved out-of-sample test sensitivity compared to the benchmark LR model. This indicates that the integration of PCA for feature reduction effectively captures essential variance in the data, which aids in improving the correct prediction of bankrupt companies without overly complicating the model structure (see **Table 4.10** for the statistical tests of model performance). Similarly, *MPLRRFE* models, which utilise RFE for feature selection, also show promising results. The *MPLRRFE-3* model which selects the top 3 features, achieves an out-of-sample test sensitivity of 73.67% and an AUC of 0.8009, significantly improves the benchmark model's performance by 8.95% and 3.47%, respectively; The *MPLRRFE-5* model which selects the top 5 profile-specific features, on the other hand, suggests a slight dip in performance with an out-of-sample sensitivity of 73.37% and an AUC of 0.7994 (see **Table 4.10** for the statistical tests of model performance).

When comparing the results, the *MPLRPCA* models generally perform slightly better in terms of AUC, suggesting that the dimensionality reduction through PCA might be slightly more effective in capturing useful bankruptcy signals from the data compared to the selective approach of RFE. Both models implementing PCA and RFE, however, show an improvement over the *MPLR-Linear* model's AUC of 0.7615 and are on par with or slightly better than the *MPLR-Exponential* model's AUC of 0.7897. This enhancement could be attributed to the more targeted approach in feature processing, which aligns well with the exponential weighting scheme that emphasises profiles that are likely to indicate higher risk of bankruptcy. Integration of profile-specific feature engineering and selection techniques within the meta-learning framework shows significant promise, as they refine input features into more capable predictors of bankruptcy, improving the models' ability to correctly detect bankrupt companies. These results underscore the effectiveness of combining feature engineering/selection techniques with exponential profile ranking to further optimise bankruptcy prediction models.

Table 4.10 Mann-Whitney U test for profiling models based on *meta-learning structure* methodologies (Comparison against Benchmark Model LR)

Metric	Group	Model	Mean	p_value (*Significance)
		Comparison against Benchmark Model (LR)		
Test Sensitivity (TPR)	Control	*LR*	64.72%	-
	6	*MPLR-Linear*	66.94%	.1205
	7	*MPLR-Exponential*	71.40%	.0001***
	8	*MPLRPCA-3*	73.63%	.0000***
	9	*MPLRPCA-5*	73.92%	.0000***
	10	*MPLRRFE-3*	73.67%	.0000***
	11	*MPLRRFE-5*	73.37%	.0000***
Test Specificity (TRN)	Control	*LR*	71.40%	-
	6	*MPLR-Linear*	77.94%	.0000***
	7	*MPLR-Exponential*	76.02%	.0000***
	8	*MPLRPCA-3*	74.91%	.0000***
	9	*MPLRPCA-5*	74.87%	.0000***
	10	*MPLRRFE-3*	75.13%	.0000***
	11	*MPLRRFE-5*	75.30%	.0000***
Test Accuracy (ACC)	Control	*LR*	71.33%	-
	6	*MPLR-Linear*	77.82%	.0000***
	7	*MPLR-Exponential*	75.97%	.0000***
	8	*MPLRPCA-3*	74.90%	.0000***
	9	*MPLRPCA-5*	74.86%	.0000***
	10	*MPLRRFE-3*	75.12%	.0000***
	11	*MPLRRFE-5*	75.28%	.0000***
Test AUC	Control	*LR*	0.7662	-
	6	*MPLR-Linear*	0.7615	.6898
	7	*MPLR-Exponential*	0.7897	.0028**
	8	*MPLRPCA-3*	0.8029	.0000***
	9	*MPLRPCA-5*	0.8025	.0000***
	10	*MPLRRFE-3*	0.8009	.0000***
	11	*MPLRRFE-5*	0.7994	.0000***

*p<.05, **p<.01, ***p<.001

Table 4.11 presents the results of the models (i.e., *MPLRREF-Reduce*, *MPLRREF-ReSelect*, and *MPLRPCA/REF-Eng*) focussing on further implementing global future engineering/selection on the prediction models, which explores how modifications in feature selection strategies impact the prediction outcomes within the meta-learning structure framework. These models represent an evolution in the approach to feature engineering by integrating both profile-specific and global feature engineering/selection techniques.

The *MPLRRFE-3-Reduce* and *MPLRRFE-5-Reduce* models, which further utilise RFE to reduce the feature sets of global prediction models (i.e., meta-learners) to 3 and 5 features, respectively, show enhanced predictive performance. The M*PLRRFE-3-Reduce* model demonstrates an out-of-sample test sensitivity of 74.55% and specificity of 75.17%, with an AUC of 0.8089. This represents a notable improvement over the benchmark model, suggesting that a targeted reduction in feature dimensions can effectively enhance models' performance (see **Table 4.12** for the statistical tests of model performance). Similarly, the *MPLRRFE-5-Reduce* model achieves an out-of-sample test sensitivity of 75.05% and an AUC of 0.8043, further supporting the efficacy of using a slightly larger but still focused set of features for bankruptcy prediction. On the other hand, the *MPLRRFE-3-ReSelect* and *MPLRRFE-5-ReSelect* models engage in a reselection process of the top 3 and 5 features for the global prediction model (i.e., the meta-leaner), respectively, after an initial profile-specific feature reduction phase. The *MPLRRFE-3-ReSelect* model records an out-of-sample test sensitivity of 75.26% and an AUC of 0.8098, showing the highest AUC among the *MPLRRFE* models discussed, and a specificity of 75.18%. This indicates that re-evaluating which features to retain in meta-learning stage, after an initial profile-specific feature reduction, can refine the predictive power of the model by focussing on the most impactful predictors (see **Table 4.12** for the statistical tests of model performance). The *MPLRRFE-5-ReSelect* model maintains this trend with a test sensitivity of 75.32% and an AUC of 0.8029.

Comparing *MPLRRFE-Reduce* and *MPLRRFE-ReSelect* models to earlier *MPLRPCA* and *MPLRRFE* models, it is evident that both approaches offer improvements in out-of-sample test sensitivity and AUC. Models with reselection strategy, in particular, demonstrate a higher performance in balancing sensitivity and specificity, suggesting that a reassessment of feature relevance in meta-learning stage contributes positively to the robustness of prediction models. The enhancement observed in these models, as compared to the benchmark LR and earlier *MPLR* models, underlines the value of adaptive feature selection processes in complex predictive modelling environments, where the selection and reduction of features based on their predictive relevance can lead to more accurate and generalisable models. This approach not only addresses the limitations of overfitting observed in profile-specific models, but also leverages the intrinsic patterns and nuances within the financial data more effectively.

Table 4.11 Results of the experiments of profiling models based on *meta-learning structure* with feature engineering

Results - Mean	Sensitivity (TPR)		Specificity (TNR)		Type-I Error Rate (FPR)		Type-II Error Rate (FNR)		Classification Accuracy (ACC)		Area Under ROC Curve (AUC)
	TPR Train	*TPR Test*	*TNR Train*	*TNR Test*	*Type-I Train*	*Type-I Test*	*Type-II Train*	*Type-II Test*	*Acc Train*	*Acc Test*	*AUC Test*
	Logistic Regression / Benchmark Model										
LR	67.01%	64.72%	71.47%	71.40%	28.53%	28.60%	32.99%	35.28%	71.42%	71.33%	0.7662
	Profiling Models Based on Meta-Learning Structure with Profile-specific/Global Feature Selection										
MPLRRFE-3-Reduce	76.62%	74.55%	75.33%	75.17%	24.67%	24.82%	23.37%	25.45%	75.34%	75.16%	0.8089
MPLRRFE-5-Reduce	78.13%	75.05%	75.48%	75.34%	24.51%	24.66%	21.87%	24.95%	75.52%	75.33%	0.8043
MPLRRFE-3-ReSelect	76.25%	75.26%	75.17%	75.18%	24.74%	24.83%	24.75%	24.28%	75.28%	75.17%	0.8098
MPLRRFE-5-ReSelect	77.94%	75.32%	75.40%	75.35%	24.60%	24.65%	22.06%	24.67%	75.43%	75.35%	0.8029
	Profiling Models Based on Meta-Learning Structure with Profile-specific/Global Feature Engineering										
MPLRPCA/RFE-3-Eng	76.98%	75.06%	75.51%	75.35%	24.49%	24.65%	23.02%	24.94%	75.53%	75.35%	0.8084
MPLRPCA/RFE-5-Eng	78.03%	74.13%	75.83%	75.78%	24.17%	24.22%	21.97%	25.87%	75.85%	75.75%	0.7988
	Profiling Models Based on Meta-Learning Structure with Voting Ensemble										
MPVOTRFE-3-Reduce	77.60%	73.61%	76.49%	76.37%	23.51%	23.62%	22.39%	26.38%	76.50%	76.35%	0.8030
MPVOTRFE-3-ReSelect	79.37%	71.83%	76.95%	76.83%	23.04%	23.16%	20.63%	28.17%	76.98%	76.78%	0.7946
MPVOTPCA/RFE-3-Eng	76.75%	75.00%	75.45%	75.36%	24.55%	24.64%	23.25%	24.99%	75.47%	75.36%	0.8074

Note: *LR* – Logistic Regression; *MPLRPCA-3/5-Reduce* – Meta-Profile Learning Model Using Logistic Regression with Profile-specific RFE Feature Selection (3 or 5 selected features) and Global Feature Reduction; *MPLRPCA-3/5-ReSelect* – Meta-Profile Learning Model Using Logistic Regression with Profile-specific RFE Feature Selection (3 or 5 selected features) and Global Feature Selection; *MPLRPCA/RFE-3/5-Eng* – Meta-Profile Learning Model Using Logistic Regression with Profile-specific RFE Feature Selection (3 or 5 selected features) and

Global PCA/RFE Feature Engineering; *MPVOTRFE-3-Reduce/ReSelect/Eng* – Meta-Profile Learning Model Using Voting Ensemble with Profile-specific RFE Feature Selection (3 selected features) and Global Feature Operations.

Building upon the insights gained from the analysis of the *MPLRRFE-Reduce* and *MPLRRFE-ReSelect* models, we can extend the discussion to include the *MPLRPCA/RFE-Eng* models, which incorporate a combination of both PCA and RFE for the feature engineering of the prediction model. These models aim to further refine the predictive capabilities through more sophisticated feature engineering techniques.

The *MPLRPCA/RFE-3-Eng* model, which uses a combination of PCA and RFE to engineer features and reduces the dimensionality with 3 profile-specific features, shows an out-of-sample test sensitivity of 75.06% and specificity of 75.35%. This model achieves an AUC of 0.8084, illustrating that integrating PCA with RFE not only helps to capture the most significant characteristics within the data, but also retains the predictive quality of the features. This dual approach of feature engineering allows the model to refine the nuanced information from financial profiles, potentially leading to a better understanding of the underlying risk factors. Similarly, the *MPLRPCA/RFE-5-Eng* model, which extends the profile-specific feature set to 5 components, shows a slightly different performance. It achieves an out-of-sample test sensitivity of 74.13% and specificity of 75.78%, with an AUC of 0.7988 (see **Table 4.12** for the statistical tests of model performance). Although this model offers the highest specificity among the discussed models, its sensitivity and AUC are slightly lower compared to the *MPLRPCA/RFE-3-Eng* model. This suggests that increasing the number of features beyond a certain threshold may introduce some complexity that does not necessarily translate to better predictive performance, possibly due to the inclusion of less discriminative features.

The discussion highlights the nuanced trade-offs between feature complexity, model interpretability, and predictive performance. Whilst the *MPLRPCA/RFE* models offer a sophisticated method for extracting and refining features, the *MPLRPCA/RFE-Eng* models demonstrate the potential benefits of a more targeted approach to feature selection. These insights suggest that the choice between these methodologies should consider the specific requirements of the predictive task at hand, including the need for simplicity of the model, the interpretability of the results, and the general predictive performance. This directly answers our research question by illustrating how feature engineering and feature selection play a crucial role in optimising profile-based prediction models for bankruptcy prediction.

Table 4.12 Mann-Whitney U test for profiling models based on *meta-learning structure* methodologies with feature engineering (Comparison against benchmark model LR)

Metric	Group	Model	Mean	p_value (*Significance)
		Comparison against Benchmark Model (LR)		
Test Sensitivity (TPR)	Control	*LR*	64.72%	-
	12	*MPLRRFE-3-Reduce*	74.55%	.0000***

	13	*MPLRRFE-5-Reduce*	75.05%	.0000***
	14	*MPLRRFE-3-ReSelect*	75.26%	.0000***
	15	*MPLRRFE-5-ReSelect*	75.32%	.0000***
	16	*MPLRPCA/RFE-3-Eng*	75.06%	.0000***
	17	*MPLRPCA/RFE-5-Eng*	74.13%	.0000***
	18	*MPVOTRFE-3-Reduce*	73.61%	.0000***
	19	*MPVOTRFE-3- ReSelect*	71.83%	.0000***
	20	*MPVOTPCA/RFE-3-Eng*	75.00%	.0000***
Test Specificity (TRN)	Control	*LR*	71.40%	-
	12	*MPLRRFE-3-Reduce*	75.17%	.0000***
	13	*MPLRRFE-5-Reduce*	75.34%	.0000***
	14	*MPLRRFE-3-ReSelect*	75.18%	.0000***
	15	*MPLRRFE-5-ReSelect*	75.35%	.0000***
	16	*MPLRPCA/RFE-3-Eng*	75.35%	.0000***
	17	*MPLRPCA/RFE-5-Eng*	75.78%	.0000***
	18	*MPVOTRFE-3-Reduce*	76.37%	.0000***
	19	*MPVOTRFE-3- ReSelect*	76.83%	.0000***
	20	*MPVOTPCA/RFE-3-Eng*	75.36%	.0000***
Test Accuracy (ACC)	Control	*LR*	71.33%	-
	12	*MPLRRFE-3-Reduce*	75.16%	.0000***
	13	*MPLRRFE-5-Reduce*	75.33%	.0000***
	14	*MPLRRFE-3-ReSelect*	75.17%	.0000***
	15	*MPLRRFE-5-ReSelect*	75.35%	.0000***
	16	*MPLRPCA/RFE-3-Eng*	75.35%	.0000***
	17	*MPLRPCA/RFE-5-Eng*	75.75%	.0000***
	18	*MPVOTRFE-3-Reduce*	76.35%	.0000***
	19	*MPVOTRFE-3- ReSelect*	76.78%	.0000***
	20	*MPVOTPCA/RFE-3-Eng*	75.36%	.0000***
Test AUC	Control	*LR*	0.7662	-
	12	*MPLRRFE-3-Reduce*	0.8089	.0000***
	13	*MPLRRFE-5-Reduce*	0.8043	.0000***
	14	*MPLRRFE-3-ReSelect*	0.8098	.0000***
	15	*MPLRRFE-5-ReSelect*	0.8029	.0000***
	16	*MPLRPCA/RFE-3-Eng*	0.8084	.0000***
	17	*MPLRPCA/RFE-5-Eng*	0.7988	.0000***
	18	*MPVOTRFE-3-Reduce*	0.7946	.0000***
	19	*MPVOTRFE-3- ReSelect*	0.9015	.0002***
	20	*MPVOTPCA/RFE-3-Eng*	0.8074	.0000***

*p<.05, **p<.01, ***p<.001

Exploring further into the profiling models based on a meta-learning structure, the following experiment results investigate into models that incorporate *voting ensemble* techniques, namely *MPVOTRFE-3-Reduce*, *MPVOTRFE-3-ReSelect*, and *MPVOTPCA/RFE-3-Eng*, which directly addressing our research question 7 regarding how ensemble methods applied to profile-based models compare to traditional models in terms of prediction performance and robustness. These models utilise profile-specific voting ensemble models (with LR and decision trees) to potentially enhance the robustness and accuracy of bankruptcy prediction by combining profile-specific predictions from multiple models.

The *MPVOTRFE-3-Reduce* model employs a voting ensemble approach with reduced features selected via RFE. It achieves an out-of-sample test sensitivity of 73.61% and a specificity of 76.37%, with an AUC of 0.8030. This model reflects a balanced approach to sensitivity and specificity, making it particularly effective in environments where avoiding both false positives and false negatives is crucial. Compared to the benchmark LR model, the *MPVOTRFE-3-Reduce* model shows an improvement in sensitivity by approximately 8.89% (see **Table 4.12** for the statistical tests of model performance). This improvement, coupled with a relatively high AUC, underscores the value of a voting ensemble in enhancing predictive performance over a single-predictor model, directly supporting our research question by demonstrating the enhanced performance and robustness achievable through ensemble methods. On the other hand, the *MPVOTRFE-3-ReSelect* model, which also uses a voting ensemble but with a reselection of features for the global prediction model, demonstrates an out-of-sample test sensitivity of 71.83% and specificity of 76.83%, achieving an AUC of 0.7946. Although this model shows a slight reduction in sensitivity compared to the *MPVOTRFE-3-Reduce* model, its specificity is higher, indicating a better performance in correctly identifying nonbankrupt cases. The drop in AUC compared to *MPVOTRFE-3-Reduce* suggests that whilst reselecting features provides a tailored approach to the model, it may not always lead to improved overall predictive accuracy. This comparison highlights the trade-offs involved in feature selection strategies within ensemble methods, suggesting better decisions are required to balance sensitivity and specificity in bankruptcy prediction models. Lastly, the *MPVOTPCA/RFE-3-Eng* model combines the profile-specific voting ensemble model with a hybrid feature engineering approach that includes both PCA and RFE. This model records an out-of-sample sensitivity of 75.00% and specificity of 75.36%, with an AUC of 0.8074. It slightly beats the other two voting ensemble models in terms of AUC, indicating a better balance between correctly predicting bankrupt and healthy companies.

Compared to the benchmark LR model, all three voting ensemble models exhibit superior performance, especially in sensitivity and prediction accuracy. This improvement highlights the effectiveness of ensemble techniques in dealing with the complexities of bankruptcy prediction, where the combination of multiple predictive signals from financial profiles can mitigate the weaknesses of individual profile-specific models and lead to a more accurate and reliable prediction performance. Integration of voting ensemble methods in the profiling models based on a meta-learning structure demonstrates reasonable benefits, not only improving upon the profile-specific LR model but also illustrating the potential of ensemble strategies in achieving more accurate and robust predictions.

4.7 Conclusion and future research directions

This study contributes to the field of bankruptcy prediction by exploring various profiling methodologies and their integration into predictive modelling using different approaches to incorporating financial profiles into bankruptcy prediction models, yielding valuable insights into their performance and effectiveness. A central focus of this research was to address the challenge of sample heterogeneity, which often undermines the accuracy of prediction models in diverse datasets. The results highlight the significant potential of prediction models that exploit profiling information and financial profiles to enhance the performance of bankruptcy prediction as well as utilising feature transformation and selection techniques.

The key findings indicate that whilst the direct utilisation of financial profiles (PDU) can improve model's performance in detecting bankrupt companies, approaches like profiling with one-hot encoding (POH) and profile ranking models (PR) provide additional benefits in generalisation and predictive power. These methods, which effectively transform and prioritise financial profiles, demonstrate the importance of refining feature processing to enhance model robustness.

The study also examined profile-specific methodologies, such as profile-specific logistic regression (PSLR) and adaptive profile-specific models (APSM), which revealed challenges related to model generalisation and overfitting. These approaches, though theoretically promising, often resulted in lower out-of-sample performance due to the complexity in the modelling and variability of financial profiles. This variability is a direct consequence of sample heterogeneity, underscoring the need for models that not only capture the nuances within individual profiles but also generalise effectively across diverse datasets.

Finally, the analysis of meta-learning structures, including linear and exponential weighting schemes, as well as with the feature engineering/selection methods like PCA and RFE, further emphasised the benefits of feature processing. By effectively addressing heterogeneity within the data, models incorporating these advanced techniques generally outperformed simpler models, highlighting the value of sophisticated feature engineering in bankruptcy prediction.

Overall, this chapter confirms that integrating advanced profiling methodologies into prediction models provides a more comprehensive and accurate approach for bankruptcy prediction. This integration is particularly beneficial for financial institutions and analysts seeking to improve their risk assessment capabilities.

Overall, this chapter reinforces the thesis's aim of enhancing bankruptcy prediction methodologies through the integration of innovative, data-driven techniques. By addressing sample heterogeneity, which is a critical challenge in predictive modelling, this study complements the contributions of **Chapter 2** and **Chapter 3**. It demonstrates how financial profiling can bridge the gap between theoretical advancements in machine learning and the practical requirements of risk assessment in real-world contexts. The findings underscore the value of leveraging financial profiles to improve model interpretability and performance, aligning with the thesis's broader goal of developing adaptable, robust, and interpretable prediction models. Furthermore, this research highlights the potential for integrating profiling methodologies with other innovative techniques, such as network analysis and advanced feature selection, as explored in previous chapters, to provide a comprehensive toolkit for corporate bankruptcy prediction. This chapter's findings solidify its role in advancing the collective contributions of this thesis to the field of bankruptcy and financial distress prediction.

Whilst the empirical results suggest our proposed methodology is promising, there are also several aspects that can be considered for future exploration.

Exploration of Hybrid Profiling Techniques: Although our study identifies *k-means* as the most effective profiling technique, future research could explore hybrid profiling methods that combine the strengths of multiple clustering algorithms. For example, combining *k-means* with density-based clustering methods (e.g., DBSCAN) could help to identify both well-separated and densely packed clusters, improving the detection of subtle financial patterns that might indicate bankruptcy risk. Hybrid methods could also address the limitations of individual algorithms, providing more robust and nuanced financial profiles.

Dynamic Financial Profiling: Our research focuses on static financial profiles based on historical data. Future research could investigate dynamic financial profiling, where profiles evolve over time as companies' financial situations alters. This could involve the use of time-series clustering or dynamic time wrapping techniques to create profiles that reflect not only a company's current state, but also its future financial trajectory. Such an approach could improve the prediction model's ability to predict bankruptcy by considering how financial behaviours develop over time.

Integration of Alternative Drivers: Whilst our study uses financial ratios and corporate governance indicators, future research could explore the integration of alternative drivers, such as sentiment analysis from news, social media, or macroeconomic indicators, such as interest rates and inflation rates. These additional data could capture external factors affecting a company's bankruptcy risk, which might not be evident from financial metrics alone. Moreover, operational and supply chain management (OM/SCM) activities represent critical areas that can significantly influence a company's financial health and stability. Metrics such as *supply chain efficiency* and *supplier reliability* could offer valuable insights into operational risks that often precede financial distress. These additional data could capture external factors affecting a company's bankruptcy risk, which might not be evident from financial metrics alone. Prediction models that incorporate these alternative data sources could provide a more holistic view of bankruptcy risk.

Exploring the Impact of Economic Cycles on Financial Profiles: Given the fact that economic conditions can significantly influence corporate bankruptcy risk, future research could investigate how economic cycles impact the effectiveness of bankruptcy prediction models. This could involve improving financial profiles and prediction models based on different phases of the economic cycle (e.g., recession, recovery).

Cross-Market Validation and Transfer Learning: As our research is focused on the UK market, future studies could test the generalisability of the findings across different markets with varying economic and regulatory environments. This could involve using *transfer learning*, where models trained on UK data are applied to other markets with minimal adjustments. This approach could validate the robustness of the models and highlight any necessary adjustments for different financial environments.

Exploring the Implementation of Managerial Prescriptions: Although the study has provided managerial prescriptions for companies based on the financial profiles to which they belonged, future research could further investigate how predictive insights derived from financial profiles can be translated into actionable managerial strategies. This includes identifying specific prescriptions for companies in different risk profiles, such as operational adjustments, governance reforms, or strategic initiatives. By developing a framework to connect profile-based bankruptcy predictions with tailored managerial actions, researchers could further bridge the gap between theoretical models and practical applications, offering a pathway to mitigate financial distress effectively.

Each mentioned future research direction could offer new avenues for enhancing the prediction performance of the proposed methodology. By addressing these areas, researchers can further improve the concept of profile analysis, refining its ability to capture the nuanced financial patterns that resulting bankruptcy. This, in turn, would contribute to the development of more robust and adaptable profile-based prediction models, ultimately providing valuable tools for financial institutions and policymakers in managing risk and ensuring financial stability.

4.A Appendix

4.A.1 The centroids of financial profiles

Profile	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
1	-0.0545	0.1783	0.0034	0.0577	0.0963	0.0495	0.0640	0.1712	0.1717	0.2251
2	-0.5519	-2.5950	0.9365	-0.2807	-1.3338	-0.7646	-0.5774	-0.3119	-0.4144	-3.0450
3	-0.1133	0.0908	0.1066	0.0454	0.0543	0.0323	0.0022	0.1517	0.1486	0.1055
4	0.6925	-1.5259	-0.1110	-5.6819	-2.8048	-3.0466	-2.0078	-4.9366	-4.8846	-2.5432
5	1.4617	-0.6916	-0.7570	3.9350	1.8569	3.0388	0.6545	-4.6094	-4.5826	-1.4679
Profile	**D11**	**D12**	**D13**	**D14**	**D15**	**D16**	**D17**	**D18**	**D19**	
1	0.2111	0.0936	-0.0351	0.2112	0.2065	-0.0422	-0.6015	0.0930	0.1830	
2	-3.2189	-0.9247	0.1615	-3.0756	-1.7304	-0.1520	0.1131	-0.4098	-0.2360	
3	0.1108	0.0485	0.0274	0.1238	0.1554	-0.0171	1.6553	-0.1955	-0.5066	
4	-2.1697	-1.4454	0.1528	-2.4082	-3.1297	-0.3056	0.1293	-0.2795	0.0168	

5	-1.1539	-0.2355	0.2434	-1.3761	-3.1065	0.5553	-0.0056	-0.3107	-0.0899

Note: *D1* – Working Capital/Total Assets; *D2* – Retained Earnings/Total Assets; *D3* – Current Liabilities/Current Assets; *D4* – Cashflow/Sales; *D5* – Cashflow/Total Assets; *D6* – Cashflow/Total Liabilities; *D7* – Cashflow/Book Value of Equity; *D8* – EBIT/Sales; *D9* – Net Income/Sales; *D10* – Operating Income/Total Assets; *D11* – Net Income/Total Assets; *D12* – Net Income/Book Value of Equity; *D13* – Net Income Growth Rate; *D14* – EBITDA/Total Assets; *D15* – EBIT/Total Liabilities; *D16* – Market Value of Equity Growth Rate; *D17* – CEO/Chairman Duality; *D18* – No. directors; *D19* – Percentage of NED;

Chapter 5

Discussion and General Conclusion

5.1 The evolution of bankruptcy prediction research

The comprehensive review and critical analysis conducted in **Chapter 2** highlights the diverse and evolving nature of this field. One of the critical findings from this review is the increasing emphasis on advanced machine learning and artificial intelligence models, such as ensemble learning techniques, which have significantly enhanced prediction accuracy. However, despite these advancements, many evidences suggest that no single methodology has emerged as superior, with each approach having its strengths and weaknesses, often influenced by design and implementation decisions. This phenomenon underscores the importance of continued innovation and diversification in developing prediction models and methodologies.

There is a growing need for models that are not only accurate but also interpretable, so that explainable AI techniques should be integrated with current methodologies to enhance the model's transparency and trustworthiness, enabling stakeholders to understand the rationale behind predictions. This trend towards interpretability is driven by increasing regulatory requirements for financial institutions, such as those outlined in the European Union's General Data Protection Regulation (GDPR) and the Basel III framework, which demand transparency and accountability in automated decision-making processes (European Commission, 2016; Basel Committee, 2010). Additionally, the ethics of financing activities necessitates that predictive models do not operate as 'black boxes' but provide clear, understandable insights into their decision-making processes, thus promoting customer protection and trust. Ethical AI in finance is crucial to prevent biases and ensure fair treatment of all customers, including companies, aligning with principles such as those advocated by the IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems (IEEE, 2019). By incorporating explainable AI into corporate bankruptcy and financial distress prediction, financial institutions can better meet these regulatory and ethical standards, protect customer interests, and maintain the integrity of financial markets. This integration ensures that all stakeholders, including regulators, auditors, and customers, can trust and verify the predictions, thereby enhancing the overall reliability and acceptance of advanced predictive models in the financial industry.

Future research in corporate bankruptcy and financial distress prediction should also focus on addressing other multiple challenges which are crucial for the general implementation of the methodologies, such as data imbalance and heterogeneity issues, cross-industry and cross-country implementation, the impact of shocks and financial crises, and the benefits for SMEs and developing economies. More comprehensive approaches, integrating advanced methodologies and addressing critical challenges, will significantly advance the field of bankruptcy and financial distress prediction, making the research invaluable for academia and industry.

5.2 Enhancing bankruptcy prediction with network analysis

The study in **Chapter 3** introduces a novel approach to bankruptcy prediction by incorporating complex network analysis focused on company relationships. This research demonstrates that integrating company relational information-based drivers with traditional financial ratios and corporate governance indicators can significantly enhance prediction accuracy. Key findings indicate that whilst financial ratios provide a solid foundation, their effectiveness is notably enhanced when combined with company relational information (CRIs) derived from board-of-director connections networks.

As businesses become increasingly interconnected, understanding the interconnectedness of businesses is crucial for improving bankruptcy and financial distress prediction. The global economy is characterised by intricate networks of corporate relationships, from supply chains and strategic partnerships to shared board members and ownership structures. Researchers should leverage this rich relational data to develop more sophisticated predictive models. Future research could focus on several key areas to advance the use of complex network analysis in bankruptcy and financial distress prediction. First, there is a need for more comprehensive datasets that capture the full spectrum of corporate relationships. This includes not only direct connections but also indirect ties that might influence a company’s risk profile. Second, network models that dynamically adapt to the temporal evolution of corporate relationships could provide more accurate representations of changing inter-company dependencies, enabling models to better reflect the shifting nature of governance structures, partnerships, and financial risks. Additionally, integrating other types of relational data, such as inter-firm transactions and social media interactions, could further enhance the predictive power of these models. However, challenges remain in terms of the privacy of data and the computational complexity of analysing large, dynamic networks.

By highlighting the importance of company relational information, our work opens new avenues for both theoretical exploration and practical application in financial risk management. It provides a framework for future studies to build upon, encouraging the development of more holistic and robust predictive methodologies. It underscores the potential for these advanced models to improve risk management strategies and regulatory policies, ultimately, contributing to greater financial stability and resilience in an interconnected world.

5.3 Advancements in financial profiling for bankruptcy prediction

Chapter 4 introduces an advancement in bankruptcy prediction by utilising a two-step methodology based on financial profiling. This approach departs from traditional prediction models by integrating financial profiles with machine learning techniques, offering a more nuanced understanding of corporate financial health. A critical aspect of this research is its focus on enhancing model interpretability without sacrificing predictive accuracy. By generating financial profiles that group companies into distinct categories based on shared financial characteristics, this methodology offers stakeholders a better understanding of the factors contributing to a company's financial risk. This aligns with the growing demand for explainable AI, particularly in the financial sector, where transparency is crucial for compliance and integrity. The ability to analyse the prediction back through specific financial profiles adds a layer of credibility to the models, making it a reliable tool for financial decision-making.

Looking forward, the potential of profile-based bankruptcy prediction approach can be further expanded into several directions. One promising direction is the development of hybrid profiling techniques that combine different profiling methods to capture both distinct and overlapping financial patterns. For instance, the integration of *k-means* with *hierarchical clustering*, or other density-based methods like DBSCAN could provide a more granular understanding of companies' financial landscape, allowing for the identification of more complex, multi-layered risk structures. Additionally, the exploration of dynamic financial profiling, where profiles evolve in accordance with changes in a company's financial condition, could offer some more interesting insights. By following how financial profiles evolves over time, prediction models could anticipate bankruptcy risks earlier, offering stakeholders valuable opportunities to take preventative action.

The advancements in financial profiling presented in this study represent a significant leap forward in the field of bankruptcy prediction. By addressing the issue of sample heterogeneity and integrating diverse data sources, the proposed methodology offers a more comprehensive and interpretable approach to risk assessment. Future research could continue refining these techniques, exploring the potential of the proposed methodology.

5.4 General conclusions

This thesis presents three comprehensive studies on the prediction of corporate bankruptcy and financial distress, with each contributing unique insights and methodologies to this vital field. Whilst each chapter focuses on distinct aspects, they collectively address the overarching research questions and advance the field through innovative approaches, rigorous analysis, and empirical validation. Together, they highlight the importance of integrating traditional and modern prediction techniques to develop more robust, interpretable, and effective bankruptcy prediction models.

Chapter 2 offers an up-to-date review, classification, and critical analysis of the extensive literature on corporate bankruptcy and financial distress prediction. Six major research streams are identified and discussed, namely, the definition and coding of bankruptcy and financial distress, the design of new prediction models/classifiers or the application of existing ones, the development and evaluation of new drivers, feature selection methods, performance evaluation methodologies for prediction models, and issues affecting model performance along with proposed solutions. This chapter establishes a comprehensive foundation for the field, as the comprehensive survey not only paints the current landscape of research but also serves as a valuable guide for future investigations. Our analysis reveals a significant trend towards advanced machine learning and artificial intelligence models, particularly ensemble learning models, in the prediction of bankruptcy and financial distress. These models, along with other methodological advances, such as DEA- and MCDA-based prediction methods, have further pushed the boundaries of research in this field. Despite these advancements, according to our survey, no single methodology emerges as superior, with each having its strengths and weaknesses influenced by the design and implementation decisions. Our survey underscores the broad research directions in this field, highlighting the importance of exploring various aspects in bankruptcy and financial distress prediction, beyond only developing new prediction models. By synthesising and critically analysing existing research, this chapter ensures that the thesis is firmly grounded in the current landscape of bankruptcy prediction research.

Chapter 3 introduces a novel approach to bankruptcy prediction by integrating complex network analysis with traditional financial ratios and corporate governance indicators. This research, focussing on UK companies listed on the London Stock Exchange, demonstrates that incorporating company relational information (CRI) significantly enhances the accuracy of bankruptcy predictions. The findings suggest that, whilst financial ratios (FRs) provide a solid foundation for predicting bankruptcy, their effectiveness is significantly enhanced when combined with CRIs derived from board-of-director connections networks. These CRIs offer valuable insights into the interconnectedness of companies, which is often overlooked in conventional bankruptcy and financial distress prediction research. Models using a combination of FRs and CRIs outperformed those relying solely on FRs or conventional corporate governance indicators, with company-company direct connection-based CRIs being particularly effective. Additionally, our study explores the impact of different network structures on prediction performance, showing that time-weighted network structures, which consider the evolving nature of company relationships over time, further improve model performance. This suggests that dynamic changes in company networks are crucial for accurately assessing corporate bankruptcy risk. Overall, our work confirms that incorporating network analysis of company relationships into bankruptcy prediction methodologies provides a more comprehensive and accurate approach, offering significant potential for financial institutions and analysts.

Chapter 4 explores various profiling methodologies and their integration into bankruptcy prediction. The empirical analysis focused on different approaches to incorporating financial profiles into bankruptcy prediction models, demonstrating the significant potential of profiling to enhance prediction performance through sophisticated feature transformation and selection techniques. The key findings indicate that whilst direct utilisation of financial profiles can improve the model's sensitivity, more sophisticated approaches, such as profiling with one-hot encoding and profile ranking models offer additional benefits in generalisation and predictive power. These methods effectively transform and prioritise the financial profiles, highlighting the importance of refining feature processing to enhance model robustness. Our research also examines profile-specific methodologies, such as profile-specific logistic regression and adaptive profile-specific models, revealing challenges related to model generalisation and overfitting. These approaches, though theoretically promising, often result in lower out-of-sample performance due to the complexity and variability of financial profiles, suggesting the need for balanced models that generalise well across diverse data sets. Furthermore, the analysis of meta-learning structures and cluster-specific feature engineering methods emphasises the benefits of feature processing. Models that incorporating advanced techniques like principal component analysis and reductive feature elimination generally outperform simpler models, suggesting the value of sophisticated feature engineering in bankruptcy prediction. Our work confirms that integrating advanced profiling and feature selection methodologies into bankruptcy prediction provides a more comprehensive and accurate approach, particularly beneficial for financial institutions and analysts seeking to improve their risk assessment capabilities. By offering a more nuanced understanding of corporate financial health, this chapter aligns with the growing demand for explainable AI in the financial sector, making this methodology both innovative and practical.

Collectively, these studies address the challenges of corporate bankruptcy and financial distress prediction by combining traditional financial analysis with cutting-edge techniques in machine learning, network analysis, and financial profiling. They provide a coherent framework that bridges gaps in the existing literature, enhances model performance, and improves interpretability, thus addressing the needs of academia, industry, and regulatory bodies. By integrating insights from diverse methodologies, the thesis demonstrates how different approaches can complement each other to provide a more comprehensive understanding of bankruptcy risk.

The contributions of this thesis extend beyond academic inquiry to offer practical tools for stakeholders, including financial institutions, regulators, and policymakers. It highlights the importance of balancing methodological sophistication with interpretability and operational feasibility, ensuring that the proposed models are both innovative and applicable in real-world contexts. By addressing key challenges, such as data imbalance, inter-company relationships, sample heterogeneity, and evolving corporate dynamics, this thesis lays a strong foundation for future research and practical advancements in bankruptcy prediction.

5.5 Future research directions

Whilst this thesis has advanced the field of corporate bankruptcy and financial distress prediction through innovative methodologies and robust empirical analysis, there remain several opportunities for future exploration that extend beyond the scope of the three studies presented. These directions aim to refine existing approaches, explore uncharted areas, and address challenges at the intersection of methodology, data, and application.

Integration of Interdisciplinary Data Sources: Future research should further explore the integration of diverse data sources beyond financial and corporate governance indicators, such as supply chain metrics and market sentiment derived from textual and social media analysis. These additional drivers could provide a more holistic view of a company's financial health and potential risks, capturing operational inefficiencies, external shocks, and market perceptions that traditional metrics may overlook. This interdisciplinary approach would enrich the predictive models and improve their robustness.

Dynamic and Adaptive Models for Evolving Contexts: Given the temporal evolution of corporate networks, financial behaviours, and economic conditions, there is a need to develop dynamic and adaptive models that can account for these changes over time. Techniques such as temporal graph networks, dynamic time-series clustering, and adaptive machine learning algorithms could enhance the ability to predict long-term bankruptcy risks. These models would better reflect the reality of shifting governance structures, economic cycles, and inter-company dependencies, providing stakeholders with timely and actionable insights.

Application Across Markets and Industries: The generalisability of the proposed methodologies across different markets and industries remains a hot area for further exploration. Cross-industry and cross-country validations would ensure that the models are not only locally effective but also globally adaptable. This specifically includes investigating the applicability of network-based approaches and financial profiling in emerging economies, which often face unique financial and governance challenges.

Exploring the Ethical Dimensions of Bankruptcy Prediction: As advanced prediction models become more prevalent, future research must address the ethical implications of their use. This includes investigating potential biases in data and algorithms, ensuring fairness across diverse company profiles, and developing guidelines for transparent and explainable AI in financial decision-making. Ethical considerations are particularly relevant in the context of regulatory compliance and customer protection, where the stakes of inaccurate or opaque predictions are high.

Enhancing Managerial Prescriptions and Practical Applications: A natural extension of this thesis would involve exploring how the insights generated by the proposed methodologies can inform managerial decision-making and corporate governance practices. Future studies could investigate ways to translate predictive findings into actionable recommendations for improving financial stability, mitigating bankruptcy risk, and optimising board structures and governance strategies. This could include frameworks for identifying early warning signs and implementing preventive measures based on profile-specific risk assessments.

Scalability and Computational Efficiency: As the complexity of predictive models increases, future research should address the challenges of scalability and computational efficiency, particularly for large datasets and real-time applications. Methods such as distributed computing, approximate kernel computations, and lightweight neural network architectures could help make advanced methodologies more accessible and practical for stakeholders in real-world settings.

Impact of Macro-Economic and Global Shocks: The resilience of prediction models under conditions of macro-economic turbulence, such as Global Financial Crises or global events like the COVID-19 pandemic, remains an open question. Future research could investigate how external shocks affect the accuracy and reliability of bankruptcy predictions, leading to the development of models that can adapt to sudden changes in the economic environment and provide robust predictions under uncertainty.

Collusion and Adverse Network Effects: Whilst the integration of company relational information has demonstrated its predictive power, further investigation is needed into the potential adverse effects of network structures, such as collusive behaviours among board members or governance inefficiencies caused by overly large boards. Future studies could explore ways to identify and account for these issues, ensuring that network-based approaches remain effective and unbiased.

Policy Implications and Stakeholder Engagement: Finally, future research could explore the policy implications of advanced bankruptcy prediction models. This includes developing guidelines for regulatory bodies on the use of predictive analytics in financial supervision, as well as fostering collaboration with stakeholders such as banks, investors, and managers to implement these models effectively. Engagement with policymakers could ensure that prediction models are aligned with broader financial stability goals and public trust.

These thesis-level future directions aim to build upon the foundational work presented in this research, encouraging interdisciplinary, adaptive, and ethical advancements in the field of corporate bankruptcy and financial distress prediction. By addressing these potential areas, future research can further enhance the robustness, applicability, and societal impact of prediction methodologies, contributing to a more stable and resilient global financial system.

Bibliography

Adler-Nissen, R., Galpin, C., & Rosamond, B. (2017). Performing Brexit: How a post-Brexit world is imagined outside the United Kingdom. *The British journal of politics and international relations*, *19(3)*, 573-591.

Afik, Z., Arad, O., & Galil, K. (2016). Using Merton model for default prediction: An empirical assessment of selected alternatives. *Journal of Empirical Finance, 35*, 43-67.

Ahmad, A.H. (2019). What factors discriminate reorganized and delisted distressed firms: Evidence from Malaysia. *Finance Research Letters, 29*, 50-56.

Ahmadi, Z., Martens, P., Koch, C., Gottron, T., & Kramer, S. (2018, October). Towards bankruptcy prediction: Deep sentiment mining to detect financial distress from business management reports. *IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA) (pp. 293-302)*. IEEE.

Ahmed, N.K., Rossi, R., Lee, J.B., Willke, T.L., Zhou, R., Kong, X., & Eldardiry, H. (2018). Learning role-based graph embeddings. *arXiv preprint* arXiv:*1802.02896*.

Ahn, H., & Kim, K.J. (2009). Bankruptcy prediction modeling with hybrid case-based reasoning and genetic algorithms approach. *Applied Soft Computing Journal, 9(2),* 599–607.

Alaka, H.A., Oyedele, L.O., Owolabi, H.A., Bilal, M., Ajayi, S.O., & Akinade, O.O. (2018). A framework for big data analytics approach to failure prediction of construction firms. *Applied Computing and Informatics, 16(1/2),* 207-222.

Alam, T.M., Shaukat, K., Mushtaq, M., Ali, Y., Khushi, M., Luo, S., & Wahab, A. (2021). Corporate bankruptcy prediction: An approach towards better corporate world. *The Computer Journal, 64(11),* 1731-1746.

Alaminos, D., Del Castillo, A., & Fernández, M.Á. (2016). A global model for bankruptcy prediction. *PloS one, 11(11),* e0166693.

Almamy, J., Aston, J., & Ngwa, L.N. (2016). An evaluation of Altman’s Z-score using cash flow ratio to predict corporate failure amid the recent financial crisis: Evidence from the UK. *Journal of Corporate Finance, 36*, 278-285.

Almaskati, N., Bird, R., Yeung, D., & Lu, Y. (2021). A horse race of models and estimation methods for predicting bankruptcy. *Advances in accounting, 52,* 100513.

Alnassar, W.I., & Chin, O.B. (2015). Why banks use credit derivatives? Review paper. *Procedia Economics and Finance, 26*, 566-574.

Altman, E.I. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *Journal of Finance, 23(4),* 589–609.

Altman, E.I., Iwanicz-Drozdowska, M., Laitinen, E.K., & Suvas, A. (2014). Distressed firm and bankruptcy prediction in an international context: A review and empirical analysis of Altman's Z-score model. Available at SSRN: https://ssrn.com/abstract=2536340 or http://dx.doi.org/10.2139/ssrn.2536340

Altman, E.I., Sabato, G., & Wilson, N. (2008). The value of non-financial information in SME risk management. Available at SSRN: https://ssrn.com/abstract=1320612 or http://dx.doi.org/10.2139/ssrn.1320612

Altman, E.I., Sabato, G., & Wilson, N. (2010). The value of non-financial information in small and medium-sized enterprise risk management. *Journal of Credit Risk, 2(6)*, 95–127.

Altman, E.I., Haldeman, R.G., & Narayanan, P. (1977). ZETA TM analysis A new model to identify bankruptcy risk of corporations. *Journal of Banking & Finance, 1(1)*, 29–54.

Amendola, A., Restaino, M., & Sensini, L. (2015). An analysis of the determinants of financial distress in Italy: A competing risks approach. *International Review of Economics & Finance, 37*, 33-41.

Andrikopoulos, P., & Khorasgani, A. (2018). Predicting unlisted SMEs' default: Incorporating market information on accounting-based models for improved accuracy. *The British Accounting Review, 50(5)*, 559-573.

Ankerst, M., Breunig, M.M., Kriegel, H.P., & Sander, J. (1999). OPTICS: Ordering points to identify the clustering structure. *ACM Sigmod Record, 28(2)*, 49-60.

Antunes, F., Ribeiro, B., & Pereira, F. (2017). Probabilistic modeling and visualization for bankruptcy prediction. *Applied Soft Computing, 60*, 831-843.

Appiah, K.O., & Chizema, A. (2016). The impact of board quality and nomination committee on corporate bankruptcy. *Advances in accounting, 35*, 75-81.

Arnold, B. & De Lange, P., (2004). Enron: an examination of agency problems. *Critical perspectives on Accounting, 15(6-7),* 751-765.

Asch, P. & Seneca, J.J. (1975). Characteristics of collusive firms. *The Journal of Industrial Economics, 22(3),* 223-237.

Ashraf, S., Félix, E.G., & Serrasqueiro, Z. (2020). Development and testing of an augmented distress prediction model: A comparative study on a developed and an emerging market. *Journal of Multinational Financial Management, 57*, 100659.

Bai, Q., & Tian, S. (2020). Innovate or die: Corporate innovation and bankruptcy forecasts. *Journal of Empirical Finance, 59,* 88-108.

Baccini, L., Brodeur, A., & Weymouth, S. (2021). The COVID-19 pandemic and the 2020 US presidential election. *Journal of Population Economics, 34(2)*, 739-767.

Barboza, F., Kimura, H., & Altman, E. (2017). Machine learning models and bankruptcy prediction. *Expert Systems with Applications, 83*, 405–417.

Basel Committee. (2010). Basel III: A global regulatory framework for more resilient banks and banking systems. *Basel Committee on Banking Supervision*, Basel. Available at: <https://www.bis.org/publ/bcbs189.htm> [Accessed 14 October 2021].

Bauer, J., & Agarwal, V. (2014). Are hazard models superior to traditional bankruptcy prediction approaches? A comprehensive test. *Journal of Banking & Finance, 40*, 432-442.

Beasley, M.S. (1996). An empirical analysis of the relation between the board of director composition and financial statement fraud. *Accounting Review, 71(4),* 443-465.

Beaver, W.H. (1966). Financial ratios as predictors of failure. *Journal of accounting research, 4*, 71-111.

Bharath, S.T., & Shumway, T. (2008). Forecasting default with the Merton distance to default model. *The Review of Financial Studies, 21(3)*, 1339-1369.

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of political economy, 81(3),* 637-654.

BoardEx. (2022). BoardEx - Company and Organizational Networks, BoardEx. https://boardex.com/ (Accessed 13 Jan 2022).

Boser, B.E., Guyon, I.M., & Vapnik, V.N. (1992, July). *A training algorithm for optimal margin classifiers*. In Proceedings of the fifth annual workshop on Computational learning theory (pp. 144-152).

Bramer, M. (2007). Avoiding overfitting of decision trees. In *Principles of Data Mining* (pp. 119-134). Springer, London.

Breiman, L. (1996). Bagging predictors. *Machine learning, 24(2),* 123-140.

Bryant, S.M. (1997). A case-based reasoning approach to bankruptcy prediction modeling. *Intelligent Systems in Accounting, Finance & Management, 6(3),* 195-214.

Burton, R.E. & Kebler, R.W. (1960). The "half-life" of some scientific and technical literatures. *American Documentation*, *11(1)*, 18-22.

Campbell, J.Y., Hilscher, J., & Szilagyi, J. (2008). In search of distress risk. *The Journal of Finance, 63(6)*, 2899-2939.

Chen, H. J., Huang, S.Y., & Lin, C.S. (2009). Alternative diagnosis of corporate bankruptcy: A neuro fuzzy approach. *Expert Systems with Applications, 36(4)*, 7710–7720.

Chen, H., Yang, B., Wang, G., Liu, J., Xu, X., Wang, S., & Liu, D. (2011a). A novel bankruptcy prediction model based on an adaptive fuzzy k-nearest neighbor method. *Knowledge-Based Systems, 24(8)*, 1348–1359.

Chen, N., Ribeiro, B., Vieira, A.S., Duarte, J., & Neves, J.C. (2011b). A genetic algorithm-based approach to cost-sensitive bankruptcy prediction. *Expert Systems with Applications, 38(10)*, 12939-12945.

Chen, T., & Guestrin, C. (2016, August). *Xgboost: A scalable tree boosting system*. In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining (pp. 785-794).

Chen, Z., Chen, W., & Shi, Y. (2020). Ensemble learning with label proportions for bankruptcy prediction. *Expert Systems with Applications*, *146*, 113155.

Cheng, M.Y., Hoang, N.D., Limanto, L., & Wu, Y.W. (2014). A novel hybrid intelligent approach for contractor default status prediction. *Knowledge-Based Systems*, *71*, 314-321.

Choi, H., Son, H., & Kim, C. (2018). Predicting financial distress of contractors in the construction industry using ensemble learning. *Expert Systems with Applications*, *110*, 1-10.

Chou, C.H., Hsieh, S.C., & Qiu, C.J. (2017). Hybrid genetic algorithm and fuzzy clustering for bankruptcy prediction. *Applied Soft Computing, 56*, 298-316.

Cielen, A., Peeters, L., & Vanhoof, K. (2004). Bankruptcy prediction using a data envelopment analysis. *European Journal of Operational Research, 154(2)*, 526-532.

Cleofas-Sánchez, L., García, V., Marqués, A.I., & Sánchez, J.S. (2016). Financial distress prediction using the hybrid associative memory with translation. *Applied Soft Computing, 44*, 144-152.

Constantin, A., Peltonen, T.A. & Sarlin, P. (2018). Network linkages to predict bank distress. *Journal of Financial Stability*, *35*, 226-241.

Cormier, D., Ledoux, M.J., Magnan, M. & Aerts, W. (2010). Corporate governance and information asymmetry between managers and investors. *Corporate Governance: The International Journal of Business in Society*, *10(5)*, 574-589.

Cornish, E.A., & Fisher, R.A. (1938). Moments and cumulants in the specification of distributions. *Revue de l'Institut international de Statistique*, *5(4),* 307-320.

Darcan, O.N., & Badur, B.Y. (2012). Student profiling on academic performance using cluster analysis. *Journal of e-Learning & Higher Education, 2012, e1-8.*

Darrat, A.F., Gray, S., Park, J.C. & Wu, Y. (2016). Corporate governance and bankruptcy risk. Journal of Accounting, *Auditing & Finance, 31(2),* 163-202.

Dasarathy, B.V., & Sheela, B.V. (1979). A composite classifier system design: Concepts and methodology. *Proceedings of the IEEE, 67(5),* 708-713.

Datastream. (2022). Datastream Macroeconomic Analysis, LSEG. https://www.lseg.com/en/data-analytics/search/datastream-macroeconomic-analysis (Accessed 21 Feb 2022).

Devery, C. (2010). Criminal profiling and criminal investigation. *Journal of Contemporary Criminal Justice, 26(4),* 393-409.

De Andres, J., Lorca, P., de Cos Juez, F.J., & Sánchez-Lasheras, F. (2011). Bankruptcy forecasting: A hybrid approach using Fuzzy c-means clustering and Multivariate Adaptive Regression Splines (MARS). *Expert Systems with Applications, 38(3),* 1866–1875.

De Bock, K.W., Coussement, K., & Lessmann, S. (2020). Cost-sensitive business failure prediction when misclassification costs are uncertain: A heterogeneous ensemble selection approach. *European Journal of Operational Research, 285(2),* 612-630.

Dimitras, A.I., Zanakis, S.H., & Zopounidis, C. (1996). A survey of business failures with an emphasis on prediction methods and industrial applications. *European journal of operational research, 90(3),* 487-513.

Du, X., Li, W., Ruan, S., & Li, L. (2020). CUS-heterogeneous ensemble-based financial distress prediction for imbalanced dataset with ensemble feature selection. *Applied Soft Computing, 97,* 106758.

Du Jardin, P. (2015). Bankruptcy prediction using terminal failure processes. *European Journal of Operational Research, 242(1),* 286-303.

Du Jardin, P. (2016). A two-stage classification technique for bankruptcy prediction. *European Journal of Operational Research, 254(1),* 236-252.

Du Jardin, P. (2017). Dynamics of firm financial evolution and bankruptcy prediction. *Expert Systems with Applications, 75,* 25-43.

Du Jardin, P. (2018). Failure pattern-based ensembles applied to bankruptcy forecasting. *Decision Support Systems, 107,* 64-77.

Du Jardin, P. (2021a). Dynamic self-organizing feature map-based models applied to bankruptcy prediction. *Decision Support Systems, 147*, 113576.

Du Jardin, P. (2021b). Forecasting bankruptcy using biclustering and neural network-based ensembles. *Annals of Operations Research, 299(1),* 531-566.

Du Jardin, P. & Séverin, E. (2011). Predicting corporate bankruptcy using a self-organizing map: An empirical study to improve the forecasting horizon of a financial failure model. *Decision Support Systems*, *51(3),* 701-711.

Edge, M.E., & Sampaio, P.R.F. (2009). A survey of signature based methods for financial fraud detection. *Computers & security*, *28(6),* 381-394.

Ester, M., Kriegel, H.P., Sander, J., & Xu, X. (1996, August). A density-based algorithm for discovering clusters in large spatial databases with noise. In *kdd*, *96(34)*, 226-231.

European Banking Authority. (2018). Implementing Basel III in Europe. [online] Available at: https://www.eba.europa.eu/regulation-and-policy/implementing-basel-iii-europe.

European Commission. (2016). Art. 22 GDPR: Automated individual decision-making, including profiling. *European Parliament and Council of the European Union*. Available at https://gdpr-info.eu/

Fejér-Király, G. (2015). Bankruptcy prediction: A survey on evolution, critiques, and solutions. *Acta Universitatis Sapientiae, Economics and Business*, *3(1)*, 93-108.

Ferraris, V., Bosco, F., Cafiero, G., D'Angelo, E., & Suloyeva, Y. (2013). Defining profiling. Available at SSRN: https://ssrn.com/abstract=2366564 or http://dx.doi.org/10.2139/ssrn.2366564

Fernandes, N. (2020). Economic effects of coronavirus outbreak (COVID-19) on the world economy. Available at SSRN: https://ssrn.com/abstract=3557504 or http://dx.doi.org/10.2139/ssrn.3557504

Fernández-Gámez, M.Á., Soria, J.A.C., Santos, J.A.C., & Alaminos, D. (2020). European country heterogeneity in financial distress prediction: An empirical analysis with macroeconomic and regulatory factors. *Economic Modelling, 88,* 398-407.

Fich, E.M., Slezak & S.L. (2008). Can corporate governance save distressed firms from bankruptcy? An empirical analysis. *Review of Quantitative Finance and Accounting, 30(2),* 225-251.

Fitzpatrick, P.J. (1932). A comparison of the ratios of successful industrial enterprises with those of failed companies. *The Certified Public Accountant*, *12,* (October, November, December) 598-605, 656-662, and 727-731, respectively.

Flath, D. (2000). Japan's business groups. In The Japanese Business and Economic System: History and Prospects for the 21st Century (pp. 281-324). London: Palgrave Macmillan UK.

Freund, Y., Schapire, R., & Abe, N. (1999). A short introduction to boosting. *Journal-Japanese Society for Artificial Intelligence, 14(771-780),* 1612.

Friedman, J.H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics*, 1189-1232.

Frydman, H., Altman, E. I., & Kao, D. L. (1985). Introducing recursive partitioning for financial classification: the case of financial distress. *The journal of finance, 40(1)*, 269-291.

Fuster, G.G., Gutwirth, S., & Ellyne, E. (2010). Profiling in the European Union: A high-risk practice. *INEX policy brief, 10,* 1-12.

Fujiwara, Y. (2006, October). Chain of Firms Bankruptcy. *In 9th Joint International Conference on Information Sciences* (JCIS-06). Atlantis Press.

Fujiwara, Y. (2008). CHAIN of FIRMS'BANKRUPTCY: a macroscopic study of link effect in a production network. *Advances in Complex Systems*, *11(05)*, 703-717.

Geng, R., Bose, I., & Chen, X. (2015). Prediction of financial distress: An empirical study of listed Chinese companies using data mining. *European Journal of Operational Research, 241(1)*, 236-247.

Godsell, J., Diefenbach, T., Clemmow, C., Towill, D., & Christopher, M. (2011). Enabling supply chain segmentation through demand profiling. *International Journal of Physical Distribution & Logistics Management*, *41(3)*, 296-314.

Gordon, M.J. (1971). Towards a theory of financial distress. *The Journal of Finance, 26(2),* 347-356.

Gorgani, M.E., Moradi, M., & Yazdi, H.S. (2010). An Empirical Modeling of Companies Using Support Vector Data Description. *International Journal of Trade, Economics and Finance, 1(2)*, 221.

Grover, A. & Leskovec, J. (2016, August). node2vec: Scalable feature learning for networks. *In Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*, 855-864.

Gupta, J., Gregoriou, A. & Healy, J. (2015). Forecasting bankruptcy for SMEs using hazard function: To what extent does size matter? *Review of Quantitative Finance and Accounting, 45(4),* 845–869.

Gupta, J., & Chaudhry, S. (2019). Mind the tail, or risk to fail. *Journal of Business Research, 99*, 167-185.

Habermann, F., & Fischer, F.B. (2023). Corporate social performance and the likelihood of bankruptcy: evidence from a period of economic upswing. *Journal of Business Ethics*, 182(1), 243-259.

He, X., Cai, D., Shao, Y., Bao, H., & Han, J. (2010). Laplacian regularized Gaussian mixture model for data clustering. *IEEE transactions on knowledge and data engineering, 23(9),* 1406-1418.

Hildebrandt, M., & Backhouse, J. (2005). D7. 2: Descriptive analysis and inventory of profiling practices. *FIDIS Future of Identity in the Information Society*, 116.

Hosaka, T. (2019). Bankruptcy prediction using imaged financial ratios and convolutional neural networks. *Expert systems with Applications, 117,* 287-299.

Howard, J.A. (1954). Collusive behavior. *The Journal of Business*, *27(3),* 196-204.

Hsieh, S.J. (1993). A note on the optimal cutoff point in bankruptcy prediction models. *Journal of Business Finance & Accounting, 20(3),* 457-464.

Htun, H.H., Biehl, M. & Petkov, N. (2023). Survey of feature selection and extraction techniques for stock market prediction. *Financial Innovation*, *9(1)*, 26.

Hu, Y.C. (2009). Bankruptcy prediction using ELECTRE-based single-layer perceptron. *Neurocomputing, 72(13),* 3150–3157.

Hu, Y.C., & Chen, C.J. (2011). A PROMETHEE-based classification method using concordance and discordance relations and its application to bankruptcy prediction. *Information Sciences*, *181(22)*, 4959-4968.

Huang, Y.P., & Yen, M.F. (2019). A new perspective of performance comparison among machine learning algorithms for financial distress prediction. *Applied Soft Computing, 83*, 105663.

Huysmans, J., Baesens, B., Vanthienen, J., Van Gestel, T. (2006). Failure prediction with self organizing maps. *Expert Systems with Applications*, *30(3),* 479-487.

IEEE. (2019). The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. *The Institute of Electrical and Electronics Engineers*. Available at https://standards.ieee.org/industry-connections/ec/autonomous-systems/

Jensen, M.C. (1986). Agency costs of free cash flow, corporate finance, and takeovers. *The American Economic Review*, *76(2),* 323-329.

Jollois, F.X., & Nadif, M. (2004). Identification de blocs homogènes sur des données continues. In EGC. 241-246.

Jones, S., & Wang, T. (2019). Predicting private company failure: A multi-class analysis. Journal of International Financial Markets, *Institutions and Money, 61*, 161-188.

Kamber, M., Winstone, L., Gong, W., Cheng, S., & Han, J. (1997, April). *Basel Committee, 2010). tion and decision tree induction: efficient classification in data mining*. In Proceedings Seventh International Workshop on Research Issues in Data Engineering. High Performance Database Management for Large-Scale Applications (pp. 111-120). IEEE.

Kim, H.J., Jo, N.O., & Shin, K.S. (2016). Optimization of cluster-based evolutionary undersampling for the artificial neural networks in corporate bankruptcy prediction. *Expert systems with applications, 59*, 226-234.

Kim, S.Y. (2018). Predicting hospitality financial distress with ensemble models: the case of US hotels, restaurants, and amusement and recreation. *Service Business, 12(3)*, 483-503.

Kirwan, G., & Power, A. 2012. *The Psychology of Cyber Crime: Concepts and Principles: Concepts and Principles*. IGI Global. https://doi.org/10.4018/978-1-61350-350-8

Knudsen, E.S. & Lien, L.B. (2023). The half-life of knowledge and strategic human capital. *Human Resource Management Review*, *33(4)*, 100989.

Kou, G., Xu, Y., Peng, Y., Shen, F., Chen, Y., Chang, K., & Kou, S. (2021). Bankruptcy prediction for SMEs using transactional data and two-stage multiobjective feature selection. *Decision Support Systems, 140*, 113429.

Koyuncugil, A.S., & Ozgulbas, N. (2008). Financial profiling for detecting operational risk by data mining. *World Academy of Science, Engineering and Technology*, *46*, 6238-6253.

Kramarz, F. & Thesmar, D. (2013). Social networks in the boardroom. *Journal of the European Economic Association*, *11(4)*, 780-807.

Krause, R., Semadeni, M. & Cannella Jr, A.A. (2014). CEO duality: A review and research agenda. *Journal of Management*, *40(1)*, 256-286.

Le, T., Le Son, H., Vo, M.T., Lee, M.Y., & Baik, S.W. (2018). A cluster-based boosting algorithm for bankruptcy prediction in a highly imbalanced dataset. *Symmetry, 10(7)*, 250.

Le, T., Vo, B., Fujita, H., Nguyen, N.T., & Baik, S.W. (2019). A fast and accurate approach for bankruptcy forecasting using squared logistics loss with GPU-based extreme gradient boosting. *Information Sciences, 494*, 294-310.

Lennox, C.S. (1999). The accuracy and incremental information content of audit reports in predicting bankruptcy. *Journal of Business Finance & Accounting, 26(5-6)*, 757-778.

Letizia, E. & Lillo, F. (2019). Corporate payments networks and credit risk rating. *EPJ Data Science*, *8(1)*, 21.

Li, H., Adeli, H., Sun, J. & Han, J. (2011). Hybridizing principles of TOPSIS with case-based reasoning for business failure prediction. *Computers & Operations Research, 38(2),* 409-419.

Li, H., Sun, J. (2009). Predicting business failure using multiple case-based reasoning combined with support vector machine. *Expert Systems with Applications, 36(6),* 10085–10096.

Li, H., Sun, J. & Sun, B. (2009). Financial distress prediction based on OR-CBR in the principle of k-nearest neighbors. *Expert Systems with Applications, 36(1),* 643–659.

Li, L., & Faff, R. (2019). Predicting corporate bankruptcy: What matters?. *International Review of Economics & Finance, 62*, 1-19.

Li, L., Niu, Z., Li, H., Griffiths, M.D., Mei, S., Jiang, H., Deng, Z., & Xin, J. (2022). Latent profile analysis of mental health among Chinese healthcare staff during the COVID-19 pandemic. *Psychology, Health & Medicine, 27(9),* 1963-1976.

Li, S.T., & Ho, H.F. (2009). Predicting financial activity with evolutionary fuzzy case-based reasoning. *Expert Systems with Applications, 36(1),* 411–422.

Li, Z., Crook, J., Andreeva, G., & Tang, Y. (2021). Predicting the risk of financial distress using corporate governance measures. *Pacific-Basin Finance Journal, 68,* 101334.

Lian, Y. (2017). Financial distress and customer-supplier relationships. *Journal of Corporate Finance, 43,* 397-406.

Liang, D., Lu, C., Tsai, C., & Shih, G. (2016). Financial ratios and corporate governance indicators in bankruptcy prediction: A comprehensive study. *European Journal of Operational Research, 252(2),* 561–572.

Liang, D., Tsai, C.F., Lu, H.Y.R., & Chang, L.S. (2020). Combining corporate governance indicators with stacking ensembles for financial distress prediction. *Journal of Business Research, 120,* 137-146.

Liang, D., Tsai, C.F., & Wu, H.T. (2015). The effect of feature selection on financial distress prediction. *Knowledge-Based Systems, 73,* 289-297.

Liao, L., Kang, H., Morris, R.D. & Tang, Q., (2013). Information asymmetry of fair value accounting during the financial crisis. *Journal of Contemporary Accounting & Economics, 9(2),* 221-236.

Lin, F.Y., Liang, D. & Chu, W.S. (2010). The role of non-financial features related to corporate governance in business crisis prediction. *Journal of Marine Science and Technology, 18(4),* 504-513.

Lin, F., Liang, D., & Chen, E. (2011a). Financial ratio selection for business crisis prediction. *Expert Systems with Applications, 38(12),* 15094-15102.

Lin, F., Liang, D., Yeh, C.C., & Huang, J.C. (2014). Novel feature selection methods to financial distress prediction. *Expert Systems with Applications, 41(5)*, 2472-2483.

Lin, F., Yeh, C.C., & Lee, M.Y. (2011b). The use of hybrid manifold learning and support vector machines in the prediction of business failure. *Knowledge-Based Systems, 24(1)*, 95–101.

Lin, K.C., & Dong, X. (2018). Corporate social responsibility engagement of financially distressed firms and their bankruptcy likelihood. *Advances in accounting, 43*, 32-45.

Lin, W.C., Lu, Y.H., & Tsai, C.F. (2019). Feature selection in single and ensemble learning-based bankruptcy prediction models. *Expert Systems, 36(1)*, e12335.

Lin, Y.K., Chen, H., Brown, R.A., Li, S.H., & Yang, H.J. (2017). Healthcare Predictive Analytics for Risk Profiling in Chronic Care. *Mis Quarterly*, *41(2)*, 473-496.

Lippmann, R. (1987). An introduction to computing with neural nets. *IEEE Assp magazine, 4(2)*, 4-22.

London Share Price Database. (2022). London Share Price lspm201912 & lspd201912 Reference Manual, Wharton Research Data Service. https://wrds-www.wharton.upenn.edu/pages/about/data-vendors/vendor-partner-lspd-london-business-school/ (Accessed 18 Jan 2022).

Long, J., Jiang, C., Dimitrov, S. & Wang, Z. (2022). Clues from networks: quantifying relational risk for credit risk evaluation of SMEs. *Financial Innovation*, *8(1)*, 91.

Lu, Y. (2019). Artificial intelligence: a survey on evolution, models, applications and future trends. *Journal of Management Analytics*, *6(1)*, 1-29.

Lu, Y.C., Shen, C.H., & Wei, Y.C. (2013). Revisiting early warning signals of corporate credit default using linguistic analysis. *Pacific-Basin Finance Journal, 24*, 1-21.

Lu, Y.C., Wei, Y.C., & Chang, T.Y. (2015). The effects and applicability of financial media reports on corporate default ratings. *International Review of Economics & Finance, 36*, 69-87.

Macqueen, J. (1967). Some methods for classification and analysis of multivariate observations. *University of California Press*. *1*, 281-297.

Macskassy, S.A. & Provost, F. (2007). Classification in networked data: A toolkit and a univariate case study. *Journal of Machine Learning Research, 8(5)*, 935-983.

Mai, F., Tian, S., Lee, C., & Ma, L. (2019). Deep learning models for bankruptcy prediction using textual disclosures. *European Journal of Operational Research, 274(2)*, 743-758.

Matin, R., Hansen, C., Hansen, C., & Mølgaard, P. (2019). Predicting distresses using deep learning of text segments in annual reports. *Expert Systems with Applications, 132*, 199-208.

Merton, R.C. (1974). On the pricing of corporate debt: The risk structure of interest rates. *The Journal of finance, 29(2)*, 449-470.

Miao, H., Ramchander, S., Ryan, P., & Wang, T. (2018). Default prediction models: The role of forward-looking measures of returns and volatility. *Journal of Empirical Finance, 46*, 146-162.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

Min, J.H., & Jeong, C. (2009). A binary classification method for bankruptcy prediction. *Expert Systems with Applications, 36(3)*, 5256-5263.

Mizruchi, M.S. (1996). What do interlocks do? An analysis, critique, and assessment of research on interlocking directorates. *Annual Review of Sociology, 22(1)*, 271-298.

Moradi, M., Salehi, M., Ghorgani, M.E., & Yazdi, H.S. (2013). Financial distress prediction of Iranian companies using data mining techniques. *Organizacija, 46(1)*, 20-27.

Mousavi, M.M., Ouenniche, J., & Xu, B. (2015). Performance evaluation of bankruptcy prediction models: An orientation-free super-efficiency DEA-based framework. *International Review of Financial Analysis, 42*, 64-75.

Mousavi, M.M., & Ouenniche, J. (2018). Multi-criteria ranking of corporate distress prediction models: empirical evaluation and methodological contributions, *Annals of Operations Research, 271(2)*, 853-886.

Mousavi, M.M., Ouenniche, J., & Tone, K. (2019). A comparative analysis of two-stage distress prediction models. *Expert Systems with Applications, 119*, 322-341.

Mousavi, M.M., Ouenniche, J., & Tone, K. (2023). A dynamic performance evaluation of distress prediction models. Journal of Forecasting, *42(4)*, 756-784.

Mousavi, M.M., & Lin, J. (2020). The application of PROMETHEE Multi-criteria Decision Aid in Financial Decision Making: Case of Distress Prediction Models Evaluation. *Expert Systems with Applications, 159*, 113438.

Muñoz-Izquierdo, N., Camacho-Miñano, M.D.M., Segovia-Vargas, M.J., & Pascual-Ezama, D. (2019). Is the external audit report useful for bankruptcy prediction? Evidence using artificial intelligence. *International Journal of Financial Studies, 7(2)*, 20.

Myles, A.J., Feudale, R.N., Liu, Y., Woody, N.A., & Brown, S.D. (2004). An introduction to decision tree modeling. *Journal of Chemometrics, 18(6)*, 275-285.

Nanni, L., & Lumini, A. (2009). An experimental comparison of ensemble of classifiers for bankruptcy prediction and credit scoring. *Expert systems with applications, 36(2)*, 3028-3033.

Ng, A., Jordan, M., & Weiss, Y. (2001). On spectral clustering: Analysis and an algorithm. *Advances in neural information processing systems*, *14*.

Ninh, B.P.V., Do Thanh, T., & Hong, D.V. (2018). Financial distress and bankruptcy prediction: An appropriate model for listed firms in Vietnam. *Economic Systems, 42(4)*, 616-624.

Nyitrai, T., & Virág, M. (2019). The effects of handling outliers on the performance of bankruptcy prediction models. *Socio-Economic Planning Sciences, 67*, 34-42.

Odom, M.D., & Sharda, R. (1990). *A neural network model for bankruptcy prediction*. In 1990 IJCNN International Joint Conference on neural networks (pp. 163-168). IEEE.

Ohlson, J. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research*, *18(1)*, 109-131.

Ouenniche, J., & Tone, K. (2017). An Out-of-Sample Evaluation Framework for DEA with Application in Bankruptcy Prediction. *Annals of Operations Research, 254(1-2)*, 235-250.

Ouenniche, J., Uvalle, O., & Ettouhami, A. (2018a). A New Classifier based on EDAS with Application in Distress and Bankruptcy Prediction. *Management Decision, 57(2)*, 314-323.

Ouenniche, J., Bouslah, B., Cabello, J.M., & Ruiz, F. (2018b). A New Classifier based on The Reference Point Method with Application in Bankruptcy Prediction. *Journal of the Operational Research Society*, *69(10)*, 1-8.

Ouenniche, J., Pérez-Gladish, B., & Bouslah, B. (2018c). An Out-of-Sample Framework for TOPSIS-based Classifiers with Application in Bankruptcy Prediction. *Technological Forecasting and Social Change*, *131*, 111-116.

Ouenniche, J., Pérez-Gladish, B., Bouslah, K., & Xu, B. (2019). A new VIKOR-based in-sample-out-of-sample classifier with application in bankruptcy prediction. *Annals of Operations Research*, *296(1)*, 495-512.

Ouenniche, J., Xu, B., & Tone, K. (2014a). Forecasting Models Evaluation using A Slacks-based Context-dependent DEA Framework, *Journal of Applied Business Research, 30 (5)*, 1477-1484.

Ouenniche, J., Xu, B., & Tone, K. (2014b). Relative performance evaluation of competing crude oil prices' volatility forecasting models: A Slacks-Based Super-Efficiency DEA Model, *American Journal of Operations Research, 4 (4)*, 235-245.

Oz, I.O., & Simga-Mugan, C. (2018). Bankruptcy prediction models' generalizability: Evidence from emerging market economies. *Advances in accounting, 41*, 114-125.

Ozgulbas, N., & Koyuncugil, A.S. (2006). A pplication of data mining method for financial profiling. *Social Responsibility Journal*, *2(3/4)*, 328-334.

Paradi, J.C., Asmild, M., & Simak, P.C. (2004). Using DEA and worst practice DEA in credit risk evaluation. *Journal of productivity analysis*, *21(2)*, 153-165.

Paraschiv, F., Schmid, M., & Wahlstrøm, R.R. (2021). Bankruptcy Prediction of Privately Held SMEs Using Feature Selection Methods. Available at SSRN: https://ssrn.com/abstract=3911490 or http://dx.doi.org/10.2139/ssrn.3911490

Pereira, J.M., Basto, M., & da Silva, A.F. (2016). The logistic lasso and ridge regression in predicting corporate failure. *Procedia Economics and Finance, 39*, 634-641.

Perozzi, B., Al-Rfou, R. & Skiena, S. (2014, August). Deepwalk: Online learning of social representations. *In Proceedings of the 20th ACM SIGKDD international conference on Knowledge Discovery and Data Mining*, 701-710.

Pinches, G.E., & Mingo, K.A. (1973). A multivariate analysis of industrial bond ratings. *The journal of Finance, 28(1)*, 1-18.

Piñeiro-Sánchez, C., de Llano Monelos, P., & Rodríguez López, M. (2013). A parsimonious model to forecast financial distress, based on audit evidence. *Contaduría y administración, 58(4)*, 151-173.

Platt, H. & Platt, M. (2012). Corporate board attributes and bankruptcy. *Journal of Business Research*, *65(8)*, 1139-1143.

Politou, E., Alepis, E., & Patsakis, C. (2019). Profiling tax and financial behaviour with big data under the GDPR. *Computer law & security review*, *35(3)*, 306-329.

Poon, S.H., Rockinger, M. & Tawn, J. (2004). Extreme value dependence in financial markets: Diagnostics, models, and financial implications. *The Review of Financial Studies*, *17(2)*, 581-610.

Pramanik, M.I., Lau, R.Y., Yue, W.T., Ye, Y., & Li, C. (2017). Big data analytics for security and criminal investigations. *Wiley interdisciplinary reviews: data mining and knowledge discovery*, *7(4)*, e1208.

Premachandra, I.M., Bhabra, G.S. & Sueyoshi, T. (2009). DEA as a tool for bankruptcy assessment: A comparative study with logistic regression technique. *European Journal of Operational Research, 193(2),* 412–424.

Putra, S.G.P., Joshi, B., Redi, J., & Bozzon, A. (2020, June). *A credit scoring model for smes based on social media data*. In International Conference on Web Engineering (pp. 113-129). Springer, Cham.

Qiu, W., Rudkin, S., & Dłotko, P. (2020). Refining Understanding of Corporate Failure through a Topological Data Analysis Mapping of Altman's Z-Score Model. *Expert Systems with Applications*, *156*, 113475.

Rafae, A., & Erritali, M. (2023). Using a profiling system to recommend employees to carry out a project. *Electronics, 12(16)*, 3388.

Raheja, C.G. (2005). Determinants of board size and composition: A theory of corporate boards. *Journal of Financial and Quantitative Analysis, 40(2)*, 283-306.

Rajput, D., Wang, W.J., & Chen, C.C. (2023). Evaluation of a decided sample size in machine learning applications. *BMC bioinformatics, 24(1),* 48.

Sagi, O., & Rokach, L. (2018). Ensemble learning: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 8(4)*, e1249.

Salvatore, D. (2020). Growth and trade in the United States and the world economy: Overview. *Journal of policy modeling, 42(4)*, 750-759.

Sarkar, S., & Sriram, R.S. (2001). Bayesian models for early warning of bank failures. *Management Science, 47(11)*, 1457-1475.

Sartori, F., Mazzucchelli, A., & Di Gregorio, A. (2016). Bankruptcy forecasting using case-based reasoning: The CRePERIE approach. *Expert Systems with Applications, 64*, 400-411.

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural networks, 61*, 85-117.

Serrano-Cinca, C., Gutiérrez-Nieto, B., & Bernate-Valbuena, M. (2019). The use of accounting anomalies indicators to predict business failure. *European Management Journal, 37(3)*, 353-375.

Seliya, N., Khoshgoftaar, T.M., & Van Hulse, J. (2009, November). *A study on the relationships of classifier performance metrics*. In 2009 21st IEEE international conference on tools with artificial intelligence (pp. 59-66). IEEE.

Shen, F., Liu, Y., Wang, R., & Zhou, W. (2020). A dynamic financial distress forecast model with multiple forecast results under unbalanced data environment. *Knowledge-Based Systems, 192*, 105365.

Shiers, D., Rapson, D., Roberts, C., & Keeping, M. (2006). Sustainable construction: the development and evaluation of an environmental profiling system for construction products. *Construction management and economics, 24(11)*, 1177-1184.

Shin, K.S., Lee, T.S., & Kim, H.J. (2005). An application of support vector machines in bankruptcy prediction model. *Expert systems with applications, 28(1)*, 127-135.

Shumway, T. (2001). Forecasting bankruptcy more accurately: A simple hazard model. *The journal of business, 74(1)*, 101-124.

Somervuori, O., 2014. Profiling behavioral pricing research in marketing. *Journal of Product & Brand Management, 23(6),* 462-474.

Son, H., Hyun, C., Phan, D., & Hwang, H.J. (2019). Data analytic approach for bankruptcy prediction. *Expert Systems with Applications, 138,* 112816.

Stankova, M., Martens, D. & Provost, F. (2015). Classification over bipartite graphs through projection. Research Paper 2015-001. *Antwerp: University of Antwerp, Faculty of Applied Economics, Department of Engineering Management*.

Sueyoshi, T., & Goto, M. (2009a). Methodological comparison between DEA (data envelopment analysis) and DEA–DA (discriminant analysis) from the perspective of bankruptcy assessment. *European Journal of Operational Research, 199(2),* 561-575.

Sueyoshi, T., & Goto, M. (2009b). DEA–DA for bankruptcy-based performance assessment: Misclassification analysis of Japanese construction industry. *European Journal of Operational Research, 199*(2), 576-594.

Sun, J., Fujita, H., Chen, P., & Li, H. (2017). Dynamic financial distress prediction with concept drift based on time weighting combined with Adaboost support vector machine ensemble. *Knowledge-Based Systems, 120,* 4-14.

Sun, J., Li, H., Fujita, H., Fu, B., Ai, W. (2020). Class-imbalanced dynamic financial distress prediction based on Adaboost-SVM ensemble combined with SMOTE and time weighting. *Information Fusion, 54*, 128-144.

Sun, L., & Shenoy, P.P. (2007). Using Bayesian networks for bankruptcy prediction: Some methodological issues. *European Journal of Operational Research, 180(2),* 738-753.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). *Going deeper with convolutions*. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1-9).

Taffler, R.J. (1984). Empirical models for the monitoring of UK corporations. *Journal of Banking & Finance, 8(2)*, 199-227.

Tax, D. M., & Duin, R. P. (2004). Support vector data description. *Machine learning, 54,* 45–66.

Tesler, L.G. (1991). Networked computing in the 1990s. *Scientific American, 265(3),* 86-93.

Tian, S., & Yu, Y. (2017). Financial ratios and bankruptcy predictions: An international evidence. *International Review of Economics & Finance, 51*, 510-526.

Tian, S., Yu, Y., & Guo, H. (2015). Variable selection and corporate bankruptcy forecasts. *Journal of Banking & Finance, 52*, 89-100.

Tinoco, M.H., & Wilson, N. (2013). Financial Distress and Bankruptcy Prediction among Listed Companies Using Accounting, Market and Macroeconomic Variables. *International Review of Financial Analysis, 30*, 394.

Tinoco, M.H., Holmes, P., & Wilson, N. (2018). Polytomous response financial distress models: The role of accounting, market and macroeconomic variables. *International Review of Financial Analysis, 59*, 276-289.

Tobback, E., Moeyersoms, J., Stankova, M., & Martens, D. (2017). Bankruptcy prediction for SMEs using relational data. *Decision Support Systems, 102*, 69–81.

Trabelsi, S., He, R., He, L., & Kusy, M. (2015). A comparison of Bayesian, Hazard, and Mixed Logit model of bankruptcy prediction. *Computational Management Science, 12(1)*, 81–97.

Tsai, C.F. (2009). Feature selection in bankruptcy prediction. *Knowledge-Based Systems, 22(2)*, 120-127.

Tsai, C.F. (2014). Combining cluster analysis with classifier ensembles to predict financial distress. *Information Fusion, 16*, 46-58.

Tsai, C.F., & Cheng, K.C. (2012). Simple instance selection for bankruptcy prediction. *Knowledge-Based Systems, 27,* 333-342.

Tsai, C., & Hsu, Y. (2013). A meta-learning framework for bankruptcy prediction. *Journal of forecasting, 32(2)*, 167–179.

Tseng, F., & Hu, Y. (2010). Comparing four bankruptcy prediction models: Logit, quadratic interval logit, neural and fuzzy neural networks. *Expert Systems with Applications, 37(3),* 1846–1853.

Tsukuda, J., & Baba, S.I. (1994). Predicting Japanese corporate bankruptcy in terms of financial data using neural network. *Computers & Industrial Engineering, 27(1-4)*, 445-448.

Uthayakumar, J., Metawa, N., Shankar, K., & Lakshmanaprabu, S.K. (2020). Financial crisis prediction model using ant colony optimization. *International Journal of Information Management, 50,* 538-556.

Valencia, C., Cabrales, S., Garcia, L., Ramirez, J., & Calderona, D. (2019). Generalized additive model with embedded variable selection for bankruptcy prediction: Prediction versus interpretation. *Cogent Economics & Finance, 7(1)*, 1597956.

Van Otterlo, M. (2013). A machine learning view on profiling. In *Privacy, Due Process and the Computational Turn*. 41-64. Routledge.

Vatousios, A., & Happonen, A. (2022). Transforming HR and improving talent profiling with qualitative analysis digitalization on candidates for career and team development efforts. *In Intelligent Computing: Proceedings of the 2021 Computing Conference, 1*, 1149-1166. Springer International Publishing.

Veganzones, D., & Séverin, E. (2018). An investigation of bankruptcy prediction in imbalanced datasets. *Decision Support Systems, 112*, 111–124.

Volkov, A., Benoit, D.F., & Van den Poel, D. (2017). Incorporating sequential information in bankruptcy prediction with predictors based on Markov for discrimination. *Decision Support Systems, 98*, 59-68.

Wang, G., Chen, G., & Chu, Y. (2018). A new random subspace method incorporating sentiment and textual information for financial distress prediction. *Electronic Commerce Research and Applications, 29*, 30-49.

Wang, G., Ma, J., & Yang, S. (2014). An improved boosting based on feature selection for corporate bankruptcy prediction. *Expert Systems with Applications, 41(5)*, 2353-2361.

Wang, G., Ma, J., Chen, G., & Yang, Y. (2020). Financial distress prediction: Regularized sparse-based Random Subspace with ER aggregation rule incorporating textual disclosures. *Applied Soft Computing, 90*, 106152.

Wang, L., & Wu, C. (2017). Business failure prediction based on two-stage selective ensemble with manifold learning algorithm and kernel-based fuzzy self-organizing map. *Knowledge-Based Systems, 121*, 99-110.

Ward Jr, J.H., 1963. Hierarchical grouping to optimize an objective function. *Journal of the American statistical association, 58(301)*, 236-244.

Warikoo, A. (2014). Proposed methodology for cyber criminal profiling. *Information Security Journal: A Global Perspective, 23(4-6)*, 172-178.

Wedel, M., & Kannan, P.K. (2016). Marketing analytics for data-rich environments. *Journal of marketing, 80(6)*, 97-121.

Weiss, L.A. (1990). Bankruptcy resolution: Direct costs and violation of priority of claims. *Journal of Financial Economics, 27(2)*, 285-314.

West, D., Dellana, S., & Qian, J. (2005). Neural network ensemble strategies for financial decision applications. *Computers & operations research, 32(10)*, 2543-2559.

Wolpert, D.H. (1992). Stacked generalization. *Neural networks, 5(2)*, 241-259.

Xu, B., & Ouenniche, J. (2012a). Performance Evaluation of Competing Forecasting Models – A Multidimensional Framework based on Multi-Criteria Decision Analysis, *Expert Systems with Applications, 39 (9),* 8312-8324.
Xu, B., & Ouenniche, J. (2012b). A Data Envelopment Analysis-based Framework for The Relative Performance Evaluation of Competing Crude Oil Prices' Volatility Forecasting Models, *Energy Economics, 34(2),* 576-583.
Xu, D. & Tian, Y. (2015). A comprehensive survey of clustering algorithms. *Annals of Data Science*, *2*, 165-193.
Xu, M., Zhang, C. (2009). Bankruptcy prediction: the case of Japanese listed companies. *Review of Accounting Studies, 14(4),* 534–558.
Yeh, C., Chi, D., & Hsu, F. (2010). A hybrid approach of DEA, rough set and support vector machines for business failure prediction. *Expert Systems with Applications, 37(2)*, 1535–1541.
Yu, F., Liu, D., Kumar, S., Tony, J., & Chang, S.F. (2013, May). *\proptoSVM for Learning with Label Proportions*. In International Conference on Machine Learning (pp. 504-512). PMLR.
Yuan, K., Chi, G., Zhou, Y., & Yin, H. (2022). A novel two-stage hybrid default prediction model with k-means clustering and support vector domain description. *Research in International Business and Finance, 59*, 101536.
Zavgren, C.V. (1985). Assessing the Vulnerability to Failure of American Industrial Firms: A Logistic Analysis. *Journal of Business Finance & Accounting, 12(1),* 19–45.
Zelenkov, Y., Fedorova, E., & Chekrizov, D. (2017). Two-step classification method based on genetic algorithm for bankruptcy forecasting. *Expert Systems with Applications, 88*, 393-401.
Zhang, X., & Hu, L. (2016). A nonlinear subspace multiple kernel learning for financial distress prediction of Chinese listed companies. *Neurocomputing, 177,* 636-642.
Zmijewski, M.E. (1984). Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting research*, 59-82.
Zhang, H., Kou, G., Peng, Y. & Zhang, B. (2024). Role-aware random walk for network embedding. *Information Sciences*, *652*, 119765.
Zhang, T., Ramakrishnan, R., & Livny, M. (1996). BIRCH: an efficient data clustering method for very large databases. *ACM sigmod record*, *25(2),* 103-114.
Zhao, J., Ouenniche, J. & De Smedt, J. (2024). Survey, classification and critical analysis of the literature on corporate bankruptcy and financial distress prediction. *Machine Learning with Applications*, *15(3),* 100527.

Zhou, L. (2013). Performance of corporate bankruptcy prediction models on imbalanced dataset: The effect of sampling methods. *Knowledge-Based Systems, 41*, 16-25.

Zhou, L., Lu, D., & Fujita, H. (2015). The performance of corporate financial distress prediction models with features selection guided by domain knowledge and data mining approaches. *Knowledge-Based Systems, 85*, 52-61.

Zięba, M., Tomczak, S.K., & Tomczak, J.M. (2016). Ensemble boosted trees with synthetic features generation in application to bankruptcy prediction. *Expert systems with applications, 58*, 93-101.

Zoričák, M., Gnip, P., Drotár, P., & Gazda, V. (2020). Bankruptcy prediction for small- and medium-sized companies using severely imbalanced datasets. *Economic Modelling, 84*, 165-176.

Zorn, M.L., Norman, P.M., Butler, F.C., & Bhussar, M.S. (2017). Cure or curse: Does downsizing increase the likelihood of bankruptcy? *Journal of Business Research, 76*, 24-33.