Machine Learning with Applications 15 (2024) 100527

Contents lists available at ScienceDirect

Machine Learning with Applications

journal homepage: www.elsevier.com/locate/mlwa

Survey, classification and critical analysis of the literature on corporate bankruptcy and financial distress prediction

Jinxian Zhao [a,*], Jamal Ouenniche [a,1], Johannes De Smedt [b]

[a] Business School, The University of Edinburgh, 29 Buccleuch Place, Edinburgh EH8 9JS, UK
[b] KU Leuven, Leuven Institute for Research on Information Systems (LIRIS), Naamsestraat 69, Leuven 3000, Belgium

ARTICLE INFO

Keywords:
Bankruptcy prediction
Financial distress prediction
Machine learning
Classifiers
Drivers

ABSTRACT

Corporate bankruptcy and financial distress prediction is a topic of interest for a variety of stakeholders, including businesses, financial institutions, investors, regulatory bodies, auditors, and academics. Various statistical and artificial intelligence methodologies have been devised to produce more accurate predictions. As more researchers are now focusing on this growing field of interest, this paper provides an up-to-date comprehensive survey, classification, and critical analysis of the literature on corporate bankruptcy and financial distress predictions, including definitions of bankruptcy and financial distress, prediction methodologies and models, data pre-processing, feature selection, model implementation, performance criteria and their measures for assessing the performance of classifiers or prediction models, and methodologies for the performance evaluation of prediction models. Finally, a critical analysis of the surveyed literature is provided to inspire possible future research directions.

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 analyze 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 analyzing 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[2] (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;

* Corresponding author.
E-mail address: Jinxian.Zhao@ed.ac.uk (J. Zhao).
[1] The African Chair of Innovation and Sustainable Management (CAIMD-UM6P).
[2] In this paper, we use the terms 'drivers', 'features', 'explanatory variables' and 'variables' as synonyms.

https://doi.org/10.1016/j.mlwa.2024.100527
Received 31 October 2023; Received in revised form 14 December 2023; Accepted 10 January 2024
Available online 11 January 2024

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 paper 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 synthesize and critically analyze 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 summarize and classify the literature and critically analyze 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 focusing 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 analyzed, 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 paper unfolds as follows. Section 2 provides a classification of research on bankruptcy and financial distress prediction into the main research streams. Section 3 investigates the definitions and criteria of bankruptcy and financial distress. Section 4 provides a painting of the landscape of research on the design of prediction models or classifiers as applied to BP & FDP. Section 5 focuses on research on the design of new drivers or the evaluation of existing ones. Section 6 looks into feature selection methodologies. Section 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 8 provides an overview of the data and markets of research in bankruptcy and financial distress prediction studies. Section 9 provides a critical analysis of the literature, and finally, Section 10 summarizes and concludes our study, as well as provides a list of potential future research directions.

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 existing ones; (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 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.

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.

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 A for our classification of definitions of corporate bankruptcy and financial distress in the bankruptcy and financial distress prediction literature surveyed in this paper. 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.

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 modeling to dynamic ones/dynamic modeling considering time process. The remainder of this section shall be organized into subsections, each focusing on a different category of models.

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 recognized 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

Table 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 Andres 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 existing ones	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), , Uthayakumar et al. (2020) Wang 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)

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 skepticism 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., *normalized standard error of 10-year EBIT/total assets*, *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 utilized 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 generalized 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 analyzing 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 generalized 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 modeling was again improved by (1985) using the categories of drivers proposed by Pinches and 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).

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 modeling* 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 and 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 and 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 generalizable 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 and 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.

4.3. Artificial intelligence / machine learning models

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 and 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 and 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 %, while that of using the numerical data only reaches 80.8 %. Although methods based on ANNs generally demonstrate better capability in modeling 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.

4.3.2. Bayesian theory-based models

Naïve Bayes (NB) and its generalization; 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 and 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 and 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. The reader is referred to Appendix B for a complete list of references on the NB and BN models in BP and FDP. While offering flexibility and simplicity, the Bayesian theory-based models, particularly the Bayesian networks, 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.

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 and Kim (2009) embedded CBR into a *genetic algorithm* (GA) which simultaneously optimizes 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 optimization 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 and Ho (2009) embedded *Fuzzy-CBR* into a *genetic algorithm* which optimizes 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 optimization* (PSO) algorithm with *time-varying acceleration coefficients* and *time-varying inertia weight* to efficiently control the local and global search ability of PSO, which optimizes the neighborhood 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 neighbor* (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 B 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. While 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.

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 B 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 optimized so as to maximize 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 extensively in BP and FDP. 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 generalization 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 neighbor 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 and 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 and 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.

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 B 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 customized 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 modeling and are good at handling different types of data. Although DTs are usually very competitive in-sample, they tend to overfit and lack generalization 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, while the others only achieve an accuracy around 93 %. On the other hand, Shen et al. (2020) proposed a new dynamic financial distress prediction modeling framework named *adaptive neighbor 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.

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 and Sheela (1979), various ensemble learning models have been proposed over the years; e.g., *stacked generalization* (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 optimization 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 generalization 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 initialization 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 and 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 and Yen (2019) claimed that *extreme gradient boosting* (XGBoost), which is a scalable *gradient boosted decision tree* that allows for penalty regularization 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 labeled, 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 B for a complete list of references on ensemble learning models.

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 B 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 and Goto (2009a, 2009b) proposed a *DEA-discriminant analysis classifier*, combining DEA's efficiency scoring with discriminant analysis for enhanced predictive capability. Ouenniche and 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 and 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.

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 B 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 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

Table 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 modeling / 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 optimize 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 generalization 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.

seen in the field of BP and FDP.

5. Design of new drivers or evaluation of existing ones

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 C 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 utilization 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 capitalization*), 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 labor 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 organizational structure of a company, has been increasingly scrutinized 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, 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 Andres 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 utilization 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 C, on the other hand.

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 optimized 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 utilizing 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 predetermined 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 D 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 optimization* (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 D 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 D for our classification of feature selection methods in the BP and FDP literature surveyed in this paper.

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 E 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 generalizability 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 paper, 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 modeling, 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 optimization 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), while 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 and 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 and 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 and 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.

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 summarized 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 organized 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 F 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 focusing on European markets mainly use *Refinitiv Datastream* (e.g., Tseng & Hu, 2010; Tinoco & Wilson, 2013; Andrikopoulos & Khorasgani, 2018) and *Moody's Bureau van Dijk* (e.g., De Andres 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 G.

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.

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 Andres et al. (2011) and Du Jardin (2016, 2017, 2021a, 2021b) have 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.

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 catalyzed 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 synthesized samples of the minority class (or algorithmically reducing oversized samples of the majority class) until the classes are almost equally distributed.

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 *winsorization* (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).

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 modeling with a largely imbalanced dataset, setting a cut-off score value so as to maximize the overall correct classification rate may result in a high correct classification rate of the majority class while sacrificing that of the minority class (Veganzones & Séverin, 2018). Attempts have been made to address this issue by, for example, minimizing 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).

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 existing ones; 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 analyzing 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.

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 focusing 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 modeling. 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 generalizability 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.

Focusing 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 emphasize 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.

CRediT authorship contribution statement

Jinxian Zhao: Conceptualization, Methodology, Investigation, Formal analysis, Writing – original draft. **Jamal Ouenniche:** Project administration, Supervision, Conceptualization, Formal analysis, Writing – review & editing. **Johannes De Smedt:** Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.mlwa.2024.100527.

Appendix

Appendix A

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	Paper / 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 and Sriram (2001), Shin et al. (2005), Min and 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 and 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 reorganized.	Veganzones and Séverin (2018)
		A company is liquidated, reorganized, or ruled by court decision as bankrupt.	Du Jardin (2018, 2021b)
		A company's assets are liquidated in order to fulfill as much debt as possible and the company is no longer a going concern, or a company is in reorganization, which involves the settlement of debt repayment between a company and its creditors while 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 and Yen (2019)
		A company has filed a bankruptcy petition under Chapter X of the US National Bankruptcy Act of 1938, or under Chapter XI (reorganization) 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 and Sharda (1990), Bryant (1997), Paradi et al. (2004), Sun and Shenoy (2007), Bharath and Shumway (2008), Chen et al. (2009), Trabelsi et al. (2015), Tian et al. (2015), Zorn et al. (2017), Tian and Yu (2017), Lin and Dong (2018), Mai et al. (2019), Bai and Tian (2020), Habermann and 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 and Chizema (2016), Andrikopoulos and 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 Andres 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 and 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 and 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 and 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 and Sun (2009), Zhou et al. (2015), Zhang and Hu (2016), Wang and Wu (2017), Choi et al. (2018), Wang et al. (2018), Sun et al. (2020), Wang et al. (2020), Shen et al. (2020), Mousavi and 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, reorganization, or failure to meet listing conditions.	Xu and 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: reorganization, bankruptcy, full-value delivery (per-share book	Lin et al. (2011a, 2011b, 2014)

(*continued on next page*)

Appendix A (*continued*)

Category of Definition	Type of Definition	Definition / Criteria	Paper / Authors & Year
		value below 5TWD), stock transaction suspension, or withdrawal from the Taiwan stock market.	
		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 and Goto (2009a, 2009b)
		A company is delisted from the Japanese stock exchange due to bankruptcy, rehabilitation, or reorganization 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, 2016), Chou et al. (2017)
Hybrid Definitions	Databases' listing criteria	*A company that is dropped from a commercial database.*	Beaver (1966), Tseng and 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 and 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 and Agarwal (2014), Mousavi et al. (2015), Ouenniche and Tone (2017), Ouenniche et al. (2018a, 2018b, 2018c, 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 and 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 characterized 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 and Baba (1994), West et al. (2005), Ahn and Kim (2009), Hu (2009), Li and Ho (2009), Premachandra et al. (2009), Nanni and Lumini (2009), Altman et al. (2010), Chen et al. (2011a), Hu and Chen (2011), Li et al. (2011), Tsai (2009), Tsai and Cheng (2012), Tsai and 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, 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)		

Appendix B

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

Methodology / Models	Paper / 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 and Zhang (2009)[B], Almamy et al. (2016)[B], Qiu et al., (2020)[B], Habermann and Fischer (2023)[B], Alam et al. (2021)[B], Oz and Simga-Mugan (2018)[D], Ninh et al. (2018)[D], Andrikopoulos and Khorasgani (2018)[D] Linear Discriminant Analysis with Regularization 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 and Chizema (2016)[B], Zorn et al. (2017)[B], Lin and Dong (2018)[B], Serrano-Cinca et al. (2019)[B], Li and Faff (2019)[B], Tinoco and 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]

(continued on next page)

Appendix B (*continued*)

Methodology / Models	Paper / Authors & Year	
		Logistic Regression with Regularization Terms: Pereira et al. (2016)[B], Volkov et al. (2017)[B]
		Quadratic Interval Logistic Regression: Tseng and Hu (2010)[B]
		Generalized Additive Model: Valencia et al. (2019)[B]
		Least-squares Probabilistic Classifier: Zoričák et al. (2020)[B]
		Multivariate Adaptive Regression Splines (MARS): De Andres et al. (2011)[B]
	Bayesian Theory-based Models	Naïve Bayes (NB): Hsieh (1993)[B], Trabelsi et al. (2015)[B], Sarkar and Sriram (2001)[D]
		Bayesian Network (BN): Sun and Shenoy (2007)[B]
Stochastic Models	Option Pricing Theory-Based Models	Black-Scholes/Merton Model: Bharath and 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 and Yu (2017)[B], Du Jardin (2017)[B], Du Jardin (2018)[B], Bai and Tian (2020)[B], Gupta and 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 Neighbor (CBR+kNN): Bryant (1997)[B], Ahn and Kim (2009)[B], Min and 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 and Sun (2009)[D], Li et al. (2009)[D]
		Case-based Reasoning with Fuzzy k-Nearest Neighbor (CBR+F-kNN): Chen et al. (2011a)[B], Li and 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 and Sriram (2001)[D]
		Classification and Regression Tree (CART): Liang et al. (2016)[B], Du Jardin (2018)[B], Nyitrai and Virág (2019)[B], Almaskati et al. (2021)[B]
		Chi-square Automatic Interaction Detection Decision Tree (CHAID): Serrano-Cinca et al. (2019)[B], Nyitrai and Virág (2019)[B]
		Logistic Model Tree (LMT): Alam et al. (2021)[B]
	Neural Networks Models	Multi-layer Perceptron (MLP): Odom and Sharda (1990)[B], Tsukuda and 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 and Hu (2010)[B], Uthayakumar et al. (2020)[D]
		Self-Organizing Maps (SOM) Neural Network: Wang and Wu (2017)[D], Du Jardin (2021b)[B]
		Deep Belief Neural Network (DBN): Huang and 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 Quantization (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 and Yen (2019)[D]
	Support Vector Methodologies	Support Vector Machines (SVM): Shin et al. (2005)[B], Nanni and Lumini (2009)[B], Tobback et al. (2017)[B], Chen et al. (2020)[B], Lin et al. (2011a)[D], Lin et al. (2011b)[D], Zhang and Hu (2016)[D], Wang and Wu (2017)[D], Huang and 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 and 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 and 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 and 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 and 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 and 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 Generalization (Stacking): Tsai and Hsu (2013)[B], Kim (2018)[D], Liang et al. (2020)[D]

(*continued on next page*)

Appendix B (*continued*)

Methodology / Models	Paper / Authors & Year	
	Data Envelopment Analysis-Based Models	Paradi et al. (2004)[B], Cielen et al. (2004)[B], Premachandra et al. (2009)[B], Sueyoshi and Goto (2009a)[B], Sueyoshi and Goto (2009b)[B], Yeh et al. (2010)[B], Ouenniche and Tone (2017)[B], Almaskati et al. (2021)[B]
	MCDA-Based Models	Elimination and Choice Translating Reality (ELECTRE): Hu (2009)[B], Hu and 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 and Chen (2011)[B]

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

Appendix C
The drivers used in the BP and FPD surveyed literature along with their percentage of use (%).

Category of information	Category of Measurement (s)	Drivers/Authors & Year
Accounting Information	Asset Utilization 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 Amortization/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);

(*continued on next page*)

Appendix C (*continued*)

Category of information	Category of Measurement (s)	Drivers/Authors & Year
		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);

(*continued on next page*)

Appendix C (*continued*)

Category of information	Category of Measurement (s)	Drivers/Authors & Year
		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 and 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 and 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);

(*continued on next page*)

Appendix C (*continued*)

Category of information	Category of Measurement (s)	Drivers/Authors & Year
		Director Internal Experience[23] (Almaskati et al., 2021); *Director Company Experience*[24] (Almaskati et al., 2021); *Nomination Committee Effectiveness*[25] (Appiah and 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); *Organization 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 Capitalization (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);

(*continued on next page*)

Appendix C (*continued*)

Category of information	Category of Measurement (s)	Drivers/Authors & Year
	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 Amortization/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 amortization (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, Organization 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, while 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}.

Appendix D
Classification of feature selection methods in BP and FDP.

Classification of Feature Selection Methods	Feature Selection Methods	Paper / Authors & Year
Domain Knowledge	Referring to (Min and Jeong, 2009)Previous Studies	Beaver (1966), Altman et al. (1977), Ohlson (1980), Zmijewski (1984), Zavgren (1985), Frydman et al. (1985), Odom and Sharda (1990), Hsieh (1993), Bryant (1997), Shumway (2001), Paradi et al. (2004), Cielen et al. (2004), West et al. (2005), Bharath and Shumway (2008), Chen et al. (2009), Hu (2009), Premachandra et al. (2009), Sueyoshi and Goto (2009a), Tseng and Hu (2010), Altman et al. (2010), Gorgani et al. (2010), De Andres et al. (2011), Hu and Chen (2011), Tinoco and 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 and Tone (2017), Lian (2017), Ouenniche et al. (2018a), Ouenniche et al. (2018b), Ouenniche et al. (2018c), Lin and Dong (2018), Miao et al. (2018), Andrikopoulos and Khorasgani (2018), Oz and Simga-Mugan (2018), Choi et al. (2018), Ahmadi et al. (2018), Alaka et al. (2018), Li and Faff (2019), Mai et al. (2019), Ouenniche et al. (2019), Muñoz-Izquierdo et al., (2019), Gupta and Chaudhry (2019), Mousavi et al. (2019), Matin et al. (2019), Qiu et al. (2020), Bai and Tian (2020), Ashraf et al. (2020), Almaskati et al. (2021)
	Expert Knowledge	Ahn and Kim (2009), Min and Jeong (2009), Li and Ho (2009), Altman et al. (2010), Bauer and Agarwal (2014)
Filter Methods	Independent Samples t-test	Shin et al. (2005), Ahn and Kim (2009), Min and 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 and Lin (2020), De Bock et al. (2020)

(continued on next page)

Appendix D (*continued*)

Classification of Feature Selection Methods	Feature Selection Methods	Paper / Authors & Year
	Fisher *F*-test	Altman (1968), Altman et al. (1977), Li and Sun (2009), Geng et al. (2015), Almamy et al. (2016), Veganzones and 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 and Sriram (2001)
	Correlation Analysis	Sarkar and Sriram (2001), Paradi et al. (2004), Xu and Zhang (2009), Tsai (2009), Sueyoshi and Goto (2009b), Lin et al. (2011a), Moradi et al. (2013), Barboza et al. (2017), Veganzones and Séverin (2018), Tinoco et al. (2018), Ninh et al. (2018), Serrano-Cinca et al. (2019), Li and Faff (2019), Hosaka (2019), Ahmad (2019), Yuan et al. (2022)
	Pearson Correlation Analysis	Sun and Shenoy (2007), Appiah and Chizema (2016), Bai and Tian (2020), Paraschiv et al. (2021)
	Mann-Whitney Test (Wilcoxon Rank Sum Test)	Tsukuda and 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 and Chizema (2016), Sartori et al. (2016), Serrano-Cinca et al. (2019), Sun et al. (2020), Li et al. (2021), Mousavi and Lin (2020)
	Stepwise Multivariate Discriminant Analysis	Altman et al. (1977), Shin et al. (2005), Min and 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 and Virág (2019)
	Stepwise Logistic Regression	Ahn and Kim (2009), Min and Jeong (2009), Li et al. (2011), Tinoco and Wilson (2013), Lu et al. (2013), Liang et al. (2015), Liang et al. (2016), Sartori et al. (2016), Nyitrai and Virág (2019), Fernández-Gámez et al. (2020), Liang et al. (2020), Ashraf et al. (2020)
	Principal Component Analysis	Taffler (1984), Tsai (2009), Sueyoshi and Goto (2009b), Nanni and Lumini (2009), Lin et al. (2011b), Mousavi et al. (2015), Mousavi et al. (2019), Mousavi and 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 and Wu (2017)
	Laplacian Eigenmaps (LE)	Wang and Wu (2017)
	Locally Linear Embedding (LLE)	Wang and 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 Optimization (PSO)-based Wrapper	Chen et al. (2011a), Liang et al. (2015), Uthayakumar et al. (2020)
	Genetic Algorithm (GA)-Based Wrapper	Ahn and Kim (2009), Min and Jeong (2009), Li and 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 and 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 Optimization (ACO)-Based Wrapper	Uthayakumar et al. (2020)
	Grey Wolf Optimization (GWO)-Based Wrapper	Uthayakumar et al. (2020)
	Deep Belief Network (DBN)-Based-Wrapper	Huang and 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 and 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 and 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 and Yu (2017), Wang et al. (2018), Du et al. (2020), Bai and Tian (2020), Paraschiv et al. (2021)
	Sparse Group LASSO	Wang et al. (2020)
	Ridge Regression	Pereira et al. (2016)
	Generalized Additive Model Selection (GAMSEL)	Valencia et al. (2019)
	Non-linear Subspace Multiple-kernel Machine (NS-MKL)	Zhang and Hu (2016)
	Artificially Synthetic Features	Zięba et al. (2016)
Not Mentioned/ Not Used		Tsai and Cheng (2012), Tsai and 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 and Fischer (2023)

Appendix E
Classification of performance criteria & measures in the BP and FDP problem.

Performance Criterion	Measures of Performance Criterion	Definition and Introduction	Paper / 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 and Baba (1994), Bryant (1997), Cielen et al. (2004), Chen et al. (2009), Tsai (2009), Premachandra et al. (2009), Sueyoshi and Goto (2009a), Sueyoshi and Goto (2009b), Nanni and Lumini (2009), Yeh et al. (2010), Chen et al. (2011a), Chen et al. (2011b), Lin et al. (2011a), Lin et al. (2011b), Tsai and Cheng (2012), Tsai and 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 and Tone (2017), Ouenniche et al. (2018a), Ouenniche et al. (2018c), Oz and Simga-Mugan (2018), Du Jardin (2018), Wang et al. (2018), Kim (2018), Serrano-Cinca et al. (2019), Huang and Yen (2019), Mousavi et al. (2019), Ouenniche et al. (2019), Lin et al. (2019), Fernández-Gámez et al. (2020), Liang et al. (2020), Mousavi and 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 and 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 and Ho (2009), Zhou (2013), Tinoco and 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 and Tone (2017), Veganzones and Séverin (2018), Ouenniche et al. (2018a), Ouenniche et al. (2018c), Huang and Yen (2019), Mousavi et al. (2019), Ouenniche et al. (2019), Fernández-Gámez et al. (2020), Sun et al. (2020), Mousavi and 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). $Specificity = \frac{True\ Positives}{True\ Positives + False\ Positives}$	Li and 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 and Lin (2020), Putra et al. (2020), Du et al. (2020), Almaskati et al. (2021)
	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 and 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	Zhou (2013), Volkov et al. (2017), Antunes et al. (2017), Ahmadi et al. (2018), Hosaka (2019), Uthayakumar et al. (2020), Sun et al.

(continued on next page)

Appendix E (*continued*)

Performance Criterion	Measures of Performance Criterion	Definition and Introduction	Paper / Authors & Year
		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.	(2020), Shen et al. (2020), Mousavi and 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 and Sharda (1990), Bryant (1997), Sarkar and Sriram (2001), Paradi et al. (2004), Cielen et al. (2004), Shin et al. (2005), West et al. (2005), Sun and Shenoy (2007), Ahn and Kim (2009), Min and Jeong (2009), Chen et al. (2009), Hu (2009), Li and Sun (2009), Li et al. (2009), Li and Ho (2009), Premachandra et al. (2009), Xu and Zhang (2009), Tsai (2009), Sueyoshi and Goto (2009a), Sueyoshi and Goto (2009b), Nanni and Lumini (2009), Yeh et al. (2010), Tseng and Hu (2010), Gorgani et al. (2010), Chen et al. (2011a), Lin et al. (2011a), Lin et al. (2011b), Li et al. (2011), Hu and Chen (2011), Tsai and Cheng (2012), Tsai and Hsu (2013), Zhou (2013), Tinoco and 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 and 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 and Wu (2017), Chou et al. (2017), Du Jardin (2017), Tinoco et al. (2018), Oz and 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 and Wang (2019), Son et al. (2019), Huang and 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 and Lin (2020)
	Matthews Correlation Coefficient (MCC)	Matthews Correlation Coefficient measures the quality of a prediction model's classifications, taking into account both true and false positives and negatives. In the context of BP and FDP, MCC assesses the model's ability to correctly identify companies in distress while minimizing incorrect classifications. It is particularly useful when dealing with imbalanced datasets and is less sensitive to class distribution.	Uthayakumar et al. (2020), Putra et al. (2020)

(*continued on next page*)

Appendix E (*continued*)

Performance Criterion	Measures of Performance Criterion	Definition and Introduction	Paper / Authors & Year
		$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$ 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 and Lumini (2009), Altman et al. (2010), Chen et al. (2011a), Tinoco and Wilson (2013), Zhou (2013), Bauer and 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 and Yu (2017), Du Jardin (2017), Veganzones and Séverin (2018), Le et al. (2018), Miao et al. (2018), Andrikopoulos and Khorasgani (2018), Ninh et al. (2018), Lin and Dong (2018), Du Jardin (2018), Choi et al. (2018), Wang et al. (2018), Alaka et al. (2018), Serrano-Cinca et al. (2019), Li and Faff (2019), Le et al. (2019), Hosaka (2019), Mai et al. (2019), Jones and Wang (2019), Gupta and Chaudhry (2019), Son et al. (2019), Nyitrai and 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 and Lin (2020), Putra et al. (2020), Du et al. (2020), De Bock et al. (2020), Bai and 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 and Wilson (2013), Bauer and Agarwal (2014), Gupta et al. (2015), Mousavi et al. (2015), Mousavi et al. (2019), Jones and Wang (2019), Mousavi and 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 and Wilson (2013), Mousavi et al. (2015), Gupta et al. (2015), Mousavi et al. (2019), Mousavi and 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 and Faff (2019), Mai et al. (2019), Almaskati et al. (2021)

(*continued on next page*)

Appendix E (*continued*)

Performance Criterion	Measures of Performance Criterion	Definition and Introduction	Paper / Authors & Year
		$Accuracy\ Ratio = \frac{Observed\ Area}{Maximum\ Possible\ Area}$ 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 and 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, Analyze and compare the rates of actual positive cases across the deciles.	Shumway (2001), Bharath and Shumway (2008), Tian et al. (2015), Tian and Yu (2017), Miao et al. (2018), Mai et al. (2019), Bai and 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 while considering the balance between the model's ability to fit the data well (goodness of fit) and its complexity. $AIC = -2 \cdot \text{loglikelihood} + 2 \cdot number\ of\ parameters$ 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 and Chizema (2016), Tian and 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 while 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. $BIC = -2 \cdot \text{loglikelihood} + \log(n) \cdot number\ of\ parameters$ Where n represents the sample size.	Appiah and Chizema (2016)

(continued on next page)

Appendix E (*continued*)

Performance Criterion	Measures of Performance Criterion	Definition and Introduction	Paper / Authors & Year
		As with the AIC, a lower BIC value indicates a better trade-off between model fit and complexity.	
	Log-likelihood Statistic	Log-likelihood statistic is a measure used in statistical modeling 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 and Chizema (2016), Jones and 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)

Appendix F

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 and Sharda (1990), Hsieh (1993), Bryant (1997), Shumway (2001), Paradi et al. (2004), West et al. (2005), Sun and Shenoy (2007), Bharath and Shumway (2008), Chen et al. (2009), Hu (2009), Premachandra et al. (2009), Tsai (2009), Tsai and 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 and Dong (2018), Li and Faff (2019), Mai et al. (2019), Qiu et al. (2020), Bai and Tian (2020), Habermann and Fischer (2023), Almaskati et al. (2021)
	UK	Taffler (1984), Altman et al. (2008), Tseng and Hu (2010), Bauer and Agarwal (2014), Gupta et al. (2015), Mousavi et al. (2015), Appiah and Chizema (2016), Almamy et al. (2016), Tobback et al. (2017), Tian and Yu (2017), Ouenniche and Tone (2017), Ouenniche et al. (2018a), Ouenniche et al. (2018b), Ouenniche et al. (2018c), Ouenniche et al. (2019)
	Japan	Tsukuda and Baba (1994), Xu and Zhang (2009), Tsai (2009), Sueyoshi and Goto (2009a), Sueyoshi and Goto (2009b), Nanni and Lumini (2009), Tsai and Cheng (2012), Zhou (2013), Antunes et al. (2017), Tian and Yu (2017), Hosaka (2019)
	France	Chen et al. (2011b), Du Jardin (2016), Tian and Yu (2017), Du Jardin (2017), Veganzones and Séverin (2018), Du Jardin (2018), De Bock et al. (2020), Du Jardin (2021a), Du Jardin (2021b)
	S. Korea	Shin et al. (2005), Ahn and Kim (2009), Min and Jeong (2009), Kim et al. (2016), Le et al. (2018), Le et al. (2019), Son et al. (2019)
	Taiwan	Yeh et al. (2010), Hu and 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 and Lumini (2009), Tsai and Cheng (2012), Antunes et al. (2017), Tian and Yu (2017), Lin et al. (2019)
	Spain	De Andres et al. (2011), Piñeiro-Sánchez et al. (2013), Muñoz-Izquierdo et al., (2019)
	Belgium	Cielen et al. (2004), Tobback et al. (2017), Jones and Wang (2019), De Bock et al. (2020)
	Australia	West et al. (2005), Tsai (2009), Nanni and Lumini (2009), Tsai and 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 and Virág (2019), Alam et al. (2021)
	Iran	Gorgani et al. (2010), Moradi et al. (2013)
	Hungary	Nyitrai and Virág (2019)
	China	Li et al. (2011), Kou et al. (2021)

(continued on next page)

Appendix F (*continued*)

Category of Market	Market	Literature
	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 and Sriram (2001), Cleofas-Sánchez et al. (2016), Lian (2017), Miao et al. (2018), Kim (2018), Gupta and Chaudhry (2019), Liang et al. (2020)
	UK	Altman et al. (2010), Tinoco and Wilson (2013), Tinoco et al. (2018), Andrikopoulos and Khorasgani (2018), Alaka et al. (2018), Mousavi et al. (2019), Ashraf et al. (2020)
	Taiwan	Li and Ho (2009), Lin et al. (2011a), Lu et al. (2013), Lin et al. (2014), Liang et al. (2015), Lu et al. (2015), Huang and 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 and Sun (2009), Li et al. (2009), Liang et al. (2015), Zhou et al. (2015), Geng et al. (2015), Zhang and Hu (2016), Sun et al. (2017), Wang and Wu (2017), Wang et al. (2018), Sun et al. (2020), Wang et al. (2020), Shen et al. (2020), Li et al. (2021), Mousavi and 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)

Appendix G
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 and 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 and Shenoy (2007), Bharath and Shumway (2008), Chen et al. (2009), Tinoco and Wilson (2013), Zhou (2013), Wang et al. (2014), Tian et al. (2015), Afik et al. (2016), Barboza et al. (2017), Lian (2017), Tian and Yu (2017), Miao et al. (2018), Lin and Dong (2018), Kim (2018), Li and Faff (2019), Mai et al. (2019), Gupta and Chaudhry (2019), Liang et al. (2020), Qiu et al. (2020), Almaskati et al. (2021)
	Refinitiv	Datastream	Xu and Zhang (2009), Tseng and Hu (2010), Bauer and Agarwal (2014), Ouenniche and Tone (2017), Ouenniche et al. (2018a), Ouenniche et al. (2018b), Ouenniche et al. (2018c), Ouenniche et al. (2019), Tinoco et al. (2018), Andrikopoulos and Khorasgani (2018), Oz and Simga-Mugan (2018), Ahmad (2019), Mousavi et al. (2019), Ashraf et al. (2020), Habermann and Fischer (2023)
		Thomson One	Tinoco and Wilson (2013), Appiah and Chizema (2016), Tinoco et al. (2018)
		WorldScope Fundamentals	Appiah and Chizema (2016), Tinoco et al. (2018)
	Bloomberg		Almamy et al. (2016), Oz and Simga-Mugan (2018), Ninh et al. (2018)
	Dun & Bradstreet	Altares	Veganzones and 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 and Baba (1994)
		Nikkei Economic Electronic Databank System (NEEDS)	Sueyoshi and 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 and Agarwal (2014), Tobback et al. (2017), Alaka et al. (2018), Andrikopoulos and Khorasgani (2018)
		Orbis	Jones and Wang (2019)
		Sabi	De Andres et al. (2011), Piñeiro-Sánchez et al. (2013), Muñoz-Izquierdo et al., (2019)

(*continued on next page*)

Appendix G (*continued*)

Type of Provider or Database	Provider	Database Name	Literature
	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 and Wu (2017), Wang et al. (2018), Sun et al. (2020), Wang et al. (2020), Shen et al. (2020), Li et al. (2021), Mousavi and 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 Andres 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 and Wilson (2013), Bauer and Agarwal (2014), Mousavi et al. (2015), Ouenniche and 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 and Lumini (2009), Chen et al. (2011a), Tsai and Hsu (2013), Liang et al. (2015), Cleofas-Sánchez et al. (2016), Antunes et al. (2017), Lin et al. (2019), Nyitrai and 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 and Tian (2020)
	University of California, San Diego (UCSD)	UCSD competition dataset	Tsai (2009), Tsai and Cheng (2012), Tsai and Hsu (2013), Cleofas-Sánchez et al. (2016)
	New York University (NYU)	Altman's Bankruptcy Database	Premachandra et al. (2009), Sueyoshi and Goto (2009a)
	Centre for Research in Security Prices (CRSP)		Shumway (2001), Tian et al. (2015), Afik et al. (2016), Li and Faff (2019), Mai et al. (2019), Bai and 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 and Hsu (2013)
	The PACAP Research Centre		Xu and 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 U.S. SEC 8K Report	Zmijewski (1984), Mai et al. (2019) Sun and Shenoy (2007), Almaskati et al. (2021)
	UK Company House		Altman et al. (2008), Altman et al. (2010), Appiah and 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 and 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 and 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 and 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 and Yen (2019)
	Financial Times (FT)	Extel Database	Tseng and Hu (2010)
Stock Exchanges	Shanghai & Shenzhen Stock Exchange		Li and Sun (2009), Li et al. (2009), Li et al. (2011), Zhang and Hu (2016)
	Karachi Stock Exchange (KSE)		Ashraf et al. (2020)

References

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). Towards bankruptcy prediction: Deep sentiment mining to detect financial distress from business management reports. In *Proceedings of the IEEE 5th international conference on Data Science and Advanced Analytics (DSAA)* (pp. 293–302). IEEE.

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), Article 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*, Article 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., Sabato, G., & Wilson, N. (2008). The value of non-financial information in SME risk management. Available at SSRN: https://ssrn.com/abstract=1320612 or 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.

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.

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*, Article 100659.

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

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

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.

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

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.

Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992). 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. InPrinciples of Data Mining (pp.119–134)*. London: Springer.

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.

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). 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*, Article 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.

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.

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

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*, Article 106758.

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*, Article 113576.

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

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

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

Flath, D. (2000). Japan's business groups. *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.

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.

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.

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.

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

Hsieh, S. J. (1993). A note on the optimal cutoff point in bankruptcy prediction models. *Journal of Business Finance & Accounting, 20*(3), 457–464.
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*, Article 105663.
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). Generalization and decision tree induction: Efficient classification in data mining. In *Proceedings of the 7th 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.
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*, Article 113429.
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.
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, 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*, Article 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.
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.
Lippmann, R. (1987). An introduction to computing with neural nets. *IEEE Assp Magazine, 4*(2), 4–22.
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.
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.
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., & Tone, K. (2023). A dynamic performance evaluation of distress prediction models. *Journal of Forecasting, 42*(4), 756–784.
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., & 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*, Article 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.
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 *Proceedings of the 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*, 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.
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 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.
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.
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). A credit scoring model for smes based on social media data. In *Proceedings of the international conference on web engineering* (pp. 113–129). Cham: Springer.
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*, Article 113475.

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). A study on the relationships of classifier performance metrics. In *Proceedings of the 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*, Article 105365.
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.
Son, H., Hyun, C., Phan, D., & Hwang, H. J. (2019). Data analytic approach for bankruptcy prediction. *Expert Systems with Applications, 138*, Article 112816.
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.
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.
Tian, S., Yu, Y., & Guo, H. (2015). Variable selection and corporate bankruptcy forecasts. *Journal of Banking & Finance, 52*, 89–100.
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., & 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), Article 1597956.
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, 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.
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, 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). \proptoSVM for learning with label proportions. In *Proceedings of the 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*, Article 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.
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.
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*, Article 106152.