Routledge
Taylor & Francis Group

Journal of Small Business Management

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ujbm20

SME default prediction: A systematic methodology-focused review

Hamid Cheraghali & Peter Molnár

To cite this article: Hamid Cheraghali & Peter Molnár (08 Dec 2023): SME default prediction: A systematic methodology-focused review, Journal of Small Business Management, DOI: 10.1080/00472778.2023.2277426

To link to this article: https://doi.org/10.1080/00472778.2023.2277426

© 2023 The Author(s). Published with license by Taylor & Francis Group, LLC.

Published online: 08 Dec 2023.

Submit your article to this journal

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=ujbm20

JOURNAL OF SMALL BUSINESS MANAGEMENT
https://doi.org/10.1080/00472778.2023.2277426

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

SME default prediction: A systematic methodology-focused review

Hamid Cheraghali [a] and Peter Molnár [a,b,c]

[a]UiS Business School, University of Stavanger, Norway; [b]Faculty of Finance and Accounting, Prague University of Economics and Business, Czech Republic; [c]Faculty of Economic Sciences and Management, Nicolaus Copernicus University, Poland

ABSTRACT

This study reviews the methodologies used in the literature to predict failure in small and medium-sized enterprises (SMEs). We identified 145 SMEs' default prediction studies from 1972 to early 2023. We summarized the methods used in each study. The focus points are estimation methods, sample re-balancing methods, variable selection techniques, validation methods, and variables included in the literature. More than 1,200 factors used in failure prediction models have been identified, along with 54 unique feature selection techniques and 80 unique estimation methods. Over one-third of the studies do not use any feature selection method, and more than one-quarter use only in-sample validation. Our main recommendation for researchers is to use feature selection and validate results using hold-out samples or cross-validation. As an avenue for further research, we suggest in-depth empirical comparisons of estimation methods, feature selection techniques, and sample re-balancing methods based on some large and commonly used datasets.

KEYWORDS

SME; default; failure; bankruptcy; methodology review

Introduction

The literature on bankruptcy prediction started in the 1930s (Bellovary et al., 2007). Altman (1968) employed multivariate analysis for predicting corporate bankruptcy. Before Altman (1968), the literature focused on univariate analysis (Bellovary et al., 2007). Small and medium-sized enterprises (SMEs) default prediction literature has begun with a study of small business failure by Edmister (1972). According to Edmister (1972), the lack of small businesses failure prediction research was due to the difficulty in obtaining data on small businesses. SMEs are considered the main block of the economy for many countries. Despite their essential role in the economy, SMEs often have no access to the capital markets when it comes to raising funds; this makes banks an imperative source of credit.

To obtain financing for SMEs, whether credit from financial institutions or funds from investors, it is crucial to understand the factors contributing to business

CONTACT Hamid Cheraghali hamid.cheraghali@uis.no UiS Business School, University of Stavanger, Arholms gate 35, 4021 Stavanger, Norway

© 2023 The Author(s). Published with license by Taylor & Francis Group, LLC.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

failure. This can help financial institutions and investors to make more informed decisions about lending and investing in these businesses. By predicting the likelihood of SME default, financial institutions can better assess the risk associated with lending to these businesses and make more accurate decisions about whether to approve a loan or not. This can help reduce the number of loan defaults, which benefits both the lender and the borrower. Moreover, studying SME default prediction can also help researchers understand the underlying factors contributing to business failure. A better understanding of the factors that affect SMEs' failures can be used to develop policies and programs that support their growth and success. This is particularly important as SMEs significantly drive economic growth and job creation in many countries. Furthermore, studying SME default prediction can also contribute to developing more accurate and effective predictive models. With the increasing availability of data and the development of machine learning and other advanced analytical techniques, there is a growing need for accurate and effective predictive models to help businesses and financial institutions make better decisions. By studying SME default prediction, researchers can develop and test new models and techniques that can be applied to other areas of finance and business.

Following the implementation of the Basel Capital Accord II in 2004, banks were required to use internal rating systems to assign ratings to their borrowers and compute their capital requirements based on those ratings. Consequently, SME failure prediction regained the interest of academics and practitioners. Using the available default prediction models on SME information for large corporations at the time might have looked like an instant solution for predicting SME failure. However, instead of using a model established for large corporations' failure prediction on SMEs' data, separating default prediction models for SMEs and large corporates will result in models with relatively higher predictive power (Altman & Sabato, 2007).

After Basel II, the global financial crisis of 2007–2009 increased the attention of academics to the topic of SME failure prediction, such that the number of studies on this topic increased significantly in 2010 compared to that in 2009. The subsequent noticeable increase in scholarly attention happened after COVID-19; the number of published articles in 2021 and 2022 are each considerably higher than in 2020.

Although the importance of effective SME failure prediction is renowned, there is only one up-to-date systematic literature review published in this field of research by Ciampi et al. (2021). However, this study does not dive deep into the methodologies. We therefore conduct a systematic methodology-focused review in this domain, with a focus on the methods and predictors used in the SME failure prediction literature. In particular, we provide a summary of all the predictors, sample re-balancing methods (undersampling/oversampling strategies), variable selection methods, estimation methods, and validation approaches used in the literature.

It is important to emphasize the relation between Ciampi et al. (2021) and our study. Our study is by no means an alternative to Ciampi et al. (2021). On the contrary, we build upon it. Ciampi et al. (2021) is a very detailed and insightful general review, and we highly recommend reading Ciampi et al. (2021) first. The main goal of our study is not a general review on SME default prediction, but a narrower review focused on methodologies used. Such a review might be useful for researchers deciding which methodology to use.

Our paper systematically reviews the existing studies about SME failure prediction from a methodological perspective. We reviewed 145 studies and identified over twelve hundred factors used in the previous literature to predict SME failures from 1973 to early 2023. Eighty estimation methods are employed in these studies. We also listed six categories of data sources the researchers have used during the past six decades, along with 54 unique feature selection techniques. We observed that more than 37% of the studies do not include or report any feature selection techniques for their models, more than 25% use in-sample validation techniques, and more than 50% of studies do not report standard measures of the predictive performance of models. It is recommended that future studies in this domain construct their model using proper feature selection techniques and test their models using either hold-out samples or cross-validation.

The rest of the paper is constructed as follows: we first define the research problem and methodology used to answer the question. Then, we summarize the studies and discuss the findings. Finally, conclusions with suggestions for future research are presented.

Background and methodology

Since the start of bankruptcy prediction literature in the 1930s, the models for predicting bankruptcy have shown great diversity. For instance, Altman (1968) uses a five-factor multivariate discriminant analysis model. In contrast, Kou et al. (2021) test one hundred indicators using seven different estimation methods: linear discriminant analysis, logistic regression, support vector machine, decision tree, random forests, XGBoost, and neural network. In a review of bankruptcy prediction studies, Bellovary et al. (2007) identified 752 variables (features/model factors), eight model types (estimation methods), and two validation methods (hold-out sample and in-sample). However, this study covers bankruptcy prediction studies from 1930 to 2004. Moreover, this study is not particularly focused on SME bankruptcy prediction and is relatively old concerning the advances in estimation techniques such as machine learning. An up-to-date systematic literature review on SMEs is carried out by Ciampi et al. (2021). This study analyzes more than one hundred peer-reviewed articles based on statistical and bibliometric characteristics.

While Ciampi et al. (2021) study is a comprehensive, systematic, and detailed review of the SME failure prediction literature, it does not address the methodologies that have been used in the literature in a similar way as the Bellovary et al. (2007) study does for bankruptcy prediction literature. A systematic methodology-focused review concerning SME failure prediction seems necessary to fill this gap. Therefore, this study focuses on the methodologies used in predicting SME default. Here, the word "methodologies" addresses estimation methods (model types), variables (features/model factors), variable selection strategies, rebalancing methods, and validating approaches. In order to provide a granular overview and analysis of the methodologies in the literature concerning SME default prediction, we applied a two-stage approach. In the first step, the relevant literature is collected and filtered using a slightly modified version of Ciampi et al. (2021) literature selection approach. Then a similar framework to Bellovary et al. (2007) is used to summarize the literature.

For creating the research query, we used a modified version of the query used by Ciampi et al. (2021). The main reason for following the Ciampi et al. (2021) query structure is that the paper is the latest literature review concerning this topic, which is reasonably up-to-date. The other reason is that we build upon their work from a different aspect, a methodology-focused review. After a slight modification, we composed this query: *("small and medium size enterprise*" OR "small enterprise*" OR "small compan*" OR "small business*" OR SMEs OR SME) AND ("credit risk*" OR "financial distress" OR default OR bankruptcy OR failure) AND (prediction OR predicting OR "credit risk*")*. We ran the above query in Scopus with the "TITLE-ABS-KEY" operator. The results were afterward limited to "articles." Contrary to Ciampi et al. (2021), we excluded "literature reviews" since the current paper aims to address the methodologies used in the previous literature. The initial list of published articles from Scopus, as the primary source of scientific database (Balzano, 2022; Ciampi et al., 2021; Falagas et al., 2008), was retrieved on December 1, 2022, and updated on January 20, 2023. There are 394 published articles included in this list.

Although the above query retrieves relevant studies, it misses some when the study uses a positive word in the keywords or title like "creditworthiness" instead of "default," "bankruptcy," or "failure." To account for this issue, the following complementary query is also used: *("small and medium size enterprise*" OR "small enterprise*" OR "small compan*" OR "small business*" OR SMEs OR SME) AND (creditworthiness OR "credit worthiness")* AND (evaluation OR analysis OR assessment). The complementary list of published articles from Scopus was retrieved on January 20, 2023, and included 39 published articles with 28 nonduplicated articles. The total number of articles is, therefore, 422 articles.[1]

[1]This number represents the articles found by the mentioned queries. However, a reviewer has suggested five additional articles to be included in this study that we added to the final set of articles.

The next step was to exclude studies outside this paper's scope. In this step, we read all the abstracts (introductions when no abstract was available). The exclusion criteria in the cleaning process were based on (1) the unit of study, (2) the default definition, and (3) the objective of the current study. For the first point, the final set of articles includes every empirical study that addresses SMEs as research units. If the unit of study is a portfolio of SMEs or a network of SMEs, the study is excluded. Regarding the definition of SME, a paper is included in the final article set as far as the definition of SME is within a regional SME definition relevant to that research.[2] Regarding the default definition, we considered four variations of failure:[3] financial distress, default, failure, and bankruptcy. A set of 217 articles is obtained at the end of this step.

The second round of screening required a more in-depth reading of each study. In this round, we identified 145 studies relevant to our methodology-focused review. A list of all the studies included in the current study is available in Table B1 in Appendix B.

While doing a systematic review, the papers included must be based on a clear and objective rule. Some systematic reviews consider all papers in a broad database, such as Scopus (Ciampi et al., 2021) or Web of Science (Marzi et al., 2017). Other systematic reviews narrow down the list of selected papers by considering only articles published in journals included in some prestigious journal ranking, such as The Academic Journal Guide of the Association of Business Schools (Balzano, 2022).

We combine these two approaches. We include all the relevant papers from the Scopus database. Still, at the same time, we divide papers into two groups: those published in journals included in the Academic Journal Guide 2021 of the Association of Business Schools (ABS) and the remaining papers. For simplicity, the studies ranked in this hierarchy are denoted as "ABS" and those not as "non-ABS" throughout this study. This allows us to review the existing literature comprehensively. At the same time, readers interested only in papers published in ABS-ranked journals can easily focus on these papers.

Figure 1 shows the distribution of the articles selected in the final stage. We included the distribution of studies investigating the "Stock Market" (the darker stair plot) to compare the trend of academic attention on this topic and a more general case. We can see that the first increase in studies that empirically investigate SME default prediction occurred in 2007. Considering the time it takes for authors to write and publish an article, we can relate this increase to implementing the Basel II accord. Following the 2007–2009 global financial crisis, an increase in 2010 and 2011 is observable. And finally,

[1]This number represents the articles found by the mentioned queries. However, a reviewer has suggested five additional articles to be included in this study that we added to the final set of articles.

[2]Ciampi et al. (2021) only included the studies where the SME definition was in line with the European definition of SME. However, for this paper, this inclusion criteria seems not necessary as the main object of the research is the method.

[3]Throughout this study, we use default and failure interchangeably for all definitions of failure we have considered.

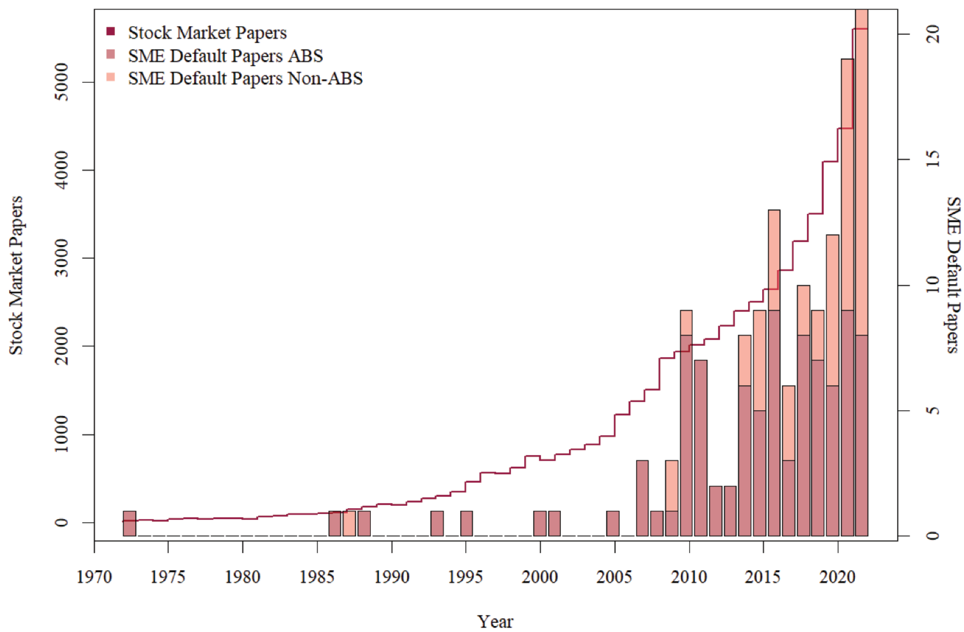


Figure 1. Distributions of the papers from 1972 to 2022. *Note*: The right axis shows the number of SME default prediction studies per year included in this review's final set of articles. The darker color represents the articles published in journals listed by the Association of Business Schools (ABS) 2021 ranking, and the lighter color represents those not published in journals listed in ABS 2021 ranking. The stair plot (left axis) shows the "Stock Market" papers distribution from 1972 to 2022, resulting from a query in Scopus with the "TITLE-ABS-KEY" operator that only contains "Stock Market."

another noticeable increase after COVID-19. However, the trend for stock market research is consistently increasing.

In the final step, we went through all the selected articles and summarized the methods used in each article. This summary includes (1) population description, such as the geographical location/locations and the time horizon of the research data, (2) data characteristics (for example, financial, nonfinancial, firm characteristics, manager characteristics, and macroeconomics), (3) generality of the study (that is, is the research focused on a specific subset of SMEs or not) (4) variables (model factors) included in the research, (5) over- and undersampling strategies, (6) variable selection strategies/methods, (7) estimation method/methods used in the analyses (model types), (8) validation method/methods used to validate the model/models prediction accuracy, (9) proposed model/models performance, and (10) the source of data. This stage uses a modified and expanded approach applied by Bellovary et al. (2007). The added overview aspects are (1), (2), (5), (6), and (10).

By including an overview of the time horizons used in previous studies, we can investigate if the data includes specific periods (for example, the 2007–2009 global financial crisis and COVID-19). Data characteristics are relevant since most research focuses on models requiring historical data, such as balance sheet information. However, these models are unsuitable for predicting failure events when financial records are not available; for example, during business planning and new venture creation. Moreover, collecting financial information from small business entrepreneurs is often difficult. Lussier (1995) developed a nonfinancial business success versus failure prediction model including only nonfinancial variables. His model later has been tested in seven other countries.

The sample re-balancing method can impact the predictive power of a model; for example, some studies use a balanced (half defaulted firms and half non-defaulted firms) sample (Altman et al. (2022) while some studies show that an imbalanced sample performs better (Kou et al. (2021)). In addition, an efficient variable selection is an essential step in modeling bankruptcy (Du Jardin, 2009). Thus, an overview of variable selection techniques from previous literature (point 6) is included in this review. Finally, data sources are one driver of errors in estimations. For example, a dataset obtained from a bank only contains the firms that applied for a loan. In worst-case scenarios, it only contains accepted loan data (which may not be a sample representing the whole population under the study). Thus, listing data sources helps distinguish previous studies' possible error sources.

Results

We present a detailed overview of each dimension of methodologies used in previous literature in this section. The studies are classified into two groups: ABS-ranked and non-ABS. Table 1 shows the distribution of studies based on ABS ranking, where 4* is the highest ranking. Among the studies reviewed in this paper, 94 are listed in ABS, and 51 are not listed.

Population description

We summarize the population distribution in terms of three dimensions: geographical location under the study, the time horizon of the data in year increments, and the sample size used for constructing the SME failure prediction models (mainly number of firm-years).

Table 1. Distribution of papers based on the journals' ranking (that is, ABS 2021 ranking).

Category	Number of Studies
Ranked by ABS	**94**
4*	1
4	6
3	48
2	22
1	17
Not Ranked by ABS	**51**

Geographical focus of the studies

A significant proportion of SME failure studies focused on European countries. For instance, 29 papers (19 ABS and 10 non-ABS) studied Italian firms, 11 (6 ABS and 5 non-ABS) investigated German firms, and 10 (6 ABS and 4 non-ABS) studied Portuguese firms. Figure 2 shows the concentration of SME failure studies worldwide. South America, Africa, and the Middle East are relatively less studied. Table A1 in Appendix A includes a table of all countries included in the studies. Some authors analyzed more than one country in their research: Cathcart et al. (2020), Filipe et al. (2016), Karas and Režňáková (2021), Karas (2022), Malakauskas and Lakstutiene (2021), Matthias et al. (2019), Muthukumaran and Hariharanath (2023), Pederzoli et al. (2013), and Tobback et al. (2017). Three studies did not explicitly specify any locations for their data (Li et al., 2021; Zhang & Song, 2022a, 2022b).

Time horizon

Periods under study vary from only one year (for example, Lee et al. (2020)) to 66 years of data (that is, Gupta and Gregoriou (2018)). The

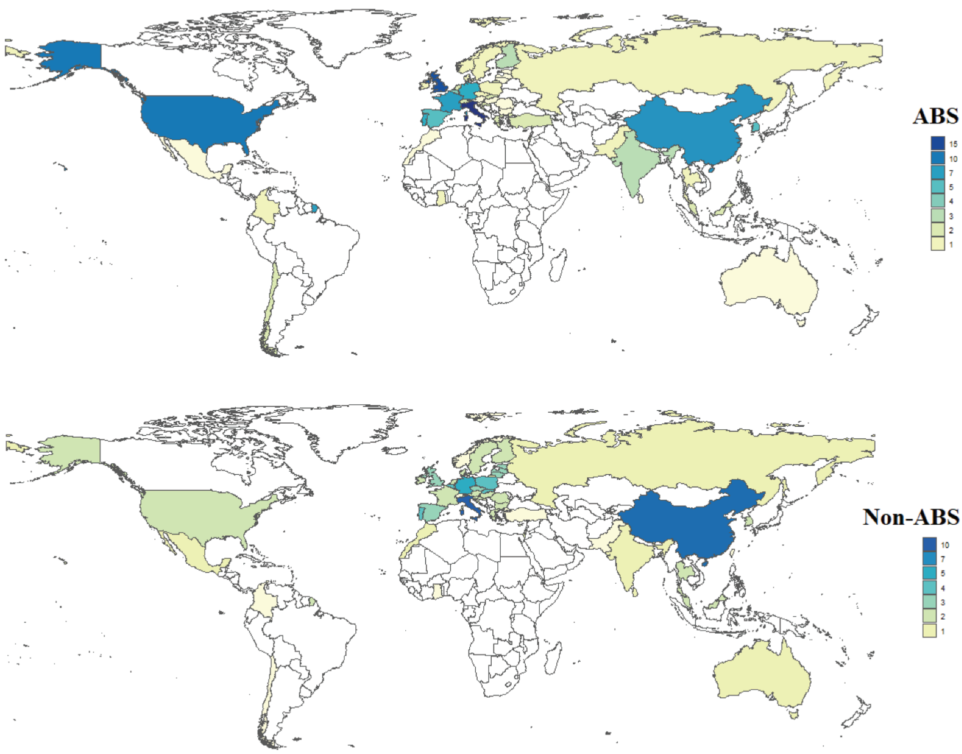


Figure 2. Geographical distribution of data used in SME default papers. The top part shows the ABS-ranked studies, and the bottom shows the non-ABS studies. *Note*: Three papers do not specify the geographical attribute of the data, and one article only defines it as "a country in South America."

average duration of the data under the study is 8.44 years (the average for ABS-ranked studies is 8.92, and for non-ABS is 7.47 years). Data from 46 studies (36 ABS and 10 non-ABS) also includes the global credit crisis of 2007–2009. Although the COVID-19 issue is relatively recent, the data in 11 studies (3 ABS and 8 non-ABS) includes this period too. Twenty-nine studies did not disclose the time horizon of their data, where 16 are ABS and 13 are non-ABS.

The forecasting horizon for studies is usually 1 year. That is, 83% of the studies predict default events within a year. Twenty-four studies, however, explicitly tried to predict defaults in longer horizons. Altman et al. (2020) models predict defaults for four different horizons: 1 year, 2–3 years, 4–5 years, and 6–10 years. DiDonato and Nieddu (2015) have studied eight different time frames for one to eight years. Lugovskaya (2010), Pacheco et al. (2022), and Pierri and Caroni (2022) test models forecasting defaults up to 5-year forecasting windows. Abdullah et al. (2016a) and Abdullah et al. (2019) built models for one to four-year prediction time frames. Glennon and Nigro (2011), Cornée (2019), and Monelos et al. (2014) studied defaults within four years. Altman et al. (2020), Ciampi et al. (2020), Laitinen (1993), Papík and Papíková (2023), Park et al. (2021), Séverin and Veganzones (2021), and Yazdanfar (2011) models forecast defaults for 1, 2, and 3-year time frames. Dewaelheyns et al. (2021), and Modina and Pietrovito (2014) studied defaults up to a 3-year horizon. The models in studies by Abdullah et al. (2016b), Ma'aji et al. (2019), Norden and Weber (2010), Svabova et al. (2020), and Zizi et al. (2021) predict financial distress within 2 years.

Sample size

Sample sizes range from four observations (Angilella & Mazzù, 2015) to over six million (Cathcart et al., 2020). Some studies defined their sample sizes regarding the number of firms, primarily when the research focuses on the default event relevant to a loan. Other studies reported their sample size in terms of the number of firm-year observations. The only study in the 1970s has a sample size equal to 42. The median for the 1980s (based on three observations) is 146. For the 1990s, there are two studies with sample sizes of 80 (Laitinen, 1993) and 216 (Lussier, 1995). The median for the 2000s is considerably higher than before, equal to 1,003. The median for the 2010s is 3,158, three times larger than the sample size median for the 2000s. The median for the 2020s is 4,039 observations. Comparing the last two decades, the 2010s and 2020s, for ABS and non-ABS, the sample size median for 2010s for ABS studies is considerably larger than for non-ABS studies; 4,262 for ABS versus 968 for non-ABS. However, in the 2020s, the median for non-ABS studies sample size is larger than ABS studies, that is, 4,354 for non-ABS studies and 2,686 for ABS.

Data sources

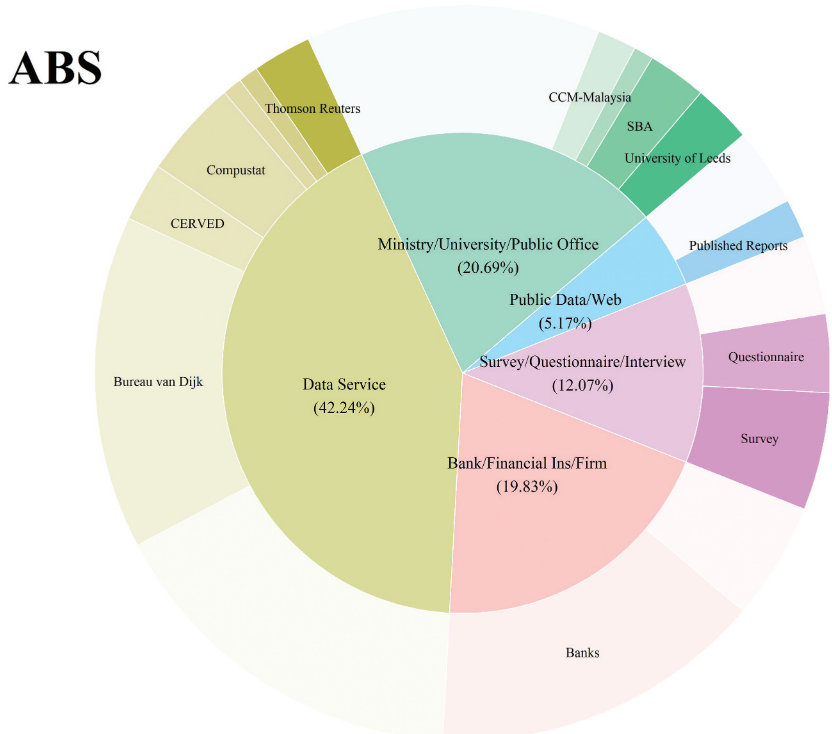
Obtaining data has been an essential aspect of SME research since Edmister (1972). However, SME data became more accessible during the two previous decades; as reported earlier, the number of observations per study has increased significantly. Regarding resources for SME failure research, more than 40% of the studies are based on data obtained from data service firms, such as Bureau van Dijk, Thomson Reuters, and Compustat. The other significant sources are ministries, public offices, and universities, which provided 20.7% of the data for ABS studies and 29.2% for non-ABS studies. Banks and financial institutions account for 19.8% of the data for ABS and 13.9% for non-ABS, where banks have the most significant share in this category in both ABS and non-ABS. Qualitative data is often collected through surveys and interviews in this domain. The percentage of surveys, questionnaires, and interviews is 12.1% for ABS and 8.3% for non-ABS. Publicly available data, like web pages or published public reports, has a share of 5.2% for ABS and 2.8% for non-ABS. Figure 3 shows the percentages of sources in SME default studies. Detailed lists of data sources and the number of studies that used those sources are available in Appendix C, see Tables C2–C6. Note that some studies have multiple data sources as they studied more than one category of factors. For example, Ciampi (2015), Ciampi (2017), and Ciampi (2018) used data from CERVED together with surveys; Ciampi et al. (2020) used CERVED along with data obtained from the Central Credit Register of Italy; and Karas and Režňáková (2021) obtained data from Amadeus (by Bureau Van Dijk), EUROSTAT, and the Transparency International Database.

Data characteristics

Financial ratios have been the most used factors in SME default models since the beginning of this research domain, such that 120 studies out of 145 used financial ratios. In contrast, 109 studies used at least one category of variables that is not in the financial ratios category. However, 93 out of the 109 studies used financial ratios and at least one type of nonfinancial ratio information. Within the nonfinancial ratio category, firm and owner/manager characteristics information is used in 43 studies, macroeconomics data in 25, and credit record information in 21 studies. Table 2 shows categories used in three or more studies. "N" denotes the number of studies (ABS plus non-ABS) in the tables where it appears.

Focus of the studies

The main focus of 91 studies is SMEs in general, while 33 studies cover small enterprises. For narrowly focused models, five studies: Abdullah et al. (2016b),



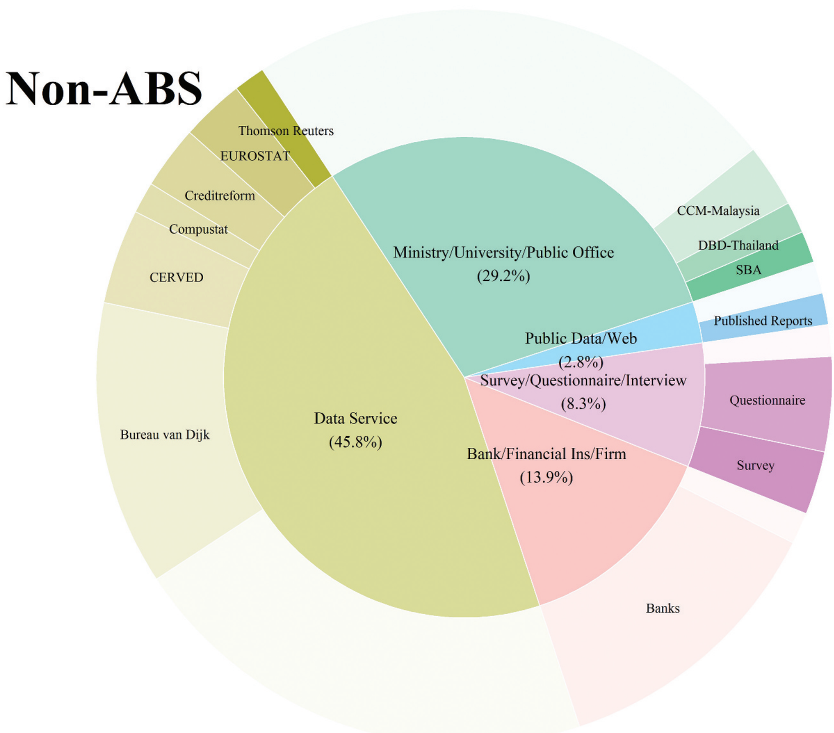


Figure 3. Distribution of data sources used in SME default papers by types. *Note*: There is one non-ABS paper that used "data from another paper," which is not presented in this figure.

Table 2. Categories of information used in three or more studies.

Data Type (Number of Studies)	ABS (94)	Non-ABS (51)	N (145)
Financial Ratios	77	43	120
Non-financial	73	41	114
Firm and Manager/Owner Characteristics	27	16	43
Macroeconomics Information	13	12	25
Credit Information	16	5	21
Relational Information	3	5	8
Loan Characteristics	4	1	5
Textual	1	2	3

Abdullah et al. (2019), Ma'aji et al. (2019), Pacheco et al. (2022), and Yin et al. (2020) cover only manufacturing SMEs. Three studies: Angilella and Mazzù (2015), Angilella and Mazzù (2019), and Pederzoli et al. (2013) study "innovative SMEs". Micro and small enterprises are studied in two papers: Bangarigadu and Nunkoo (2022), and Li et al. (2021). Mittal et al. (2011) investigate only micro-enterprises. Conversely, Dewaelheyns et al. (2021) study SMEs excluding micro-enterprises. Small industrial enterprises are studied by Sun et al. (2022), and small manufacturing enterprises are investigated by Ciampi et al. (2020). Table C1 in Appendix C shows a list of all studies' focuses.

Sample imbalance

Regarding the imbalance proportion of defaults and nondefaults, 28 studies (19 ABS and 9 non-ABS) used balanced samples, seven (4 ABS and 3 non-ABS) used almost balanced samples (the split is not exactly but close to 1:1), and 105 papers (67 ABS and 38 non-ABS) used imbalanced samples.[4] The imbalance problem is often addressed using undersampling techniques, either random undersampling or stratified random undersampling. The main difference between these two methods is that random undersampling only arbitrarily matches the number of nondefaults with the existing defaults in the sample. For instance, Altman and Sabato (2007) used undersampling by selecting the firms over the same period to match the average default rate in the sample to the expected average default rate for SMEs in the USA. However, stratified random undersampling matches the observation based on the similarity of some characteristics in both defaults and nondefaults. For example, Abdullah et al. (2016b) matched the distressed and nondistressed firms based on the asset size and industry group.

The other possible solution for the imbalance problem is to over sample. Synthetic minority oversampling technique (SMOTE), which is the most used oversampling technique in SMEs default prediction studies, is an oversampling method in that each minority class (defaults) observation creates a

[4]Four ABS studies and one non-ABS study did not disclose their sample composition.

percentage of artificial observations comparable to the majority class observations (nondefaults). This augmentation in the minority observations may enhance the trained model's classification accuracy.

Random undersampling is used by seven studies (all ABS), while 10 studies (6 ABS and 4 non-ABS) employ stratified random undersampling. However, SMOTE is used by six studies (3 ABS and 3 non-ABS[5]). Some authors used these techniques to balance the sample to 1:1 splits (Altman et al., 2022; Ciampi et al., 2020; Lee et al., 2020), while other authors tried to replicate the distribution of defaults to nondefaults in the population under study (Altman & Sabato, 2007; Calabrese et al., 2016).

Factors, features or variables included in the main models

We identified 1,205 unique factors, excluding the studies' time, sector, and location identifiers. These factors denote those that are used in the studies' final models. Among the factors, 971 are employed in one study, 124 are used in two studies, and 110 are utilized in three or more studies. A complete list of the variables used by three or more studies is available in Appendix A, see Tables A3–A5. We presented the variables under four main classes, general features (balance sheet items, financial ratios, and time-varying firm performance measures), firm characteristics (for example, size, legal form, and age), owner/manager characteristics (for instance, gender, age, and education of the manager/owner), and macroeconomics variables (such as GDP growth and interest rate). Among the main class, the quick ratio is utilized in 30 studies, the current ratio in 26 studies, net income to total assets in 26 studies, retained earnings to total assets in 22 studies, and sales to total assets in 20 studies. The most used firm characteristics are firm age (31 studies), the natural logarithm of total assets (17 studies), the number of partners (11 studies), the natural logarithm of firm age (10 studies), and number of employees (nine studies). It is worth noting that the natural logarithm of the total assets, the number of employees, and the natural logarithm of sales are taken as size measures among studies. The top four employed owner/manager characteristic features are education (14 studies), management experience (12 studies), age of owner/debtor/legal representative (11 studies) of the owner/manager, and if the owner's parents owned a business (11 studies). The most used macroeconomics variable is GDP growth, utilized in six studies.

Moreover, time, sector, and location identifiers are used in several studies. Time dummies are utilized in nine studies, varying from year dummies to dummies for a certain period. Sector dummies are used in 34 studies, mainly identifiers for a firm's business sector. Geographical location identifiers are

[5]One non-ABS study used SMOTE and six variations of under and oversampling (nWSMOTE, nWSMOTE-ensemble, nRUS, nMChanUS, nUSOS, and nRUSSMOTE); that is, Abedin et al. (2022).

utilized in 27 studies, such as country identifiers, county identifiers, city identifiers, dummies for cardinal locations within a country, and district identifiers.

Number of factors in the main models

The number of predictor factors used in the final models per study ranges from one to 52 for ABS studies and one to 79 for non-ABS studies. On average, 13 factors are presented in the primary models among all ABS studies and 12 for non-ABS studies. Dividing the period under the review into two sub-periods, that is, pre- and post-2000, the average features per model for the first period is approximately seven. The second-period models utilize 13 features per model. Table 3 shows the number of variables per model grouped by decades. The maximum number of variables per model shows an upward trend.[6] This can be due to the ability of the advanced estimation methods to work with highly correlated variables and having relatively fewer assumptions about the distribution of indicators such that models can accommodate more features.

Variable transformation and winsorization

Regarding dealing with outliers, 35 studies (24 ABS and 11 non-ABS) used a technique to deal with them. Fifteen studies utilized winsorization. Winsorizing at the 1st and 99th percentiles is the most popular method, which is used by 11 studies; for example, Altman et al. (2020), Gupta et al. (2014a), Karas and Režňáková (2021), and Karas (2022). Three studies winsorized their factors at 5th and 95th percentiles, that is, Andrikopoulos and Khorasgani (2018), El Kalak and Hudson (2016), and Wilson et al. (2016). Another transformation like logarithmic transformation is used by four

Table 3. Number of variables that are used in SME default models.

	Minimum	Maximum	Average (rounded)
ABS (94 Studies)			
1970s	7	7	7
1980s	5	6	6
1990s	3	15	8
2000s	4	29	13
2010s	1	50	12
2020s	5	52	16
Overall	1	52	13
Non-ABS (51 Studies)			
1970s			
1980s	5	9	7
1990s			
2000s	16	16	16
2010s	3	24	11
2020s	1	79	12
Overall	1	79	12

[6]Please note that there is only one paper for the 1970s and two studies for the 1990s.

studies: Altman and Sabato (2007), Angelini et al. (2008), Lextrait (2023), and Sigrist and Hirnschall (2019); hyperbolic tangent transformation is utilized in two studies: Inekwe (2016), and Piatti et al. (2015). Removing outliers is used in five studies; for example, Pacheco et al. (2022), and Grishunin et al. (2021).

Feature selection techniques

Although more advanced estimation methods might be able to deal with more input features without causing severe problems, collecting information involves considerable costs. Moreover, some studies show that having fewer variables in models results in better predictive performance or the same performance level with the models including more variables (Kou et al. 2021). Only 88 out of 145 studies utilized at least one feature selection method, which accounts for 56 out of 94 ABS and 33 out of 51 for non-ABS. Furthermore, 33 ABS and 14 non-ABS studies used more than one technique. In total, 54 unique variable selection techniques are used in the studies. Table 4 shows variable selection methods used in at least three studies. The most utilized techniques are the forward stepwise method and correlation analysis. In the forward stepwise approach, one starts with an empty model, and the model is constructed by adding the most significant features. In correlation analysis, when a pair of variables are highly correlated, the one with the highest significance level is kept in the model. The third most utilized approach is removing a variable with strong multicollinearity.

Estimation methods

The first model for small business bankruptcy prediction in a study by Edmister (1972) employed multiple discriminant analysis (MDA). Keasey

Table 4. Feature selection techniques utilized in two or more studies.

Variable Selection Method	ABS	Non-ABS	N
(Number of Studies)	(94)	(51)	(145)
Correlation Analysis	18	8	26
Forward Stepwise Selection	17	9	26
VIF	9	4	13
Backward Stepwise Elimination	4	5	9
Stepwise Method[a]	6	3	9
PCA	4	3	7
Univariate Analysis	4	2	6
Average Marginal Effect (AME)	4	0	4
Significance	3	1	4
Wrapper Method	2	2	4
LASSO	2	1	3
RF Feature Selection Method	3	0	3

VIF stands for variance inflation factor, which shows the multicollinearity between the features. PCA stands for principal component analysis. It is often used to find a subset of variables that explains most of the variation in the data. LASSO stands for least absolute shrinkage and Selection operator. Significance stands for keeping only (the most) significant variables in the model. RF in RF feature selection method stands for random forests.

[a]It is not specified if the method is forward or backward.

and Watson (1986) and Keasey and Watson (1988) also used discriminant analysis in this subject. However, after the conditional logit model was applied to the default prediction studies by Ohlson (1980) for the first time, logit became and remained the most utilized estimation method in this research domain. Table 5 shows estimation methods utilized in two or more SMEs failure studies grouped by decades. Logit is used in 77 studies (46 ABS and 31 non-ABS), neural network (NN) and discriminant analysis (DA) each in 14 studies, support vector machine (SVM) in 13 papers, and random forests (RF) in 11 research. We identified 80 unique estimation methods utilized in SME failure studies as the primary estimation method, where 59 of the model types were only employed by one study. Most of the studies have only one primary estimation method. However, 39 studies have more than one primary estima-

Table 5. Estimation methods that are used in two or more SME default papers as the primary models.

	1970s	1980s	1990s	2000s	2010s	2020s	Overall
ABS (94 Studies)							
Logit	0	0	2	5	26	13	46
SVM	0	0	0	0	4	6	10
NN	0	0	0	2	2	5	9
DA	1	2	0	0	4	2	9
Disc-t H.	0	0	0	1	4	1	6
Probit	0	0	0	0	5	1	6
RF	0	0	0	0	1	5	6
XGBoost	0	0	0	0	0	6	6
DT	0	0	0	0	1	4	5
Cox P. H.	0	0	0	0	3	1	4
k-NN	0	0	0	0	2	2	4
ELECTRE-TRI	0	0	0	0	2	0	2
MURAME	0	0	0	0	1	1	2
Panel Logit	0	0	0	0	2	0	2
LightGBM	0	0	0	0	0	2	2
CatBoost	0	0	0	0	0	2	2
MLP	0	0	0	0	1	0	1
L-Reg	0	0	0	0	0	1	1
Other	0	2	0	0	17	12	31
Non-ABS (51 Studies)							
Logit		1		0	13	17	31
RF		0		0	0	5	5
DA		0		0	4	1	5
NN		0		0	1	4	5
XGBoost		0		0	0	4	4
SVM		0		0	0	3	3
LightGBM		0		0	0	3	3
Cox P. H.		0		0	1	2	3
CART		0		0	1	1	2
RSF		0		1	0	1	2
Probit		0		0	2	0	2
CNN		0		0	0	2	2
Disc-t H.		0		0	0	1	1
Panel Logit		0		0	1	0	1
DT		0		0	0	1	1
MLP		0		0	0	1	1
L-Reg		0		0	0	1	1
Other		0		1	5	29	35

NN, neural network; SVM, support vector machines; RF, random forests; XGBoost, eXtreme gradient boosting; DA, discriminant analysis; Cox P. H., Cox proportional hazards model; Disc-t H., for discrete-time hazards model; DT, decision tree; LightGBM, light gradient boosting machine; k-NN, k-nearest neighbors algorithm; ELECTRE-TRI, elimination and choice translating reality – tree; MURAME, multicriteria ranking method; MLP, multilayer perceptron; CART, classification trees, RSF, random survival forests; L-Reg, linear regression; CatBoost, categorical boosting; CNN, convolutional neural network.

tion method; for example, Altman et al. (2020), Ciampi et al. (2020), Figini et al. (2017), Kou et al. (2021), and Zhang and Song (2022b). Table A2 in Appendix A shows the number of papers, grouped by decade, which studied more than one primary estimation method, along with studies that tested more than one method but not as their primary methods (66 papers); for instance, a different method tested as a robustness check of the primary estimation method. We distinguished the studies as follows: in the case of multiple model types, if a study compares model types and does not a priori take a model as its primary model type, the study has more than one primary estimation method.

Validation methods

Keasey and Watson (1987) used a hold-out sample to test their results for the first time in this research domain. Although using hold-out samples was known very early in this research subject, 36 studies used in-sample validation. Some may justify their usage of in-sample validation by their small sample size. However, Isaksson et al. (2008) suggest that even cross-validation and bootstrapping are unreliable as validation approaches when the sample is small and they suggested using a simple holdout test. Kim (2009) compared bootstrapping and cross-validation (as a traditional validation method) and concluded that cross-validation outperforms bootstrapping. Thus, using in-sample validation when the goal is to predict failure is not justifiable in today's research.

Table 6 presents the validation methods used in SME default studies in each decade. Seventy-seven studies (54 ABS and 23 non-ABS) utilize the hold-out sample, and 15 (5 ABS and 10 non-ABS) use cross-validation. Using cross-validation increased during the past 22 years, such that out of 10 studies in the 2000s, no study used cross-validation; in 2010s, only four out of 64 studies utilized cross-validation, and from 2020 to early in 2023, 10 studies out of 49 employed cross-validations. A total of 12 studies either did not report their validation strategies or validation was not necessary based on the nature of their studies.

Model performance

Model performance is measured based on various measures. The two most repeated measures are error rates, in terms of type I and type II errors, and the area under the receiver operating characteristic (ROC) curve, known as AUC in short or AUC(ROC). Table A6 in Appendix A shows the most repeated performance measures. Type I and type II errors are reported in 74 studies, and AUC(ROC) in 68 studies. Type I errors show the false-positive or when a nondefaulted firm is classified as defaulted. Type II errors denote the false-

Table 6. Validation methods that are used in SME default studies.

	Hold out sample	Cross-validation	In-sample
ABS (94 Studies)			
1970s	0	0	1
1980s	0	0	2
1990s	0	1	1
2000s	6	0	2
2010s	34	2	15
2020s	14	2	4
Overall	54	5	25
Non-ABS (51 Studies)			
1970s			
1980s	1	0	0
1990s			
2000s	2	0	0
2010s	8	2	7
2020s	12	8	9
Overall	23	10	16

Seven ABS studies and one non-ABS study either did not report their validation strategies or the validation was not required based on their study scope.
Two ABS studies did not use any validation methods.
Two studies (one ABS and one non-ABS) used another dataset with different variables to validate their results.

negative or when the model misclassifies the defaulted firm as nondefaulted. Usually, studies report an "accuracy ratio" based on these two errors, which is a weighted average of one minus each type of error. However, such an accuracy ratio is subject to a cutoff point designated to the prediction model. The cutoff point refers to a value between 0 and 1, separating defaulted and nondefaulted firms based on the probability of defaults (PD). While the accuracy ratio might seem to be sufficient on most occasions, AUC(ROC) is superior since it is not impacted by the cutoff points; it reports the area under the imaginary curve where all the cutoff points are accounted for.

We ranked the best models of each study, regardless of whether the study has more than one primary estimation method, based on AUC(ROC) and accuracy ratio; each study has only one model in the ranking. Only models validated with hold-out samples or cross-validation are considered. The main reason is that in-sample accuracy is often higher than hold-out sample accuracy, and in-sample validation is not an appropriate measure to report prediction accuracy. Table 7 shows the top 10 ranked studies based on AUC (ROC) and accuracy ratio grouped by the estimation method, such that the best measure for each method is reported. Tables A7 and A8 in Appendix A show the top 20 performing models grouped by studies. The top three models based on AUC(ROC) are gradient boosting decision tree-convolutional neural network-logistic regression (GBDT-CNN-LR), random forests (RF), support vector machine (SVM) where the AUC(ROC)s are almost the same in the first two (0.992 vs. 0.991) and not that lower in the third model (0.988). Based on the accuracy ratio, the best-performing model is RF, with an overall accuracy of 99.1% in a study by Abedin et al. (2022). The discrepancy between the two

Table 7. Top 10 performing models based on the AUC(ROC) and accuracy ratio by estimation methods.

Rank	AUC(ROC)	Estimation Method	Study	ABS
Panel A				
1	0.992	GBDT-CNN-LR	Zhang and Song (2022a)	No
2	0.991	RF	Abedin et al. (2022)	No
3	0.988	SVM	Sun et al. (2022)	Yes
4	0.984	Soft voting	Gao et al. (2021)	No
5	0.973	CNN-Logistic-Stacking	Zhang and Song (2022b)	No
6	0.970	XGBoost (Focal Loss)	Sun and Jiao (2022)	No
7	0.963	Elman network	Corazza et al. (2021)	Yes
8	0.959	Logit	Zizi et al. (2021)	No
9	0.956	Cox Proportional Hazards	Gupta and Gregoriou (2018)	Yes
10	0.949	NN	Da and Peng (2022)	Yes
Rank	Accuracy Ratio	Estimation Method	Study	
Panel B				
1	99.1%	RF	Abedin et al. (2022)	No
2	97.7%	MDA	Terdpaopong and Mihret (2011)	Yes
3	97.1%	NN	Da and Peng (2022)	Yes
4	96.8%	CNN	Zeng (2022)	No
6	96.2%	Logit	Abdullah et al. (2016a)	Yes
8	93.8%	XGBoost (Focal Loss)	Sun and Jiao (2022)	No
9	93.6%	Elman network	Corazza et al. (2021)	Yes
11	91.6%	CART	DiDonato and Nieddu (2015)	No
12	91.5%	LPM	Figini and Giudici (2011)	Yes
15	90.8%	CatBoost	Papík and Papíková (2023)	Yes

Panel A shows the AUC(ROC) results, and Panel B corresponds to accuracy ratio. studies that used hold-out samples and cross-validation as their validation methods are shown in this table. The first column shows the model's ranking grouped by studies. The AUC(ROC) ranking grouped by studies and estimation methods is the same.
GBDT-CNN-LR, gradient boosting decision tree-convolutional neural network-logistic regression; LPM, parametric longitudinal models.

panels in Table 7 is due to some studies only reporting AUC(ROC) and some only accuracy ratios.

Discussion

Potential bias from data sources

Data has been an issue from the beginning of SME default prediction research area. However, technological development has changed the situation considerably. Data collection and management became less complicated with the introduction of computers and the Internet, and data service firms started to serve data to researchers and practitioners. Nearly half of the papers we reviewed obtained the data from data service firms. Moreover, technological advances made it easier for public authorities and the private sector to store information in a more efficient and accessible way.

Although data is relatively more readily available these days, it does not necessarily mean that data accurately represents the population under study. For example, data collected through a bank or financial institution often includes firms that applied for financing and even only those firms that received the funding. By the Basel Capital Accord II implementation in

2004, banks were mandated to use internal rating systems. This means the bank's internal rating system has initially filtered firms in the bank portfolios. Therefore, the sample of firms in the bank database may not represent the whole population accurately due to this selection bias. This bias is less pronounced when data is obtained from data service firms or public authorities since those databases often include firms with various financing sources. For example, Lussier (1995) introduced a two-step sampling process that can potentially reduce the selection bias; that is, the positive cases (failures) were collected at the first stage from bankruptcy court records, and the negative cases (nonfailures) were matched based on industry and geographical location in the second stage.

Dealing with sample imbalance

While sample imbalance has not been addressed at all or appropriately in more than 70% of the studies on SME failure prediction, it appears to be a source of inaccuracy. Veganzones and Séverin (2018) show that the performance of the models, which are built on samples with the minority class representing less than or equal to 20% of the whole population, are significantly inferior. However, they reported that SVM is less sensitive to the imbalance than the other estimation methods, only showing a noticeable decrease in performance when the minority class is equal to or less than 10%. They also suggested that oversampling is the optimal choice for dealing with sample imbalance as it is the most suitable technique for all estimation methods and sample sizes. Abedin et al. (2022) tested six different sample re-balancing strategies and demonstrated the same conclusion that oversampling outperforms nonsampling and undersampling. They also reported that the SVM does not significantly benefit from sample re-balancing, confirming the result of Veganzones and Séverin (2018). Moreover, Yin et al. (2020) show that oversampling (SMOTE) significantly improves the performance of RF while it does not impact the performance of XGBoost. Contrary to the abovementioned studies in favor of oversampling, Piatt and Piatt (2002) concluded that oversampling might cause choice-based sample bias due to a nonrandom sample created from oversampling of the defaulted firms.

Feature selection

About 40% of the studies did not use any statistical variable selection methods. Although Du Jardin (2009) discussed the importance of the variable selection techniques in failure models performance, more than 85% of papers without reporting any variable selection techniques were written after Du Jardin (2009) study, that is, 2010 to early 2023. While some new estimation techniques account for potential problems sourcing from collinearity, multicollinearity, and irrelevancy of a subset of variables, this does not mean that having hundreds of variables does not impose the cost of collecting them. Moreover, leaving the variable selection to some internal hidden features of some advanced estimation methods reduces the generality of a model. That is, a model assumes to be an appropriate prediction model if it works equally well with the same subset of variables on another sample.

Validation methods

Although the issue of improper in-sample validation is discussed by Bellovary et al. (2007), about 30% of studies published afterward still used in-sample validation. Even though one may argue that the goal of a study may not be to construct an excellent predictive model but a descriptive model, this argument cannot wholly justify not using a hold-out sample or cross-validation.

Model performance

Regarding reporting predictive model performance, type I and type II errors and AUC(ROC) are standard measures to report. However, about 50% of the studies did not report type I and type II errors; this also holds for AUC(ROC).

Among the top five estimation methods used in the studies,[7] SVM has the highest AUC(ROC) on average (0.9175), NN comes right after with AUC (ROC) averages at 0.9143, RF with AUC(ROC) on average equal to (0.8288), and logit is the last in the list by average equal to (0.8225). Models constructed using discriminant analysis did not report AUC(ROC). For the accuracy ratio, RF comes first (99.1% based on one study), discriminant analysis second (88.35%), NN third (83.86%), and logit fourth (82.86%). Only one study reported accuracy ratio for SVM, which is relatively lower than logit (74.3%). While logit has been used the most, it has the lowest accuracy of the most used models overall, the last under AUC(ROC), and fourth under accuracy ratio. In general, RF, NN, CNN, and stacked models (CNN-LR, for instance) are observed to outperform logit.

ABS versus non-ABS

We checked the difference between articles published in ABS-ranked journals and non-ABS journals for each dimension of the methodology we presented earlier in the results. We show pronounced differences between these two groups in this subsection.

The largest difference between ABS and non-ABS studies is in the country they study. ABS studies usually investigate Western coutries,

[7]Only models tested with hold out samples or cross-validation are considered.

while non-ABS studies are usually focused on China and Eastern Europe. About 15% of ABS-published articles have studied the UK, about 10.5% studied the USA, and 7.5% studied France. These percentages are 6%, 4%, and 4% for non-ABS published articles, respectively. Furthermore, 17.5% of non-ABS articles studied China, while only 8.5% of ABS articles focused on China. Moreover, non-ABS studies have focused relatively more on Slovakia, Poland, the Czech Republic, Estonia, Latvia, and Lithuania.

The median size of the samples used in ABS and non-ABS studies are almost similar, 2,681 for ABS versus 2,558 for non-ABS. The time horizon under study is longer for ABS papers, approximately 9 years on average, versus 7.5 years for non-ABS.

ABS and non-ABS studies obtained more than 40% of their data from data service providers. However, the proportion of the data obtained from ministries, public offices, and universities is 29% for non-ABS while 20% for ABS articles. Banks, financial institutions, and firms provided data to 20% of ABS studies and 14% of non-ABS studies. The sample imbalance problem has been addressed in 17% of ABS studies and 13.5% of non-ABS studies. At least one statistical feature selection method has been used by 60% of ABS studies and 65% of non-ABS articles. Comparing primary estimation methods, logit is used in 49% of ABS studies and 61% in non-ABS studies. Although logit has more pronounced domination in non-ABS studies, the newer estimation methods, like XGBoost and LightGBM, are also used relatively more frequently in non-ABS studies.

The hold-out sample has been used in 57.5% of ABS studies to validate the results. However, only 45% of the non-ABS studies have used out-of-sample validation. Cross-validation and in-sample validation are less used in ABS studies compared to non-ABS studies. Finally, AUC has been reported more often in non-ABS studies (55% versus 42.5% for ABS studies).

Limitations of this review and future research

In this section, we discuss the limitations of this review, and based on these limitations, we suggest avenues for further research.

Limitations of this review

In this review, we summarized which methods and variables (features) are used in existing studies, and how often. While in some cases, the number of appearances of a variable in the final model of a study shows that the variable is potentially a good predictor of failure, this is not necessarily applicable to the number of appearances of an estimation method. The number of times an estimation method has been used in previous studies does not represent whether it is superior. A particular method can be used frequently due to the ease of implementation or because it was introduced relatively early to the research community. Even though some methods were introduced for a long time, some barriers existed to apply them. Those barriers can be, for example, a lack of expertise to implement them and the limited computational power of computers available to researchers at the time. Thus, the frequency that an estimation method appears as the primary estimation method in previous studies does not support that the estimation method is the best choice today.

Although we also have presented the performance of estimation methods reported by previous studies, the degree to which the conclusion about the method superiority is limited across different studies since the models are not constructed uniformly, data compositions are dissimilar, and even the definition of failure among studies is different.[8] In some instances, conclusions can be drawn relatively easily; for example, the superiority of the hold-out sample validation approach is obvious. However, some suggestions are weaker. For instance, we know that re-balancing the sample increases the model performance. However, which strategy is superior combined with a specific estimation method is still a question for future research.

In summary, the main limitation of our study is that it is a review. Therefore, we can conclude which methods are used most frequently, but we are not able to conclude which methods work the best. We elaborate more on this in the next subsection.

Avenues for future research

Since researchers and practitioners need to know which method(s) work the best, our main recommendation for future research is that future studies should consider in-depth empirical comparisons of estimation methods, feature selection techniques, and sample re-balancing methods over large and commonly used datasets so that it can be concluded which methods are expected to work the best. This can be done by combining estimation methods, feature selection techniques, and sample re-balancing strategies. This way, one can systematically check which sample re-balancing strategy is more suitable for a specific estimation method. The same is true for the feature selection technique and estimation method. Especially systematic studies that compare re-balancing (under and oversampling) strategies can greatly contribute to the field, as the number of bankrupt or defaulted firms, both in existing datasets and in the real world, is usually significantly smaller than the number of operational firms.

[8]It is worth mentioning that there are some studies within the reviewed papers that specifically compare various aspects of methodologies. While the deduction can be relatively more straightforward in this case, the conclusion is still not terminal. This is evident in the lack of consensus between the results of studies comparing estimation methods.

The Omega score (Altman et al., 2022, 2023) has lately been developed using data from Croatia. This approach should be applied to the other nations' data. Moreover, its out-of-sample predictive performance should be compared with the performance of optimized predictive models employing various estimations, feature selection, and re-balancing methods.

An additional avenue for future research on default prediction methodology is to investigate how the data source bias in this research domain impacts the predictive performance and generality of the models. This can be implemented by validating the predictive power of models on independently collected data from various sources. For example, within the same economy, if the model is constructed based on data obtained from a bank, the predictive power is tested on data collected from data service providers. Then, a comparison of how well the model performs on predicting default on a test sample from the same data source and the data from a different data source can reveal the potential discrepancies.

Lastly, we note that the listed avenues for further research are from a methodological perspective. For a reader searching for a broader set of possible research topics within the field of SME default prediction, we recommend Ciampi et al. (2021).

Conclusion

An up-to-date methodology-focused review of the SME's failure prediction can help give researchers a clear overview of the methodologies used and the methods' constituents. Such an overview can save considerable time in this field's early stages of research. The present study contributes to the previous research on review of the literature (Balcaen & Ooghe, 2006; Bellovary et al., 2007; Du Jardin, 2009). The review is also narrowed down to only SME studies. Moreover, the added contribution is that this paper summarized data sources and imbalance problem solutions. This paper also contributes to Ciampi et al. (2021) by reviewing the studies from the methodological aspects.

Compared to six decades ago, data availability is significantly higher for SME studies, from less than 50 observations in the relatively low dimensional dataset in Edmister (1972) to studies based on high dimensional datasets with millions of observations. However, academics and practitioners should be aware that some data sources may have idiosyncratic characteristics that might not necessarily represent the intended population under study, such as data that only includes accepted loans. Since the dimension of available data has grown noticeably, the variable selection techniques are more critical than before. A model with many factors may not be easily generalized. Furthermore, having more data available sometimes means more majority (nondefaults) cases and less proportion of minority cases (defaults). Traditionally, random undersampling is employed in such cases. However, oversampling, particularly SMOTE, appears to be more appropriate for this type of study (Abedin et al., 2022; Veganzones & Séverin, 2018), while Piatt and Piatt (2002) argued that it might cause a choice-based sample bias. Another benefit of the larger available datasets is that models can be tested on hold-out samples or using cross-validation. This makes in-sample validation less credible than it used to be.

We observed that over one-third of studies do not report utilizing feature selection methods. Thus, we recommend that future default forecasting research and similar event prediction studies consider appropriate feature selection methods. Although the necessity of using hold-out sample validation was known from the very beginning of this research domain, more than one-quarter of the studies use in-sample validation. Using hold-out samples or cross-validation to validate the results is highly recommended since in-sample validation often falsely shows higher predictive performance.

Moreover, technological advances bring about new estimation techniques, often available to everyone. Although we identified 80 unique estimation methods used as primary estimation methods in studies, logit is still used in more than half of them, often as the sole primary estimation method. Even though logit is the most popular model in this research area, its popularity does not justify using it without testing at least one proven better-performing estimation method. While the logit model's predictive ability is often acceptable, we recommend trying machine learning methods to improve the accuracy of the predictive models.

In addition, more than half of the studies do not report either type I and type II errors or AUC(ROC). This makes it difficult, if not impossible, for readers to compare the results of the studies that do not report standard predictive performance measures with other studies' results. Moreover, not reporting type I and type II errors gives no information on how a model misclassifies the defaults as healthy firms and vice versa. Therefore, the predictive ability of the models should be reported in a comparable form in terms of AUC(ROC), type I, and type II errors.

We also investigate this topic separately for journals ranked in the ABS ranking and remaining journals. From a methodological perspective, we do not observe large differences between these two groups. The largest difference is that Western countries are relatively more frequently studied in ABS journals, while non-ABS studies are more often focused on China and Eastern Europe.

Our review of SME default prediction methodologies gives researchers a comprehensive gateway to potential data sources and commonly used techniques, together with an overview of the most common methodological shortcomings in the existing literature. This will help researchers avoid these shortcomings and contribute to faster development of this critical field.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Hamid Cheraghali http://orcid.org/0000-0002-9171-3473
Peter Molnár http://orcid.org/0000-0003-1354-0113

References

Abdullah, N. A. H., Ahmad, A. H., Zainudin, N., & Rus, R. M. (2016a). Modelling small and medium-sized enterprises' failure in Malaysia. *International Journal of Entrepreneurship & Small Business*, *28*, 101–116. https://doi.org/10.1504/IJESB.2016.075686

Abdullah, N. A. H., Ahmad, A. H., Zainudin, N., & Rus, R. M. (2019). Predicting financially distressed small- and medium-sized enterprises in Malaysia. *Global Business Review*, *20*, 627–639. https://doi.org/10.1177/0972150919837053

Abdullah, N. A. H., Ma'aji, M. M., & Khaw, K. L.-H. (2016b). The value of governance variables in predicting financial distress among small and medium-sized enterprises in Malaysia. *Asian Academy of Management Journal of Accounting and Finance*, *12*(Suppl. 1), 75–88. https://doi.org/10.21315/aamjaf2016.12.S1.4

Abedin, M. Z., Guotai, C., Hajek, P., & Zhang, T. (2022). Combining weighted SMOTE with ensemble learning for the class-imbalanced prediction of small business credit risk. *Complex & Intelligent Systems*, *9*(4), 3559–3579. https://doi.org/10.1007/s40747-021-00614-4

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, *23*, 589–609. https://doi.org/10.1111/j.1540-6261.1968.tb00843.x

Altman, E. I., Balzano, M., Giannozzi, A., & Srhoj, S. (2022). Revisiting SME default predictors: The omega score. *Journal of Small Business Management*, *61*(6), 2383–2417. https://doi.org/10.1080/00472778.2022.2135718

Altman, E. I., Balzano, M., Giannozzi, A., & Srhoj, S. (2023). The omega score: An improved tool for sme default predictions. *Journal of the International Council for Small Business*, *4*(4), 362–373. https://doi.org/10.1080/26437015.2023.2186284

Altman, E. I., Esentato, M., & Sabato, G. (2020). Assessing the credit worthiness of Italian SMEs and mini-bond issuers. *Global Finance Journal*, *43*. https://doi.org/10.1016/j.gfj.2018.09.003

Altman, E. I., Iwanicz-Drozdowska, M., Laitinen, E. K., & Suvas, A. (2020). A race for long horizon bankruptcy prediction. *Applied Economics*, *52*, 4092–4111. https://doi.org/10.1080/00036846.2020.1730762

Altman, E. I., & Sabato, G. (2007). Modelling credit risk for SMEs: Evidence from the U.S. market. *Abacus*, *43*, 332–357. https://doi.org/10.1111/j.1467-6281.2007.00234.x

Andrikopoulos, P., & Khorasgani, A. (2018). Predicting unlisted SMEs' default: Incorporating market information on accounting-based models for improved accuracy. *The British Accounting Review*, *50*, 559–573. https://doi.org/10.1016/j.bar.2018.02.003

Angelini, E., DiTollo, G., & Roli, A. (2008). A neural network approach for credit risk evaluation. *The Quarterly Review of Economics & Finance*, *48*(4), 733–755. https://doi.org/10.1016/j.qref.2007.04.001

Angilella, S., & Mazzù, S. (2015). The financing of innovative SMEs: A multicriteria credit rating model. *European Journal of Operational Research*, *244*, 540–554. https://doi.org/10.1016/j.ejor.2015.01.033

Angilella, S., & Mazzù, S. (2019). A credit risk model with an automatic override for innovative small and medium-sized enterprises. *Journal of the Operational Research Society*, *70*, 1784–1800. https://doi.org/10.1080/01605682.2017.1411313

Arslan, Ö., & Karan, M. B. (2009). Credit risks and internationalization of SMEs. *Journal of Business Economics and Management*, *10*, 361–368. https://doi.org/10.3846/1611-1699.2009.10.361-368

Baidoun, S. D., Lussier, R. N., Burbar, M., & Awashra, S. (2018). Prediction model of business success or failure for Palestinian small enterprises in the west bank. *Journal of Entrepreneurship in Emerging Economies*, *10*, 60–80. https://doi.org/10.1108/JEEE-02-2017-0013

Baixauli, J. S., & Módica-Milo, A. (2010). The bias of unhealthy SMEs in bankruptcy prediction models. *Journal of Small Business and Enterprise Development*, *17*, 60–77. https://doi.org/10.1108/14626001011019134

Balcaen, S., & Ooghe, H. (2006). 35 years of studies on business failure: An overview of the classic statistical methodologies and their related problems. *The British Accounting Review*, *38*(1), 63–93. https://doi.org/10.1016/j.bar.2005.09.001

Balzano, M. (2022). Serendipity in management studies: A literature review and future research directions. *Management Decision*, *60*, 130–152. https://doi.org/10.1108/md-02-2022-0245

Bangarigadu, K., & Nunkoo, R. (2022). A success versus failure prediction model for small firms. *Journal of African Business*, *24*(4), 529–545. https://doi.org/10.1080/15228916.2022.2119015

Behr, P., & Güttler, A. (2007). Credit risk assessment and relationship lending: An empirical analysis of German small and medium-sized enterprises. *Journal of Small Business Management*, *45*, 194–213. https://doi.org/10.1111/j.1540-627X.2007.00209.x

Bellovary, J. L., Giacomino, D. E., & Akers, M. D. (2007). A review of bankruptcy prediction studies: 1930 to present. *Journal of Financial Education*, *33*, 1–42.

Calabrese, R., Marra, G., & Osmetti, S. A. (2016). Bankruptcy prediction of small and medium enterprises using a flexible binary generalized extreme value model. *Journal of the Operational Research Society*, *67*, 604–615. https://doi.org/10.1057/jors.2015.64

Caselli, S., Corbetta, G., Cucinelli, D., & Rossolini, M. (2021). A survival analysis of public guaranteed loans: Does financial intermediary matter? *Journal of Financial Stability*, *54*. https://doi.org/10.1016/j.jfs.2021.100880

Castillo, J. A., Mora-Valencia, A., & Perote, J. (2018). Moral hazard and default risk of SMEs with collateralized loans. *Finance Research Letters*, *26*, 95–99. https://doi.org/10.1016/j.frl.2017.12.010

Cathcart, L., Dufour, A., Rossi, L., & Varotto, S. (2020). The differential impact of leverage on the default risk of small and large firms. *Journal of Corporate Finance*, *60*. https://doi.org/10.1016/j.jcorpfin.2019.101541

Chai, N., Wu, B., Yang, W., & Shi, B. (2019). A multicriteria approach for modeling small enterprise credit rating: Evidence from China. *Emerging Markets Finance and Trade*, *55*, 2523–2543. https://doi.org/10.1080/1540496X.2019.1577237

Chen, Y., Huang, R. J., Tsai, J., & Tzeng, L. Y. (2015). Soft information and small business lending. *Journal of Financial Services Research*, *47*, 115–133. https://doi.org/10.1007/s10693-013-0187-x

Chi, G., & Meng, B. (2019). Debt rating model based on default identification: Empirical evidence from Chinese small industrial enterprises. *Management Decision*, *57*, 2239–2260. https://doi.org/10.1108/MD-11-2017-1109

Chi, G., & Zhang, Z. (2017). Multi criteria credit rating model for small enterprise using a nonparametric method. *Sustainability*, *9*(10), 1834. https://doi.org/10.3390/su9101834

Ciampi, F. (2015). Corporate governance characteristics and default prediction modeling for small enterprises. An empirical analysis of Italian firms. *Journal of Business Research*, *68*, 1012–1025. https://doi.org/10.1016/j.jbusres.2014.10.003

Ciampi, F. (2017). The potential of top management characteristics for small enterprise default prediction modelling. *WSEAS Transactions on Business and Economics*, *14*, 397–408.

Ciampi, F. (2018). Using corporate social responsibility orientation characteristics for small enterprise default prediction. *WSEAS Transactions on Business and Economics*, *15*, 113–127.

Ciampi, F., Cillo, V., & Fiano, F. (2020). Combining kohonen maps and prior payment behavior for small enterprise default prediction. *Small Business Economics*, *54*, 1007–1039. https://doi.org/10.1007/s11187-018-0117-2

Ciampi, F., Giannozzi, A., Marzi, G., & Altman, E. I. (2021). Rethinking SME default prediction: A systematic literature review and future perspectives. *Scientometrics*, *126*, 2141–2188. https://doi.org/10.1007/s11192-020-03856-0

Ciampi, F., & Gordini, N. (2013). Small enterprise default prediction modeling through artificial neural networks: An empirical analysis of Italian small enterprises. *Journal of Small Business Management*, *51*, 23–45. https://doi.org/10.1111/j.1540-627X.2012.00376.x

Corazza, M., De March, D., & DiTollo, G. (2021). Design of adaptive Elman networks for credit risk assessment. *Quantitative Finance*, *21*, 323–340. https://doi.org/10.1080/14697688.2020.1778175

Corazza, M., Fasano, G., Funari, S., & Gusso, R. (2021). MURAME parameter setting for creditworthiness evaluation: Data-driven optimization. *Decisions in Economics and Finance*, *44*(1), 295–339. https://doi.org/10.1007/s10203-021-00322-1

Corazza, M., Funari, S., & Gusso, R. (2016). Creditworthiness evaluation of Italian SMEs at the beginning of the 2007–2008 crisis: An MCDA approach. *The North American Journal of Economics & Finance*, *38*, 1–26. https://doi.org/10.1016/j.najef.2016.05.008

Cornée, S. (2019). The relevance of soft information for predicting small business credit default: Evidence from a social bank. *Journal of Small Business Management*, *57*, 699–719. https://doi.org/10.1111/jsbm.12318

Costa, M., Lisboa, I., & Gameiro, A. (2022). Is the financial report quality important in the default prediction? SME Portuguese construction sector evidence. *Risks*, *10*(5), 98. https://doi.org/10.3390/risks10050098

Crosato, L., Domenech, J., & Liberati, C. (2021). Predicting sme's default: Are their websites informative? *Economics Letters*, *204*. https://doi.org/10.1016/j.econlet.2021.109888

Cultrera, L., & Brédart, X. (2016). Bankruptcy prediction: The case of Belgian SMEs. *Review of Accounting and Finance*, *15*, 101–119. https://doi.org/10.1108/RAF-06-2014-0059

Da, F., & Peng, Y. (2022). Non-financial indicators for credit risk analysis of Chinese technology-oriented micro and small enterprises. *Journal of the Operational Research Society*, *74*(4), 1198–1210. https://doi.org/10.1080/01605682.2022.2072781

Dereliolu, G., & Gürgen, F. (2011). Knowledge discovery using neural approach for sme's credit risk analysis problem in Turkey. *Expert Systems with Applications*, *38*(8), 9313–9318. https://doi.org/10.1016/j.eswa.2011.01.012

Dewaelheyns, N., Van Hulle, C., Van Landuyt, Y., & Verreydt, M. (2021). Labor contracts, wages and SME failure. *Sustainability*, *13*(14), 7864. https://doi.org/10.3390/su13147864

DiDonato, F., & Nieddu, L. (2015). The effects of performance ratios in predicting corporate bankruptcy: The Italian case. *Lecture Notes in Business Information Processing*, *216*, 61–72. https://doi.org/10.1007/978-3-319-18533-0_6

Du Jardin, P. (2009). Bankruptcy prediction models: How to choose the most relevant variables. *Bankers, Markets & Investors*, *98*, 39–46. https://mpra.ub.uni-muenchen.de/44380/

Du, M., Ma, Y., & Zhang, Z. (2022a). A meta-path-based evaluation method for enterprise credit risk. *Computational Intelligence and Neuroscience*, *2022*, 1–14. https://doi.org/10.1155/2022/1783445

Du, M., Ma, Y., & Zhang, Z. (2022b). A path-based feature selection algorithm for enterprise credit risk evaluation. *Computational Intelligence and Neuroscience*, *2022*, 1–11. https://doi.org/10.1155/2022/7650207

Duarte, F. D., Gama, A. P. M., & Gulamhussen, M. A. (2018). Defaults in bank loans to SMEs during the financial crisis. *Small Business Economics*, *51*, 591–608. https://doi.org/10.1007/s11187-017-9944-9

Edmister, R. O. (1972). An empirical test of financial ratio analysis for small business failure prediction. *Journal of Financial and Quantitative Analysis*, *7*, 1477–1493. https://doi.org/10.2307/2329929

El Kalak, I., & Hudson, R. (2016). The effect of size on the failure probabilities of SMEs: An empirical study on the US market using discrete hazard model. *International Review of Financial Analysis*, *43*, 135–145. https://doi.org/10.1016/j.irfa.2015.11.009

Falagas, M. E., Pitsouni, E. I., Malietzis, G. A., & Pappas, G. (2008). Comparison of PubMed, Scopus, web of science, and Google scholar: Strengths and weaknesses. *The FASEB Journal*, *22*(2), 338–342. https://doi.org/10.1096/fj.07-9492LSF

Fantazzini, D., & Figini, S. (2009a). Default forecasting for small-medium enterprises: Does heterogeneity matter? *International Journal of Risk Assessment and Management*, *11*, 138–163. https://doi.org/10.1504/IJRAM.2009.022202

Fantazzini, D., & Figini, S. (2009b). Random survival forests models for SME credit risk measurement. *Methodology and Computing in Applied Probability*, *11*, 29–45. https://doi.org/10.1007/s11009-008-9078-2

Figini, S., Bonelli, F., & Giovannini, E. (2017). Solvency prediction for small and medium enterprises in banking. *Decision Support Systems*, *102*, 91–97. https://doi.org/10.1016/j.dss.2017.08.001

Figini, S., & Giudici, P. (2011). Statistical merging of rating models. *Journal of the Operational Research Society*, *62*, 1067–1074. https://doi.org/10.1057/jors.2010.41

Filipe, S. F., Grammatikos, T., & Michala, D. (2016). Forecasting distress in European SME portfolios. *Journal of Banking and Finance*, *64*, 112–135. https://doi.org/10.1016/j.jbankfin.2015.12.007

Gabbianelli, L. (2018). A territorial perspective of SME's default prediction models. *Studies in Economics and Finance*, *35*, 542–563. https://doi.org/10.1108/SEF-08-2016-0207

Gabbi, G., Giammarino, M., & Matthias, M. (2020). Die hard: Probability of default and soft information. *Risks*, *8*(2), 46. https://doi.org/10.3390/risks8020046

Gama, A. P. M., & Geraldes, H. S. A. (2012). Credit risk assessment and the impact of the New Basel Capital Accord on small and medium-sized enterprises: An empirical analysis. *Management Research Review*, *35*, 727–749. https://doi.org/10.1108/01409171211247712

Gao, G., Wang, H., & Gao, P. (2021). Establishing a credit risk evaluation system for SMEs using the soft voting fusion model. *Risks*, *9*(11), 202. https://doi.org/10.3390/risks9110202

Glennon, D., & Nigro, P. (2005). An analysis of SBA loan defaults by maturity structure. *Journal of Financial Services Research*, *28*, 77–111. https://doi.org/10.1007/s10693-005-4357-3

Glennon, D., & Nigro, P. (2011). Evaluating the performance of static versus dynamic models of credit default: Evidence from long-term small business administration-guaranteed loans. *The Journal of Credit Risk*, *7*(2), 3–35. https://doi.org/10.21314/JCR.2011.125

Grishunin, S., Suloeva, V., Shiryakina, S., & Burova, E. (2021). Analyzing insolvency drivers and developing credit rating system for small and medium-sized enterprises in Russia. *International Journal of Technology*, *12*(7), 1479–1487. https://doi.org/10.14716/IJTECH.V12I7.5349

Gupta, J., Barzotto, M., & Khorasgani, A. (2018). Does size matter in predicting SMEs failure? *International Journal of Finance and Economics*, *23*, 571–605. https://doi.org/10.1002/ijfe.1638

Gupta, J., & Gregoriou, A. (2018). Impact of market-based finance on SMEs failure. *Economic Modelling*, *69*, 13–25. https://doi.org/10.1016/j.econmod.2017.09.004

Gupta, J., Gregoriou, A., & Ebrahimi, T. (2018). Empirical comparison of hazard models in predicting SMEs failure. *Quantitative Finance*, *18*, 437–466. https://doi.org/10.1080/14697688.2017.1307514

Gupta, J., Gregoriou, A., & Healy, J. (2015). Forecasting bankruptcy for SMEs using hazard function: To what extent does size matter? *Review of Quantitative Finance & Accounting*, *45*, 845–869. https://doi.org/10.1007/s11156-014-0458-0

Gupta, J., Wilson, N., Gregoriou, A., & Healy, J. (2014a). The effect of internationalisation on modelling credit risk for SMEs: Evidence from UK market. *Journal of International Financial Markets, Institutions and Money*, *31*, 397–413. https://doi.org/10.1016/j.intfin.2014.05.001

Gupta, J., Wilson, N., Gregoriou, A., & Healy, J. (2014b). The value of operating cash flow in modelling credit risk for SMEs. *Applied Financial Economics*, *24*, 649–660. https://doi.org/10.1080/09603107.2014.896979

Guzmán, J. B., & Lussier, R. N. (2015). Success factors for small businesses in guanajuato, mexico. *International Journal of Business & Social Science*, *6*, 1–7.

Gyimah, P., Appiah, K. O., & Lussier, R. N. (2020). Success versus failure prediction model for small businesses in Ghana. *Journal of African Business*, *21*, 215–234. https://doi.org/10.1080/15228916.2019.1625017

Halabi, C. E., & Lussier, R. N. (2014). A model for predicting small firm performance: Increasing the probability of entrepreneurial success in Chile. *Journal of Small Business and Enterprise Development*, *21*, 4–25. https://doi.org/10.1108/JSBED-10-2013-0141

Hyder, S., & Lussier, R. N. (2016). Why businesses succeed or fail: A study on small businesses in pakistan. *Journal of Entrepreneurship in Emerging Economies*, *8*(1), 82–100. https://doi.org/10.1108/jeee-03-2015-0020

Inekwe, J. N. (2016). Financial distress, employees' welfare and entrepreneurship among SMEs. *Social Indicators Research*, *129*, 1135–1153. https://doi.org/10.1007/s11205-015-1164-6

Isaksson, A., Wallman, M., Göransson, H., & Gustafsson, M. G. (2008). Cross-validation and bootstrapping are unreliable in small sample classification. *Pattern Recognition Letters*, *29*, 1960–1965. https://doi.org/10.1016/j.patrec.2008.06.018

Káčer, M., Ochotnický, P., & Alexy, M. (2019). The Altman's revised Z'-score model, non-financial information and macroeconomic variables: Case of Slovak SMEs. *Ekonomicky casopis*, *2019*, 335–366.

Karas, M. (2022). The hazard model for European SMEs: Combining accounting and macro-economic variables. *Journal of Competitiveness*, *14*, 76–92. https://doi.org/10.7441/joc.2022.03.05

Karas, M., & Reznakova, M. (2020). Cash flows indicators in the prediction of financial distress. *Engineering Economics*, *31*, 525–535. https://doi.org/10.5755/j01.ee.31.5.25202

Karas, M., & Režňáková, M. (2021). The role of financial constraint factors in predicting SME default. *Equilibrium Quarterly Journal of Economics and Economic Policy*, *16*, 859–883. https://doi.org/10.24136/eq.2021.032

Keasey, K., & Watson, R. (1986). The prediction of small company failure: Some behavioural evidence for the UK. *Accounting and Business Research*, *17*, 49–57. https://doi.org/10.1080/00014788.1986.9729781

Keasey, K., & Watson, R. (1987). Non-financial symptoms and the prediction of small company failure: A test of argenti's hypotheses. *Journal of Business Finance & Accounting*, *14*, 335–354. https://doi.org/10.1111/j.1468-5957.1987.tb00099.x 3

Keasey, K., & Watson, R. (1988). The non-submission of accounts and small company financial failure prediction. *Accounting and Business Research*, *19*, 47–54. https://doi.org/10.1080/00014788.1988.9728835

Keenan, S. C., & Sobehart, J. R. (1999). Performance measures for credit risk models. *Moody's Risk Management Services*, *10*.

Khermkhan, J., & Chancharat, S. (2015). Performance of minority data in financial distress prediction models. application of multiple discriminate analysis, Logit, probit and artificial neural network methods. *Journal of Applied Economic Sciences*, *10*, 954–960.

Kim, H. S., & Sohn, S. Y. (2010). Support vector machines for default prediction of SMEs based on technology credit. *European Journal of Operational Research*, *201*, 838–846. https://doi.org/10.1016/j.ejor.2009.03.036

Kim, J.-H. (2009). Estimating classification error rate: Repeated cross-validation, repeated hold-out and bootstrap. *Computational Statistics and Data Analysis*, *53*, 3735–3745. https://doi.org/10.1016/j.csda.2009.04.009

Kosmidis, K., & Stavropoulos, A. (2014). Corporate failure diagnosis in SMEs: A longitudinal analysis based on alternative prediction models. *International Journal of Accounting & Information Management*, *22*, 49–67. https://doi.org/10.1108/IJAIM-01-2013-0001

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*. https://doi.org/10.1016/j.dss.2020.113429

Kumar Roy, P., Shaw, K., & Ishizaka, A. (2022). Developing an integrated fuzzy credit rating system for SMEs using fuzzy-BWM and fuzzy-TOPSIS-Sort-C. *Annals of Operations Research*. https://doi.org/10.1007/s10479-022-04704-5

Laitinen, E. K. (1993). The use of information contained in annual reports and prediction of small business failures. *International Review of Financial Analysis*, *2*, 155–176. https://doi.org/10.1016/1057-5219(93)90015-A

Lee, S., Choi, K., & Yoo, D. (2020). Predicting the insolvency of SMEs using technological feasibility assessment information and data mining techniques. *Sustainability*, *12*, 1–17. https://doi.org/10.3390/su12239790

Lextrait, B. (2023). Scaling up SMEs' credit scoring scope with LightGBM. *Applied Economics*, *55*, 925–943. https://doi.org/10.1080/00036846.2022.2095340

Li, J., Liu, H., Yang, Z., & Han, L. (2021). A credit risk model with small sample data based on G-XGBoost. *Applied Artificial Intelligence*, *35*(15), 1550–1566. https://doi.org/10.1080/08839514.2021.1987707

Li, K., Niskanen, J., Kolehmainen, M., & Niskanen, M. (2016). Financial innovation: Credit default hybrid model for SME lending. *Expert Systems with Applications*, *61*, 343–355. https://doi.org/10.1016/j.eswa.2016.05.029

Li, Z., & Guo, L. (2021). Construction of credit evaluation index system for two-stage bayesian discrimination: An empirical analysis of small Chinese enterprises. *Mathematical Problems in Engineering*, *2021*, 1–12. https://doi.org/10.1155/2021/8837419

Lin, S.-M., Ansell, J., & Andreeva, G. (2012). Predicting default of a small business using different definitions of financial distress. *Journal of the Operational Research Society*, *63*, 539–548. https://doi.org/10.1057/jors.2011.65

Long, J., Jiang, C., Dimitrov, S., & Wang, Z. (2022). Clues from networks: Quantifying relational risk for credit risk evaluation of SMEs. *Financial Innovation*, *8*. https://doi.org/10.1186/s40854-022-00390-1

Lu, Y., Yang, L., Shi, B., Li, J., & Abedin, M. Z. (2022). A novel framework of credit risk feature selection for SMEs during industry 4.0. *Annals of Operations Research*. https://doi.org/10.1007/s10479-022-04849-3

Lugovskaya, L. (2010). Predicting default of Russian SMEs on the basis of financial and non-financial variables. *Journal of Financial Services Marketing*, *14*, 301–313. https://doi.org/10.1057/fsm.2009.28

Luo, Z., Hsu, P., & Xu, N. (2020). SME default prediction framework with the effective use of external public credit data. *Sustainability*, *12*(18), 7575. https://doi.org/10.3390/su12187575

Lussier, R. N. (1995). A nonfinancial business success versus failure prediction model for young firms. *Journal of Small Business Management*, *33*, 8.

Lussier, R. N., Bandara, C., & Marom, S. (2016). Entrepreneurship success factors: An empirical investigation in sri lanka. *World Journal of Entrepreneurship, Management and Sustainable Development*, *12*(2). https://doi.org/10.1108/wjemsd-10-2015-0047

Lussier, R. N., & Halabi, C. E. (2010). A three-country comparison of the business success versus failure prediction model. *Journal of Small Business Management*, *48*, 360–377. https://doi.org/10.1111/j.1540-627X.2010.00298.x

Lussier, R. N., & Pfeifer, S. (2001). A crossnational prediction model for business success. *Journal of Small Business Management*, *39*, 228–239. https://doi.org/10.1111/0447-2778.00021

Ma'aji, M. M., Abdullah, N. A. H., & Khaw, K. L.-H. (2019). Financial distress among SMEs in Malaysia: An early warning signal. *International Journal of Business and Society*, *20*, 775–792.

Malakauskas, A., & Lakstutiene, A. (2021). Financial distress prediction for small and medium enterprises using machine learning techniques. *Engineering Economics*, *32*, 4–14. https://doi.org/10.5755/j01.ee.32.1.27382

Marom, S., & Lussier, R. N. (2014). A business success versus failure prediction model for small businesses in israel. *Business and Economic Research*, *4*, 63. https://doi.org/10.5296/ber.v4i2.5997

Marzi, G., Dabić, M., Daim, T., & Garces, E. (2017). Product and process innovation in manufacturing firms: A 30-year bibliometric analysis. *Scientometrics*, *113*, 673–704. https://doi.org/10.1007/s11192-017-2500-1

Matthias, M., Giammarino, M., & Gabbi, G. (2019). Modeling hard and soft facts for SMEs: Some international evidence. *Journal of International Financial Management & Accounting*, *30*, 203–222. https://doi.org/10.1111/jifm.12108

Meng, B., Kuang, H., Lv, L., Fan, L., & Chen, H. (2022). A novel credit rating model: Empirical analysis from Chinese small enterprises. *Emerging Markets Finance and Trade*, *58*, 2368–2387. https://doi.org/10.1080/1540496X.2021.1984226

Mittal, S., Gupta, P., & Jain, K. (2011). Neural network credit scoring model for micro enterprise financing in India. *Qualitative Research in Financial Markets*, *3*, 224–242. https://doi.org/10.1108/17554171111176921

Modina, M., & Pietrovito, F. (2014). A default prediction model for Italian SMEs: The relevance of the capital structure. *Applied Financial Economics*, *24*, 1537–1554. https://doi.org/10.1080/09603107.2014.927566

Monelos, P. D. L., Sánchez, C. P., & López, M. R. (2014). DEA as a business failure prediction tool application to the case of Galician SMEs. *Contaduría y administración*, *59*(2), 65–96. https://doi.org/10.1016/S0186-1042(14)71255-0

Moon, T., & Sohn, S. (2010). Technology credit scoring model considering both SME characteristics and economic conditions: The Korean case. *Journal of the Operational Research Society*, *61*, 666–675. https://doi.org/10.1057/jors.2009.7

Muthukumaran, K., & Hariharanath, K. (2023). Deep learning enabled financial crisis prediction model for small-medium sized industries. *Intelligent Automation & Soft Computing*, *35*, 521–536. https://doi.org/10.32604/iasc.2023.025968

Norden, L., & Weber, M. (2010). Credit line usage, checking account activity, and default risk of bank borrowers. *Review of Financial Studies*, *23*, 3665–3699. https://doi.org/10.1093/rfs/hhq061

Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, *18*, 109–131. https://doi.org/10.2307/2490395

Oliveira, M. D. N. T., Ferreira, F. A. F., Pérez-Bustamante Ilander, G. O., & Jalali, M. S. (2017). Integrating cognitive mapping and MCDA for bankruptcy prediction in small-and medium-sized enterprises. *Journal of the Operational Research Society*, *68*(9), 985–997. https://doi.org/10.1057/s41274-016-0166-3

Pacheco, L., Madaleno, M., Correia, P., & Maldonado, I. (2022). Probability of corporate bankruptcy: Application to Portuguese manufacturing industry SMEs. *International Journal of Business and Society*, *23*, 1169–1189. https://doi.org/10.33736/IJBS.4863.2022

Papík, M., & Papíková, L. (2023). Impacts of crisis on SME bankruptcy prediction models' performance. *Expert Systems with Applications*, *214*. https://doi.org/10.1016/j.eswa.2022.119072

Park, S. B., Kim, S.-K., & Lee, S. (2021). Earnings management of insolvent firms and the prediction of corporate defaults via discretionary accruals. *International Journal of Financial Studies*, *9*(2), 17. https://doi.org/10.3390/ijfs9020017

Pederzoli, C., Thoma, G., & Torricelli, C. (2013). Modelling credit risk for innovative SMEs: The role of innovation measures. *Journal of Financial Services Research*, *44*, 111–129. https://doi.org/10.1007/s10693-012-0152-0

Pederzoli, C., & Torricelli, C. (2010). A parsimonious default prediction model for Italian SMEs. *Banks and Bank Systems*, *5*, 5–9. https://doi.org/10.25431/11380_1197426

Piatt, H. D., & Piatt, M. B. (2002). Predicting corporate financial distress: Reflections on choice-based sample bias. *Journal of Economics & Finance*, *26*(2), 184–199. https://doi.org/10.1007/bf02755985

Piatti, D., Cincinelli, P., & Castellani, D. (2015). Efficiency and corporate bankruptcy prediction. *Academy of Accounting & Financial Studies Journal*, *19*, 319–354.

Pierri, F., & Caroni, C. (2017). Bankruptcy prediction by survival models based on current and lagged values of time-varying financial data. *Communications in Statistics Case Studies, Data Analysis and Applications*, *3*(3–4), 62–70. https://doi.org/10.1080/23737484.2018.1431816

Pierri, F., & Caroni, C. (2022). Identifying changes in predictors of business failures during and after the economic crisis in Italy. *Electronic Journal of Applied Statistical Analysis*, *15*, 40–49. https://doi.org/10.1285/i20705948v15n1p40

Ptak-Chmielewska, A., & Matuszyk, A. (2018). The importance of financial and non-financial ratios in SMEs bankruptcy prediction. *Bank i Kredyt*, *49*, 45–62.

Ptak-Chmielewska, A., & Matuszyk, A. (2020). Application of the random survival forests method in the bankruptcy prediction for small and medium enterprises. *Argumenta Oeconomica*, *2020*(1), 127–142. https://doi.org/10.15611/aoe.2020.1.06

Rikkers, F., & Thibeault, A. E. (2011). Default prediction of small and medium-sized enterprises with industry effects. *International Journal of Banking, Accounting and Finance*, *3*, 207–231. https://doi.org/10.1504/IJBAAF.2011.041455

Roy, P. K., & Shaw, K. (2021a). A credit scoring model for SMEs using AHP and TOPSIS. *International Journal of Finance and Economics*, *7*(1). https://doi.org/10.1002/ijfe.2425

Roy, P. K., & Shaw, K. (2021b). A multicriteria credit scoring model for SMEs using hybrid BWM and TOPSIS. *Financial Innovation*, *7*. https://doi.org/10.1186/s40854-021-00295-5

Schalck, C., & Yankol-Schalck, M. (2021). Predicting French SME failures: New evidence from machine learning techniques. *Applied Economics*, *53*, 5948–5963. https://doi.org/10.1080/00036846.2021.1934389

Séverin, E., & Veganzones, D. (2021). Can earnings management information improve bankruptcy prediction models? *Annals of Operations Research*, *306*, 247–272. https://doi.org/10.1007/s10479-021-04183-0

Shetty, S., Musa, M., & Brédart, X. (2022). Bankruptcy prediction using machine learning techniques. *Journal of Risk and Financial Management*, *15*(1), 35. https://doi.org/10.3390/jrfm15010035

Sigrist, F., & Hirnschall, C. (2019). Grabit: Gradient tree-boosted tobit models for default prediction. *Journal of Banking and Finance*, *102*, 177–192. https://doi.org/10.1016/j.jbankfin.2019.03.004

Sohn, S. Y., & Jeon, H. (2010). Competing risk model for technology credit fund for small and medium-sized enterprises. *Journal of Small Business Management*, *48*, 378–394. https://doi.org/10.1111/j.1540-627X.2010.00299.x

Sohn, S. Y., & Kim, H. S. (2007). Random effects logistic regression model for default prediction of technology credit guarantee fund. *European Journal of Operational Research*, *183*, 472–478. https://doi.org/10.1016/j.ejor.2006.10.006

Stevenson, M., Mues, C., & Bravo, C. (2021). The value of text for small business default prediction: A deep learning approach. *European Journal of Operational Research*, *295*, 758–771. https://doi.org/10.1016/j.ejor.2021.03.008

Sun, Y., Chai, N., Dong, Y., & Shi, B. (2022). Assessing and predicting small industrial enterprises' credit ratings: A fuzzy decision-making approach. *International Journal of Forecasting*, *38*, 1158–1172. https://doi.org/10.1016/j.ijforecast.2022.01.006

Sun, J., & Jiao, Y. (2022). Enterprise financial risk analysis based on improved model C-means clustering algorithm. *Security and Communication Networks*, *2022*, 1–12. https://doi.org/10.1155/2022/1109813

Svabova, L., Michalkova, L., Durica, M., & Nica, E. (2020). Business failure prediction for Slovak small and medium-sized companies. *Sustainability*, *12*(11), 4572. https://doi.org/10.3390/su12114572

Terdpaopong, K., & Mihret, D. G. (2011). Modelling SME credit risk: Thai empirical evidence. *Small Enterprise Research*, *18*, 63–79. https://doi.org/10.5172/ser.18.1.63

Tobback, E., Bellotti, T., Moeyersoms, J., Stankova, M., & Martens, D. (2017). Bankruptcy prediction for SMEs using relational data. *Decision Support Systems*, *102*, 69–81. https://doi.org/10.1016/j.dss.2017.07.004

Veganzones, D., & Séverin, E. (2018). An investigation of bankruptcy prediction in imbalanced datasets. *Decision Support Systems*, *112*, 111–124. https://doi.org/10.1016/j.dss.2018.06.011

Wilson, N., & Altanlar, A. (2014). Company failure prediction with limited information: Newly incorporated companies. *Journal of the Operational Research Society*, *65*, 252–264. https://doi.org/10.1057/jors.2013.31

Wilson, N., Ochotnický, P., & Káčer, M. (2016). Creation and destruction in transition economies: The SME sector in Slovakia. *International Small Business Journal: Researching Entrepreneurship*, *34*, 579–600. https://doi.org/10.1177/0266242614558892

Wu, C., & Wang, X.-M. (2000). A neural network approach for analyzing small business lending decisions. *Review of Quantitative Finance & Accounting*, *15*, 259–276. https://doi.org/10.1023/A:1008324023422

Yazdanfar, D. (2011). Predicting bankruptcy among SMEs: Evidence from Swedish firm-level data. *International Journal of Entrepreneurship & Small Business*, *14*, 551–565. https://doi.org/10.1504/IJESB.2011.043475

Yin, C., Jiang, C., Jain, H. K., & Wang, Z. (2020). Evaluating the credit risk of SMEs using legal judgments. *Decision Support Systems*, *136*. https://doi.org/10.1016/j.dss.2020.113364

Yoon, J. S., & Kwon, Y. S. (2010). A practical approach to bankruptcy prediction for small businesses: Substituting the unavailable financial data for credit card sales information. *Expert Systems with Applications*, *37*, 3624–3629. https://doi.org/10.1016/j.eswa.2009.10.029

Yoshino, N., Taghizadeh-Hesary, F., Charoensivakorn, P., & Niraula, B. (2016). Small and medium-sized enterprise (SME) credit risk analysis using bank lending data: An analysis of Thai SMEs. *Journal of Comparative Asian Development*, *15*, 383–406. https://doi.org/10.1080/15339114.2016.1233821

Zeng, H. (2022). Credit risk evaluation in enterprise financial management by using convolutional neural network under the construction of smart city. *Security and Communication Networks*, *2022*, 1–8. https://doi.org/10.1155/2022/8532918

Zhang, L., & Song, Q. (2022a). Credit evaluation of SMEs based on GBDT-CNN-LR hybrid integrated model. *Wireless Communications and Mobile Computing*, *2022*, 1–8. https://doi.org/10.1155/2022/5251228

Zhang, L., & Song, Q. (2022b). Multimodel integrated enterprise credit evaluation method based on attention mechanism. *Computational Intelligence and Neuroscience*, *2022*, 1–12. https://doi.org/10.1155/2022/8612759

Zhang, J., & Thomas, L. C. (2015). The effect of introducing economic variables into credit scorecards: An example from invoice discounting. *The Journal of Risk Model Validation*, *9*(1), 57–78. https://doi.org/10.21314/JRMV.2015.134

Zizi, Y., Jamali-Alaoui, A., El Goumi, B., Oudgou, M., & El Moudden, A. (2021). An optimal model of financial distress prediction: A comparative study between neural networks and logistic regression. *Risks*, *9*(11), 200. https://doi.org/10.3390/risks9110200

Appendix A

Table A1. Geographical distribution of data used in SME default papers. Some articles studied more than one country. Thus, the total number of country-study listed is more than the number of papers.

Country	ABS	Non-ABS	N	Studies
Italy	19	10	29	Altman et al. (2020); Angelini et al. (2008); Angilella and Mazzù (2015); Angilella and Mazzù (2019); Calabrese et al. (2016); Caselli et al. (2021); Cathcart et al. (2020); Ciampi and Gordini (2013); Ciampi (2015); Ciampi (2017); Ciampi (2018); Ciampi et al. (2020); Corazza et al. (2016); Corazza et al. (2021); Corazza et al. (2021); DiDonato and Nieddu (2015); Figini et al. (2017); Filipe et al. (2016); Gabbi et al. (2020); Gabbianelli (2018); Karas and Režňáková (2021); Karas (2022); Matthias et al. (2019); Modina and Pietrovito (2014); Pederzoli and Torricelli (2010); Pederzoli et al. (2013); Piatti et al. (2015); Pierri and Caroni (2017); Pierri and Caroni (2022);
China	8	9	17	Abedin et al. (2022); Chai et al. (2019); Chi and Zhang (2017); Chi and Meng (2019); Da and Peng (2022); Du et al. (2022a); Du et al. (2022b); Gao et al. (2021); Kou et al. (2021); Li and Guo (2021); Long et al. (2022); Lu et al. (2022); Luo et al. (2020); Meng et al. (2022); Sun et al. (2022); Yin et al. (2020); Zeng (2022);
UK	14	3	17	Andrikopoulos and Khorasgani (2018); Cathcart et al. (2020); Filipe et al. (2016); Gupta et al. (2014b); Gupta et al. (2014a); Gupta et al. (2015); Karas and Režňáková (2021); Karas (2022); Keasey and Watson (1986); Keasey and Watson (1987); Keasey and Watson (1988); Lin et al. (2012); Matthias et al. (2019); Pederzoli et al. (2013); Tobback et al. (2017); Wilson and Altanlar (2014); Zhang and Thomas (2015);
USA	10	2	12	Altman and Sabato (2007); Edmister (1972); El Kalak and Hudson (2016); Glennon and Nigro (2005); Glennon and Nigro (2011); Gupta and Gregoriou (2018); Gupta et al. (2018); Gupta et al. (2018) Inekwe (2016); Sun and Jiao (2022); Wu and Wang (2000); Lussier (1995);
Germany	6	5	11	Behr and Güttler (2007); Fantazzini and Figini (2009b); Fantazzini and Figini (2009a); Figini and Giudici (2011); Filipe et al. (2016); Karas and Režňáková (2021); Karas (2022); Matthias et al. (2019); Muthukumaran and Hariharanath (2023); Norden and Weber (2010); Pederzoli et al. (2013);
Portugal	6	4	10	Cathcart et al. (2020); Costa et al. (2022); Duarte et al. (2018); Filipe et al. (2016); Gama and Geraldes (2012); Karas and Režňáková (2021); Karas (2022); Oliveira et al. (2017); Pacheco et al. (2022); Pederzoli et al. (2013);
France	7	2	9	Cathcart et al. (2020); Cornée (2019); Filipe et al. (2016); Karas and Režňáková (2021); Karas (2022); Lextrait (2023); Pederzoli et al. (2013); Schalck and Yankol-Schalck (2021); Séverin and Veganzones (2021);
Belgium	4	4	8	Cathcart et al. (2020); Cultrera and Brédart (2016); Dewaelheyns et al. (2021); Karas and Režňáková (2021); Karas (2022); Pederzoli et al. (2013); Shetty et al. (2022); Tobback et al. (2017);
Spain	5	3	8	Baixauli and Módica-Milo (2010); Cathcart et al. (2020); Crosato et al. (2021); Filipe et al. (2016); Karas and Režňáková (2021); Karas (2022); Monelos et al. (2014); Pederzoli et al. (2013);
South Korea	5	2	7	Kim and Sohn (2010); Lee et al. (2020); Moon and Sohn (2010); Park et al. (2021); Sohn and Kim (2007); Sohn and Jeon (2010); Yoon and Kwon (2010);
Slovakia	2	4	6	Káčer et al. (2019); Karas and Režňáková (2021); Karas (2022); Papík and Papíková (2023); Svabova et al. (2020); Wilson et al. (2016);
Finland	3	2	5	Altman et al. (2020); Karas and Režňáková (2021); Karas (2022); Laitinen (1993); Li et al. (2016);
Poland	1	4	5	Filipe et al. (2016); Karas and Režňáková (2021); Karas (2022); Ptak-Chmielewska and Matuszyk (2018); Ptak-Chmielewska and Matuszyk (2020);
Croatia	2	2	4	Altman et al. (2022); Karas and Režňáková (2021); Karas (2022); Lussier and Pfeifer (2001);
Czech Republic	1	3	4	Filipe et al. (2016); Karas and Reznakova (2020); Karas and Režňáková (2021); Karas (2022);
Greece	2	2	4	Karas and Režňáková (2021); Karas (2022); Kosmidis and Stavropoulos (2014); Pederzoli et al. (2013);
India	3	1	4	Mittal et al. (2011); Roy and Shaw (2021a); Roy and Shaw (2021b); Kumar Roy et al. (2022);

(*Continued*)

Table A1. (Continued).

Country	ABS	Non-ABS	N	Studies
Malaysia	2	2	4	Abdullah et al. (2016a); Abdullah et al. (2016b); Abdullah et al. (2019); Ma'aji et al. (2019);
Netherlands	2	2	4	Karas and Režňáková (2021); Karas (2022); Pederzoli et al. (2013); Rikkers and Thibeault (2011);
Denmark	1	2	3	Karas and Režňáková (2021); Karas (2022); Pederzoli et al. (2013);
Ireland	1	2	3	Karas and Režňáková (2021); Karas (2022); Pederzoli et al. (2013);
Luxembourg	1	2	3	Karas and Režňáková (2021); Karas (2022); Pederzoli et al. (2013);
Sweden	1	2	3	Karas and Režňáková (2021); Karas (2022); Yazdanfar (2011);
Thailand	1	2	3	Khermkhan and Chancharat (2015); Terdpaopong and Mihret (2011); Yoshino et al. (2016);
Estonia	0	3	3	Karas and Režňáková (2021); Karas (2022); Malakauskas and Lakstutiene (2021);
Latvia	0	3	3	Karas and Režňáková (2021); Karas (2022); Malakauskas and Lakstutiene (2021);
Lithuania	0	3	3	Karas and Režňáková (2021); Karas (2022); Malakauskas and Lakstutiene (2021);
Chile	2	0	2	Halabi and Lussier (2014); Lussier and Halabi (2010);
Russia	1	1	2	Grishunin et al. (2021); Lugovskaya (2010);
Switzerland	2	0	2	Pederzoli et al. (2013); Sigrist and Hirnschall (2019);
Turkey	2	0	2	Arslan and Karan (2009); Dereliolu and Gürgen (2011);
Austria	0	2	2	Karas and Režňáková (2021); Karas (2022);
Bulgaria	0	2	2	Karas and Režňáková (2021); Karas (2022);
Cyprus	0	2	2	Karas and Režňáková (2021); Karas (2022);
Hungary	0	2	2	Karas and Režňáková (2021); Karas (2022);
Malta	0	2	2	Karas and Režňáková (2021); Karas (2022);
Romania	0	2	2	Karas and Režňáková (2021); Karas (2022);
Slovenia	0	2	2	Karas and Režňáková (2021); Karas (2022);
Colombia	1	0	1	Castillo et al. (2018);
Ghana	1	0	1	Gyimah et al. (2020);
Mauritius	1	0	1	Bangarigadu and Nunkoo (2022);
Norway	1	0	1	Pederzoli et al. (2013);
Pakistan	1	0	1	Hyder and Lussier (2016);
Palestine	1	0	1	Baidoun et al. (2018);
Taiwan	1	0	1	Chen et al. (2015);
Australia	0	1	1	Muthukumaran and Hariharanath (2023);
Israel	0	1	1	Marom and Lussier (2014);
Mexico	0	1	1	Guzmán and Lussier (2015);
Morocco	0	1	1	Zizi et al. (2021);
Sri Lanka	0	1	1	Lussier et al. (2016);

Table A2. An overview of papers that studied more than one country, tested more than one estimation method and had more than one primary estimation method.

	Country > 1	Estimation Methods Tested > 1	Primary Estimation Methods > 1
ABS (94 Studies)			
1970s	0	0	0
1980s	0	2	2
1990s	0	1	0
2000s	0	2	0
2010s	4	17	8
2020s	1	16	12
Overall	5	38	22
Non-ABS (51 Studies)			
1970s			
1980s	0	0	0
1990s			
2000s	0	2	0
2010s	0	7	5
2020s	4	19	12
Overal	4	28	17

Table A3. General factors included in three or more studies.

Factor	ABS	Non-ABS	N
Quick Ratio	13	17	30
Current Ratio	15	11	26
Net Income/Total Assets	15	11	26
Retained Earnings/Total Assets	19	3	22
Sales/Total Assets	11	9	20
Cash/Total Assets	15	1	16
Net Income/Equity	8	8	16
Working Capital/Total Assets	11	5	16
Earnings Before Interest and Taxes/Total Assets	9	6	15
Total Liabilities/Total Assets	11	4	15
EBITDA/Interest Expenses	12	2	14
Total Debt/Total Assets	9	4	13
Equity/Total Assets	7	5	12
Capital	8	3	11
EBITDA/Total Assets	9	2	11
Net Income/Sales	7	4	11
Cash Ratio	4	6	10
Product/Service Timing	7	3	10
Cash Flow/Total Debt	7	2	9
Marketing	6	3	9
Operating Profit/Sales	4	5	9
Planning	6	3	9
Trade Creditors/Total Assets	8	1	9
Cash Flow/Total Assets	2	6	8
Record Keeping And Financial Control	5	3	8
Short-term Debt/Equity Book Value	8	0	8
Staffing	5	3	8
Total Debt/Equity	4	4	8
Earnings Before Interest and Taxes/Interest Expenses	4	3	7
Taxes/Total Assets	6	1	7
Intangible Assets/Total Assets	6	0	6
Liquidity Ratio	1	5	6
Return on Investment	4	2	6
Short-term Debt/Equity	4	2	6
Capital Employed/Total Liabilities	5	0	5
Equity/Total Debt	3	2	5
Financial Expenses/Total Assets	5	0	5
Gross Profit Margin	2	3	5
Interest Expenses/Sales	2	3	5
Ln(Current Ratio)	5	0	5
Professional Advice	3	2	5
Current Liabilities/Total Assets	4	0	4
Profit After Tax/Sales	3	1	4
Tangible Assets/Total Assets	2	2	4
Total Assets Turnover	3	1	4
Trade Debtors/Total Assets	4	0	4
Cash Flow From Operations/Total Assets	0	4	4
Equity/Total Liabilities	0	4	4
Account Payable/Sales	3	0	3
Account Receivable/Total Liabilities	3	0	3
Capital Growth	3	0	3
Capital Tied Up	1	2	3
Cash And Short-term Investments/Total Assets	3	0	3
Cash Flow/Sales	1	2	3
Cash Flow/Total Liabilities	1	2	3
Current Liabilities/Sales	2	1	3
Debt Service Coverage	2	1	3
Earnings Before Interest and Taxes/Current Liabilities	1	2	3
Earnings Before Interest and Taxes/Sales	2	1	3
Earnings Before Taxes/Total Assets	1	2	3
EBITDA/Sales	2	1	2
Financial Expenses/Sales	1	1	3

(*Continued*)

Table A3. (Continued).

Factor	ABS	Non-ABS	N
Fixed Assets/Total Assets	2	1	3
Legal Dispute Number	2	1	3
Level of Brand Products	2	1	3
Leverage	2	1	3
Ln(Share Capital)	1	2	3
Long-term Liabilities/Total Assets	2	1	3
Net Profit/Equity	1	2	3
Non-current Liabilities/Total Assets	2	1	3
Outside Capital Structure	1	2	3
Patent Condition	2	1	3
Short-term Debt/Total Assets	3	0	3
Supplier Target Days	1	2	3
Total Assets Growth Rate	2	1	3
Total Liabilities/Equity	2	1	3
Total Revenue/Total Assets	1	2	3

Table A4. Firm and owner/manager characteristics features used in three or more studies.

Factor	ABS	Non-ABS	N
Firm Characteristics			
Age	16	15	31
Ln(Total Assets)	12	5	17
Partners	8	3	11
Ln(Age)	7	3	10
Number of Employees	7	2	9
Ln(Sales)	7	1	8
Economic Timing	5	2	7
Legal Form	3	4	7
New Business	6	0	6
Number of Directors	2	3	5
Ln(Total Assets) Squared	4	0	4
Registered Capital	1	3	4
Date of Establishment	2	1	3
Family Ownership	1	2	3
Foreign Ownership	1	2	3
Management-Owner	1	2	3
Owner/Manager Characteristics			
Education	10	4	14
Management Experience	7	5	12
Age (Owner/Debtor/Legal Representative)	7	4	11
Parents Owned a Business	9	2	11
Industry Experience	6	4	10
Gender	6	3	9
Minority	6	2	8
Residence Status	2	2	4
CEO Duality	1	2	3
Marital Status	1	2	3

Table A5. Macroeconomics factors included in three or more studies.

	ABS	Non-ABS	N
GDP Growth	5	1	6
Engel Coefficient	2	2	4
An Industry Weight of Evidence[a]	3	0	3
Consumer Price Index	1	2	3
Industry Sentiment Index	1	2	3
Insolvency Rate[b]	3	0	3
Interest Rate	1	2	3

[a]Expresses the previous year's sector failure rate as a log odds of failure in each of the industrial sectors.
[b]Denotes the previous year's sector insolvency rate within the firm's industrial sectors.

Table A6. Models performance measure utilized in 10 or more studies.

Performance Measure	ABS	Non-ABS	N
(Number of Studies)	(94)	(51)	(145)
Error I and II	49	25	74
AUC (ROC)	40	28	68
−2 Log L	14	5	19
Hosmer–Lemeshow	10	5	15
McFadden *R*2	6	5	11
Nagelkerke *R*2	9	2	11
Likelihood Ratio (LR)	7	3	10

Table A7. Top 20 performing models based on the AUC(ROC) by study.

	AUC(ROC)	Estimation Method	Study	ABS
1	0.992	GBDT-CNN-LR	Zhang and Song (2022a)	No
2	0.991	RF	Abedin et al. (2022)	No
3	0.988	SVM	Sun et al. (2022)	Yes
4	0.984	Soft voting	Gao et al. (2021)	No
5	0.973	CNN-Logistic-Stacking	Zhang and Song (2022b)	No
6	0.970	XGBoost (Focal Loss)	Sun and Jiao (2022)	No
7	0.963	Elman network	Corazza et al. (2021)	Yes
8	0.959	Logit	Zizi et al. (2021)	No
9	0.956	Cox Proportional Hazards	Gupta and Gregoriou (2018)	Yes
10	0.949	NN	Da and Peng (2022)	Yes
11	0.941	CatBoost	Papík and Papíková (2023)	Yes
12	0.940	NN	Altman et al. (2020)	Yes
13	0.923	Logit	Baixauli and Módica-Milo (2010)	Yes
14	0.922	LightGBM	Lextrait (2023)	Yes
15	0.920	LightGBM	Luo et al. (2020)	No
16	0.909	LPM	Figini and Giudici (2011)	Yes
17	0.903	Logit	Altman et al. (2022)	Yes
18	0.902	Two-stage Nonparametric Bayesian Discriminant	Li and Guo (2021)	No
19	0.900	RF	Figini et al. (2017)	Yes
20	0.893	Logit	Abdullah et al. (2016a)	Yes

Studies that used hold-out samples and cross-validation as their validation methods are shown in this table.

Table A8. Top 20 performing models based on the accuracy ratio by study.

	Accuracy	Estimation Method	Study	ABS
1	99.1%	RF	Abedin et al. (2022)	No
2	97.7%	MDA	Terdpaopong and Mihret (2011)	Yes
3	97.1%	NN	Da and Peng (2022)	Yes
4	96.8%	CNN	Zeng (2022)	No
5	96.7%	NN	Angelini et al. (2008)	Yes
6	96.2%	Logit	Abdullah et al. (2016a)	Yes
7	95.0%	Logit	Zizi et al. (2021)	No
8	93.8%	XGBoost (Focal Loss)	Sun and Jiao (2022)	No
9	93.6%	Elman network	Corazza et al. (2021)	Yes
10	93.0%	Logit	Laitinen (1993)	Yes
11	91.6%	CART	DiDonato and Nieddu (2015)	No
12	91.5%	LPM	Figini and Giudici (2011)	Yes
13	91.2%	Logit	Ma'aji et al. (2019)	No
14	91.1%	NN	Meng et al. (2022)	Yes
15	90.8%	CatBoost	Papík and Papíková (2023)	Yes
16	90.0%	Logit	Abdullah et al. (2019)	Yes
17	89.3%	Logit	Baixauli and Módica-Milo (2010)	Yes
18	88.2%	Logit	Abdullah et al. (2016b)	No
19	87.9%	Logit	Ciampi (2015)	Yes
20	87.0%	LightGBM	Luo et al. (2020)	No

Studies that used hold-out samples and cross-validation as their validation methods are shown in this table.

Appendix B

Table B1. List of all papers included in the present review. Best est. method shows the best primary estimation method, B. AUC stands for the best primary model AUC(ROC), and B.ACC is the best primary model accuracy based on error type I and error type II.

Study	Focus	Sample Size	Validation	Best Est. Method	B.AUC	B.ACC	ABS Level
Edmister (1972)	Small enterprises	42	In-sample				4
Keasey and Watson (1986)	Small enterprises	36	In-sample	LDA		0.750	3
Keasey and Watson (1987)	Small enterprises	146	Hold out sample	Logit		0.650	N/A
Keasey and Watson (1988)	Small enterprises	146	In-sample	Combination		0.678	3
Laitinen (1993)	Small enterprises	80	Cross-validation	Logit		0.930	3
Lussier (1995)	Small enterprises	216	In-sample	Logit		0.692	3
Wu and Wang (2000)	Small enterprises	182	Hold out sample				3
Lussier and Pfeifer (2001)	Small enterprises	120	In-sample				3
Glennon and Nigro (2005)	Small enterprises	21,301	Hold out sample				3
Altman and Sabato (2007)	SMEs	2,010	Hold out sample	Logit		0.802	3
Behr and Güttler (2007)	SMEs	88,402	Hold out sample	Logit	0.852		3
Sohn and Kim (2007)	SMEs	660	Hold out sample				4
Angelini et al. (2008)	Small enterprises	76	Hold out sample	NN		0.967	2
Arslan and Karan (2009)	SMEs	1,166	In-sample	Logit		0.854	2
Fantazzini and Figini (2009b)	SMEs	1,003	Hold out sample	RSF	0.841		N/A
Fantazzini and Figini (2009a)	SMEs	1,003	Hold out sample	Bayesian Logit	0.809		N/A
Baixauli and Módica-Milo (2010)	SMEs	34	Hold out sample	Logit	0.923	0.893	2
Kim and Sohn (2010)	SMEs	3,827	Hold out sample				4
Lugovskaya (2010)	SMEs	520	Hold out sample	LDA		0.790	1
Lussier and Halabi (2010)	Small enterprises	234	In-sample	Logit		0.634	3
Moon and Sohn (2010)	SMEs	4,262	Hold out sample	Logit	0.661		3
Norden and Weber (2010)	Large firms, small firms, and individuals	67,215	N/A				4
Pederzoli and Torricelli (2010)	SMEs		In-sample	Logit		0.668	N/A
Sohn and Jeon (2010)	SMEs	4,482	In-sample				3
Yoon and Kwon (2010)	Small enterprises	10,000	Hold out sample	SVM		0.742	1
Dereliolu and Gürgen (2011)	SMEs	512	Hold out sample	MLP		0.762	1
Figini and Giudici (2011)	SMEs	1,003	Hold out sample	LPM	0.909	0.915	3

(*Continued*)

Table B1. (Continued).

Study	Focus	Sample Size	Validation	Best Est. Method	B.AUC	B.ACC	ABS Level
Glennon and Nigro (2011)	Small enterprises	20,050	Hold out sample				1
Mittal et al. (2011)	Micro enterprises	2,864	Hold out sample	NN		0.717	1
Rikkers and Thibeault (2011)	SMEs	1,894	Hold out sample	Logit	0.853		2
Terdpaopong and Mihret (2011)	SMEs	266	Hold out sample	MDA		0.977	1
Yazdanfar (2011)	SMEs	4,496	In-sample	Logit		0.967	2
Gama and Geraldes (2012)	SMEs operating in the food or beverage manufacturing sector	2,496	In-sample				1
Lin et al. (2012)	SMEs	429	Cross-Validation	Logit	0.877		3
Ciampi and Gordini (2013)	Small enterprises	7,113	Hold out sample	NN		0.684	3
Pederzoli et al. (2013)	Innovative SMEs	5,449	Hold out sample	Logit		0.710	3
Gupta et al. (2014b)	SMEs	118,878	Hold out sample	Logit	0.688	0.661	2
Gupta et al. (2014a)	SMEs	344,205	Hold out sample	Logit	0.692	0.636	3
Halabi and Lussier (2014)	Small enterprises	403	In-sample				2
Kosmidis and Stavropoulos (2014)	SMEs	58	In-sample	MDA		0.845	2
Modina and Pietrovito (2014)	SMEs	9,208	In-sample				2
Monelos et al. (2014)	SMEs	75,640	In-sample	Data Envelopment Analysis (DEA)		0.800	N/A
Wilson and Altanlar (2014)	SMEs	4,427,896	Hold out sample	Logit	0.784		3
Marom and Lussier (2014)	Small enterprises	205	In-sample	Logit		0.854	N/A
Angilella and Mazzù (2015)	Innovative SMEs	4	N/A				4
Chen et al. (2015)	Small enterprises	2,682	In-sample				3
Ciampi (2015)	Small enterprises	934	Hold out sample	Logit		0.879	3
DiDonato and Nieddu (2015)	SMEs	100	Cross-validation	CART		0.916	N/A
Gupta et al. (2015)	SMEs	393,895	Hold out sample	Discrete-time hazards	0.776		3
Khermkhan and Chancharat (2015)	SMEs	30,463	Hold out sample	NN		0.806	N/A
Piatti et al. (2015)	SMEs	8,145	Cross-validation	Logit	0.882	0.789	N/A
Zhang and Thomas (2015)	SMEs	5,826	Hold out sample				1
Guzmán and Lussier (2015)	Small enterprises	303	In-sample	Logit		0.663	N/A
Abdullah et al. (2016a)	SMEs	132	Hold out sample	Logit	0.893	0.962	2
Abdullah et al. (2016b)	Manufacturing SMEs	172	Hold out sample	Logit		0.882	N/A
Calabrese et al. (2016)	SMEs	49,738	Hold out sample	BGEVA	0.811		3
Corazza et al. (2016)	SMEs	39,742	N/A				2
Cultrera and Brédart (2016)	SMEs	7,152	Hold out sample	Logit		0.792	2
El Kalak and Hudson (2016)	SMEs	11,117	Hold out sample				3
Filipe et al. (2016)	SMEs	2,721,861	Hold out sample	Multiperiod Logit	0.875		3

(*Continued*)

Table B1. (Continued).

Study	Focus	Sample Size	Validation	Best Est. Method	B.AUC	B.ACC	ABS Level
Inekwe (2016)	SMEs	7,294	In-sample	Panel Logit	0.910		N/A
Li et al. (2016)	SMEs	2,681	Hold out sample	ANN-logistic hybrid model		0.846	1
Wilson et al. (2016)	SMEs	44,597	In-sample	Logit	0.774		3
Yoshino et al. (2016)	SMEs	3,272	N/A				N/A
Hyder and Lussier (2016)	Small enterprises	143	In-sample	Logit		0.818	1
Lussier et al. (2016)	Small enterprises	450	In-sample	Logit		0.784	N/A
Chi and Zhang (2017)	Small enterprises	1,231	In-sample	Entropy weighting method	0.863	0.810	N/A
Ciampi (2017)	Small enterprises	423	Hold out sample	Logit		0.851	N/A
Figini et al. (2017)	SMEs	38,036	Hold out sample	RF	0.900		3
Oliveira et al. (2017)	SMEs	5	N/A				3
Pierri and Caroni (2017)	Small enterprises	8,999	Hold out sample	Logit	0.795		N/A
Tobback et al. (2017)	SMEs	2,400,000	Cross-validation	SVM	0.847		3
Andrikopoulos and Khorasgani (2018)	SMEs	196,807	Hold out sample	Hybrid (Logit and Merthon-KMV)	0.875	0.859	3
Baidoun et al. (2018)	Small enterprises	246	In-sample	Logit		0.970	1
Castillo et al. (2018)	SMEs	10,603	No-validation				2
Ciampi (2018)	Small enterprises	382	Hold out sample	Logit		0.828	N/A
Duarte et al. (2018)	SMEs	5,898	Hold out sample	Probit	0.772	0.741	3
Gabbianelli (2018)	Small enterprises	141	In-sample	Logit		0.865	1
Gupta and Gregoriou (2018)	SMEs	50,269	Hold out sample	Cox Proportional Hazards	0.956		3
Gupta et al. (2018)	SMEs	40,171	Hold out sample	Multivariate hazards model	0.822		3
Gupta et al. (2018)	SMEs	26,229	Hold out sample	Panel Logit	0.846		2
Ptak-Chmielewska and Matuszyk (2018)	SMEs	968	Hold out sample				N/A
Abdullah et al. (2019)	Manufacturing SMEs	260	Hold out sample	Logit		0.900	1
Angilella and Mazzù (2019)	Innovative SMEs	194	N/A				3
Chai et al. (2019)	Small wholesale and retail enterprises	687	In-sample	Entropy-weighting TOPSIS with fuzzy C-means (FCM)	0.917	0.801	2
Chi and Meng (2019)	Small enterprises	3,045	Hold out sample				2
Cornée (2019)	Small enterprises	389	In-sample	Probit	0.733		3

(*Continued*)

Table B1. (Continued).

Study	Focus	Sample Size	Validation	Best Est. Method	B.AUC	B.ACC	ABS Level
Káčer et al. (2019)	SMEs	661,622	Hold out sample	Logit	0.844		N/A
Chi and Meng (2019)	Manufacturing SMEs	172	Hold out sample	Logit		0.912	N/A
Ma'aji et al. (2019)	SMEs	17,248	In-sample	Logit		0.872	2
Matthias et al. (2019)	SMEs	850	Hold out sample	Grabit	0.830		3
Sigrist and Hirnschall (2019)	SMEs	59,099	Hold out sample	NN	0.940		2
Altman et al. (2020)	SMEs	14,420	Hold out sample	Logit		0.821	2
Altman et al. (2020)	SMEs and large firms separately	6,200,000	In-sample	Logit		0.700	4
Ciampi et al. (2020)	Small Manufacturing enterprises	1,200	Hold out sample	Trajectory-based		0.854	3
Gabbi et al. (2020)	SMEs	16,850	In-sample	Logit		0.783	N/A
Gyimah et al. (2020)	Small enterprises	208	In-sample	Logit	0.939	0.865	1
Karas and Reznakova (2020)	SMEs	4,350	Hold out sample	Logit	0.820		N/A
Lee et al. (2020)	SMEs	4,358	Hold out sample	Boosting(DT)		0.827	N/A
Luo et al. (2020)	SMEs	15,605	Hold out sample	LightGBM	0.920	0.870	N/A
Ptak-Chmielewska and Matuszyk (2020)	SMEs	806	Cross-validation				N/A
Svabova et al. (2020)	SMEs	75,652	In-sample	MDA	0.934		N/A
Yin et al. (2020)	Manufacturing SMEs	1,091	Cross-validation	Logit	0.723		3
Caselli et al. (2021)	SMEs	14,917	No-validation				3
Corazza et al. (2021)	SMEs	76	Hold out sample	Elman network	0.963	0.936	3
Corazza et al. (2021)	SMEs	23,034	N/A				1
Crosato et al. (2021)	SMEs	319	Hold out sample	Logit		0.845	3
Dewaelheyns et al. (2021)	SMEs (excluding micro)	153,826	In-sample	Discrete-time hazards	0.902		N/A
Gao et al. (2021)	SMEs	123	Hold out sample	Soft voting	0.984		N/A
Grishunin et al. (2021)	SMEs	885	Cross-validation	Logit	0.820	0.840	N/A
Karas and Režňáková (2021)	SMEs	213,731	Hold out sample	Cox Proportional Hazards	0.853		N/A
Kou et al. (2021)	SMEs	672,116	Hold out sample	XGBoost	0.785		3
Li and Guo (2021)	Small enterprises	3,111	Hold out sample	Two-Stage Nonparametric Bayesian Discriminant Model	0.902	0.779	N/A
Li et al. (2021)	Micro and small enterprises	2,000	Hold out sample	G-XGBoost	0.748		N/A
Malakauskas and Lakstutiene (2021)	SMEs	12,000	Cross-validation	RF	0.680		N/A
Park et al. (2021)	SMEs and large firms separately	134,724	In-sample	Logit	0.710		N/A
Roy and Shaw (2021a)	SMEs	6	N/A				3

(*Continued*)

Table B1. (Continued).

Study	Focus	Sample Size	Validation	Best Est. Method	B.AUC	B.ACC	ABS Level
Roy and Shaw (2021b)	SMEs	31	In-sample	BWM-TOPSIS		0.903	N/A
Schalck and Yankol-Schalck (2021)	SMEs	579,892	Hold out sample				2
Séverin and Veganzones (2021)	SMEs	6,000	Hold out sample	NN	0.854	0.815	3
Stevenson et al. (2021)	SMEs	60,000	Hold out sample	Deep Learning BERT	0.880		4
Zizi et al. (2021)	SMEs	180	Hold out sample	Logit	0.959	0.950	N/A
Abedin et al. (2022)	SMEs	2,005	Cross-validation	RF	0.991	0.991	N/A
Altman et al. (2022)	SMEs	2,040	Hold out sample	Logit	0.903	0.826	3
Bangarigadu and Nunkoo (2022)	Micro and small enterprises	129	In-sample	Logit		0.837	1
Costa et al. (2022)	SMEs in the construction sector	7,790	In-sample	RF		0.877	N/A
Da and Peng (2022)	Technology-oriented micro and small enterprises	2,327	Hold out sample	NN	0.949	0.971	3
Du et al. (2022a)	SMEs	1,547	In-sample	Logit	0.791		N/A
Du et al. (2022b)	SMEs	1,547	Hold out sample	Logit	0.838		N/A
Karas (2022)	SMEs	202,209	Hold out sample	Cox Proportional Hazards	0.881		N/A
Lextrait (2023)	SMEs	1,610,419	Hold out sample	LightGBM	0.922		2
Long et al. (2022)	SMEs	7,943	Cross-validation	RF	0.741		N/A
Lu et al. (2022)	SMEs	2,044	Different data sets with different variables!	GBDT	0.650	0.890	3
Meng et al. (2022)	Small enterprises	3,045	Hold out sample	NN		0.911	2
Pacheco et al. (2022)	Manufacturing SMEs	208	In-sample				N/A
Pierri and Caroni (2022)	Small enterprises	21,667	In-sample	Logit	0.733		N/A
Kumar Roy et al. (2022)	SMEs	51	In-sample	FBWM TOPSIS-Sort-C		0.941	3
Shetty et al. (2022)	SMEs	3,728	Hold out sample	Deep Feedforward NN	0.850		N/A
Sun and Jiao (2022)	SMEs	900,000	Cross-validation	XGB (Focal Loss)	0.970	0.938	N/A
Sun et al. (2022)	Small industrial enterprises	1,820	Cross-validation	SVM	0.988		3
Zeng (2022)	SMEs	100	Hold out sample	CNN		0.968	N/A
Zhang and Song (2022b)	SMEs	14,366	Cross-validation	CNN-Logistic-Stacking	0.973		N/A
Zhang and Song (2022a)	SMEs	14,366	Cross-validation	GBDT-CNN-LR	0.992		N/A
Muthukumaran and Hariharanath (2023)	SMEs	1,740	Different data sets with different variables!	Optimal deep learning based FCP (ODL-FCP)		0.949	N/A
Papík and Papíková (2023)	SMEs	306,828	Hold out sample	CatBoost	0.941	0.908	1

Table B2. List of all papers included in the present review. Tested est. method stands for tested estimation method(s) which refers to the estimation method(s) the study tested but not necessarily considered as the primary estimation method(s), and primary est. method stands for primary estimation method(s). The variable category "governance" refers to firm and owner/manager characteristics features.

Study	Variable Category	Feature Selection	Tested Est. Method	Primary Est. Method
Edmister (1972)	Financial	Stepwise Method\Correlation Analysis	MDA	MDA
Keasey and Watson (1986)	Financial	Forward Stepwise Selection	Judgmental Consensus\LDA	Judgmental Consensus\LDA
Keasey and Watson (1987)	Financial\Non-financial	Forward Stepwise Selection	Logit	Logit
Keasey and Watson (1988)	Financial	N/A	LDA\Combination of discriminant\analysis and the simple decision rule	LDA\Combination of discriminant\analysis and the simple decision rule
Laitinen (1993)	Financial\Non-financial	Forward Stepwise Selection	Kruskal–Wallis nonparametric test\Logit	Logit
Lussier (1995)	Non-Financial\Governance	N/A	Logit	Logit
Wu and Wang (2000)	Financial\Non-financial \Credit	N/A	NN\k nearest neighbor method (kNN) \linear discriminant\quadratic discriminant	NN
Lussier and Pfeifer (2001)	Non-financial\Governance	N/A	Logit	Logit
Glennon and Nigro (2005)	Loan Characteristics \Lender Characteristics \Non-financial	N/A	Discrete-time hazard model (stacked-Logit)	Discrete-time hazards
Altman and Sabato (2007)	Financial	Accuracy ratio – as defined by Keenan and Sobehart (1999)	Logit\Altman Z?-Score\MDA	Logit
Behr and Güttler (2007)	Financial\Non-financial	N/A	Logit	Logit
Sohn and Kim (2007)	Financial\Non-financial	Significance	Logit	Logit
Angelini et al. (2008)	Financial\Non-financial \Credit	Removing variables containing more than 30% of missing and wrong values	Neural networks	NN
Arslan and Karan (2009)	Financial	VIF	Logit	Logit
Fantazzini and Figini (2009b)	Financial\Non-financial	N/A	Random Survival Forests (RSF)\Logit	RSF
Fantazzini and Figini (2009a)	Financial\Non-financial	N/A	Bayesian Logit\Logit	Bayesian Logit
Baixauli and Módica-Milo (2010)	Financial	N/A	Logit	Logit
Kim and Sohn (2010)	Financial\Non-financial	Stepwise Method	SVM\Logit\back-propagation neural networks (BPNs)	SVM
Lugovskaya (2010)	Financial\Non-financial	PCA\Forward Stepwise Selection	LDA	LDA
Lussier and Halabi (2010)	Non-financial\Governance	N/A	Logit	Logit
Moon and Sohn (2010)	Financial\Non-financial \Technology \Macroeconomic	Factor Analysis\Forward Stepwise Selection	Logit	Logit

(*Continued*)

Table B2. (Continued).

Study	Variable Category	Feature Selection	Tested Est. Method	Primary Est. Method
Norden and Weber (2010)	Credit	N/A	Probit	Probit
Pederzoli and Torricelli (2010)	Financial	Backward Stepwise Elimination	Logit	Logit
Sohn and Jeon (2010)	Non-financial\Governance	N/A	Weibull Competing Risk Model	Weibull Competing Risk Model
Yoon and Kwon (2010)	Credit card sales information\Transaction	t-test	SVM\backpropagation neural networks (BPN)\CART\C5.0\MDA\Linear Regression analysis (LRA)	SVM
Dereliolu and Gürgen (2011)	Financial\Non-financial	Decision Tree (DT)\SVM-RFE techniques for feature selection	k-Nearest Neighbor (k-NN)\Multilayer perceptron (MLP)\Support vector machine(SVM)	k-NN\MLP\SVM
Figini and Giudici (2011)	Financial\Non-financial \Prior payment behavior \Relational	N/A	Parametric longitudinal predictive models (LPM)\Semi-parametric duration models (SDM)	Parametric longitudinal predictive models (LPM)\Semi-parametric duration models (SDM)
Glennon and Nigro (2011)	Loan Characteristics \Lender Characteristics \Non-financial	N/A	Logit\Parametric survival\Discrete-time hazard\split-population survival-time model	Logit\Parametric survival\Discrete-time hazards\split-population survival-time model
Mittal et al. (2011)	Non-fFinancial\Credit	N/A	NN	NN
Rikkers and Thibeault (2011)	Financial\Non-financial	Forward Stepwise Selection	Logit	Logit
Terdpaopong and Mihret (2011)	Financial	Correlation Analysis	MDA	MDA
Yazdanfar (2011)	Financial\Non-financial	Forward Stepwise Selection	Logit	Logit
Gama and Geraldes (2012)	Financial\Non-financial \Relational\Governance \Credit	Forward Stepwise Selection	Logit	Logit
Lin et al. (2012)	Financial	Correlation Analysis\SAS stepwise selection	Logit\WoE-LR	Logit
Ciampi and Gordini (2013)	Financial	VIF\Stepwise Method	artificial neural networks\Logit\MDA	NN
Pederzoli et al. (2013)	Financial\Non-financial	Backward Stepwise Elimination	Logit	Logit
Gupta et al. (2014b)	Financial	Forward Stepwise Selection\Correlation Analysis	Logit	Logit
Gupta et al. (2014a)	Financial	Forward Stepwise Selection\Correlation Analysis	Logit	Logit
Halabi and Lussier (2014)	Non-financial\Governance	N/A	Probit	Probit
Kosmidis and Stavropoulos (2014)	Financial	Univariate discriminant Analysis\only one variable from each cluster is selected	MDA\Logit\Probit	MDA\Logit\Probit
Modina and Pietrovito (2014)	Financial	PCA\Forward Stepwise Selection	Logit	Logit

(*Continued*)

Table B2. (Continued).

Study	Variable Category	Feature Selection	Tested Est. Method	Primary Est. Method
Monelos et al. (2014)	Financial	Factor Analysis	Data Envelopment Analysis (DEA)\Logit \MDA	Data Envelopment Analysis (DEA)\Logit \MDA
Wilson and Altanlar (2014)	Non-financial\Governance \Credit\Macroeconomic	N/A	Logit	Logit
Marom and Lussier (2014)	Non-Financial\Governance	N/A	Logit	Logit
Angilella and Mazzù (2015)	Financial\Non-financial \Governance	N/A	Multicriteria model (ELECTRE-TRI)	Multicriteria model (ELECTRE-TRI)
Chen et al. (2015)	Financial\Non-financial \Macroeconomic \Governance	N/A	Logit\3SLS	Logit
Ciampi (2015)	Financial\Non-financial \Governance	VIF\Stepwise Method	Logit	Logit
DiDonato and Nieddu (2015)	Financial	N/A	Linear Discriminant Analysis (LDA) \Quadratic Discriminant Analysis (QDA) \Classification Trees (CARTs)	LDA\QDA\CART
Gupta et al. (2015)	Financial\Non-financial	VIF\Average Marginal Effect-AME	Discrete time duration-dependent hazard model	Discrete-time hazards
Khermkhan and Chancharat (2015)	Financial	Significance\Correlation Analysis	MDA\Logit\Probit\NN	MDA\Logit\Probit\NN
Piatti et al. (2015)	Financial\Non-financial	PCA	Logit	Logit
Zhang and Thomas (2015)	Financial\Non-financial \Credit	Univariate Analysis\Stepwise Method	Logit	Logit
Guzmán and Lussier (2015)	Non-financial\Governance	N/A	Logit	Logit
Abdullah et al. (2016a)	Financial\Non-financial	Forward Stepwise Selection\VIF	Logit	Logit
Abdullah et al. (2016b)	Financial\Non-financial \Governance	Forward Stepwise Selection\VIF	Logit	Logit
Calabrese et al. (2016)	Financial	Backward Stepwise Elimination\VIF	BGEVA\log-log\Logit	BGEVA
Corazza et al. (2016)	Financial	Correlation Analysis	Multicriteria Ranking Method (MURAME)	Multicriteria Ranking Method (MURAME)
Cultrera and Brédart (2016)	Financial\Non-financial	VIF\Correlation Analysis	Logit	Logit
El Kalak and Hudson (2016)	Financial\Non-financial \Macroeconomic	Univariate regression Analysis \Correlation Analysis	Discrete-time duration-dependent hazard model	Discrete-time duration-dependent hazard model
Filipe et al. (2016)	Financial\Non-financial \Macroeconomic	AUC for each ratio\Correlation Analysis \Forward Stepwise Selection	Multi-period Logit \Cox proportional hazard model)	Multi-period Logit\Cox proportional hazard
Inekwe (2016)	Financial\Non-financial \Macroeconomic	N/A	Panel conditional Logit	Panel Logit
Li et al. (2016)	Financial\Non-financial	PCA	Logit\ANN\ANN/logistic hybrid model	ANN-logistic hybrid model

(*Continued*)

Table B2. (Continued).

Study	Variable Category	Feature Selection	Tested Est. Method	Primary Est. Method
Wilson et al. (2016)	Financial\Non-financial \Governance	Significance	Logit\Cox Proportional Hazards	Logit\Cox Proportional Hazards
Yoshino et al. (2016)	Loan information	PCA	Probit	Probit
Hyder and Lussier (2016)	Non-financial\Governance	N/A	Logit	Logit
Lussier et al. (2016)	Non-financial\Governance	N/A	Logit	Logit
Chi and Zhang (2017)	Financial\Non-financial \Macroeconomic \Governance	Rank Sum Test-Mann-Whitney \Correlation Analysis	Entropy weighting method\MDA	Entropy weighting method
Ciampi (2017)	Financial\Non-financial \Governance	VIF\Stepwise Method	Logit	Logit
Figini et al. (2017)	Financial\Non-financial \Credit	N/A	Classification Tree\k-NN\LDA\GLM\BGEV \Gradient Boosting Machines (GBM) \Random Forest (RF)	DT\k-NN\LDA\GLM\BGEV\GBM\RF
Oliveira et al. (2017)	Non-financial	N/A	Cognitive mapping with MACBETH	Cognitive mapping with MACBETH
Pierri and Caroni (2017)	Financial\Non-financial \Macroeconomic	Forward Stepwise Selection	Cox semi-parametric proportional hazards regression\Logit	Cox Proportional Hazards\Logit
Tobback et al. (2017)	Financial\Non-financial \Relational	N/A	Weighted-vote relational\neighbor (wvRN) classifier (SVM with a linear kernel)	SVM
Andrikopoulos and Khorasgani (2018)	Financial\Market (using listed SMEs information)	Forward Stepwise Selection	Hybrid (Logit and Merthon-KMV)	Hybrid (Logit and Merthon-KMV)
Baidoun et al. (2018)	Non-financial\Governance	Stepwise Method	Logit	Logit
Castillo et al. (2018)	Financial	N/A	Logit	Logit
Ciampi (2018)	Financial\Non-financial \Governance\corporate social responsibility (CSR)	Altman (1968) method\Forward Stepwise Selection\Backward Stepwise Elimination\Filter Methods-zero and first order\Wrapper Method	Logit	Logit
Duarte et al. (2018)	Financial\Non-financial \Macroeconomic\Credit \Loan Characteristics	N/A	Probit	Probit
Gabbianelli (2018)	Financial\Non-financial	Variables characterized by a greater predictive power\VIF	Logit	Logit
Gupta and Gregoriou (2018)	Financial\Non-financial	Significance\Average Marginal Effect-AME\Correlation Analysis	Logit\Discrete-time hazard\Continuous-time Cox Proportional Hazards model	Discrete-time hazards\Cox Proportional Hazards
Gupta et al. (2018)	Financial\Non-financial	Average Marginal Effect-AME \Correlation Analysis	Multivariate hazards model	Multivariate hazards model

(*Continued*)

Table B2. (Continued).

Study	Variable Category	Feature Selection	Tested Est. Method	Primary Est. Method
Gupta et al. (2018)	Financial\Non-financial \Market	Average Marginal Effect-AME \Correlation Analysis	Panel Logit	Panel Logit
Ptak-Chmielewska and Matuszyk (2018)	Financial\Non-financial	Forward Stepwise Selection	Gradient boosting RF\Logit\DT\NN	GB(RF)
Abdullah et al. (2019)	Financial\Non-financial	Forward Stepwise Selection\VIF	Logit	Logit
Angilella and Mazzù (2019)	Financial\Non-financial \Governance	N/A	Multicriteria model (ELECTRE-TRI)	Multicriteria model (ELECTRE-TRI)
Chai et al. (2019)	Financial\Non-financial \Macroeconomic \Governance	Correlation Analysis\Probit regression	Entropy-weighting TOPSIS with fuzzy C-means (FCM)	Entropy-weighting TOPSIS with fuzzy C-means (FCM)
Chi and Meng (2019)	Financial\Non-financial \Macroeconomic \Governance	F-test\Correlation Analysis	A weighting method	A weighting method
Cornée (2019)	Non-financial\soft-information	N/A	Probit\Cox proportional hazard model	Probit
Káčer et al. (2019)	Financial\Non-financial \Macroeconomic	Stepwise Method	Logit	Logit
Ma'aji et al. (2019)	Financial\Non-financial \Governance	Forward Stepwise Selection	MDA\Logit	MDA\Logit
Matthias et al. (2019)	Financial\Non-financial	N/A	Logit	Logit
Sigrist and Hirnschall (2019)	Financial\Non-financial \Credit\Loan Characteristics\Social media platform rating \online user behavior data	N/A	Logit\classification tree\RF\boosted Logit \NN\Tobit\boosted multiclass Logit \Grabit	Grabit
Altman et al. (2020)	Financial\Non-financial \Governance	Linear R-squared method\Forward Stepwise Selection	Decision tree (DT)\Generalized boosting (GB)\Logit (LR)\neural network with multi-layer perceptron (NN)\support vector machine (SVM)	DT\GB\Logit\NN\SVM
Altman et al. (2020)	Financial	Forward Stepwise Selection	Logit	Logit
Cathcart et al. (2020)	Financial\Non-financial \Macroeconomic	N/A	Logit	Logit

(*Continued*)

Table B2. (Continued).

Study	Variable Category	Feature Selection	Tested Est. Method	Primary Est. Method
Ciampi et al. (2020)	Financial\Non-financial \Governance\Prior payment behavior	Altman (1968) method\Forward Stepwise Selection\Backward Stepwise Elimination\Filter Methods-zero and first order\Wrapper Method	Logit\Discrete-time hazard\Trajectory-based	Logit\Discrete-time hazard\Trajectory-based
Gabbi et al. (2020)	Financial\Nonffinancial \Relational\Governance	N/A	Logit	Logit
Gyimah et al. (2020)	Non-financial\Governance	N/A	Logit	Logit
Karas and Reznakova (2020)	Financial	PCA\Classification and Regression Trees-CART\Kaiser-Meyer-Olkin test-KMO	Logit	Logit
Lee et al. (2020)	Non-financial\Governance \Technical ability	Backward Stepwise Elimination \Wrapper Method	Logit\DT\NN\Boosting (Logit)\Boosting (DT)\Boosting (ANN)	Logit\DT\NN\Boosting(Logit)\Boosting(DT) \Boosting(ANN)
Luo et al. (2020)	Non-financial\Credit\Court verdict	Gini impurity index	Logit\CART\LightGBM	Logit\CART\LightGBM
Ptak-Chmielewska and Matuszyk (2020)	Financial\Non-financial	Relative variable importance	Random Survival Forests (RSF) \semiparametric Cox regression survival model	RSF
Svabova et al. (2020)	Financial\Non-financial	Forward Stepwise Selection	Logit\MDA	Logit\MDA
Yin et al. (2020)	Financial\Non-financial	chi-squared test\Correlation Analysis	Logit\RF\XGBoost	Logit\RF\XGBoost
Caselli et al. (2021)	Financial\Non-fFinancial	N/A	Cox proportional hazards\The two-step Heckman model (Probit to estimate inverse Mills' ratio "IMR". Second stage Cox proportional hazards including the estimated IMR)	Cox proportional hazards
Corazza et al. (2021)	Financial\Non-financial \Credit\Macroeconomic	N/A	Elman network\NN\Logit	Elman network
Corazza et al. (2021)	Financial	N/A	MUlticriteria RAnking MEthod (MURAME)	MUlticriteria RAnking MEthod (MURAME)
Crosato et al. (2021)	Financial\Non-financial \Website quality	N/A	Logit\Kernel Discriminant Analysis (KDA)	Logit\Kernel Discriminant Analysis (KDA)
Dewaelheyns et al. (2021)	Financial\Non-financial \Governance	N/A	Discrete time hazard models\Multinomial Logit models	Discrete-time hazards
Gao et al. (2021)	Financial\Non-financial \Relational\Credit	Bayesian optimization	Logit\SVM\RF\XGBoost\LightGBM\Soft voting	Logit\SVM\RF\XGBoost\LightGBM\Soft voting
Grishunin et al. (2021)	Financial\Non-financial \Macroeconomic \Governance	Weight of evidence method-\VIF	Logit	Logit

(*Continued*)

Table B2. (Continued).

Study	Variable Category	Feature Selection	Tested Est. Method	Primary Est. Method
Karas and Režňáková (2021)	Financial\Non-fFinancial \Macroeconomic	Univariate Cox proportional hazard model\Correlation Analysis\VIF \Forward Stepwise Selection	Cox semiparametric model	Cox Proportional Hazards
Kou et al. (2021)	Financial\Non-financial \Network\Payment and Transaction	the optimum-seeking method\NSGA-II	LDA\Logical Regression\SVM\DT\RF\XGB \NN	LDA\Logical Regression\SVM\DT\RF\XGB \NN
Li and Guo (2021)	Financial\Non-financial \Governance \Macroeconomic\Credit	NonParametric Bayesian discrimination \Parametric Bayesian discrimination	Two-stage Nonparametric Bayesian Discriminant Model\Two-stage Parametric Bayesian Discriminant Model \Two-stage Logit model	Two-stage Nonparametric Bayesian Discriminant Model\Two-stage Parametric Bayesian Discriminant Model
Li et al. (2021)	Financial\Non-financial \Governance \Macroeconomic\Credit	N/A	XGBoost\GAN and XGBoost (G-XGBoost)	G-XGBoost
Malakauskas and Lakstutiene (2021)	Financial\Non-financial	N/A	Logit\ANN\RF	Logit\NN\RF
Park et al. (2021)	Financial\Non-financial	N/A	Logit	Logit
Roy and Shaw (2021a)	Financial\Non-financial	N/A	AHP-TOPSIS	AHP-TOPSIS
Roy and Shaw (2021b)	Financial\Non-fFinancial \Governance	N/A	BWM-TOPSIS	BWM-TOPSIS
Schalck and Yankol-Schalck (2021)	Financial\Non-financial \Governance	N/A	Dynamic Probit model\Logistic LASSO regression\XGBoost	Probit\Logistic LASSO Regression\XGBoost
Séverin and Veganzones (2021)	Financial\Non-fFinancial	t-test\Kruskal-Wallis tests\Correlation Analysis\Wrapper Method	LDA\Logit\NN\ELM\SVM	LDA\Logit\NN\ELM\SVM
Stevenson et al. (2021)	Financial\Non-financial \loan\Governance \Textual	RF Feature Selection	Logit\RF\Deep Learning BERT	Deep Learning BERT
Zizi et al. (2021)	Financial	Stepwise Method\LASSO	Logit\NN	Logit\NN
Abedin et al. (2022)	Financial\Non-financial \Macroeconomic \Governance\Credit	N/A	C4.5\k-NN\SVM\Bagging\Boosting\LB (Logit boost)\RC (random committee) \RTF (rotation forest)\RF	RF
Altman et al. (2022)	Financial\Non-financial \Governance\Credit	LASSO	Logit\MDA (Omega Score)\RF\XGBoost	Logit
Bangarigadu and Nunkoo (2022)	Non-financial\Governance	N/A	Logit	Logit

(*Continued*)

Table B2. (Continued).

Study	Variable Category	Feature Selection	Tested Est. Method	Primary Est. Method
Costa et al. (2022)	Financial\Non-financial \Financial Report Quality Measures \Macroeconomic	Backward Stepwise Elimination method \Pearson Correlation Analysis	Logit\Random Forest	Logit\RF
Da and Peng (2022)	Financial\Non-financial \Innovation\Business model	N/A	Logit (LR)\support vector machine (SVM) \decision tree (DT)\random forest (RF) \neural network (NN)	Logit\SVM\DT\RF\NN
Du et al. (2022a)	Financial\Non-fFinancial \Textual\Relational	N/A	Logit	Logit
Du et al. (2022b)	Financial\Non-financial \Textual\Relational	mRMR\FCBF\mIMR\RCDFS\FS-RRC\The Proposed Feature Selection Algorithm-introduce in this paper	Logit	Logit
Karas (2022)	Financial\Non-financial \Macroeconomic	Univariate regression Analysis \Correlation Analysis\Forward Stepwise Selection	Cox semiparametric model	Cox Proportional Hazards
Lextrait (2023)	Financial\Non-financial	PCA	LightGBM\CatBoost\XGBoost\SVM\Logit	LightGBM\CatBoost\XGBoost
Long et al. (2022)	Financial\Non-financial \Relational	N/A	Logit\RF\XGBoost	Logit\RF\XGBoost
Lu et al. (2022)	Financial\Non-financial \Governance \Macroeconomic\Credit	BOWOA\KS statistic\RF Feature Selection	K-Nearest Neighbor (KNN)\Linear Regression\SVM\DT\RF\GBDT\XGBoost	k-NN\Linear Regression\SVM\DT\RF\GBDT \XGBoost
Meng et al. (2022)	Financial\Non-financial \Governance \Macroeconomic\Credit	Neural network stepwise screening \Correlation Analysis	NN	NN
Pacheco et al. (2022)	Financial\Non-financial	Correlation Analysis	Logit	Logit
Pierri and Caroni (2022)	Financial\Non-financial	Backward elimination	Logit	Logit
Kumar Roy et al. (2022)	Financial\Non-financial \Governance\Credit	N/A	FBWM TOPSIS-Sort-C	FBWM TOPSIS-Sort-C
Shetty et al. (2022)	Financial	N/A	XGBoost\SVM\Deep Feedforward NN	XGBoost\SVM\Deep Feedforward NN
Sun and Jiao (2022)	Non-financial\Loan Characteristics \Macroeconomic	Correlation Analysis	Logit\SVM\MLP\RF\XGBoost\XGB (Focal Loss)	XGB(Focal Loss)

(*Continued*)

Table B2. (Continued).

Study	Variable Category	Feature Selection	Tested Est. Method	Primary Est. Method
Sun et al. (2022)	Financial\Non-financial \Governance\Credit\loan \Macroeconomic	Correlation Analysis\Univariate method \Backward Stepwise Elimination \LASSO\RF Feature Selection \neighborhood\rough set-NRS feature selection method	Logit\k-NN\SVM\NB\RF	Logit\k-NN\SVM\NB\RF
Zeng (2022)	Financial	N/A	Convolutional neural network (CNN) \multivariate linear model\Logit\BPNN	CNN
Zhang and Song (2022b)	Financial\Non-financial	Correlation Analysis	Linear Regression\Logit\Bayesian ridge regression\CNN\NN\LGB\XGB\NN-ATT \NN- LGB\NN -XGB\CNN-LGB\CNN-XGB \NN-ATT-LGB\NN-ATT-XGB\NN-ATT-LGB-XGB	Linear Regression\Logit\Bayesian ridge regression\CNN\NN\LightGBM\XGBoost \NN-ATT\NN-LGB\NN-XGB\CNN-LGB \CNN-XGB\NN-ATT-LGB\NN-ATT-XGB \NN-ATT-LGB-XGB
Zhang and Song (2022a)	Financial\Non-financial	GBDT model	GDBT-LR\GDBT-CNN-LR\RF\DT\Logit\SVM \MLP\GaussianNB\KNN	GBDT-CNN-LR
Muthukumaran and Hariharanath (2023)	N/A	Arithmetic Optimization Algorithm-AOA	Optimal deep learning based FCP (ODL-FCP)\QABOLSTM\LSTM-RNN\ACO Model \MLP Model\SVM Model\AdaBoost Model	ODL-FCP\QABOLSTM\LSTM-RNN\ACO \MLP\SVM\AdaBoost
Papík and Papíková (2023)	Financial\Non-financial	N/A	CatBoost\LightGBM\XGBoost	CatBoost\LightGBM\XGBoost

Appendix C

Table C1. Focus of SMEs failure studies.

Focus	ABS	Non-ABS	N
SMEs	57	34	91
Small enterprises	23	10	33
Manufacturing SMEs	2	3	5
Innovative SMEs	3	0	3
Micro and small enterprises	1	1	2
SMEs and large firms separately	1	1	2
Large firms, small firms, and individuals	1	0	1
Micro-enterprises	1	0	1
Small industrial enterprises	1	0	1
Small Manufacturing enterprises	1	0	1
Small wholesale and retail enterprises	1	0	1
SMEs operating in the food or beverage manufacturing sector	1	0	1
Technology-oriented micro and small enterprises	1	0	1
SMEs (excluding micro)	0	1	1
SMEs in the construction sector	0	1	1

Table C2. Detailed distribution of data sources used in SME default papers obtained by surveys, questionnaires, and interviews.

Source	ABS	Non-ABS	N
Survey	6	2	8
Questionnaire	4	3	7
Personal interview survey	2	1	3
Interview	1	0	1
Panel interview	1	0	1

Table C3. Detailed distribution of data sources used in SME default papers obtained from public sources and web-pages.

Source	ABS	Non-ABS	N
Published reports	2	1	3
Data was crawled using python programming from multiple platforms and sources	1	0	1
Online data obtained by web scraping companies websites	1	0	1
Rossiyskaya gazeta (newspaper that publishes announcements on all bankruptcy-related news)	1	0	1
www.qcc.com	1	0	1
Publicly available external credit data	0	1	1

Table C4. Detailed distribution of data sources used in SME default papers obtained from data services.

Source	ABS	Non-ABS	N
AIDA (by Bureau Van Dijk)	6	2	8
Amadeus (by Bureau Van Dijk)	4	4	8
CERVED	3	3	6
Compustat	5	1	6
Datastream (by Thomson Reuters)	3	1	4
Capitaline Database	2	1	3
Creditreform	1	2	3
EUROSTAT	1	2	3
SABI (by Bureau Van Dijk)	1	2	3
BelFirst (by Bureau Van Dijk)	1	1	2
Diane (by bureau Van Dijk)	2	0	2
FAME (by Bureau Van Dijk)	2	0	2
Finstat	1	1	2
ORBIS (by Bureau Van Dijk)	2	0	2
Suomen Asiakastieto	2	0	2
CSMAR	0	2	2
Affärsdata	1	0	1
ATO	1	0	1
Bureau Van Dijk	1	0	1
Dun & Bradstreet (D&B)	1	0	1
ECB	1	0	1
EPO BULLETIN	1	0	1
ICAP	1	0	1
INPI opendata service	1	0	1
inter alia	1	0	1
K-VAN service	1	0	1
PATSTAT	1	0	1
SABI (by Informa SA)	1	0	1
SPARK	1	0	1
Albertina Platinum	0	1	1
AnalcatData	0	1	1
Australian Credit	0	1	1
CNINF	0	1	1
German Credit	0	1	1
KOSME	0	1	1
NEEQ SMEs dataset	0	1	1
Pordata	0	1	1
SPARK-Interfax	0	1	1
Transparency International databases	0	1	1
Wind database	0	1	1

Table C5. Detailed distribution of data sources used in SME default papers obtained from banks, financial institutions, and firms.

Source	ABS	Non-ABS	N
A Chinese bank	1	2	3
AChinese commercial bank	1	2	3
World Bank	1	2	3
A bank in China	1	0	1
A bank in Italy	1	0	1
A commercial bank	1	0	1
A commercial bank in a Chinese city	1	0	1
A commercial credit reference database	1	0	1
A credit card provider	1	0	1
A Dutch bank	1	0	1
A French social bank	1	0	1
A German universal bank	1	0	1
A leading bank in Central New York	1	0	1
A major Chinese city commercial bank	1	0	1
A major commercial bank operating in Portugal	1	0	1
A major German promotional bank	1	0	1
A specialized Micro and SME lender	1	0	1
A Taiwanese finance company	1	0	1
A UK Credit Reference Agency (CRA)	1	0	1
Advanon, a Swiss start-up company	1	0	1
An Italian bank	1	0	1
UniCredit bank	1	0	1
Yapı ve Kredi Bankası A.S.	1	0	1
A consultancy firm	0	1	1
A large Italian commercial bank	0	1	1
Major banks	0	1	1
Swedbank AB	0	1	1

Table C6. Detailed distribution of data sources used in SME default papers obtained from ministries, public offices, and universities.

Source	ABS	Non-ABS	N
SBA (Small Business Administration)	3	1	4
Credit Management Research Center of the University of Leeds	3	0	3
the Companies Commission of Malaysia (CCM) database	2	1	3
China Judgements Online	1	1	2
Department of Business Development (DBD) Thailand	1	1	2
the Centres for Urban and Regional Development Studies – The University of Newcastle-upon-Tyne (England)	1	1	2
GEM (the Growth Enterprise Market from Shenzhen Stock Exchange)	0	2	2
SB (the Small and Medium-Sized Enterprise Board from Shenzhen Stock Exchange)	0	2	2
STAR (the Science and Technology Innovation Board from Shanghai Stock Exchange)	0	2	2
The Chamber of Commerce of Perugia	0	2	2
Bankruptcy court records, obtained in person at 6 bankruptcy courts (see Lussier (1995))	1	0	1
Census datasets	1	0	1
Central Credit Register (Italy)	1	0	1
Centrale Rischi Finanziari (CRIF)	1	0	1
Fondo Nacional de Garantías (FNG)	1	0	1
Legal Execution Department, Ministry of Justice (Thailand)	1	0	1
Office for National Statistics (UK)	1	0	1
Superintendencia de Sociedades	1	0	1
Technology credit guarantee fund recipient data	1	0	1
Technology credit loan recipient data	1	0	1
The Chamber of Commerce	1	0	1
The National Board of Patents and Registration of Trademarks	1	0	1
The SIRENE system of INSEE (French National Statistics Office)	1	0	1
The Small Business Administration and Robert Morris Associate	1	0	1
International Monetary Fund (IMF)	0	1	1
Italian National Institute of Statistics (ISTAT)	0	1	1
Manufacturing SMEs dataset (China)	0	1	1
NSFC (National Natural Science Foundation of China)	0	1	1
Statistics of U.S. Businesses-Consensus Bureau	0	1	1
The Companies Commission of Malaysia (CCM)	0	1	1
The Korea Credit Guarantee Fund	0	1	1
The National Credit Bureau (Thailand)	0	1	1