Bank failure prediction models: Review and outlook☆

Alberto Citterio

Department of Economics, University of Insubria, Via Monte Generoso 71, Varese, 21100, Italy

ABSTRACT

This paper presents a literature review of recent empirical contributions on bank default prediction. The topic has always been important in the banking and finance literature, but it gained increasing interest especially after the 2008–2009 global financial crisis. The significant consequences of bankruptcy cases have highlighted the need for managers and regulators to develop and adopt appropriate early warning systems.

Previous studies are analysed in this review according to three different aspects: definition of default and financial distress, application of statistical and intelligent techniques, and the selection of variables. The review also proposes some possible upgrades to promote future research on the topic, that is, pointing out the potential role of non-financial information and ESG (Environmental, Social and Governance) variables as a determinant in the default prediction of banking institutions. While the predictive accuracy of macroeconomic information has been tested in several works, there is only limited evidence of the influence of non-financial performance on a bank's probability of default.

1. Introduction

Corporate failure prediction represents a key issue in managerial decision-making for financial and non-financial firms [1]. A business failure, or financial distress in general, is usually anticipated by several warning signs of different intensities. The proper identification and interpretation of such symptoms might be useful for managers to promptly adopt the most appropriate measures to reduce the adverse consequences of a crisis. Prediction models can therefore be used to predict whether a business will suffer financial distress and to define the related determinants using mathematical and statistical methods.

In the banking sector, this issue is particularly relevant because of the consequences that financial distress (and eventually bank failure) can have on market stability and on depositors and investors, and on the economic growth. Given its role in injecting money and credit into the economy, the banking system represents a major conduit through which instability may be transmitted to other financial and non-financial sectors and, ultimately, to the real economy. This transmission may be driven by different mechanisms, such as disruption of the interbank lending channel and payments services, reduction of credit to the economy, and collapse in the value of collaterals [2]. First, unlike non-financial industries, banking crises generate negative externalities for other banks in the market in the form of a decline in confidence in the stability of the financial system and losses from interbank exposures to

failed banks [3]. Given the strong interconnectedness between banks, idiosyncratic failures may then degenerate in systemic bank crises, causing national and international contagions [4].

Second, the reduction in bank credit may lead to underutilization and misallocation of funds, which in turn can affect aggregate consumption and investment [5]. International Monetary Fund [6] and World Bank [7] estimates place the effective fiscal cost of a banking distress as high as half of a nation's annual GDP. The cost of banking crises can therefore be quantified in terms of output losses, increases in unemployment, fiscal costs associated with bank support measures, and increases in public debt.

Regulation has also moved to a tighter approach to detect prodromes of crisis in the banking sector and anticipate potential difficulties by credit institutions to act in a timely manner, to monitor the stability of the financial system, and to guide macroprudential policy ([8]). If bank examiners can detect problems early enough, regulatory actions can be taken either to prevent a bank from failing or minimize the cost to authorities, taxpayers, and society as a whole [9]. In other words, the implementation of an early warning system enables better identification of problematic institutions and a more efficient allocation of governmental and authority resources.

However, bank regulators and supervisors cannot prevent bank from assuming risk because risk taking is a natural part of banking transactions. The unique functions of a bank to transform risk and create

liquidity by financing themselves with highly liquid, low risk deposits and investing in higher-risk, illiquid assets expose the bank to several types of risks, such as interest-rate risk, market risk, credit risk, liquidity risk, off-balance-sheet risk, foreign-exchange risk, and others [10,11]. Bank failure prediction models help in this respect, as they generate a better understanding of a bank's business and provide a tool for managers to identify areas where the bank is more vulnerable to failure risk.

The construction of an appropriate prediction model could be particularly important in the banking sector, both for managers and supervisors. Therefore, literature has extensively focused on developing increasingly precise and sophisticated models to enhance the precision and reliability of these instruments.

Bank distress prediction models first drew on the general bankruptcy prediction models designed for manufacturing firms. The most well-known model is built on work by Altman [12] on publicly traded firms. Using multiple discriminant analysis, Altman provided an original Z-Score model based on five financial statement ratios: working capital to total assets, retained earnings to total assets, earnings before interest and taxes to total assets, market value equity to book value of total debt, and sales to total assets. The Z-Score model has therefore been widely applied as a measure of distance to default. Consequently, Sinkey [13] was the first to adopt the same model to predict bank failures in the US from 1969 to 1972. Through implementation of the Z-score model, the author was able to correctly predict 72% of problem banks and find that asset composition, loan characteristics, capital adequacy, source of revenues, efficiency and profitability are the best predictor for bank distress. Since then, the use of discriminant analysis has evolved through different published articles (e.g., Ref. [14–16]). To overcome the drawbacks of discriminant analysis, several authors have then started to use alternative statistical methods, such as logistic, probit, and hazard models (e.g., Ref. [15,17,18]). Starting from the 90s, with the development of information sciences, artificial intelligence models such as artificial neural networks, evolutionary approaches, support vector machines, and hybrid intelligent methods have become popular for predicting financial problems (e.g., Ref. [19–22]). Bell et al. [19] were the first to apply an artificial intelligence method to predict problems in the banking sector. In their analysis, they demonstrated the superiority of artificial intelligence on traditional statistical models. Similarly, Tam and King [21] compared the discriminant analysis, logit model, K-nearest neighbour and artificial neural network and found that the latter outperform other techniques. Numerous comparative studies were subsequently conducted (e.g., Ref. [23–27]). However, although several studies highlighted the superior predictive ability of machine learning techniques, the supremacy of one method over another remained subject to various controversies. First, if non-parametric techniques do not require researchers to define the structure of the model a priori and do not depend on strict assumptions like normal distribution and no correlations between independent variables, they usually offer less interpretable output. More precisely, non-parametric techniques do not offer the possibility to determine which variables are the most useful in predicting the event, limiting the analysis to the effectiveness of the variable in discriminating between the two groups (i.e., failed and non-failed banks). This is the reason why majority of artificial intelligence methods function as black box [28].

Second, the comparison between models tuned out to be strictly influenced by heterogeneity of data in the validation (period, sample size, sample selection, nature, and number of explanatory variable). In particular, there is a big debate in the current literature regarding the superiority of certain indicators in predicting bank failures. Historically, the literature on bank failures has drawn on the Uniform Financial Institutions Rating System (UFIRS), also known as CAMELS rating, introduced by US regulators in 1979 [29]. The UFIRS is an internal supervisory tool for evaluating the soundness of financial institutions on a uniform basis. The CAMELS rating relies on balance sheet indicators which measure capital adequacy, asset quality, management quality, earnings, liquidity, and sensitivity to market risk. While different studies have shown the usefulness of the CAMELS rating [9,30], other authors have highlighted the rapid decay of this indicator [31], and the opportunity to include other variables that contain useful predictive information, such as macroeconomic and market price-base indicators [32, 33]. Furthermore, the CAMELS rating is limited to representing only the current condition of a bank [34], it fails to account for a change in the behaviour of agents [35], and it may miss extreme events, since CAMEL rating is built on a going-concern paradigm [36]. The literature has therefore stressed how advanced statistical prediction tools allow the relationship between financial indicators and bank failures to be determined and, more importantly, high-risk banks to be identified reasonably well in advance.

Despite the extensive literature and debate on this topic, the approaches still remain very diverse in terms of definition of default and financial distress, application of statistical and artificial intelligence techniques, and variable selection.

In addition, recent defaults episodes have emphasized the necessity to reassess the structure of prediction models in light of emerging concepts and phenomena come into the area of financial default prediction. A very explicative episode is the recent failure of Silicon Valley Bank (SVB), a California state-chartered bank that ranked as the sixteenth largest in the US by the end of 2022. It boasted $211 billion in assets, $175 billion in deposits, operated 47 branches, and employed around 8500 individuals. The bank primarily engaged with Californian startups, managing their liquidity and private equity funds.

Following a period of rapid expansion that effectively doubled the bank's size, shifts in the macroeconomic landscape swiftly plunged the institution into a crisis. The Federal Reserve's decision to hike its policy rate on March 17, 2022, in response to unprecedent levels of inflation (increasing the policy rate from 0% to 4.75% since March 2022), had a far-reaching repercussion. While these rate increases were initially anticipated to benefit the banking sector by allowing higher net interest margins for new loans [37], they introduced vulnerabilities for existing fixed-rate assets and led to increasing difficulty for Silicon Valley startups in accessing sources of financing at sustainable costs. Consequently, these startups progressively withdrew liquidity from the bank at an unexpectedly rapid pace to sustain their operations.

As a main consequence, to deal with the outflow of liquidity, SVB found itself compelled to liquidate a significant portion of its securities portfolio, resulting in substantial losses and a weakening of its capital ratios. This situation contributed to a "bank run" among uninsured depositors and a rapid devaluation of the bank's shares (between March 8 and 9, the share price more than halved, dropping from $267 to $105 per share), necessitating intervention by the Federal Deposit Insurance Corporation (FDIC). On the same days, other US banks witnessed sharp declines in share prices and experienced substantial withdrawals by depositors. On March 12, the Signature Bank of New York (SBNY) also failed, followed shortly by the acquisition of First Republic Bank by JP Morgan Chase on May 1. The FDIC estimated the total cost of its intervention, net of recoveries, in favor of these three banks at $35.5 billion.

This scenario has underscored the presence of new factors prompting banking crises and highlighted the inadequacy of the old CAMELS ratings, which had assigned an overall positive supervisory rating to SBNY until March 10—just two days before the bank's bankruptcy [38].

Among the newly identified factors following these events, one main concern revolves around the spillover effects stemming from unrealized losses in highly liquid securities. Unlike traditional bank run scenarios where liquid assets act as buffers, the realization of losses on these securities due to funding outflows posed significant problems. Additionally, while bank runs have been extensively studied in banking theory and empirical research [39], new factors have accelerated such crises. These include a highly concentrated deposit base, extensive connections between depositors, the influence of social media disseminating real-time negative (sometimes false) news about the bank, and the new technologies enabling immediate fund withdrawals from the bank. Third, despite the implementation of regulatory reforms after the 2008

financial crisis, banks run, and contagions remains important features of the banking system. For instance, Jiang et al. [40] recently demonstrated that if even just half of uninsured depositors decide to withdraw, almost 190 US banks could face potential risks that extend to insured depositors, endangering approximately $300 billion in insured deposits.

These recent episodes have reignited interest in financial default prediction, highlighting the need for an updated review. While there have been several reviews in this literature, many are either outdated [41,42] or narrowly focused on corporate prediction [43–47].

This paper aims to provide an updated overview by analyzing the primary prediction models proposed in literature spanning from 2000 to 2022.[1] Specifically, it delves into how researchers approach crucial concepts in constructing bankruptcy prediction models, with a dedicated focus on the banking sector. Furthermore, it contributes to the literature by offering insights into the role and potential contributions of non-financial information in enhancing the predictive capabilities of early-warning models.

Through this analysis, this work lays the groundwork for outlining possible future developments and delineates practical implications for both managers and policymakers, in light of the growing importance of non-financial information.

The remainder of this paper is organized as follows: Firstly, we introduce the methodology. In Section 3, we discuss the various definitions of business failure and financial distress in the literature. Section 4 outlines the most frequently used models and methods in this area, comparing their predictive accuracy. Moving to Section 5, we present an overview of the variables used in the literature as predictors of bank default, while Section 6 delves into the role of ESG variables and their impact on performance, reputation, and consequently, bank risk. Finally, the last section provides concluding remarks.

2. Methodology

To answer our research question and to gain insights into the current debate and identify leading trends in the literature, we collect the key papers published within the period of 2000–2022, following three steps. Firstly, we chose the database for data extraction. To gather the necessary data, the study uses the Scopus database, a reliable tool commonly employed in previous research (e.g., Ref. [48]), known for its lack of bias toward publishers. Then we selected the keywords able to restrict our analysis on papers covering the topic under consideration. We use three set of terms: the first is composed by the terms "fail*", "fragility", "distress", "stress", "insolvenc*", "default", "bankrupt*", "disapp*", and "early warning system"; the second includes the words "predict*", "anticipat*" and "forecast*"; the third contains the terms "bank", "banks" or "banking". To be included in the analysis, papers must include in the title, keywords or abstract at least one keyword of each set of terms. The results were then filtered to include only English peer-reviewed articles published between 2000 and 2022 and belonging to the following subject areas: "Economics, Econometrics and Finance"; "Business, Management and Accounting" and "Computer sciences". The prediction of corporate default has indeed traditionally been an active area of research, not only in finance but also in fields like data analytics and operations research/management science [41,46]. This search query yields total 1120 articles published between 2000 and 2022. The final step consists in a cursory examination of each paper, to confirm or reject its relevance and remove papers focused exclusively on corporate default prediction. The finale sample consists of 76 articles (the complete list of papers is available in Table B1). We then conducted a comprehensive examination of each paper to determine various characteristics of the default prediction models utilized in these studies. Specifically, we identified the country of study, the definition of default, predictors used in the model, and the statistical techniques employed. Additionally, we gathered information on the predictive ability of each model, considering indicators such as overall predictive accuracy, type I and type II errors, and the Area under the ROC Curve.

3. Bank failure and financial distress: a discussion of definitions and terminology

In the numerous publications on bank failure and financial distress, the definition of the latter is varied and non-unique. It is therefore important to understand how the literature interprets failure distress with reference to the banking sector.

From a theoretical perspective, financial distress could indeed have different degrees of severity. Mild financial distress may just be temporary, while serious financial distress corresponds to business failure or bankruptcy [46], although especially in the banking system, elements of distress are perceived well before failure or bankruptcy, which remain very rare events. An exact definition of financial distress has therefore not yet been determined. As a result, different authors define the failure of a business in different ways in their studies. The variation in this definition can make the comparison of models and results extremely difficult [49].

3.1. Failure and financial distress for non-financial firms

Balcaen and Ooghe [44] summarized previous research and concluded that three different definitions of financial distress can be adopted. The first refers to legal bankruptcy, consisting in a formal declaration in court. Once a company has become insolvent, there are different possible courses of action: administration, voluntary arrangement, receivership, liquidation, and dissolution [50]. The second type relies on the concept of financial distress. While the juridical definition of bankruptcy provides an objective criterion, the financial difficulty of an enterprise strongly depends on the financial criteria chosen by the authors.[2] The application of the criteria entails an additional issue, since some non-distressed companies might engage in such activities for strategic purposes [51] or due to external factors.[3] The third classification includes failure-related events such as cash insolvency, loan default, capital reconstructions, informal government support, and loan covenant renegotiation with bankers.

It is reasonable to conclude that the definition of financial distress can have strong influence on the output of the predictive model. As mentioned above, the legal definition of bankruptcy provides an objective criterion that allows the population under examination to be easily classified [50] and the moment of failure can be objectively defined. On the contrary, financial distress represents an arbitrary criterion that reduces the ability to compare models and could reduce the classification power of the model. According to Platt and Platt [51] the lack of a commonly recognized definition of financial distress could imply the misclassification of a non-distressed company as financially distressed and vice versa. A possible partial solution relies on

---

[1] For a comprehensive review of the papers up to the 2000s, the reader can refer to Kumar and Ravi [41].

[2] The most frequently used definitions of financial distress are missed dividend payments, low interest coverage ratio, cash flow less than current maturities of long-term debt, change in equity prices, negative EBIT, negative net income before special items [51], company stock price less than 10 cents [220], losses, and selling shares to private investors [221].

[3] According to Platt and Platt [51], layoffs may occur in specific divisions of an otherwise healthy enterprise, restructuring may occur at a different state of decline, and there are many explanations for missed divided payments. As an example, it should be noted that the European Central Bank, with a Recommendation issued on 20 March 2020, strongly suggested that the significant banks under its supervision and the less significant institutions under the supervision of competent national authorities not pay dividends for the financial years 2019 and 2020 until at least 1 October 2020 in response to the economic impact of the COVID-19 pandemic.

multidimensional screening for financial distress than combines several financial metrics.

However, the legal definition of bankruptcy could also lead to specific errors. Balcaen and Ooghe [44] reported the most recurring problems. First, there may be a long-time lag between the moment when a firm suffers financial distress and the final juridical exit in the form of bankruptcy. Second, failure could affect companies that were not actually in financial distress. Firms could declare bankruptcy for strategic reasons (e.g., rid themselves of debt) or due to an unexpected event, such as natural disasters, changing government regulations, or legal judgments. Third, bankruptcy does not represent the only possible ending of the failure process. Alternative measures could be mergers or absorption. Juridical bankruptcy is thus characterized by an extremely low frequency, which is even more evident within the banking industry. According to Adnan Aziz and Dar [43], corporate bankruptcy prediction that relies on the legal definition of bankruptcy is thus vulnerable to problems arising from small samples, which appear to be an inevitable limitation.

3.2. Failure and financial distress in the banking industry

Studies on the banking sector should consider that bankruptcy could also be avoided by the intervention of the government or authorities. According to Stolz and Wedow [52] the justification behind the support of financially distressed banks can be found in the need to restore stability in the financial system, maintaining lending to the real economy and avoiding phenomena such as bank runs. In addition to the macroeconomic costs, another fundamental pillar of authorities' action can be found in the need to protect savings and guarantee the overall confidence in the banking system. The first issue to address is that authorities could take several different measures, such as asset supports, recapitalisation, or guarantees on bank bonds. Moreover, the way in which authorities deal with this issue differs among countries. For example, an important difference between American and European authorities is that the former has historically adopted a range of non-standard measures to grant liquidity to the financial system and specific measures to support financial institutions directly. In contrast, the European Central Bank has adopted monetary policies in response to financial market tensions while euro-area governments have simultaneously implemented several additional ad hoc measures for individual banks in the case of a financial distress. We avoid any specific comment on the different ability of the two approaches to preserve financial stability and bank soundness, which lies outside the scope of this review, but note that the first approach allows for a higher data homogeneity or bank default while the second strongly depends on the action of each government. This conclusion can be drawn looking at the different definition of default used in the papers under analysis and summarized in Table B1.

Upon examining disparities between US and EU studies, the definition of default used in research papers reflects divergent approaches adopted by respective authorities.

When looking at the differences between US and EU studies, the definition of default used in the papers seems to reflect the different approach of the two authorities.

With regards to US studies, almost all the papers (30 studies, representing about the 90% of the sample) defined all the banks included in the Federal Deposit Insurance Corporation's failed bank list as 'failed'. The FDIC is an independent federal agency insuring deposits in US banks and has the authority to manage the resolution of a distressed bank. The FDIC lists all banks that went into FDIC conservatorship and that closed owing to bankruptcy, merger, or acquisition. It emerges that in the US, almost the only one in the international context, the authority is responsible to carry out functions of either banking supervision and deposit guarantee and management of crisis procedures. This kind of approach has undoubtedly the advantage to reduce multiplicity in the dependent variable among studies and increase the comparability of models. Owever, it confines its application solely to US banks, limiting

international evaluations, crucial considering interconnected global banking systems and internationally active banking groups.

In contrast, European and non-US studies manifest diverse approaches to identifying banking crises, showcasing the subjectivity associated with crisis identification:

1. **Market information-based approaches**: Distinguin et al. [53] relied on the ability of market indicators to predict bank financial distress. In detail, the authors identified changes in a bank's financial condition by checking downgrading announcements by the three major rating agencies (Fitch, Moody's, and Standard and Poor's). Once any of the three rating agencies announced a downgrade during the accounting year, the related bank was considered as distressed at the end of the corresponding period. Similarly, Switzer et al. [54] created distress events using Credit Default Swaps (CDS), namely swaps on senior debt with a maturity of five years that could be considered a measure of credit risk. The amount paid by creditors is in fact usually proportionate to the risk and the debtor's probability of default. Cipollini and Fiordelisi [55] used the shareholder value ratio, defined as the ratio between the economic values added (EVA) and the shareholder capital invested at time $t - 1$. The threshold values considered in the study were the 25th, 20th, and 15th percentile of the empirical probability distribution of the SHVR.

2. **News and media coverage**: Some studies, like Poghosyan and Čihak [56], employ global news databases to detect distressed banks based on specific keywords in news articles. In detail, the authors defined as distressed any bank for which the combination of bank name and specific keywords was found among the global news contained in the NewPlus database, powered by Factiva. The keywords designed to capture references to failing banks were the following: 'rescue'; 'bailout', 'financial support', 'liquidity support', 'government guarantee', and 'distressed merger'.

3. **Accounting variable-based approaches**: Halteh et al. [57] and Zaki et al. [58] use accounting variables to detect prodromes of crisis. In detail, Halteh et al. [57] use the Altman Z-score and the standardized profits to identify distressed banks, while Zaki et al. [58] computed percentage changes in annual equity, annual return on average equity (ROAE), and annual net interest margin (NIM) for each year from 2000 to 2008 for UAE banks and classified them as "distressed" if they experienced an annual variation in such index below the median.

4. **Expansive definition of distress:** Some studies (e.g., Ref. [29, 59–62]) opt for broader definitions encompassing various distress events beyond mere bank failures or receiverships This approach implies the inclusion of additional events, such as state support or government intervention, removal of management body, breach of minimum capital requirements, merger (or distressed merger) and acquisition, failure to meet obligations.

However, all these studies suffer from various limitations. Papers relying on stress test outcomes, agency ratings, CDS prices, or any market-dependent information significantly restrict the initial sample size. Similarly, studies considering EVA as a distress measure must make assumptions to include non-listed banks in their samples. While determining the cost of equity capital for listed banks is relatively straightforward, authors dealing with non-listed banks often resort to assuming that the cost of equity capital equals the mean of comparable domestic quoted banks. Papers relying on accounting variables offer simplicity but utilize data reflecting past performance, often indicating financial distress only after significant escalation.

Moreover, studies providing comprehensive distress definitions encompassing various distress events (e.g., bankruptcy, state support, distressed mergers, special administration, absorption by another entity) suffer due to the lack of a unified and recognized set of information and data source.

The analysis emphasizes the potential challenge of using different

definition of defaults, which can result in heterogeneous samples and diminish the comparability of various models. Moreover, regardless of the crisis definition employed, careful handling of crisis duration is essential to prevent endogeneity issues. Once a crisis occurs, it tends to exacerbate any economic downturn and impact the explanatory variables significantly.

Furthermore, it's evident that employing simplistic definitions such as 'bankruptcy' or 'liquidation' as proxies for bank default prediction as may not be adequate. Relying solely on these definitions might lead to missing cases of distress managed through different solutions or identifying distressed banks too late, thereby diminishing the chance for early interventions to prevent and manage crises effectively. For an early warning system on bank distress to be practically beneficial for supervisors and policymakers, the identification of distress events must indeed occur with sufficient timeliness to provide a window for intervention. This becomes particularly crucial in light of new European directives like the BRRD (Banking Recovery and Resolution Directive), which establish preparatory measures and early intervention systems to avert defaults.

Identifying a range of distress events that encompass direct failures, state support, and private sector solutions for bank distress appears to be a more comprehensive approach. This approach allows for a more holistic understanding of a bank's vulnerability to distress and better captures the multifaceted nature of potential crises.

4. Analysis and comparison of prediction models and methods for distress

In this section we present a brief description of models frequently used in the empirical literature on bank default prediction and we contextually provide an analysis and a comparison of their predictive ability. To facilitate the comparison between models, the analysis presented in this section is limited to papers for which was possible to obtain information about overall prediction accuracy (OPA), type I error and type II error of models one year before the distress event. When available, data about the predictive ability of models two and three years before the failure were collected. As a result, the analysis consisted of 31 articles (over a total sample of 76 papers) comparing a total of 88 different models.

The indicators being analysed were derived from the predictive classification table, also referred to as the confusion matrix or error matrix:

		PREDICTED	
		DISTRESSED	NON-DISTRESSED
ACTUAL	DISTRESSED	**TP**	**FN**
	NON-DISTRESSED	**FP**	**TN**

where:

- TP (True Positive): Number of distressed banks correctly predicted as distressed.
- FP (False Positive): Number of healthy banks incorrectly predicted as distressed.
- TN (True Negative): Number of healthy banks correctly predicted as healthy.
- FN (False Negative): Number of distressed banks incorrectly predicted as healthy.

The overall prediction accuracy represents the proportion of records that are correctly predicted by the model. This is computed as:

$$OPA = \frac{TP + TN}{TP + FP + TN + FN} \tag{1.1}$$

All results in the papers being analysed were systematized in order to

consider type I error as the percentage of healthy banks classified as distressed and type II error as the percentage of distressed banks classified as healthy. Therefore, the two errors can be expressed as:

$$Type\ I = \frac{FP}{FP + TN} \tag{1.2}$$

$$Type\ II = \frac{FN}{TP + FN} \tag{1.3}$$

Although a good model should provide low percentage of both errors, the literature agrees that, in the case of financial distress prediction of a bank, false negatives are considered to be far more costly than false positives [56,62–66]. From the prudence point of view, false positives lead to additional bank examination costs for the misclassified healthy banks, but missing failures typically imply higher resolution costs or delayed resolutions. Consequently, the models should be primarily evaluated for their OPA and type II error.

Table 1 reports the fundamental information from the studies including — in addition to the above-mentioned measures — the definition of distress, sample size, period, region, and the validation method applied to each model.

According to Table 1, the sample size ranges from 37 to 8293 banks with an average of 1202, which increases to 1670 considering only EU and US samples. In addition, populations with less than 100 financial firms represent 23% of papers. However, circumscribing the analysis to the US and EU samples, the minimum increases to 155 and the percentage of populations smaller than 100 goes to zero. The minimum time window is equal to 1 year, while the maximum is 18 years, with an average of about 8.1 years.

The main focus of this section regards the description and comparison of methodologies used in the papers under analysis. Fig. 1 shows that nearly 40% of the prediction techniques are statistical models. Among statistical models, logit model is the preferred method (Fig. 2). While these models have the highest percentage, this predominance is lower than the one identified in similar studies on bank bankruptcy.

In their review, Adnan Aziz and Dar [43] found that statistical methods represent 64% of techniques used in a sample of studies on bankruptcy prediction published between 1968 and 2003. Similarly, Bellovary et al. [45] showed that discriminant, logit, and probit analysis alone represent 61% of their sample, defined by the bankruptcy studies published from 1930 to the early 2000s.

Starting in 1968, statistical models such as multiple discriminant analysis (MDA), logit analysis, and probit analysis developed as the earliest methods for financial distress prediction and were then widely applied to the construction of corporate prediction models. From a general perspective, statistical methods have been extensively used in the literature due to their high interpretability and low computational complexity. However, they rely on a series of assumptions that have important limitations. Discriminant analysis, which represented the leading technique for failure prediction [42], requires a series of restrictive assumptions that are often not met and, in particular, (i) the normal distribution of regressors and (ii) the equal variance-covariance matrices across groups. Similarly, the logistic regression, which is not based on the assumption of normal distribution and equal covariance, still depends on the existence of a linear relationship.

As shown in Table 2, the need to relax such assumptions and the increasing availability of large economic datasets have driven the development of Artificial intelligence techniques (AI) as a tool for economic prediction and credit score modelling since the 1990s [91]. Popular tools such as neural networks [71,72,77,87], support vector machines (e.g. Ref. [74,76,81]), and decision trees (e.g. Ref. [61,76,80]) have evolved as systems that 'learn' from previous experience in order to improve their problem-solving performance and to deal with non-linearity. The last decade has also seen the diffusion of Ensemble and Hybrid models, which have progressively become more attractive.

The popularity of these methods is mainly motivated by their

Table 1
Summary of previous research findings.

#	Study	Sample	Region	Period	Definition of Financial Distress	Validation method	Model(s)	OPA			Type 1			Type 2		
1	Swicegood and Clark [67]	1741	US	1993	Banks in the lowest profit quintile (bottom 20%) in terms of ROA	TS-TR (60-40)	MDA	82.42			6.46			61.87		
							BPNN	79.59			16.16			37.41		
2	Bongini et al. [68]	43	East Asia	1996–1998	Closed, recapitalized, suspendend or mergered	N/A	Logit	86.05			7.69			23.53		
3	Canbas et al. [69]	40	Turkey	1994–2001	Transfer to SDIF (M&A, liquidation, revoked)	N/A	MDA	Year(s) before failure: 1 — 90	2 — 85	3 — 72.5	Year(s) before failure: 1 — 5	2 — 15	3 — 25	Year(s) before failure: 1 — 15	2 — 15	3 — 15
							Logit	87.5	70	72.5	5	30	25	10	30	20
							Probit	87.5	70	72.5	5	30	25	15	30	20
4	Mannasoo and Mayes (2009) [70]	600	EU	1995–2004	Bankruptcy, dissolved, in liquidation or negative net worth (Bankscope database)	TS-TR (TS: 1997–2001; TR: 2002–2004)	Survival Model	Training 89	Validation 87		Training 10	Validation 13		Training 19	Validation 50	
5	Poghosyan and Čihak [56]	5708	EU	1996–2007	Presence in the NewPlus/Factiva database of the combination "bank name" and keywords designed to capture references to failing banks, i.e. "rescue", "bailout", "financial support", "liquidity support", "government guarantee", "distressed merger"	N/A	Logit	99.06			0.83			37.97		
6	Cole and White [64]	6900	US	2009	FDIC failed bank list (M&A, liquidation, revoked)	TS-TR	Logit	95			5			3.7		
7	Serrano-Cina and Gutierrez-Nieto [71]	8293	US	2009–2011	FDIC failed bank list (M&A, liquidation, revoked)	TS-TR (by time)	LDA	Training 91.79	Validation 95.56		Training 4.29	Validation 3.59		Training 12.14	Validation 41.67	
							Logit (full model)	96.07	94.07		2.86	5.32		5.00	32.22	
							Logit (Stepwise)	95.36	95.42		4.29	3.80		5.00	38.33	
							MLP	93.93	95.31		5.71	3.88		6.43	40.00	
							k-NN	93.57	94.02		7.14	5.27		5.71	36.67	
							Naive Bayes	95.00	94.03		2.86	5.32		7.14	33.89	
							SVM	90.71	95.92		5.71	3.18		12.86	43.33	
							Boosting C4.5	100.00	93.45		0.00	5.85		0.00	37.22	
							Bagging RT	99.64	93.41		0.71	5.77		0.00	42.22	
							PLS	91.07	95.02		6.43	4.26		11.43	36.11	
8	Ecer [72]	68	Turkey	1994–2001	Transfer to SDIF (M&A, liquidation, revoked)	TS-TR (50-50)	MLP	Training 94.12	Validation 97.06		Validation 0			Validation 5.88		
							RBF	88.24	91.18		5.88			11.76		
							SVM	88.24	85.29		11.76			17.65		
9	Fiordelisi and Mare [73]	476	Italy	1997–2009	1) Extraordinary administration 2) Liquidation	TS-TR (TR: 1997–2006 TS: 2007–2009)	Survival Model	Training 83.2	Validation 78.9		Training 17	Validation 20.7		Training 21	Validation 50	
10	Papadimitriou et al. [74]	300	US	2003–2011	FDIC failed bank list (M&A, liquidation, revoked)	TS-TR	SVM	90			0			29.41		

Table 1 (*continued*)

#	Study	Sample	Region	Period	Definition of Financial Distress	Validation method	Model(s)	OPA			Type 1			Type 2		
								Year(s) before failure: 1	2	3	Year(s) before failure: 1	2	3	Year(s) before failure: 1	2	3
11	Betz et al. [29]	546	EU	2000–2013	1) Bankruptcy, liquidation and default 2) Receipt of state support 3) Distress merger (if a parent receives state aid within 12 months after the merger or if a merged entity has a coverage ratio smaller than 0 within 12 months before the merger	TS-TR	Logit	83	72	64	16	29	39	30	24	16
12	Cox and Wang [75]	644	US	2007–2010	FDIC failed bank list (M&A, liquidation, revoked)	Leave one out classification	LDA	LDA 80.04	74.68	72.73	6.93	20.97	28.17	33.00	29.68	26.38
								LDA-LOO 79.4	74.02	71.48	6.98	21.02	28.21	34.23	30.95	28.83
						QDA	QDA 80.37	70.49	66.36	20.72	44.02	52.49	18.55	15.01	14.80
								QDA-LOO 79.61	68.52	63.86	20.8	44.07	52.56	19.98	18.91	19.73
13	Li et al. [76]	316	US	n.a.	FDIC failed bank list (M&A, liquidation, revoked)	10-fold CV	BN	91.77			8.75			7.7		
							Naive Bayes	92.4			8.13			7.06		
							RBF	88.92			15.26			5.76		
							DT (J48)	89.87			14.61			4.35		
							DT (alternating DT)	88.29			0			18.98		
							Logistic	89.84			8.09			13.12		
							SVM	91.77			10.16			5.37		
							TR_SVM	93.8			5.73			8.54		
14	López Iturriaga and Sanz [77]	876	US	2002–2013	FDIC failed bank list (M&A, liquidation, revoked)	TS-TR (TS: 2002–2009; TR: 2010–2013)	MDA	77.88	74.04	70.19	21.15	23.08	28.85	23.08	28.85	30.77
							Logit	81.73	81.73	75.00	17.31	15.38	23.08	19.23	21.15	26.92
							RF	87.50	78.85	75.96	11.54	23.08	25.00	13.46	19.23	23.08
							MLP	93.27	85.58	82.69	7.69	13.46	15.38	5.77	15.38	19.23
							SVM	89.42	87.50	82.69	9.62	11.54	17.31	11.54	13.46	17.31
							MLP_SOM	96.15	90.38	84.62	5.77	7.69	11.54	1.92	11.54	19.23
15	Cleary and Hebb [78]	638	US	2002–2011	FDIC failed bank list (M&A, liquidation, revoked)	TS-TR (TS: 2002–2009; TR: 2010–2011)	MDA	Training 91.7	Validation 89.5		Training 5			Training 11.4		
16	Chiaramonte et al. [65]	8478	US	2003–2012	FDIC failed bank list (M&A, liquidation, revoked)	N/A	Logit	68.22			31.94			19.62		
17	Momparler et al. [79]	155	EU	2006–2012	Receipt of financial support	TS-TR	Boosted RT	98.13			12.12			1.4		
							Logit	77.87			22.58			22.08		
							Lasso	72.13			25.8			28.07		

Table 1 (*continued*)

#	Study	Sample	Region	Period	Definition of Financial Distress	Validation method	Model(s)	OPA			Type 1			Type 2		
18	Ekinci and Erdal [80]	37	Turkey	1997–2001	Transfer to SDIF (M&A, liquidation, revoked)	10-fold CV	Logit (L)	75.67			21.6			24.3		
							DT (J48)	72.93			26.5			27		
							VP	67.56			34.6			32.4		
							Multiboost J48	81.08			19.6			18.9		
							Multiboost VP	75.67			25.1			24.3		
							Multiboost L	75.67			21.6			24.3		
							Bagging J48	78.37			21			21.6		
							bagging L	83.78			16.4			16.2		
							Bagging VP	70.27			31.4			29.7		
							Rnd subspace B J48	83.78			16.4			16.2		
							Rnd subspace B L	78.37			21.9			21.6		
							Rnd subspace B VP	72.97			30			27		
							Rnd subspace MJ48	81.08			19.6			18.9		
							Rnd subspace M L	75.67			22.4			24.3		
							Rnd subspace M VP	78.37			21.9			21.6		
19	Gogas et al. [81]	1443	US	2007–2013	FDIC failed bank list (M&A, liquidation, revoked)	TS-TR (75-25)	SVM	Year(s) before failure: 1	2	3	Year(s) before failure: 1	2	3	Year(s) before failure: 1	2	3
								98.25	87.17	79.6	1.31	11.40	18.86	2.68	15.65	23.48
20	Jing and Fang [34]	586	US	2002–2010	FDIC failed bank list (M&A, liquidation, revoked)	TS-TR (TS: 2002–2009; TR: 2010)	BPNN	Training		Validation	Training		Validation	Training		Validation
								100		54.1	0		48.6	0		43.4
							SVM	99		76.3	0.3		39.8	1.6		7.7
							Logit	90.4		91.6	8.3		13.9	10.9		2.9
21	Le and Viviani [66]	3000	US	2001–2006	Inactive banks (Bankscope database)	TS-TR (70 - 30)	BPNN	75.2			35.1			15.5		
							k-NN	73.1			41.2			14.1		
							LDA	72			32.9			23.6		
							Logit	73.9			30.9			21.8		
							SVM	71.6			31.2			25.8		
22	Affes and Hentati-Kaffel [82]	928	US	2008–2013	FDIC failed bank list (M&A, liquidation, revoked)	TS-TR (80-20)	Logit	95.63			1.49			35.24		
							QDA	92.75			6.07			21.94		
23	Bräuning et al. [61]	3000	EU	2014–2016	1) Bankruptcy, liquidation, default 2) Breach of the minumum capital requirements 3) Removal of management body (art. 28 BRRD) 4) Special administration	TS-TR (75-25)	DT (C5.0)	Training		Validation	Training		Validation	Training		Validation
								90		91	11		10	1		3
							Logit	84		83	18		19	1		2

(*continued on next page*)

Table 1 (*continued*)

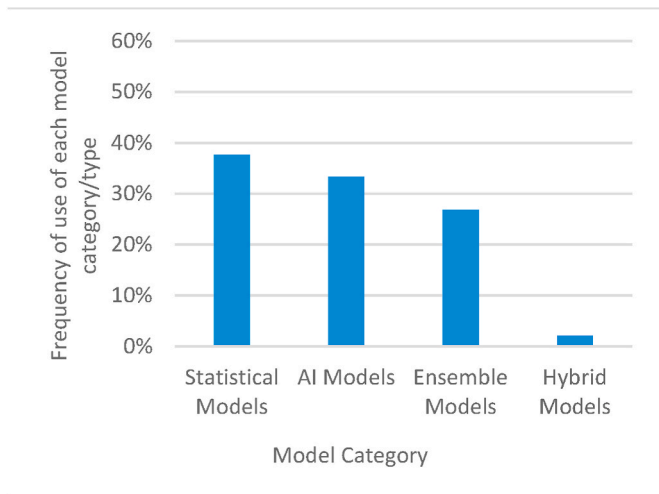| #	Study	Sample	Region	Period	Definition of Financial Distress	Validation method	Model(s)	OPA			Type 1			Type 2		
					5) Rapid deterioration of the financial situation (Art. 29 BRRD)											
24	Carmona et al. [83]	156	US	2001–2015	FDIC failed bank list (M&A, liquidation, revoked)	TS-TR (75-25)	Xgboost Logit RF	94.74 84.21 92.11			10.53 21.05 10.53			0 10.53 5.26		
25	Climent et al. [84]	168	EU	2006–2013	Receipt of financial support	10-fold CV	XgBoost	Year(s) before failure: 1 70	2 81	3 73	Year(s) before failure: 1 25	2 18	3 26	Year(s) before failure: 1 30	2 10	3 28
26	Kolari et al. [85]	273	EU	2010–2014	Failure of ECB stress test	TS-TR (70-30)	AdaBoost	Training 98.4		Validation 90.4	Training 1.8		Validation 8.1	Training 0.0		Validation 22.2
27	Filippopoulou et al. [86]	360	EU	1999–2007	1) Liquidation 2) Receipt of financial support	TS-TR	Logit	Year(s) before failure: 1 96	2 93	3 92	Year(s) before failure: 1 1	2 2	3 2	Year(s) before failure: 1 45	2 38	3 30
28	Paule-Vianez et al. [87]	148	Spain	2012–2016	1) Bankruptcy, liquidation 2) absorbed by another entity 3) intervention of the DGF 4) Distress merger (if it has a coverage ratio smaller than 0) 5) Receipt of financial support 6) failure to meet obligations	TS-TR (75-25)	MLP	Training 98.2		Validation 93.8	Training 2.2		Validation 8.3	Training 0.0		Validation 0.0
29	Shrivastava et al. [88]	56	India	2000–2017	Bankruptcy, liquidation, negative total assets, state intervention, merger or acquisition	TS-TR (80-20)	Logit RF AdaBoost	68.65 71.8 98.8			3.64 2.91 0.7			64.34 58.26 1.73		
30	Mishra et al. [89]	75	India	2014–2019	Bankruptcy, liquidation, negative total assets, state intervention, merger or acquisition	TS-TR (80-20)	Logit RF AdaBoost	77.33 77.33 86.66			25 25 20			27.27 27.27 10		
31	Pham and Ho [90]	180	US	2009–2019	FDIC failed bank list (M&A, liquidation, revoked)	TS-TR	AdaBoost Gradient Boosting Xgboost	92.59 94.44 96.3			12 8 8			3.45 3.45 0		

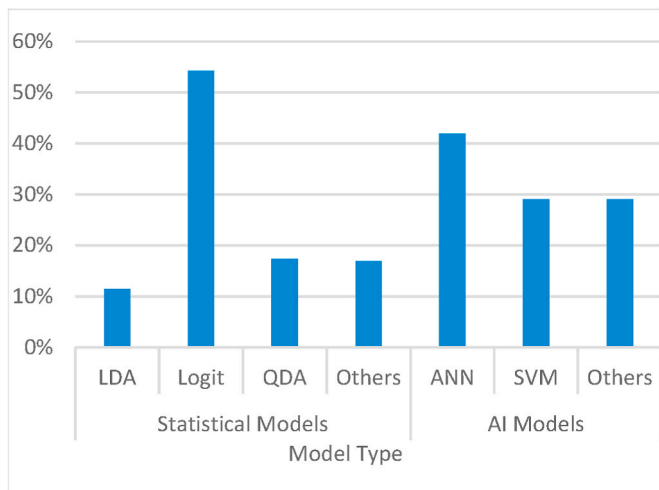**Fig. 1.** Proportion of model categories.

Fig. 2. Proportion of models among Statistical and AI models.

Table 2
Proportion of model categories over time.

Period	Statistical Models	AI Models	Ensemble Models	Hybrid Models
<1970	100.0%	0.0%	0.0%	0.0%
1970–1990	98.6%	1.4%	0.0%	0.0%
1990–2010	47.4%	52.6%	0.0%	0.0%
>2010	32.5%	33.75%	31.25%	2.5%

Source: Bellovary et al. [45] data and Fig. 1. Data processed by the author.

capability to increase the performance of base models in terms of model accuracy and quality of reasoning, especially when dealing with high-dimensional complex regression and classification problems [92]. For these reasons, they have been applied to numerous real-world problems, from medical diagnosis to financial forecasting [93].

Ensemble methods construct a set of learners and combine them to achieve lower variance, better accuracy, and model stability than each classifier in alone. The unique feature of such methods is that they usually combine multiple but homogenous weak models (base learners) with boosting and bagging approaches. A classifier is defined as 'weak' when small changes in the data imply big changes in the classification model. The logic behind these methods is that the combination of several simple classifiers into a more complex classifier is easier than

building a single complex learner [94]. On the other side, hybrid methods combine different individual machine learning approaches, which lead to overcoming several limitations of the single classifiers in cost of higher complexity of final solution.[4]

Overall, while the use of AI in forecasting offers benefits, it also poses challenges. The adoption of models based on Machine Learning is indeed often limited by the complexity and the limited interpretability of their functional form, which make predictions difficult to explain and validate [95]. The "black box" nature of AI models [96,97] represents a major drawback in many applications where rationale for model's decision is a requirement for trust and where explanation of individual forecasts are needed. This is particularly true when forecasts are used in policy making, where misguided correlations could have powerful implications and explainability of results may be more critical than predictive power [98].

Model explainability, and thus transparency and interpretability, is a key factor for policy makers and supervisors, because it allows to correctly inform the customers and stakeholders directly and indirectly affected, to ensure the consistency of process wherein humans take decisions, to verify whether a model suggests unethical results, and to facilitate the validation and monitoring of predictive models. A model can be considered as "explainable" when "it is possible to generate explanations that allow humans to understand how a result is reached or on what grounds the result is based" [99].

Despite the extensive research, AI models are still far from offering a completely explainable result, given the difficulties to rank and weight each factor considered in the analysis and to understand their internal mechanisms [100]. In addition, the absence of a clear causal relationships between the explanatory variables and the dependent variables makes more difficult to understand when a model is beginning to perform less well and to respond to this by respecifying the model [101]. Finally, AI models exhibit potential sources of bias (i.e., sample bias, bias by association, and algorithm bias), especially given the risk for machines to learn from training data that reflect human biases and then to exhibit and perhaps amplify such biases [102].

Therefore, stakeholders are currently proceeding with caution in the adoption of machine learning models in favor of simpler statistical models [103]. In addition, evidence suggests that as the algorithmic complexity increases, the gains in predictive ability of AI models are non-monotonic and uneven, strongly depending on the dataset and model used in each research [104].

To compare the predictive ability of different models, Fig. 3 firstly summarizes the mean average accuracy for each model category in three different time windows: one, two, and three years before the distress event. The hybrid model was excluded from the analysis since it has

---

[4] More in detail, these methods combine clustering and classification techniques. This structure mainly allows for four different strategies:

- Classification + clustering: since clustering is an unsupervised learning technique, the classifier is first trained to distinguish data more accurately. The related output is thus used as input for the cluster to improve the results [217].
- Clustering + classification: to improve the quality of classification, isolated clusters can be investigated and then eliminated from training samples. The training sample filtered by clustering is thus used to train the classification model in order to improve the classification output. The cluster technique is therefore used to pre-process samples into homogenous clusters and the classification model is used to build the classifier [218].
- Classification + classification and clustering + clustering: in both techniques, the first algorithm is used for data reduction. The data correctly classified/clustered by the first algorithm are used to train the second technique, which is thus computed on a sample that is smaller than the initial one. Given a new testing set, the second algorithm could increase the model accuracy with respect to the single learner trained by the original dataset [219].
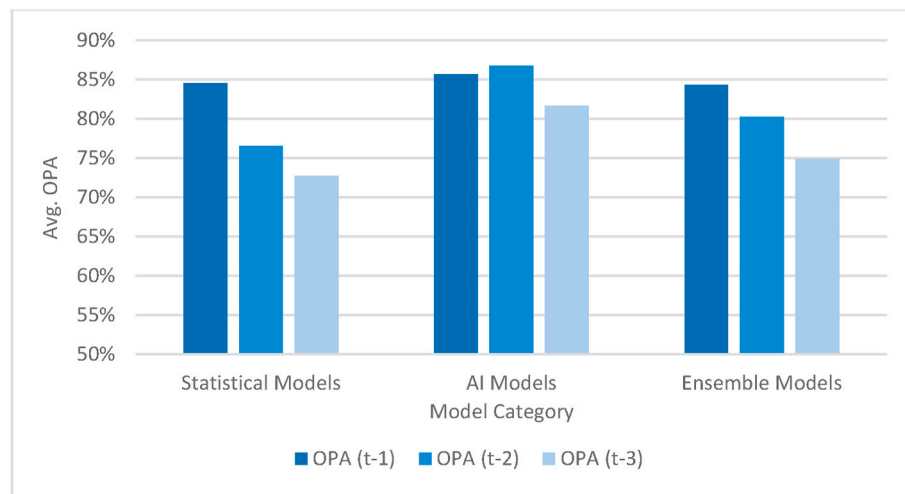
Fig. 3. Overall average accuracy for each model category.

been analysed only once. Although the chart suggests that AI models provide the best prediction accuracy one year before the failure, the performance of all the categories is similar.

However, the figure clearly shows that AI models perform better as the time horizon increases. At periods *t-2* and *t-3*, AI models show an accuracy of 86.75% and 81.66%, respectively. Statistical models show a lower accuracy with respect to AI techniques: about 10 points two years before the event and 9 points three years before. Ensemble methods seem to perform slightly better than statistical models but worse than AI techniques. However, only two observations were available two years and three years before the distress event. Other estimates will be thus necessary to better understand the predictive accuracy of ensemble methods in wider time horizons.

Table 3 shows the mean average accuracy and related statistics for each type of model calculated one year before the failure. The table includes only models that appeared at least in 3 different papers.

The table clearly shows that logit outperforms other statistical models in terms of OPA and standard deviation, while SVM is the most accurate AI model. On the other hand, QDA appears to be the least accurate statistical model and ANN is the most inaccurate in terms of overall accuracy and the least reliable technique, given its standard deviation. Ensemble methods seem to be promising techniques since they outperform most of the base learners and exhibit low standard deviations.

For accuracy ratings, it is important to also assess the misclassification rates for each category and type of model. Figs. 4 and 5 show the frequency of type I and type II errors, respectively. As already mentioned, type II error — the percentage of distressed banks misclassified as healthy — plays a primary role in evaluating predictive accuracy since it could be far more costly than type I error. Fig. 5 shows that on average, AI models perform better in each time window, while statistical models always show the highest error percentage.
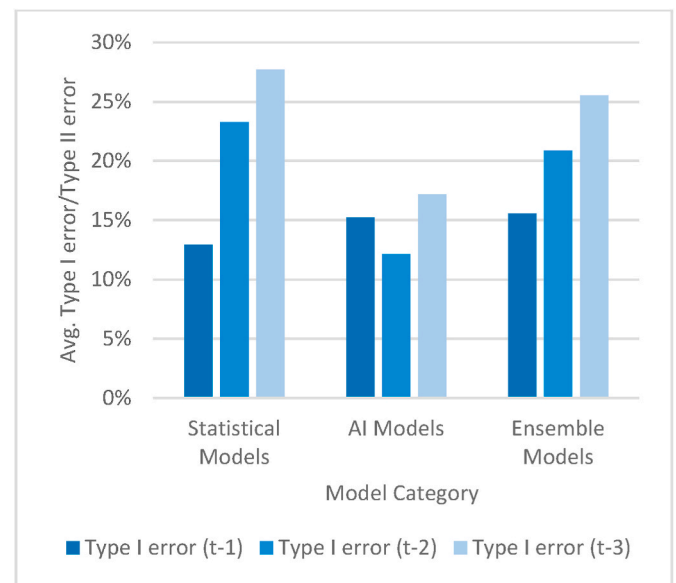

Fig. 4. Type I error for each model category.

Surprisingly, the increase in time horizon does not seem to affect the level of type II error. This conclusion cannot be drawn for type I error, which tends to increase in *t-2* and *t-3*, especially in the case of statistical and ensemble models. From a general perspective, statistical models seem to perform better one year before the distress event, but for larger horizons, AI models again seem to be more reliable. As stated in Fig. 3, more observations are needed to better understand the performance of ensemble models, since only one observation was available in period *t-1*

Table 3
Summary statistics (model type).

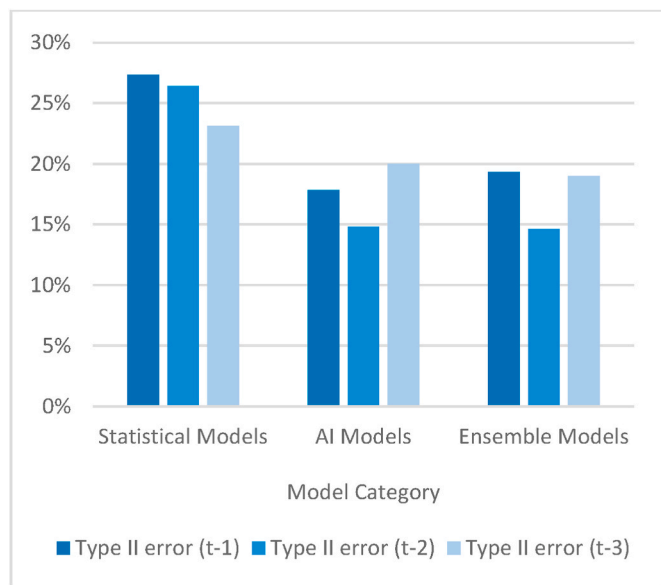| Model Category	Model type	Number of applications	OPA (t-1)	Lowest Accuracy	Highest Accuracy	Standard Deviation
Statistical Models	LDA	4	83,1%	72,0%	95,6%	8,79
	Logit	19	86,0%	68,7%	99,1%	8,6
	QDA	6	82,9%	68,2%	92,8%	8.6
AI Models	ANN	13	82,8%	54,1%	100,0%	14,1
	SVM	9	87,7%	71,6%	98,3%	8,2
	DT	3	83,7%	72,9%	89,9%	7,6
	K-NN	3	83,4%	73,1%	94,0%	10,2
Ensemble Models	Bagging	4	81,5%	70,3%	93,4%	8,4
	Boosting	12	91,1%	75,7%	98,8%	6.4

Fig. 5. Type II error for each model category.

and *t-2*.

Fig. 6 shows the average type I and type II errors for methods applied in at least 2 different papers.

With respect to false negatives, QDA shows the lowest frequency among statistical models, even though in Table 3, the same model was the least precise in term of overall accuracy. This difference can be explained by the high frequency of false positives. Logit models seem to provide the best compromise between the two errors, which is not surprising since this model shows the best overall accuracy among statistical models. ANN is confirmed to be one of the least reliable AI methods, since it shows the highest misclassification rates after *k*-NN models. SVM, which was the best AI model in term of OPA, shows the second-best level of type II error and the lowest type I error among its own category. SVM thus seems to be the best AI method overall. Fig. 6 confirms the hypothesis drawn from Table 3: boosting methods seem to outperform both statistical and AI methods in terms of accuracy and misclassification rates.

To further confirm this result, in Fig. 7 we show the area under the ROC curve (AUC) to visualize the performance of predictive models. The ROC curve represents the set of corresponding specificity rates and sensitivity rates for all possible thresholds of classification. The AUC thus represents a measure of the model's ability to correctly classify the observations in their respective category. The higher the AUC, the better the model is at distinguishing between classes. The AUC varies between 0.5 and 1, where 1 represents perfect predictive capacity and 0.5 means that the classifier is no better than a random guess. The figure shows that ensemble classifiers have the best predictive ability. In addition, AI models underperform statistical models both in terms of absolute value and in terms of performance,[5] as highlighted also in Table 3. Future research may be able to test the validity of this hypothesis.

This analysis overall supports the popular opinion that on average, ensemble classifiers outperform individual techniques (e.g., Ref. [41, 42]), since ensembles integrate several predictions. However, the performance of one model compared to others strongly depends on various factors such as the choice of dependent variable, sample size, validation method, and the control variables considered in the analysis. Consequently, the comparison of results obtained in different studies could be challenging and difficult to generalize. Indeed, it is worth noting that

---

[5] Untabulated results show that Statistical, AI, and ensemble models have a standard deviation equal to 11.59, 12.43, and 8,5, respectively.

performance is not the only relevant parameter to consider; the interpretability and complexity of the model should also be evaluated. In particular, while artificial intelligence methods usually require fewer assumptions and allow nonlinear functions to be approximated, the determination of various parameters is often complex and arbitrary. Furthermore, the configuration and processing of AI models are very time-consuming, and interpreting the contributions of the individual variables is often extremely complex. This analysis appears to corroborate the findings of Jones et al. [105]. In their study, the authors demonstrate that statistical models, such as logit and LDA, perform quite well and are sometimes even on par with more flexible structures. Simple classifiers still represent a reliable alternative to more advanced approaches, particularly when interpretability of the analysis holds significant importance.

5. Variable selection

The third aspect in question concerns the choice of variables to include in the models. This study follows and updates the analysis conducted by Mayes and Stremmel [106], who found considerable similarity among papers using the CAMELS category to classify the choices made in previous work.

Table B1 presents an overview of the main papers on bank default prediction and describes the number and categories of variables used in each analysis. Table B1 shows that, in total, the studies used 1340 variables, with an average of about 18 variables per study. It should be noted that while the choice of factors undoubtably plays a pivotal role in structuring a good predictive model, the number of variables does not guarantee per se higher accuracy of the model [45].

5.1. Accounting and financial variables

Despite the large number of variables, the analysis confirms the existence of homogeneity among the papers. Most of the financial ratios identified in the analysis can be classified into the categories of the CAMEL rating framework. The remaining variables can instead be classified into three main clusters: macroeconomic variables (e.g., GDP growth, GDP per capita, inflation rate, public debt, Herfindahl–Hirschman Index), bank characteristics (e.g., location, age, corporate governance characteristics, bank model), and market information (e.g., market capitalization, annual stock returns, abnormal returns, price-to-book value, credit ratings).

Below, a brief description of the CAMELS categories and the main variables used in the studies with an analysis of the related effects on bank default probability is presented.

Capital adequacy: Capital has a primarily role in explaining the soundness of a bank. Capital serves as a cushion to absorb losses, deal with risk from a managerial and regulatory point of view, and it is a fundamental source for financing and pursuing growth strategies. As a result, capital ratios were included in almost all the studies presented in Table B1. Nevertheless, the definition of capital and its computation method is not unique, and it is strongly influenced by accounting issues. In particular, a distinction can be made between traditional risk-insensitive indicators and risk-weighted measures, as stated in the Basel framework (i.e., regulatory capital). Even if risk-weighted measures consider the riskiness of assets, their computation could suffer from arbitrariness and manipulation [56]. With this assumption, most of the studies employ non-risk-weighted capital ratios, among which the ratio of total equity to total asset is the most widespread, while only few analyses consider risk-weighted measures, such as the tier 1 capital ratio or total capital ratio. Nevertheless, most of the studies suggest a negative relationship, both between probability of default and total equity to total assets ratio [29,56,78,107–109] and between probability of default and risk-weighted measures [110]. Only a few studies find no significant relationship (e.g., Ref. [59]). The results suggest that the higher the capital ratios, the less the probability that losses will cause a financial
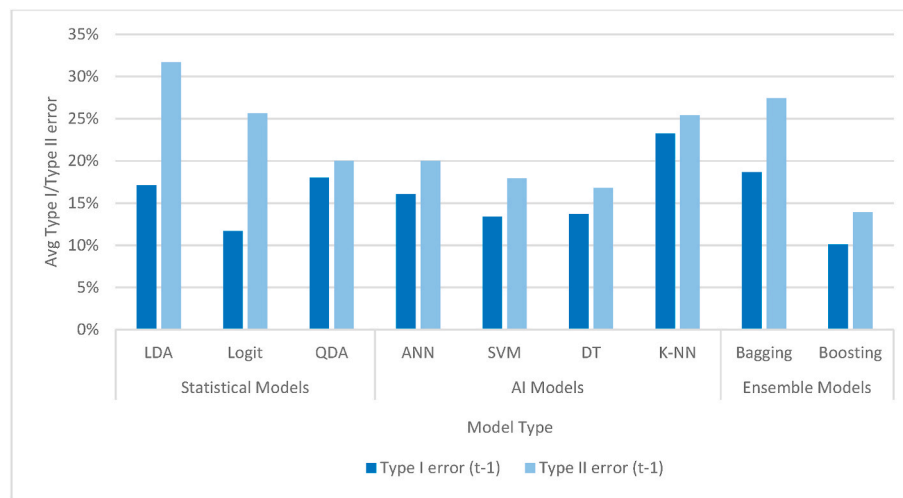
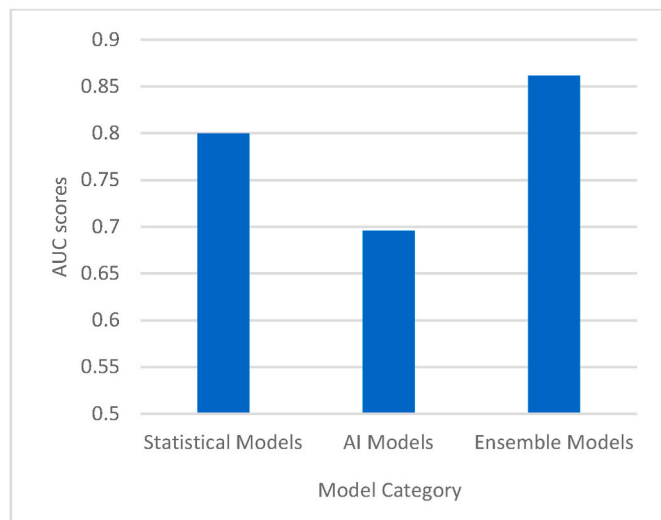**Fig. 6.** Type I and type II errors for each model type.

Fig. 7. AUC scores for each model category.

default.

Asset quality: This category reflects the quality of asset categories of an institution. The analysis suggests that the composition of the loan portfolio is an important determinant of the overall risk profile of a bank since Table B1 shows wide application of these measures. In fact, asset quality has a considerable impact on liquidity and costs. The value of high-risk assets can decrease rapidly and may generate losses and a reduction in the capital cushion, in turn increasing the overall risk of failure.

A potential measure of loan quality is the ratio between the amount of provision to loan losses and total loans (or total assets). This indicator could, however, be contradictory. Whereas higher reserves should imply higher coverage to expected losses, they could also reveal a higher expectation for losses [29]. Despite the potential ambiguity, almost all studies find a positive relationship with the probability of default [55, 56,78,110]. A second widespread measure is represented by the ratio between total loans and total assets. A higher share of risky assets could indeed increase the probability of losses. Although several authors find a positive relationship with the probability of default (e.g., Ref. [107]), other works consider the composition of the loan portfolio, assuming a different impact among borrower classes. Interestingly, Berger et al. [109] verify the influence of construction and development (C&D) loans

in increasing the probability of failure as in Cole and White [64] and Hong et al. [108] and confirm the insignificance of loans to total assets ratio when cleansed of the effect of C&D components. A more reliable indicator could therefore be the NPL[6] ratio, since it is considered in the literature as a relatively nondiscretionary and timely source of information about loan default [110,111]. The analysis confirms the strong positive influence of the NPL ratio on default probability [64,65,109, 110,112].

Management efficiency: This category assesses the competence of bank management and its ability to react to financial distress. Despite the relevance of this aspect, it is difficult to identify a unique measure based on data from balance sheets or income statements. As a result, Table B1 shows less frequent application than other components. Most of the studies applied the cost-to-income ratio to reflect the cost efficiency and, indirectly, management quality. Despite the frequent application of this ratio, most of the works find no significant relationship [29,56,65],[7] while only few studies show a positive relationship with default probability (e.g., Ref. [62]), thus confirming the ability of the ratio to represent a proxy of bank operational efficiency. Among others, the average salary per employee represents another potential indicator of management quality.

Earning ability: This category captures the ability to produce sustainable earnings and profits. A bank needs to be profitable in the long term in order to sustain and foster its competitiveness. Intuitively, a higher level of profitability allows banks to improve capital and economic performance. Most of the studies confirm this assumption, finding

---

[6] According to the definition by the IMF [6], a loan is non-performing when 'payments of interest and/or principal are past due by 90 days or more, or interest payments equal to 90 days or more have been capitalized, refinanced, or delayed by agreement, or payments are less than 90 days overdue, but there are other good reasons —such as a debtor filing for bankruptcy—to doubt that payments will be made in full'.

[7] Because of its simplicity, the cost to-income-ratio is frequently used to assess managerial quality. A low value of such ratio should indicate better managerial efficiency and, consequently, the relationship between the cost-to-income ratio and probability of bank failure is expected to be positive. However, most of the studies find no significant relationship among the two variables: several authors argue that this measure is not a perfect measure of bank efficiency because it does not consider input prices and output of banks [56]. Furthermore, lower cost-to-income ratio does not necessarily imply better efficiency, but it might be the result of high lending margins and volatile trading income, which leads to an increase in the denominator [70]. More direct measures of efficiency would be thus preferable, such as the stochastic frontier analysis, but at the cost of higher implementation and explanatory complexity.

a negative relationship between default probability and return on assets [54,56,64,78,107–109], Return on equity [53,56], and/or interest income on total assets [108]. Interestingly, some authors find no significant relationship (e.g., Ref. [59,70]) and, even more surprisingly, Betz et al. [29] showed a positive association. According to the authors, the positive relationship could potentially be explained by risks in the business model of a bank that are not represented in the selected measure of profitability.

Finally, the analysis shows how diversification in bank income sources plays an important role in achieving sustainable earnings. The negative association of non-interest income to total operating income [55,59] suggests that the expansion of non-interest income revenues (e. g., fees, commissions, trading) helps to obtain additional sources of revenue and reduces volatility in overall income [113].

Liquidity: Liquidity captures the ability of a bank to repay its obligations to depositors and meet customer loan demand. Liquidity risk may occur as a difficulty to raise the necessary funds (funding liquidity risk) or to sell portions of the portfolio (market liquidity risk) to meet both predictable and unpredictable liquidity demands. This issue can lead to higher financial pressure or, in extreme cases, to illiquidity. The amount of illiquid assets should therefore be positively correlated to the probability of failure. The analysis confirms this assumption. Banks with higher liquid assets to total assets ratio show a lower probability of observing financial distress [55,65,107]. Furthermore, most of the studies show that healthy banks have a higher proportion of customer deposits to total loans [109], since this source of funding is commonly considered more stable and the least costly [33]. On the other hand, banks with a higher level of short-term borrowings and securities tend to have higher risk of default [29]. According to the studies, these funding sources might make banks more exposed to funding liquidity shocks.

Sensitivity to market risks: This category reflects the ability of a bank to withstand fluctuations in the financial market. Sensitivity thus highlights the risk of losses arising from changes in interest rates, asset prices, and exchange rates. Due to computation difficulties, this category has been widely ignored, as confirmed in Table B1, or replaced by the bank size variable (e.g., Ref. [114]). The analysis confirms inconsistent results. Betz et al. [29] test the share of trading income, revealing a weak negative relationship, while Männasoo and Mayes [70] find this relationship to be insignificant. Conversely, the authors show a positive association of equity investment to asset ratio with default probability. Finally, Kerstein and Kozberg [115] find a weak positive association with interest-bearing deposits to total assets.

As shown in Table B1, the role of accounting information as a key component of an efficient early warning system has been well documented in the literature. Starting from contributions by Altman [12], Beaver [116], and Libby [117], a wide range of studies have indeed attempted to estimate the probability of default using accounting-based information. Although accounting variables are usually easily available through the balance sheets and income statements of banks, these metrics suffer from several limitations.

Firstly, accounting reports usually depict the past performance of a firm and may not be informative for predicting the future [34]. According to Hillegeist et al. [36], financial statements are defined on the assumption of the going-concern principle, which presupposes that companies will not go bankrupt. Consequently, their capacity to accurately and dependably evaluate the likelihood of default will be inherently constrained. In addition, accounting information could reflect the recorded book value of an asset and not the true value [118].

Secondly, accounting discretion may be used by managers to manipulate data and provide investors inaccurate financial information regarding the true value of illiquid assets.

Furthermore, detecting troubled banks by comparing financial ratios to predefined regulatory thresholds using historical data may fail to identify exceptional circumstances. This includes situations when banks suffer catastrophic losses that, despite an adequate capitalization (based on past thresholds), lead to financial distress [119]. As a result, these

limits may lead to higher misclassification rates (e.g., Ref. [120]),

Several articles in the sample have thus emphasized the role of market-based information in monitoring banks. Although past studies on US banks have found evidence that market information can add value to bank distress prediction compared to the CAMELS indicators (e.g., Ref. [32,121]), our analysis shows mixed results.

On the one hand, Curry et al. [122] demonstrate that the introduction of market-related variables can improve model's predictive ability, especially when combined with accounting data. Similarly, Poghosyan and Čihak [56] find a positive correlation between deviations from stock market trends and probability of default. The authors argue that the deviation of a bank stock from the general stock market trend in one year led to a correction in the next year, which could highlight the banks' vulnerabilities.

On the other hand, Distinguin et al. [53] show that the predictive power of market price, cumulative excess return, change of standard deviation in daily returns doesn't convey useful information when banks rely heavily on insured deposits. Avkiran and Cai [114] observe that neither annual stork return nor the market capitalization variable allows for a better discrimination between distressed and non-distressed banks. For a sample of individual banks in the East Asian nations during the years 1996–1997, Bongini et al. [68] compare the performance of accounting data, stock market prices, and credit ratings. The analysis demonstrates that stock market information underperforms in comparison to balance sheet data but reacts to changing financial conditions more quickly than credit risk agency ratings, emphasizing the greater ability of market variables to incorporate new sources of information than accounting variables.

Overall, the analysis seems to indicate that market-based information might improve the predictive power of models by serving not as a replacement for the accounting variables but rather as a complement to them. If market-based information indeed seems to be more appealing since it does not suffer from accounting-based variables' limitations, it tends to have a short time horizon [29] and relies strongly on the market efficiency assumption, according to which markets are always able to reflect all available information [123]. On the contrary, financial market are often influenced by information asymmetries and speculative/irrational behaviour, rather than fundamentals [124], raising doubts about the reliability of market evaluations.

Moreover, Campbell et al. [125] demonstrate that market-based information has little forecasting power after controlling for other variables and Distinguin et al. [126] conclude that accounting and market indicators are more capable to detect upgrade in rating changes rather than downgrades, especially for large banks.

Because of the above, several authors highlight the opportunity to include additional variables such as macroeconomic indicators (e.g., Ref. [32,33]) and non-financial information (e.g., Ref. [109,127,128]) to improve the predictive ability of the model.

5.2. The role of macroeconomic and non-financial variables

The wide adoption of macroeconomic and non-financial variables is well shown in Table B1, since more than half of papers in our analysis see the adoption of additional non-accounting variables. In particular, while a first strand of papers has adopted a micro-approach through the use of individual banks' balance sheet data (sometime integrated with financial market data), a second strand of papers has implemented a macro-approach, where a set of macroeconomic and industry level variables have been introduced in the predictive model. Based on this perspective, the fragility level of individual banks is intricately connected to both the propensity to banking crisis and to contagion.

From a macroeconomic perspective, several papers have estimated default prediction models using various indicators, such as economic growth, interest rates, and inflation. The analysis demonstrates that deteriorating macroeconomic situation may enhance bank distress and failure. Despite some inconclusive results (e.g., Ref. [55,129]), the large

majority of papers demonstrates that failures and distresses came from countries with lower GDP growth and higher inflation rate (e.g., Ref. [29,65,109,130]). Other variables, such as unemployment rate [130], terms of trade [131], house prices [29], and subprime exposure [109] have been identified as leading indicators of banking crisis.

Literature has also investigated the relationship between several country-specific banking sector indicators and probability of default. Among the others, concentration, regulation, and government intervention have been identified as primarily predictors of bank default. Competition in the banking sector is usually proxied by the Lerner index, which represents the extent to which market power allows companies to fix a price above the marginal cost (e.g., Ref. [59,129]), and the Herfindal–Hirschman index, which measures banking system concentration as the sum of the squared market share of each bank competing in a country in a given period (e.g., Ref. [109]). Literature offers two opposite views: while some authors support the "concentration-fragility view" [56], according to which banks located in more concentrated banking sector are found to be more vulnerable, others defend the "concentration-stability" theory [55,59,129], which suggests a negative relationship between market concentration and probability of default. On the one hand, concentrated banking markets generate greater incentives for bank to raise interest rate, at the cost of higher levels of default among borrowers [132]. On the other hand, bank failure is less likely in concentrated markets, where higher profits and robust capital reserves reduce the assumption of excessive risk, and more probable in more competitive markets, where the need to increase profitability may lead banks to finance riskier projects [133].

Furthermore, the analysis points out the effect of banking regulation and government intervention on bank's probability of default. Arena [107] analyses bank crises in East Asia and Latin America and find that regulatory and institutional environment need to be taken into account for bankruptcy analysis. In detail, Arena demonstrates that lower standard in the institutional environment increase the probability of individual bank failure. Similarly, Maghyereh and Awartani [129] shows that supervision and regulatory interventions reduce the probability of distress. To evaluate such relationships, the authors use the KKZ institutional index [134], which controls for country level differences in institutional development in terms of voice accountability, political stability, government effectiveness, rule of law, quality of regulation, and control of corruption.

Furthermore, Davis and Karim [131] assess the effect of financial liberalization on banks' stability on a heterogenous sample of banks from 105 different countries during the period 1979–2003. Results reveal that within liberalized markets, increased competition can erode the inherent value of bank charters. Without effective supervision and regulation, banks may opt to forego prudent credit risk assessment in their pursuit of attracting borrowers. This effect is enhanced by the presence of governments safety nets: deposit insurance may increase the risk of moral hazard when institutions are weak.

Overall, literature shows that macroeconomic variables may help to understand when problems are most likely to occur, while bank specific variables determine which banks are most likely to be affected. Similarly, the structure of the banking industry plays a role in explaining which countries are likely to encountering problems under a given set of macroeconomic and bank conditions. The combination of micro and macro perspective thus allow to increase the accuracy of predictions of bank financial distress.

Extending this line of reasoning, the analysis shows that innovative approaches have recently emerged in the literature. Constantin et al. [60] proposes a new method to address systemic risk surveillance by introducing bank networks into a bank's prediction model. Examining a dataset of 171 listed European banks between 1999 and 2012, the authors approximate the effect of interbank contagion using estimated network linkages or location of banks' incorporation. Results show that such approach is particularly effective because it allows to provide information of potentially vulnerable banks following a distress episode.

The existence of robust interconnections among banks in situation of elevated financial distress confirms and emphasizes the necessity for a macro-prudential perspective to supplement the micro-prudential analysis of individual bank risks, facilitating the monitoring of systemic risk and contagion risk.

A second innovative approach has been identified in the work of Gandhi et al. [119] and Cerchiello et al. [135], who have developed models for text extraction and sentiment analysis from different sources. Gandhi et al. [119] use sentiment analysis of annual reports to predict and measure financial institution instability. The analysis reveals that banks with a higher percentage of negative words in the annual report are more likely to suffer a distress episode. According to the authors, litigation risk and reputational costs reduce managers' propensity to not disclose negative news to shareholders in a timely manner and to be overly optimistic in their reports. The sentiment analysis of annual reports could therefore be viewed as a supplement to existing financial variables in assessing financial distress. Similarly, Cerchiello et al. [135] employs a sentiment analysis of news to detect intentions, emotional states, and personality traits. The underlying hypothesis posits that news and social media serve as channels through which individuals create narratives about the possible outcomes of their actions. Within financial markets, the creation and dissemination of narratives such as excitement about expected gains and concern about potential losses may have an impact on asset prices [136]. As a result, if models including only textual data achieve an average relative usefulness of only 13%, the inclusion of both news data and financial numerical data improves the performance of models (43,2%, compared to 31,1% when using only numerical data).

The overall findings suggest that it's the whole pattern of variables that conveys the narrative, rather than any specific indicators. Bank distress is a multifaceted and complex event, typically triggered by not just single factors, but by a series of mutually reinforcing bank-specific, macroeconomic, and structural variables that exhibit considerable dynamic variation as the event unfolds.

6. ESG variables: could they predict banks' financial distress?

This section is devoted to a literature review of the contributions aimed to shed light on the role of ESG variables as a potential predictor of bank's financial distress.

It should be noted that, despite the increasing relevance of ESG topic among media and scholars, almost no studies have empirically investigated the relationship between ESG and bank financial distress and to what extent ESG may improve the predictive ability of early warning systems. This poor knowledge about how the corporate social performance of a bank influences default risk is surprising for several reasons.

First, non-financial performance and corporate social performance have progressively increased in importance among firms and their stakeholders (e.g., Ref. [137,138]), as well as media and regulators [139,140]. The loss in trust and credibility suffered during the global financial crisis induced banks to adopt social responsibility actions and increase their transparency and compliance with sustainability standards and guidelines [141,142]. The behaviour of bank managers — identified in the literature as one of the underlying causes of the global crisis [142,143] — has highlighted the need to introduce new business manager tools to regain and reinforce credibility toward stakeholders [144]. Civil society has progressively expressed the need for 'moral capitalism' [145], which has prompted managers to shift from maximizing shareholders' wealth to maximizing stakeholders' value. This managerial revolution has been supported by several international bodies since the 1990s when the Kyoto protocol first recognized the relevance of environmental issues and proposed a plan to reduce the emissions contributing to global warming. More recently, a growing number of guiding principles and global authoritative standards (GRI – Global Reporting Initiative, SASB – Sustainability Accounting Standards Boards, PRI – Principles for Responsible Investment, EMAS, Global Compact) have been defined internationally to lead this transformation

through the worldwide spread of best practices.

In the banking sector, the role of corporate social responsibility now represents not only a matter of ethical behaviour towards stakeholders, but also a problem of prudential supervision [146]. Internationally, several initiatives have been set out to enhance disclosure and the integration of ESG risks withing traditional regulatory framework. For example, the "Principle for Responsible Banking", which have proposed a new framework to ensure that banks' strategy and practice align with the objectives of the UN Sustainable Development Goals and the 2015 Paris Climate Agreement; and the Basel Accords, which is promoting a common disclosure baseline for climate-related financial risks across internationally active banks.

At the European level, regulators are following a systematic approach toward climate transition and sustainability disclosure. In particular, Directive no. 2014/95/EU introduced a set of rules to regulate sustainability disclosures, especially for larger companies or groups, requiring them to report non-financial information with the goal of encouraging firms to develop a responsible approach and helping investors and stakeholders to evaluate non-financial performance and thus better estimate risks and opportunities for potential investments.

Within the banking sector, the Action Plan on Sustainable Finance published by the European Banking Authority on December 6, 2019 set out a series of principles aimed to encourage banks to include ESG factors in their business plans and identify tools and approaches to measure and monitor risks related to environmental, social, and governance aspects, looking at the stability of each individual entity and the financial system as a whole [147]. More recently, the European Banking Authority issued a Consultation Paper on March 2021 aimed at developing new standards to improve disclosure among end-investors [148]. The sustainable development goals designed by several international institutions and organizations now represent a new challenge on which scholars have only recently started to concentrate.

In addition, there is growing consensus about the ability of the banking sector to promote sustainable behaviour, whether directly or indirectly. Through the direct channel, banks could achieve savings in the use of electricity, water, and paper [149], and gain social performance in terms of relations with employees, the community, and clients [142,150] by adopting the proper measures. On the other hand, banks could play a pivotal role in the sustainability of other industries through lending [151,152]. Banks can select investment projects and direct funds according to the non-financial results of the target companies, thereby promoting sustainable approaches among borrowers [153–155].

It should be noted that among the three ESG dimensions, a strand of literature has paid particular attention to the role of corporate governance as a non-financial source of information for bankruptcy prediction. Amid studies on the banking sector, the first contribution was provided by Simpson and Gleason [127]. In their study, the authors analyse the influence of board structure and ownership on the probability of financial distress. The following variables were investigated: ownership by directors and officers, ownership by the CEO, number of directors, percentage of inside directors, and CEO duality. Testing a sample of 283 banking firms listed in the SNL Quarterly Bank Digest, Simpson and Gleason [127] found that the combination of CEO and chairman of the board in a single position has a negative relationship with the probability of financial distress. According to the authors, a CEO-chairman duality in the board may be able to better separate personal interests from shareholder interests, which conversely tend to prefer greater risks. Nevertheless, the other factors did not have a significant effect.

The relevance and attention to this topic among scholars has increased consistently, especially after the 2008–2009 financial crisis. Indeed, recent papers have continued to confirm the relationship between corporate governance and bank default risk. Using CDS spread as a measure of credit risk, Switzer and Wang [156] explore this relationship for 228 US banks over the period 2001–2010. They showed that

commercial banks with larger and more independent boards are associated with lower credit risk. In addition, the authors found that corporate governance structure has a greater impact on commercial banks than on saving institutions. Berger et al [109] provided evidence in support of this view, considering ownership, management, and compensation variables for 341 US banks in the period 2007–2010. The authors showed that ownership structure plays a primarily role in explaining the probability of failure. A higher shareholding of lower-level managers and non-CEO higher-level managers is positively associated with the risk of default. More recent evidence is provided by Switzer et al. [54], who analyse this relationship for a sample of financial firms in 28 countries outside of North America in the post-financial crisis period. The results show that institutional ownership and board independence are negatively related to default probabilities, while insider ownership, CEO duality, and board size increase the likelihood of default. Interestingly, the paper highlights the higher impact of such variables on Asian companies than on European firms. This suggests that cultural factors and different norms that tie governance to cultural constriction may have some sort of influence. Even if the contribution of each variable is sometimes controversial and often differs among types of institutions and countries of origin, the role of corporate governance in explaining the probability of failure seems to be clear. These results are also confirmed in the broader range of studies on non-financial firms.[8]

However, despite the number of contributions on corporate governance in banking, almost no studies have empirically investigated the relationship between bank financial distress and the remaining two ESG components, i.e., environmental, and social factors.

To shed light on such issue, we focus our attention on the current debates among scholar and researcher. Literature has indeed investigated the relationship between responsible conduct and banks performance extensively. While the findings are sometimes inconclusive [157], several studies show that sustainable initiatives could multiply the value of a firm. Indeed, ESG performance can directly influence financial performance through an increase in financial returns, market value [158,159], or a reduction in the cost of capital [160,161]. Indirectly, non-financial results can lead to a better reputation [143], which in turn can increase the value of the firm. A better reputation improves the image of firms on the market [162], allowing it to recruit and retain the best talent [163], and attract more creditworthy borrowers [164]. All these benefits can contribute to higher profits and better asset quality in the medium-long term and, accordingly, higher financial conditions [165]. Corporate reputation and improved economic performance thus exert influence on the perceived risk profiles of companies within the market [166]. This association becomes particularly remarkable when significant operational or reputational losses lead to a deterioration in the risk profile of financial institutions, as observed by Gillet et al. [167] and Fiordelisi et al. [168]. This underscores the importance of exploring the adverse effects associated with poor corporate social performance.

6.1. Relationship between corporate social performance and corporate financial performance

In this section, we aim to shed light on the impact of corporate social performance on economic performance by providing a concise overview of the latest empirical studies, specifically focusing on the banking sector. Table 4 describes the sample used, the main measures of corporate social performance (CSP), and corporate financial performance (CFP), as well as the most relevant results and the sign of the impact of sustainability initiatives on financial results.

Most of the studies find a positive relationship between ESG performance and accounting variables [144,145,154,169,173,175, 179–182] or market indicators [174,183]. Nevertheless, other studies

---

[8] For a comprehensive review of previous papers, see Fernando et al. [128].

Table 4

Main empirical studies on the impact of Corporate Social Performance (CSP) on Corporate Financial Performance (CFP) in the banking sector.

Study	Sample and period	Measure of CSP	Measure of CFP	Finding
Simpson and Kohers [169]	385 US banks 1993–1994	Community Reinvestment Act rating (dummy variable 0–1 if CRA rating needs improvement or not)	ROA, Loan losses/total loans	-strong positive relation with ROA -strong negative relation with Loan losses
Chih et al. [170]	520 financial institutions from 34 countries 2003–2005	dummy variable (1-0 in case the financial institution appears in the Down Jones World Index or not)	ROA	No relation
Soana [171]	21 international banks and 31 Italian banks 2005	Ethibel, AXIA, AEI ratings (aggregated and disaggregated form)	-Accounting: ROAE, ROAA, cost-to-income -Market: Market to book value, price to book value, P/E	-Ethibel: "internal social policy" shows a negative relation with EOAA; PTB, P/E ratio. Other parameters show no relation -Axia: positive link between "CG" and ROAE, "employees" and cost/income, "international operations" and market to book value, "international operations" and PBV. "transparency" shows negative relation with cost to income. Other parameters show no relation AEI: "CG" shows negative relations with ROAE and ROAA → no significant relations
Saxena and Kohli [172]	14 Indian banks 2007–2010	Score defined in Karmayog website (0–5 score constructed with surveys)	Profit after Tax and Earning Per Share	No relation
Wu and Shen [173]	162 banks from 22 countries 2003–2009	EIRIS database	net interest income (nii); non-interest income (non-ii), overhead costs, ROA, ROE, NPL	-Positive relation with nii; non-ii, ROA, ROE: -Negative relation with NPL
Carnevale and Mazzucca [174]	176 European banks 2002–2011	dummy variable (1-0 if banks publish a sustainability report or not)	Book value per share, earning per share	positive relation
Jo et al. [175]	4924 firm-year observations from 29 countries 2002–2011	Environment costs (provided by Trucost Plc)	ROA	negative relation (thus, positive link between financial performance and environmentally friendly performance)
Cornett et al. [141]	235 US banks 2003–2011	MSGI ESG STATS	ROA, ROE	positive relation (for larger banks)
Nobanee and Ellili [176]	16 Indian banks 2003–2013	sustainability disclosure (energy and natural environment disclosure) performed by a content analysis	Growth of deposits	positive relation (only for conventional banks)
Dell'Atti et al. [177]	75 international banks 2008–2012	Asset 4 database	Reputation (Global RepTrak Pulse)	-Social score: positive relation -environmental score: negative relation -corporate governance score: negative relation
Esteban-Sanchez et al. [178]	154 financial entities from 22 countries 2005–2010	Asset 4 database (no environmental scores)	ROE, ROA	-CG and employees variables: positive relation -community and responsibility dimensions: no relation (financial crisis has reduced the positive effect of sustainable activities)
Abou Fayad et al. [179]	7 Lebanese banks 2012–2015	Economic development ($), Community development ($), environmental development ($), human development ($)	ROA, ROE	positive relation
Forcadell and Aracil [180]	18 European banks listed in the DJSI 2003–2013	Dummy variable (0–1 if banks belong to the Down Jones Sustainability Index or not)	ROAA	Positive relation (the effect is reduced during the crisis period)
García-Sánchez and García-Meca [154]	159 banks from 9 countries 2004–2010	EIRIS database/Spencer & Stuart Board Index	Earnings persistence (EBT)	positive relation
Weber [181]	46 Chinese banks 2009–2013	Dummy variable (1-0 in case they found corporate social services or processes implemented or not) constructed with a content analysis	Net profits, ROA, ROE, NPL	-Strong positive relation with net profits -Weak positive relation with ROA and ROE
Laguir et al. [182]	68 French banks 2008–2011	Vigeo database (proxy measures for the extent to which banks engage in ESG actions)	ROA, EBIT, EBITDA	positive relation
Maqbool and Zameer [183]	28 Indian banks 2007–2016	comprehensive score (community, environment, workplace, and diversity) assessed by a content analysis	ROE, ROA, NP (profitability measures) SR and PE (market return indicators)	positive relation with both profitability variables and market return indicators
Brogi and Lagasio [144]	3476 US companies (848 banks) 2000–2016	MSGI ESG STATS	ROA, ROAt+1	positive relation (comprehensive score, social score, governance score and environmental score)
Miralles-Quirós et al. [142]	51 banks listed in 20 stock markets 2002–2015	Thomson Reuters Eikon ratings (aggregated and disaggregated form)	Price, book value per share; earnings per share	- ESG score: no relation - social score: negative relation - environmental score: positive relation - governance score: positive relation
Nizam et al. [145]	713 banks from 75 countries 2013–2015	MSGI ESG STATS	ROE	Positive relation

find mixed results [141,142,171,176,178] or no significant correlations [170,172]. Despite the controversial results, some conclusions can be drawn.

Discordant results could in fact be explained by the adoption of heterogeneous factors in the studies: according to Table 4, the authors adopt different samples in term of countries and historical series, as well as different CSP measures. Several authors suggest that bank origin and characteristics can influence the significance of the relationship being analysed. Cornett et al. [141] and Chih et al. [170] show that the relationship is stronger for larger banks, while Nobanee and Ellili [176]

show that social initiatives positively affect the deposit growth rate only for conventional banks. In their study, the authors compare conventional and Islamic banks and find that the former have a higher degree of sustainability disclosure, higher leverage, and more constraints, and therefore better reactivity to stakeholder and government demands regarding environmental issues. In addition, the authors that consider international samples and control of this aspect find that the relationship among CSP and CFP measures may vary among countries and markets. In detail, Jo et al. [175] showed that lowering environmental costs enhances firm performance more significantly in well-developed financial markets such as Europe and North America than in the Asia Pacific region. Similar results were obtained in the work of García-Sánchez and García-Meca [154] and Miralles-Quirós et al. [142]. The authors observed that the value relevance of ESG performance is higher in markets with higher levels of investor protection and bank regulation; more specifically, Miralles-Quirós et al. [142] find that the relationship is stronger in common-law countries. Interestingly, the same conclusion can be drawn from papers that consider only European or US banks: in these cases, the relationship between sustainable performance and financial performance is always positive. Conversely, papers that consider banks from the rest of the world show discordant conclusions. These results suggest that the varying incidence of sustainable activities and the inconsistent results obtained in the literature can be explained primarily by endogeneity problems, considering countries with different accounting standards, different levels of investor protection [184], or a different recognition of environmental problems [175].

Furthermore, the analysis revealed that the choice of time period can influence the strength of the relationship. More precisely, several papers that control for the crisis period found that during the last financial crisis, the positive effect of sustainable initiatives was reduced or wiped out. In particular, Esteban-Sanchez et al. [178] showed that the last financial crisis had a negative effect on the CSP-CFP interaction and that the effects mainly flow through corporate governance variables and community and responsibility aspects.

Likewise, Forcadell and Aracil [180] found that one's reputation for sustainability strategies did not improve returns during the 2008–2013 period and Miralles- Quirós et al. [142] demonstrated that financial stakeholders gave more value to ESG performance after the financial crisis since the relationship with corporate performance became stronger and more positive after subprime period. According to the authors, the crisis revealed failures in CSP, and transparency and integrity were questioned [180], with the consequent effect of mistrust towards the banking sector [185]. Nevertheless, Carnevale and Mazzucca [174] found that the crisis had a minimal negative effect on the relevance of CSP since it continued to have a positive effect on stock price. Wu and Shen [173] in turn found no significant difference in the effect on return on equity, net interest income, or non-performing loans before and after the crisis, while the effect on return on assets and non-interest income turned out to be insignificant during the 2007–2009 period. Finally, another difficulty revealed during the analysis is the insufficient standardization of ESG assessment methods. The analysis shows that different authors measure CSP with different methods, such as content analysis of annual reports, surveys, measures of reputation, and agency ratings.

Several authors (e.g., Ref. [186–188]) have suggested that the measurement of CSP with agency ratings provides various benefits that derive from the application of multiple and objective criterion systematically applied across companies by independent assessors. However, some disadvantages emerged from the analysis. First, some studies used an aggregate measure of CSP, while other studies considered each dimension individually. This finding indicates that individual dimensions are not equally important or relevant and should be considered separately. Among the studies, Miralles-Quirós et al. [142] found that a comprehensive ESG score has no relation to performance, while individual environmental and governance scores have a positive impact, and social score has a negative relationship with performance. Similarly,

Dell'Atti et al. [177] and Esteban-Sanchez et al. [178] showed that dimensions have discordant effects on reputation and performance. Furthermore, even if agencies use objective criteria for all companies, this choice differs among agencies. King et al. [189] show that CSR proposal by activist shareholders influences risk and returns according to the proposal category and the identity of the activist sponsor. In particular, whereas proposals centred on human issues or backed by elite groups are rewarded, proposals centred on the environment or backed by narrower internal interests are penalised. The lack of both standardized rules and formal auditing processes adds a subjective nature to ratings [190] that could lead to different evaluations of the same banks [171]. As validation, Berg et al. [191] argued that ESG ratings differ mainly due to measurement deviations and, residually, to deviations in weight and scope. In fact, the study confirms the poor correlation among ESG ratings (0.61), in contrast to the correlation of credit ratings, which is stronger at 0.9 [192].

Although the results are sometimes controversial, the analysis shows that after controlling for sample characteristics and definition of CSP, there is evidence of a correlation between sustainable performance and CFP and reputation that could allow ESG indicators to be considered a potential good predictor of financial default.

7. Concluding remarks

This study contributes to the literature by summarizing the empirical studies of bank bankruptcy prediction published after 2000. This work, which represented increasing interest in the topic especially after the 2008–2009 global financial crisis, reviewed and investigated previous studies focusing on three main issues: definition of financial distress, choice of failure prediction models, and selection of the best predictor variables. The main conclusions are illustrated below.

The absence of a precise definition of default and the existence of different degrees of severity among financial distress concepts could lead to an arbitrary definition of failure. This condition has primarily created great heterogeneity among models, reducing the level of comparability and increasing the risk of unstable and sample-specific results. Furthermore, the difficult distinction between failed and non-failed companies could cause misclassification that may result in an inappropriate application of classical statistical techniques which require the use of dichotomous variables.

Second, the review reveals an increasing use of AI techniques, despite the persistent predominance of statistical models. Although all techniques show a promising predictive accuracy and no specific models have emerged above the others, this study highlights how ensemble classifiers could outperform individual techniques. Further research is needed to confirm this conclusion.

Finally, the study confirms the relevance of the CAMELS factors in predicting bank distress. However, the existing focus has been limited concerning non-accounting variables that could encompass significant non-financial information. Specifically, the study found that bank distress is a complex event caused by multiple factors - specific to the bank, influenced by the economic conditions, and structural - that evolve as the situation unfolds.

Additionally, the integration of new technologies and big data is poised to significantly enhance the model's predictive capacity. This advancement allows for a deeper evaluation, such as assessing the sentiment of media and stakeholders toward the bank, which played a pivotal role in the SBV default.

Moreover, the final section highlights the escalating attention toward corporate social performance among investors and the media. This attention is due to its potential influence on financial performance, reputation, and consequently, the overall risk profile of the company.

Therefore, future research will further investigate if and to what extent such non-financial variables will be effectively relevant and suitable for consideration as component of an efficient early-warning system.

7.1. Practical implications and limitations of the study

Despite the extensive literature on predicting banking crises through Early Warning Systems (EWS), their practical adoption by policymakers has historically been limited, even within international financial institutions. This presents a paradox considering the evolving nature of banking risks due to the liberalization and development of financial systems across economies, alongside ongoing innovations, which make the use of EWS imperative in guiding policies aimed at preventing crises.

The analysis has shed light on the insufficient adaptability of well-established models like the CAMELS predictive models to swiftly evolving systems. Accounting indicators primarily reflect past performance, often signalling financial distress only after it has significantly escalated, as evidenced in the case of SBNY. Furthermore, the current scenario is marked by a series of rapid changes and diverse crises (such as the Covid-19 pandemic, inflation, and the recent Ukraine invasion) with profound impacts on the economy. These phenomena intersect with rapid technological innovations and societal demands for transitioning towards a more sustainable economy.

This study provides insights into recent attempts documented in the literature to encompass these multifaceted aspects. These efforts include leveraging new technologies or considering new predictors and variables. The aim is to empower supervisors, managers, and investors to proactively address these challenges, enabling the identification of financial distress well in advance and facilitating timely actions to restore the bank to a state of normalcy.

While this study has made significant contributions, it is important to acknowledge its limitations. Firstly, the reliance on a relatively small sample of papers (n = 76), further narrowed down to 31 for model comparison, might constrain definitive evaluations regarding the optimal predictive model for application.

Second, the paper was structured to address individual focal points separately. Nevertheless, for an accurate assessment of predictive models, a holistic approach encompassing all these facets is essential. Additionally, the structural design of predictive models significantly hinges on the context and expertise of the evaluator (e.g., managers vs authorities vs investors). However, this study did not incorporate these contextual variations.

Lastly, the proposed innovations in this paper necessitate access to diverse data sources. Regrettably, most of these data are only available for larger institutions. Consequently, traditional models continue to be the preferable choice for analysing smaller banks due to limited access to such data.

CRediT authorship contribution statement

Alberto Citterio: Conceptualization, Formal analysis, Methodology, Project administration, Writing – original draft, Writing – review & editing.

Data availability

No data was used for the research described in the article.

Appendix A

The abbreviations used in Table 1 are described in Table A1.

Table A1
List of abbreviations used in Table 1

Notation	Meaning
BPNN	Back-Propagation Neural Networks
CA	K-means Cluster Analysis
CL	Competitive Learning
DT	Decision Tree
k-NN	K-Nearest Neighbour
LDA	Linear Discriminant Analysis
LOO	Leave One Out Classification
LVQ	Learning Vector Quantization
MDA	Multiple Discriminant Analysis
MLP	Multi Layer Perceptron
n-fold CV	n-fold Cross Validation
OPA	Overall Prediction Accuracy
PLS	Partial Least Squares Discriminant Analysis
QDA	Quadrati Discriminant Analysis
RBF	Radial Basis Function
RF	Random Forest
Rnd	Random
RT	Random Tree
SOM	Self-Organizing map
SVM	Support Vector Machine
TS-TR	Training set - Test set
VP	Voted Perceptron

Appendix B

Table B1
List of studies

Study	Country	Definition of default/financial distress	Variables used							
			Used	C	A	M	E	L	S	Others
US studies										

Table B1 (*continued*)

Study	Country	Definition of default/financial distress	Variables used							
			Used	C	A	M	E	L	S	Others
Alam et al. [112]	US	FDIC failed bank list (M&A, liquidation, revoked)	5		X		X			
Sarkar and Sriram [193]	US	FDIC failed bank list (M&A, liquidation, revoked)	9	X	X	X	X			
Swicegood and Clark [67]	US	Banks in the lowest profit quintile (bottom 20%) in terms of ROA	25	X	X		X	X		
Kolari et al. [194]	US	FDIC failed bank list (M&A, liquidation, revoked)	28	X	X		X	X		Size
Tung et al. [195]	US	FDIC failed bank list (M&A, liquidation, revoked)	9	X	X	X	X	X		
Curry et al. [122]	US	FDIC failed bank list (M&A, liquidation, revoked)	15	X	X		X	X		market variables and risk measures
Ng et al. [196]	US	FDIC failed bank list (M&A, liquidation, revoked)	9	X	X	X	X	X		
Zhao et al. [197]	US	FDIC failed bank list (M&A, liquidation, revoked)	25	X	X	X	X	X		
Jin et al. [110]	US	FDIC failed bank list (M&A, liquidation, revoked)	10	X	X					Size, financial and audit quality variables
Avkiran and Cai [114]	US	1) FDIC failed bank list (M&A, liquidation, revoked) 2) Recipient of substantial federal government bailout funds 3) Keyword search combining BHC and 'rescue', 'bailout', 'distressed merger', 'financial support'	8	X	X	X	X	X		market variables
Cole and White [64]	US	FDIC failed bank list (M&A, liquidation, revoked)	15	X	X		X			Portfolio characteristics
Serrano-Cina and Gutierrez-Nieto [71]	US	FDIC failed bank list (M&A, liquidation, revoked)	15	X			X	X	X	
Kerstein and Kozberg [115]	US	FDIC failed bank list (M&A, liquidation, revoked)	12	X	X	X	X	X	X	
Papadimitriou et al. [74]	US	FDIC failed bank list (M&A, liquidation, revoked)	6	X	X		X			
Cox and Wang [75]	US	FDIC failed bank list (M&A, liquidation, revoked)	19	X	X		X			Size and loan portfolio characteristics
Li et al. [76]	US	FDIC failed bank list (M&A, liquidation, revoked)	12	X	X		X	X		Size and macroeconomic variables
Hong et al. [108]	US	FDIC failed bank list (M&A, liquidation, revoked)	20	X	X		X	X		Size and diversification
López Iturriaga and Sanz [77]	US	FDIC failed bank list (M&A, liquidation, revoked)	32	X	X	X	X			Loan portfolio characteristics and country-specific banking sector indicators
Kimmel et al. [130]	US	FDIC failed bank list (M&A, liquidation, revoked)	20	X	X	X	X			Size, macroeconomic variables and bank characteristics
Berger et al. [109]	US	FDIC failed bank list (M&A, liquidation, revoked)	41	X	X		X	X		Corporate governance variables, market competition, macroeconomic variables, and federal regulator indicators
Chiaramonte et al. [65]	US	FDIC failed bank list (M&A, liquidation, revoked)	10		X	X	X	X		Diversification and macroeconomic variables
Cleary and Hebb [78]	US	FDIC failed bank list (M&A, liquidation, revoked)	12	X	X		X	X		
Papanikolaou [198]	US	FDIC failed bank list (M&A, liquidation, revoked)	18	X	X	X	X	X	X	Size, country-specific banking sector indicators and macroeconomic variables
Gogas et al. [81]	US	FDIC failed bank list (M&A, liquidation, revoked)	2	X			X			
Jing and Fang [34]	US	FDIC failed bank list (M&A, liquidation, revoked)	16	X	X	X	X	X		
Le and Viviani [66]	US	Inactive banks (Bankscope database)	31	X	X	X	E	X		
Affes and Hentati-Kaffel [82]	US	FDIC failed bank list (M&A, liquidation, revoked)	10	X	X		X	X		
Audrino et al. [199]	US	FDIC failed bank list (M&A, liquidation, revoked)	15	X	X		X	X		Size and loan portfolio characteristics
Gandhi et al. [119]	US	Distressed delisting (if the bank has a distressed delisting code within 3 subsequent years of the 10-K filing date)	5	X						Financial market information, size, and sentiment analysis
Carmona et al. [83]	US	FDIC failed bank list (M&A, liquidation, revoked)	30	X	X	X	X	X		
Manthoulis et al. [200]	US	FDIC failed bank list (M&A, liquidation, revoked)	14	X	X	X	X	X	X	Size and diversification
Petropoulos et al. [201]	US	FDIC failed bank list (M&A, liquidation, revoked)	44	X	X	X	X	X	X	
Pham and Ho [90]	US	FDIC failed bank list (M&A, liquidation, revoked)	21	X	X		X	X		
EU studies										
Distinguin et al. [53]	EU	Downgrading announcements by the three major rating agencies (Fitch, Moody's, and Standard and Poor's)	35	X	X		X	X		Market variables
Mannasoo and Mayes (2009) [70]	EU	Bankruptcy, dissolved, in liquidation or negative net worth (Bankscope database)	21	X	X		X	X	X	Macroeconomic and structural variables
Poghosyan and Cihak [56]	EU	Presence in the NewPlus/Factiva database of the combination "bank name" and keywords designed to capture references to failing banks, i.e., "rescue", "bailout", "financial support", "liquidity support", "government guarantee", "distressed merger"	14	X	X	X	X	X		Macroeconomic variables, structural variables, and market discipline

Table B1 (*continued*)

Study	Country	Definition of default/financial distress	Variables used							
			Used	C	A	M	E	L	S	Others
Cipollini and Fiordelisi [55]	EU	Banks in the lowest percentiles (bottom 25%, 20%, and 15%) in terms of shareholder value ratio	8		X		X	X		Diversification, macroeconomic and structural variables
Fiordelisi and Mare [73]	Italy	1) Extraordinary administration 2) Liquidation	12	X	X	X	X	X		Macroeconomic variables and size
Betz et al. [29]	EU	1) Bankruptcy, liquidation, and default 2) Receipt of state support 3) Distress merger (if a parent receives state aid within 12 months after the merger or if a merged entity has a coverage ratio smaller than 0 within 12 months before the merger	26	X	X	X	X	X	X	Country-specific banking sector indicators and macro-financial indicators
Momparler et al. [79]	EU	Receipt of financial support	25	X	X	X	X	X		Size
Chiaramonte and Casu [59]	EU	1) Bankruptcy, dissolved, or in liquidation 2) Receipt of state support 3) Distress merger (if a parent receives state aid within 12 months after the merger or if a merged entity has a coverage ratio smaller than 0 within 12 months before the merger)	13	X	X	X	X	X		Diversification and macroeconomic variables
Constantin et al. [60]	EU	1) Bankruptcy, liquidation, and default 2) Receipt of state support 3) Distress merger (if a parent receives state aid within 12 months after the merger or if a merged entity has a coverage ratio smaller than 0 within 12 months before the merger)	24	X	X	X	X	X	X	Country-specific banking sector indicators, network linkage and macro-financial indicators
Climent et al. [84]	EU	Receipt of financial support	26	X	X	X	X	X		Size
Kolari et al. [85]	EU	Failure of ECB stress test	14		X		X	X		Macroeconomic variables
Bräuning et al. [61]	EU	1) Bankruptcy, liquidation, 2) Breach of the minimum capital requirements 3) Removal of management body 4) Special administration 5) Rapid deterioration of the financial situation	12	X	X		X	X	X	Macroeconomic variables
Filippopoulou et al. [86]	EU	1) Liquidation 2) Receipt of financial support	18		X			X		Size, macroeconomic variables, risk variables and country-specific banking sector indicators
Paule-Vianez et al. [87]	Spain	1) Bankruptcy or liquidation 2) Absorbed by another entity 3) Intervention of the DGF 4) Distress merger (if it has a coverage ratio smaller than 0) 5) Receipt of financial support 6) Failure to meet obligations	53	X	X	X	X	X	X	Macroeconomic variables
Cerchiello et al. [135]	EU	1) Bankruptcies, liquidations and defaults 2) State support 3) Forced mergers (if a parent receives state aid within 12 months after a merger or if a merged entity exhibits a negative coverage ratio within 12 months before the merger)	13	X	X	X	X	X		Macroeconomic variables and financial news
Kristóf and Virág [202]	EU	1) Bankruptcy, reorganization, liquidation, payment default, and dissolution due to negative equity 2) State aid measures, distressed mergers, bail-ins and bailouts from the European Commission	31	X	X	X	X	X	X	
Other studies										
Molina [203]	Venezuela	Bank stocks taken by FOGADE (government regulators)	13	X		X	X	X		Size and bank characteristics
Bongini et al. [68]	East Asia	Closed, recapitalized, suspended, or merged	7	X	X	X	X	X		Size, macroeconomic, credit rating agencies information and market variables
Canbas et al. [69]	Turkey	Transfer to SDIF (M&A, liquidation, revoked)	12	X		X	X	X		
Lanine and Vennet [204]	Russia	License revocation from the Central Bank of Russia	7	X	X		X	X		Size
Nur Ozkan-Gunay and Ozkan [205])	Turkey	Transfer to SDIF (M&A, liquidation, revoked)	21	X	X	X	X	X		Bank characteristics
Davis and Karim [131]	International	1) The proportion of non-performing loans to total banking system assets exceeded 10% 2) public bailout cost exceeded 2% of GDP 3) systemic crisis caused large scale bank nationalization 4) extensive bank runs	12		X			X		Macroeconomic variables
Arena [107]	International	1) Either the central bank or a government agency, specifically created to address the crisis, recapitalized the financial institution or the institution required a liquidity injection from the monetary authority.	12	X	X		X	X		Macroeconomic variables and bank characteristics

Table B1 (*continued*)

Study	Country	Definition of default/financial distress	Variables used							
			Used	C	A	M	E	L	S	Others
		2) The government temporarily suspended ("froze'') the financial institution's operations. 3) The government closed the financial institution								
Ravi and Pramodh [206]	Turkey	Transfer to SDIF (M&A, liquidation, revoked)	10	X		X	X	X		
Boyacioglu et al. [207]	Turkey	Transfer to SDIF (M&A, liquidation, revoked)	20	X	X	X	X	X	X	
Ravisankar and Ravi [208]	Spain, Turkey, UK and US	Bankruptcy	13	X		X	X	X		
Distinguin et al. [209]	African countries	Downgrading announcements by the three major rating agencies (Fitch, Moody's, and Standard and Poor's)	39	X	X		X	X		
Peresetsky et al. [210]	Russia	1) The license was withdrawn 2) The bank is under the administration of ARCO (Agency for Restructuring Credit Organizations) 3) The bank is merged with another bank and was in poor financial shape at the time of the merger	15	X	X		X	X		Size, macroeconomic variables
Zaki et al. [58]	UAE	1) Annual equity change less than or equal to 65.66 per cent 2) Change in NIM less than or equal to 1.11 per cent 3) Change in ROAE less than or equal to 130.48 per cent	10	X	X		X	X		Structural variables and macroeconomic variables
Zaghdoudi [211]	Tunisia	Index of banking weakness measure inspired by the works of Kibritcioglu (2002)	7		X	X	X	X		
Ecer [72]	Turkey	Transfer to SDIF (M&A, liquidation, revoked)	35	X	X	X	X	X		
Maghyereh and Awartani [129]	Gulf Cooperation Council countries	1) The bank's operation was temporarily suspended 2) The bank was recapitalized, or it received any liquidity support from the monetary authority 3) The bank eventually merged with another bank due to financial distress (i.e. distressed mergers) 4) The bank was closed by the government 5) The ratio of non-performing loans to total loans during two subsequent years belongs to the fourth quartile of the sample empirical distribution of this ratio	24	X	X	X	X	X		Diversification, size, macroeconomic variables, market structure, regulations, and institutional variables
Ekinci and Erdal [80]	Turkey	Transfer to SDIF (M&A, liquidation, revoked)	34	X	X	X	X	X		
Halteh et al. [57]	International	Altman Z-Score	18	X	X	X	X	X		Size
Ristolainen [212]	International	1) Bank runs that lead to the closure, merging, or a takeover by the public sector of one or more financial institutions 2) Closure, merging, takeover, or large-scale government assistance of an important financial institution	13							macro-financial indicators and market structure
Switzer et al. [54]	International	1) Average five-years CDS spread 2) Five-year default probability (Merton-type distance-to-default model)	9		X	X	X	X		Size and market information
Rosa and Gartner [213]	Brazil	1) Intervention or receivership processes 2) Merger and acquisition events with the assumption of financial distress	12	X	X	X	X	X		Loan portfolio characteristics and macroeconomics variables
Shrivastav and Janaki Ramudu [88]	India	1) Merger or acquisition 2) Bankruptcy 3) Dissolution 4) Negative assets	25	X	X	X	X	X		Size
Shrivastava et al. [88]	India	Bankruptcy, liquidation, negative total assets, state intervention, merger, or acquisition	24	X	X	X	X	X		Z-score and macroeconomic variables
de Haan et al. [214]	International	Bank runs and banking policy interventions	9	X				X		Macroeconomic variables and country-specific banking sector indicators
Mishra et al. [89]	India	Bankruptcy, liquidation, negative total assets, state intervention, merger, or acquisition	6	X			X	X		Size
Wang et al. [215]	International	1) Huge financial distress as critical bank runs, losses in the banking system, and/or bank liquidations within its system. 2) Significant policy intervention during huge losses in a banking system (liquidity support exceeds 5% of total non-resident deposits and debt; Costs of bank restructuring exceed 3% of GDP in that year; bank nationalization; Government proposes guarantee to banks; The government's purchase of bank assets exceeds 5% GDP in that year; Mandatory actions such as deposit freeze or bank holiday)	12	X				X		Macroeconomic variables and country-specific banking sector indicators
Chen et al. [216]	International	1) Bank run or bailout (bailout, capital injection, government guarantee) 2) Failure (taken over, acquired or nationalized)	13	X	X	X	X	X		Size, macroeconomic variables and diversification

References

[1] Meyer PA, Pifer HW. Prediction of bank failures. J Finance 1970;25(4):853–68.

[2] Laeven L. Banking crises: a review. Annual Review of Financial Economics 2011;3 (1):17–40.

[3] Bhattacharya S, Boot AWA, Thakor AV. The economics of bank regulation. J Money Credit Bank 1998;30(4):745.

[4] Benston GJ, Kaufman GG. Is the banking and payments system fragile? J Financ Serv Res 1995;9(3–4):209–40.

[5] Rajan RG, Zingales L. Power in a theory of the firm. Q J Econ 1998;113(2): 387–432.

[6] IMF. World economic outlook financial crises: causes and indicators. Washington DC: International Monetary Fund; 1998. 8 May 1998.

[7] World Bank. Global economic prospects, June 2020. Washington, DC: World Bank; 2020.

[8] Drehmann M, Juselius M. Evaluating early warning indicators of banking crises: satisfying policy requirements. Int J Forecast 2014;30(3):759–80.

[9] Thomson JB. Predicting bank failures in the 1980s. Federal Reserve Bank of Cleveland Economic Review 1991;27(1):9–20.

[10] Bhattacharya S, Thakor AV. Contemporary banking theory. J Financ Intermediation 1993;3(1):2–50.

[11] Diamond DW. Financial intermediation and delegated monitoring. Rev Econ Stud 1984;51(3):393.

[12] Altman EI. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. J Finance 1968;23(4):589–609.

[13] Sinkey JF. A multivariate statistical analysis of the characteristics of problem banks. J Finance 1975;30(1):21–36.

[14] Bilderbeek J. An empirical study of the predictive ability of financial ratios in The Netherlands. Zeitschrift fur Betriebswirtschaft 1979;5:388–407.

[15] Ohlson JA. Financial ratios and the probabilistic prediction of bankruptcy. J Account Res 1980;18(1):109.

[16] Altman EI. A further empirical investigation of the bankruptcy cost question. J Finance 1984;39(4):1067–89.

[17] Martin D. Early warning of bank failure. J Bank Finance 1977;1(3):249–76.

[18] Lau AH-L. A five-state financial distress prediction model. J Account Res 1987;25 (1):127.

[19] Bell TB, Ribar GS, Verichio J. Neural nets versus logistic regression: a comparison of each model's ability to predict commercial bank failures. X Auditing Symposium Deloitteand Touche. KS: University of KS; 1990.

[20] Tam K. Neural network models and the prediction of bank bankruptcy. Omega 1991;19(5):429–45.

[21] Tam KY, Kiang MY. Managerial applications of neural networks: the case of bank failure predictions. Manag Sci 1992;38(7):926–47.

[22] Olmeda I, Fernández E. Hybrid classifiers for financial multicriteria decision making: the case of bankruptcy prediction. Comput Econ 1997;10(4):317–35.

[23] Dimitras AI, Zanakis SH, Zopounidis C. A survey of business failures with an emphasis on prediction methods and industrial applications. Eur J Oper Res 1996; 90(3):487–513.

[24] Altman EI, Narayanan P. An international survey of business failure classification models. Financ Mark Inst Instrum 1997;6(2):1–57.

[25] Wong BK, Bodnovich TA, Selvi Y. Neural network applications in business: a review and analysis of the literature (1988–1995). Decis Support Syst 1997;19(4): 301–20.

[26] Adya M, Collopy F. How effective are neural networks at forecasting and prediction? A review and evaluation. J Forecast 1998;17(5–6):481–95.

[27] O'Leary DE. Using neural networks to predict corporate failure. Int J Intell Syst Account Finance Manag 1998;7(3):187–97.

[28] Bernd T, Kleutges M, Kroll A. Nonlinear black box modelling – fuzzy networks versus neural networks. Neural Comput Appl 1999;8(2):151–62.

[29] Betz F, Oprică S, Peltonen TA, Sarlin P. Predicting distress in European banks. J Bank Finance 2014;45:225–41.

[30] Barker D, Holdsworth D. The causes of bank failures in the 1980s. New York: Federal Reserve Bank of; 1993. Research Paper, No. 9325.

[31] Cole RA, Gunther J. A CAMEL rating's shelf life. Federal Reserve Bank of Dallas Review; 1995. p. 13–20.

[32] Flannery MJ. Using market information in prudential bank supervision: a review of the US empirical evidence. J Money Credit Bank 1998:273–305.

[33] Jagtiani J, Lemieux C. Market discipline prior to bank failure. J Econ Bus 2001;53 (2–3):313–24.

[34] Jing Z, Fang Y. Predicting US bank failures: a comparison of logit and data mining models. J Forecast 2018;37(2):235–56.

[35] Lucas RE. Econometric policy evaluation: a critique. Carnegie-Rochester Conf Ser Public Policy 1976;1:19–46.

[36] Hillegeist SA, Keating EK, Cram DP, Lundstedt KG. Assessing the probability of bankruptcy. Rev Account Stud 2004;9:5–34.

[37] Drechsler I, Savov A, Schnabl P. The deposits channel of monetary policy. Q J Econ 2017;132(4):1819–76.

[38] FDIC. FDIC's supervision of Signature Bank. Federal Deposit Insurance Corporation; 2023. April 28.

[39] Diamond DW, Dybvig PH. Bank runs, deposit insurance, and liquidity. J Polit Econ 1983;91(3):401–19.

[40] Jiang EX, Matvos G, Piskorski T, Seru A. Monetary tightening and U.S. Bank fragility in 2023: mark-to-market losses and uninsured depositor runs? National Bureau of Economic Research working paper; 2023. No. 31048.

[41] Kumar PR, Ravi V. Bankruptcy prediction in banks and firms via statistical and intelligent techniques–A review. Eur J Oper Res 2007;180(1):1–28.

[42] Demyanyk Y, Hasan I. Financial crises and bank failures: a review of prediction methods. Omega 2010;38(5):315–24.

[43] Adnan Aziz M, Dar HA. Predicting corporate bankruptcy: where we stand? Corp Govern 2006;6(1):18–33.

[44] Balcaen S, Ooghe H. 35 years of studies on business failure: an overview of the classic statistical methodologies and their related problems. Br Account Rev 2006; 38(1):63–93.

[45] Bellovary J, Giacomino D, Akers M. A review of bankruptcy prediction studies: 1930 to present. Journal of Financial Education 2007;33:1–42.

[46] Sun J, Li H, Huang QH, He KY. Predicting financial distress and corporate failure: a review from the state-of-the-art definitions, modeling, sampling, and featuring approaches. Knowl Base Syst 2014;57:41–56.

[47] Altman EI, Iwanicz-Drozdowska M, Laitinen EK, Suvas A. Financial distress prediction in an international context: a review and empirical analysis of altman's Z- score model. J Int Financ Manag Account 2017;28(2):131–71.

[48] Khan MA. ESG disclosure and Firm performance: a bibliometric and meta analysis. Res Int Bus Finance 2022;61:101668.

[49] Karels GV, Prakash AJ. Multivariate normality and forecasting of business bankruptcy. J Bus Finance Account 1987;14(4):573–93.

[50] Charitou A, Neophytou E, Charalambous C. Predicting corporate failure: empirical evidence for the UK. Eur Account Rev 2004;13(3):465–97.

[51] Platt HD, Platt MB. Predicting corporate financial distress: reflections on choice-based sample bias. J Econ Finance 2002;26(2):184–99.

[52] Stolz SM, Wedow M. Government measures in support of the financial sector in the EU and the United States. Intereconomics 2011;46(1):53–60.

[53] Distinguin I, Rous P, Tarazi A. Market discipline and the use of stock market data to predict bank financial distress. J Financ Serv Res 2006;30(2):151–76.

[54] Switzer LN, Tu Q, Wang J. Corporate governance and default risk in financial firms over the post-financial crisis period: international evidence. J Int Financ Mark Inst Money 2018;52:196–210.

[55] Cipollini A, Fiordelisi F. Economic value, competition and financial distress in the European banking system. J Bank Finance 2012;36(11):3101–9.

[56] Poghosyan T, Čihak M. Determinants of bank distress in Europe: evidence from a new data set. J Financ Serv Res 2011;40(3):163–84.

[57] Halteh K, Kumar K, Gepp A. Financial distress prediction of Islamic banks using tree-based stochastic techniques. Manag Finance 2018;44(6):759–73.

[58] Zaki E, Bah R, Rao A. Assessing probabilities of financial distress of banks in UAE. Int J Manag Finance 2011;7(3):304–20.

[59] Chiaramonte L, Casu B. Capital and liquidity ratios and financial distress. Evidence from the European banking industry. Br Account Rev 2017;49(2): 138–61.

[60] Constantin A, Peltonen TA, Sarlin P. Network linkages to predict bank distress. J Financ Stabil 2018;35:226–41.

[61] Bräuning M, Malikkidou D, Scricco G, Scalone S. A new approach to Early Warning Systems for small European banks. ECB Working Paper Series; 2019. No. 2348.

[62] Shrivastava S, Jeyanthi PM, Singh S. Failure prediction of Indian Banks using SMOTE, Lasso regression, bagging and boosting. Cogent Economics & Finance 2020;8(1):1729569.

[63] Persons O. Using financial information to differentiate failed vs. surviving finance companies in Thailand: an implication for emerging economies. Multinatl Finance J 1999;3(2):127–45.

[64] Cole RA, White LJ. Déjà vu all over again: the causes of US commercial bank failures this time around. J Financ Serv Res 2012;42(1–2):5–29.

[65] Chiaramonte L, Liu H, Poli F, Zhou M. How accurately can Z-score predict bank failure? Financ Mark Inst Instrum 2016;25(5):333–60.

[66] Le HH, Viviani JL. Predicting bank failure: an improvement by implementing a machine-learning approach to classical financial ratios. Res Int Bus Finance 2018; 44:16–25.

[67] Swicegood P, Clark JA. Off-site monitoring systems for predicting bank underperformance: a comparison of neural networks, discriminant analysis, and professional human judgment. Intell Syst Account Finance Manag 2001;10(3): 169–86.

[68] Bongini P, Laeven L, Majnoni G. How good is the market at assessing bank fragility? A horse race between different indicators. J Bank Finance 2002;26(5): 1011–28.

[69] Canbas S, Cabuk A, Kilic SB. Prediction of commercial bank failure via multivariate statistical analysis of financial structures: the Turkish case. Eur J Oper Res 2005;166(2):528–46.

[70] Männasoo K, Mayes DG. Explaining bank distress in Eastern European transition economies. J Bank Finance 2009;33(2):244–53.

[71] Serrano-Cinca C, Gutiérrez-Nieto B. Partial least square discriminant analysis for bankruptcy prediction. Decis Support Syst 2013;54(3):1245–55.

[72] Ecer F. Comparing the bank failure prediction performance of neural networks and support vector machines: the Turkish case. Econ Res 2013;26(3):81–98.

[73] Fiordelisi F, Mare DS. Probability of default and efficiency in cooperative banking. J Int Financ Mark Inst Money 2013;26:30–45.

[74] Papadimitriou T, Gogas P, Plakandaras V, Mourmouris JC. Forecasting the insolvency of US banks using support vector machines (SVMs) based on local learning feature selection. Int J Comput Econ Econom 2013;3(1–2):83–90.

[75] Cox RA, Wang GWY. Predicting the US bank failure: a discriminant analysis. Econ Anal Pol 2014;44(2):202–11.

[76] Li Y, Shi Y, Huang A, Yang H. Failure prediction in commercial banks with a hybrid prediction model. Annals of Data Science 2014;1(2):209–20.

[77] López-Iturriaga FJ, Sanz IP. Bankruptcy visualization and prediction using neural networks: a study of US commercial banks. Expert Syst Appl 2015;42(6):2857–69.

[78] Cleary S, Hebb G. An efficient and functional model for predicting bank distress: in and out of sample evidence. J Bank Finance 2016;64:101–11.

[79] Momparler A, Carmona P, Climent F. Banking failure prediction: a boosting classification tree approach. Revista Española De Financiación Y Contabilidad 2016;45(1):63–91.

[80] Ekinci A, Erdal Hİ. Forecasting bank failure: base learners, ensembles and hybrid ensembles. Comput Econ 2017;49(4):677–86.

[81] Gogas P, Papadimitriou T, Agrapetidou A. Forecasting bank failures and stress testing: a machine learning approach. Int J Forecast 2018;34(3):440–55.

[82] Affes Z, Hentati-Kaffel R. Predicting US banks bankruptcy: logit versus canonical discriminant analysis. Comput Econ 2019;54(1):199–244.

[83] Carmona P, Climent F, Momparler A. Predicting failure in the US banking sector: an extreme gradient boosting approach. Int Rev Econ Finance 2019;61:304–23.

[84] Climent F, Momparler A, Carmona P. Anticipating bank distress in the eurozone: an extreme gradient boosting approach. J Bus Res 2019;101:885–96.

[85] Kolari JW, López-Iturriaga FJ, Sanz IP. Predicting European bank stress tests: survival of the fittest. Global Finance J 2019;39:44–57.

[86] Filippopoulou C, Galariotis E, Spyrou S. An early warning system for predicting systemic banking crises in the Eurozone: a logit regression approach. J Econ Behav Organ 2020;172:344–63.

[87] Paule-Vianez J, Gutiérrez-Fernández M, Coca-Pérez JL. Prediction of financial distress in the Spanish banking system. Applied Economic Analysis 2019;28(82):69–87.

[88] Shrivastav SK, Ramudu PJ. Bankruptcy prediction and stress quantification using support vector machine: evidence from Indian banks. Risks 2020;8(2):52.

[89] Mishraz N, Ashok S, Tandon D. Predicting financial distress in the Indian banking sector: a comparative study between the logistic regression, LDA and ANN models. Global Bus Rev 2021:097215092110267.

[90] Pham XTT, Ho TH. Using boosting algorithms to predict bank failure: an untold story. Int Rev Econ Finance 2021;76:40–54.

[91] Lin WY, Hu YH, Tsai CF. Machine learning in financial crisis prediction: a survey. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) 2011;42(4):421–36.

[92] Verikas A, Kalsyte Z, Bacauskiene M, Gelzinis A. Hybrid and ensemble-based soft computing techniques in bankruptcy prediction: a survey. Soft Comput 2010;14(9):995–1010.

[93] Kazienko P, Lughofer E, Trawinski B. Editorial on the special issue "Hybrid and ensemble techniques in soft computing: recent advances and emerging trends". Soft Comput 2015;19(12):3353–5.

[94] Ferreira AJ, Figueiredo MA. Boosting algorithms: a review of methods, theory, and applications. In: Ensemble machine learning. New York: Springer; 2012. p. 35–85.

[95] Dastile X, Celik T, Potsane M. Statistical and machine learning models in credit scoring: a systematic literature survey. Appl Soft Comput 2020;91:106263.

[96] Adadi A, Berrada M. Peeking inside the black-box: a survey on explainable artificial intelligence (xai). IEEE Access 2018;6:52138–60.

[97] Shin D. The effects of explainability and causability on perception, trust, and acceptance: implications for explainable AI. Int J Hum Comput Stud 2021;146:102551.

[98] Cascarino G, Moscatelli M, Parlapiano F. Explainable artificial intelligence: interpreting default forecasting models based on machine learning. Occasional papers n. 674. Banca d'Italia; 2022. March 2022.

[99] EBA. Report on big data and advanced analytics. Paris: European Banking Authority; 2020. EBA/REP/2020/01. 13 January 2020.

[100] Guo W. Explainable artificial intelligence for 6G: improving trust between human and machine. IEEE Commun Mag 2020;58(6):39–45.

[101] Bank of England. Machine learning explainability in finance: an application to default risk analysis. 2019. Staff Working Paper 816. August 2019.

[102] IIF. IIF machine learning recommendations for policymakers. Institute of International Finance; 2019. 25 September 2019.

[103] Kim D, Kim S-H, Kim T, Kang BB, Lee M, Park W, Ku S, et al. Review of machine learning methods in soft robotics. PLoS One 2021;16(2):e0246102.

[104] Alonso A, Carbo JM. Machine learning in credit risk: measuring the dilemma between prediction and supervisory cost. In: Banco de Espana Working Paper; 2020. No. 2032.

[105] Jones S, Johnstone D, Wilson R. An empirical evaluation of the performance of binary classifiers in the prediction of credit ratings changes. J Bank Finance 2015;56:72–85.

[106] Mayes DG, Stremmel H. The effectiveness of capital adequacy measures in predicting bank distress: SUERF Study 2014/1. Brussels: Larcier; 2014.

[107] Arena M. Bank failures and bank fundamentals: a comparative analysis of Latin America and East Asia during the nineties using bank-level data. J Bank Finance 2008;32(2):299–310.

[108] Hong H, Huang JZ, Wu D. The information content of Basel III liquidity risk measures. J Financ Stabil 2014;15:91–111.

[109] Berger AN, Imbierowicz B, Rauch C. The roles of corporate governance in bank failures during the recent financial crisis. J Money Credit Bank 2016;48(4):729–70.

[110] Jin JY, Kanagaretnam K, Lobo GJ. Ability of accounting and audit quality variables to predict bank failure during the financial crisis. J Bank Finance 2011;35(11):2811–9.

[111] Liu CC, Ryan SG. Income smoothing over the business cycle: changes in banks' coordinated management of provisions for loan losses and loan charge-offs from the pre-1990 bust to the 1990s boom. Account Rev 2006;81(2):421–41.

[112] Alam P, Booth D, Lee K, Thordarson T. The use of fuzzy clustering algorithm and self-organizing neural networks for identifying potentially failing banks: an experimental study. Expert Syst Appl 2000;18(3):185–99.

[113] Stiroh KJ. Diversification in banking. In: Berger AN, Molyneux P, Wilson JOS, editors. Oxford handbook of banking. Oxford: Oxford University Press; 2010.

[114] Avkiran NK, Cai LC. Predicting bank financial distress prior to crises. In: New Zealand finance colloquium; 2012.

[115] Kerstein J, Kozberg A. Using accounting proxies of proprietary FDIC ratings to predict bank failures and enforcement actions during the recent financial crisis. J Account Audit Finance 2013;28(2):128–51.

[116] Beaver WH. Financial ratios as predictors of failure. J Account Res 1966;4:71.

[117] Libby R. Accounting ratios and the prediction of failure: some behavioral evidence. J Account Res 1975;13(1):150.

[118] Agarwal V, Taffler R. Comparing the performance of market-based and accounting-based bankruptcy prediction models. J Bank Finance 2008;32(8):1541–51.

[119] Gandhi P, Loughran T, McDonald B. Using annual report sentiment as a proxy for financial distress in U.S. Banks. J Behav Finance 2019;20(4):424–36.

[120] Begley J, Ming J, Watts S. Bankruptcy classification errors in the 1980s: an empirical analysis of Altman's and Ohlson's models. Rev Account Stud 1996;1(4):267–84.

[121] Curry TJ, Elmer PJ, Fissel GS. Using market information to help identify distressed institutions: a regulatory perspective. FDIC Bank Rev 2003;15(3):1–16.

[122] Curry TJ, Elmer PJ, Fissel GS. Equity market data, bank failures and market efficiency. J Econ Bus 2007;59(6):536–59.

[123] Fama EF. Efficient capital markets: a review of theory and empirical work. J Finance 1970;25(2):383–417.

[124] Calabrese R, Giudici P. Estimating bank default with generalised extreme value regression models. J Oper Res Soc 2015;66(11):1783–92.

[125] Campbell JY, Hilscher JD, Szilagyi J. Predicting financial distress and the performance of distressed stocks. J Invest Manag 2011;9(2):14–34.

[126] Distinguin I, Hasan I, Tarazi A. Predicting rating changes for banks: how accurate are accounting and stock market indicators? Ann Finance 2013;9:471–500.

[127] Simpson WG, Gleason AE. Board structure, ownership, and financial distress in banking firms. Int Rev Econ Finance 1999;8(3):281–92.

[128] Fernando JMR, Li L, Hou Y. Corporate governance and default prediction: a reality test. Appl Econ 2019;51(24):2669–86.

[129] Maghyereh AI, Awartani B. Bank distress prediction: empirical evidence from the gulf cooperation council countries. Res Int Bus Finance 2014;30:126–47.

[130] Kimmel RK, Thornton JH, Bennett SE. Can statistics-based early warning systems detect problem banks before markets? N Am J Econ Finance 2016;37:190–216.

[131] Davis EP, Karim D. Comparing early warning systems for banking crises. J Financ Stabil 2008;4(2):89–120.

[132] Boyd JH, De Nicolò G. The theory of bank risk taking and competition revisited. J Finance 2005;60(3):1329–43.

[133] Allen F, Gale D. Competition and financial stability. J Money Credit Bank 2004;36(3b):453–80.

[134] Kaufmann D, Kraay A, Mastruzzi M. The worldwide governance indicators: a summary of methodology, data and analytical issues. World Bank 2010;5430 (September).

[135] Cerchiello P, Nicola G, Rönnqvist S, Sarlin P. Assessing banks' distress using news and regular financial data. Frontiers in Artificial Intelligence 2022;5.

[136] Nyman R, Kapadia S, Tuckett D. News and narratives in financial systems: exploiting big data for systemic risk assessment. J Econ Dynam Control 2021;127:104119.

[137] Gramlich D, Finster N. Corporate sustainability and risk. J Bus Econ 2013;83(6):631–64.

[138] Brooks C, Oikonomou I. The effects of environmental, social and governance disclosures and performance on firm value: a review of the literature in accounting and finance. Br Account Rev 2018;50(1):1–15.

[139] Al-Hadi A, Chatterjee B, Yaftian A, Taylor G, Monzur Hasan M. Corporate social responsibility performance, financial distress and firm life cycle: evidence from Australia. Account Finance 2017;59(2):961–89.

[140] Sassen R, Hinze A-K, Hardeck I. Impact of ESG factors on firm risk in Europe. J Bus Econ 2016;86(8):867–904.

[141] Cornett MM, Erhemjamts O, Tehranian H. Greed or good deeds: an examination of the relation between corporate social responsibility and the financial performance of US commercial banks around the financial crisis. J Bank Finance 2016;70:137–59.

[142] Miralles-Quirós MM, Miralles-Quirós JL, Redondo-Hernández J. The impact of environmental, social, and governance performance on stock prices: evidence from the banking industry. Corp Soc Responsib Environ Manag 2019;26(6):1446–56.

[143] Branco MC, Rodrigues LL. Corporate social responsibility and resource-based perspectives. J Bus Ethics 2006;69(2):111–32.

[144] Brogi M, Lagasio V. Environmental, social, and governance and company profitability: are financial intermediaries different? Corp Soc Responsib Environ Manag 2019;26(3):576–87.

[145] Nizam E, Ng A, Dewandaru G, Nagayev R, Nkoba MA. The impact of social and environmental sustainability on financial performance: a global analysis of the banking sector. J Multinatl Financ Manag 2019;49:35–53.

[146] EBA. Consultation paper. Draft implementing standards on prudential disclosures on ESG risks in accordance with article 449a CRR. Paris: European Banking Authority; 2021. EBA/CP/2021/06. 1 March 2021.

[147] EBA. EBA Action plan on sustainable finance. Paris: European Banking Authority; 2019.

[148] EBA. EBA Report on management and supervision of ESG risks for credit institutions and investment firms. Paris: European Banking Authority; 2021. EBA/REP/2021/18. 23 June 2021.

[149] Jeucken M. Sustainable finance and banking: the financial sector and the future of the planet. London: Earthscan; 2010.

[150] Gangi F, Meles A, D'Angelo E, Daniele LM. Sustainable development and corporate governance in the financial system: are environmentally friendly banks less risky? Corp Soc Responsib Environ Manag 2019;26(3):529–47.

[151] Scholtens B, van't Klooster S. Sustainability and bank risk. Palgrave Communications 2019;5(1):1–8.

[152] Beck T, Demirgüç-Kunt A, Levine R. Financial institutions and markets across countries and over time: the updated financial development and structure database. World Bank Econ Rev 2010;24(1):77–92.

[153] Thompson P, Cowton CJ. Bringing the environment into bank lending: implications for environmental reporting. Br Account Rev 2004;36(2):197–218.

[154] García-Sánchez IM, García-Meca E. CSR engagement and earnings quality in banks. The moderating role of institutional factors. Corp Soc Responsib Environ Manag 2017;24(2):145–58.

[155] Kemp-Benedict E. Investing in a green transition. Ecol Econ 2018;153:218–36.

[156] Switzer LN, Wang J. Default risk and corporate governance in financial vs. non-financial firms. Risk Decis Anal 2013;4(4):243–53.

[157] Weber O. The financial sector's impact on sustainable development. Journal of Sustainable Finance & Investment 2014;4(1):1–8.

[158] Margolis JD, Walsh JP. Misery loves companies: rethinking social initiatives by business. Adm Sci Q 2003;48(2):268–305.

[159] Friede G, Busch T, Bassen A. ESG and financial performance: aggregated evidence from more than 2000 empirical studies. Journal of Sustainable Finance & Investment 2015;5(4):210–33.

[160] Bassen A, Meyer K, Schlange J. The influence of corporate responsibility on the cost of capital. SSRN Network Papers; 2006.

[161] Ciciretti R, Kobeissi N, Zhu Y. Corporate social responsibility and financial performance: an analysis of bank community responsibility. Int J Bank Account Finance 2014;5(4):342–73.

[162] Carreras E, Alloza Á, Carreras A. Corporate reputation, vol. 1. LID Editorial; 2013.

[163] Glavas A, Kelley K. The effects of perceived corporate social responsibility on employee attitudes. Bus Ethics Q 2014;24(2):165–202.

[164] Linthicum C, Reitenga AL, Sanchez JM. Social responsibility and corporate reputation: the case of the Arthur Andersen Enron audit failure. J Account Publ Pol 2010;29(2):160–76.

[165] Cheng B, Ioannou I, Serafeim G. Corporate social responsibility and access to finance. Strat Manag J 2014;35(1):1–23.

[166] Krueger TM, Wrolstad Mark, Shane Van Dalsem. Contemporary relationship between corporate reputation and return. Managerial Finance 2010;36(June): 482–90. https://doi.org/10.1108/03074351011042964.

[167] Gillet R, Hübner G, Plunus S. Operational risk and reputation in the financial industry. J Bank Finance 2010;34(1):224–35.

[168] Fiordelisi F, Soana M-G, Schwizer P. The determinants of reputational risk in the banking sector. J Bank Finance 2013;37(5):1359–71.

[169] Simpson WG, Kohers T. The link between corporate social and financial performance: evidence from the banking industry. J Bus Ethics 2002;35(2): 97–109.

[170] Chih HL, Chih HH, Chen TY. On the determinants of corporate social responsibility: international evidence on the financial industry. J Bus Ethics 2010; 93(1):115–35.

[171] Soana MG. The relationship between corporate social performance and corporate financial performance in the banking sector. J Bus Ethics 2011;104(1):133–48.

[172] Saxena M, Kohli AS. Impact of corporate social responsibility on corporate sustainability: a study of the Indian banking industry. IUP J Corp Gov 2012;11(4).

[173] Wu MW, Shen CH. Corporate social responsibility in the banking industry: motives and financial performance. J Bank Finance 2013;37(9):3529–47.

[174] Carnevale C, Mazzuca M. Sustainability report and bank valuation: evidence from European stock markets. Bus Ethics Eur Rev 2014;23(1):69–90.

[175] Jo H, Kim H, Park K. Corporate environmental responsibility and firm performance in the financial services sector. J Bus Ethics 2015;131(2):257–84.

[176] Nobanee H, Ellili N. Corporate sustainability disclosure in annual reports: evidence from UAE banks: Islamic versus conventional. Renew Sustain Energy Rev 2016;55:1336–41.

[177] Dell'Atti S, Trotta A, Iannuzzi AP, Demaria F. Corporate social responsibility engagement as a determinant of bank reputation: an empirical analysis. Corp Soc Responsib Environ Manag 2017;24(6):589–605.

[178] Esteban-Sanchez P, de la Cuesta-Gonzalez M, Paredes-Gazquez JD. Corporate social performance and its relation with corporate financial performance: international evidence in the banking industry. J Clean Prod 2017;162:1102–10.

[179] Abou Fayad A, Ayoub R, Ayoub M. Causal relationship between CSR and FB in banks. Arab Economic and Business Journal 2017;12(2):93–8.

[180] Forcadell FJ, Aracil E. European banks' reputation for corporate social responsibility. Corp Soc Responsib Environ Manag 2017;24(1):1–14.

[181] Weber O. Corporate sustainability and financial performance of Chinese banks. Sustainability Accounting, Management and Policy Journal 2017;8(3):358–85.

[182] Laguir I, Marais M, El Baz J, Stekelorum R. Reversing the business rationale for environmental commitment in banking. Manag Decis 2018;56(2):358–75.

[183] Maqbool S, Zameer MN. Corporate social responsibility and financial performance: an empirical analysis of Indian banks. Future Business Journal 2018;4(1):84–93.

[184] Mechelli A, Cimini R, Mazzocchetti F. The usefulness of the business model disclosure for investors' judgements in financial entities. A European study. Rev Contab - Spanish Accounting Rev 2017;20(1):1–12.

[185] Fassin Y, Gosselin D. The collapse of a European bank in the financial crisis: an analysis from stakeholder and ethical perspectives. J Bus Ethics 2011;102(2): 169–91.

[186] Turban DB, Greening DW. Corporate social performance and organizational attractiveness to prospective employees. Acad Manag J 1997;40(3):658–72.

[187] Dyer Jr WG, Whetten DA. Family firms and social responsibility: preliminary evidence from the S&P 500. Entrep Theory Pract 2006;30(6):785–802.

[188] Liston-Heyes C, Ceton G. An investigation of real versus perceived CSP in S&P-500 firms. J Bus Ethics 2009;89(2):283–96.

[189] King T, Bozos K, Koutmos D. Shareholder activism and equity price reactions. Economics Letters 160, 2017:100–4. https://doi.org/10.1016/j.econlet.2017.09.012.

[190] LaBella MJ, Russell J, Novikov D. The devil is in the details: the divergence in ESG data and implications for responsible investing. Research paper QS Investor; 2019.

[191] Berg F, Koelbel JF, Rigobon R. Aggregate confusion: the divergence of ESG ratings. MIT Working Paper; 2019.

[192] Kerber R, Flaherty M. Investing with green ratings? A gray area. Reuters; 2017.

[193] Sarkar S, Sriram RS. Bayesian models for early warning of bank failures. Manag Sci 2001;47(11):1457–75.

[194] Kolari J, Glennon D, Shin H, Caputo M. Predicting large US commercial bank failures. J Econ Bus 2002;54(4):361–87.

[195] Tung WL, Quek C, Cheng P. GenSo-EWS: a novel neural-fuzzy based early warning system for predicting bank failures. Neural Network 2004;17(4):567–87.

[196] Ng G, Quek C, Jiang H. FCMAC-EWS: a bank failure early warning system based on a novel localized pattern learning and semantically associative fuzzy neural network. Expert Syst Appl 2008;34(2):989–1003.

[197] Zhao H, Sinha AP, Ge W. Effects of feature construction on classification performance: an empirical study in bank failure prediction. Expert Syst Appl 2009;36(2):2633–44.

[198] Papanikolaou NI. A dual early warning model of bank distress. Econ Lett 2018; 162:127–30.

[199] Audrino F, Kostrov A, Ortega J-P. Predicting U.S. Bank failures with MIDAS logit models. J Financ Quant Anal 2019;54(6):2575–603.

[200] Manthoulis G, Doumpos M, Zopounidis C, Galariotis E. An ordinal classification framework for bank failure prediction: Methodology and empirical evidence for US banks. Eur J Oper Res 2020;282(2):786–801.

[201] Petropoulos A, Siakoulis V, Stavroulakis E, Vlachogiannakis NE. Predicting bank insolvencies using machine learning techniques. Int J Forecast 2020;36(3): 1092–113.

[202] Kristóf T, Virág M. EU-27 bank failure prediction with C5.0 decision trees and deep learning neural networks. Res Int Bus Finance 2022;61:101644.

[203] Molina CA. Predicting bank failures using a hazard model: the Venezuelan banking crisis. Emerg Mark Rev 2002;3(1):31–50.

[204] Lanine G, Vennet R Vander. Failure prediction in the Russian bank sector with logit and trait recognition models. Expert Syst Appl 2006;30(3):463–78.

[205] Nur Ozkan-Gunay E, Ozkan M. Prediction of bank failures in emerging financial markets: an ANN approach. J Risk Finance 2007;8(5):465–80.

[206] Ravi V, Pramodh C. Threshold accepting trained principal component neural network and feature subset selection: application to bankruptcy prediction in banks. Appl Soft Comput 2008;8(4):1539–48.

[207] Boyacioglu MA, Kara Y, Baykan ÖK. Predicting bank financial failures using neural networks, support vector machines and multivariate statistical methods: a comparative analysis in the sample of savings deposit insurance fund (SDIF) transferred banks in Turkey. Expert Syst Appl 2009;36(2):3355–66.

[208] Ravisankar P, Ravi V. Financial distress prediction in banks using Group Method of Data Handling neural network, counter propagation neural network and fuzzy ARTMAP. Knowl Base Syst 2010;23(8):823–31.

[209] Distinguin I, Hasan I, Tarazi A. The use of accounting data to predict bank financial distress in MENA countries. Int J Bank Account Finance 2010;2(4):332.

[210] Peresetsky AA, Karminsky AA, Golovan SV. Probability of default models of Russian banks. Econ Change Restruct 2011;44(4):297–334.

[211] Zaghdoudi T. Bank failure prediction with logistic regression. Int J Econ Financ Issues 2013;3(2):537.

[212] Ristolainen K. Predicting banking crises with artificial neural networks: the role of nonlinearity and heterogeneity. Scand J Econ 2018;120(1):31–62.

[213] Rosa PS, Gartner IR. Financial distress in Brazilian banks: an early warning model. Revista Contabilidade & Finanças 2017;29(77):312–31.

[214] de Haan J, Fang Y, Jing Z. Does the risk on banks' balance sheets predict banking crises? New evidence for developing countries. Int Rev Econ Finance 2020;68: 254–68.

[215] Wang T, Zhao S, Zhu G, Zheng H. A machine learning-based early warning system for systemic banking crises. Appl Econ 2021;53(26):2974–92.

[216] Chen T-H, Lee C-C, Shen C-H. Liquidity indicators, early warning signals in banks, and financial crises. N Am J Econ Finance 2022;62:101732.

[217] Huysmans J, Baesens B, Vanthienen J, Van Gestel T. Failure prediction with self organizing maps. Expert Syst Appl 2006;30(3):479–87.

[218] Hsieh NC. Hybrid mining approach in the design of credit scoring models. Expert Syst Appl 2005;28(4):655–65.

[219] Tsai CF, Chen ML. Credit rating by hybrid machine learning techniques. Appl Soft Comput 2010;10(2):374–80.

[220] Bose I. Deciding the financial health of dot-coms using rough sets. Inf Manag 2006;43(7):835–46.

[221] McLeay S, Omar A. The sensitivity of prediction models to the non-normality of bounded and unbounded financial ratios. Br Account Rev 2000;32(2):213–30.

Dr. Alberto Citterio is a postdoctoral researcher in the Department of Economics at the University of Insubria (Italy). He received his Ph.D. degree in Method and Models for Economic Decisions at the University of Insubria in 2022. He visited the Bank of Italy in Rome in 2020 and the Católica Porto Business School in 2021 as a visiting researcher.

He presented papers at national and international conferences. He received the "Cost FinAI award" in 2022 for the best paper presented at the EFCF conference in London.

His research interests and activities are mainly focused on the financial structure of financial and non-financial firms, the corporate governance of financial intermediaries, and fintech.