

Article

Predicting Bank Failures: A Synthesis of Literature and Directions for Future Research

Li Xian Liu ¹, Shuangzhe Liu ² and Milind Sathye ^{3,*}

¹ College of Business, Law and Governance, James Cook University, Douglas, QLD 4811, Australia; li.liu1@jcu.edu.au

² Faculty of Science and Technology, University of Canberra, Bruce, ACT 2617, Australia; shuangzhe.liu@canberra.edu.au

³ Faculty of Business, Government and Law, University of Canberra, Bruce, ACT 2617, Australia

* Correspondence: Milind.Sathye@canberra.edu.au

Abstract: Risk management has been a topic of great interest to Michael McAleer. Even as recent as 2020, his paper on risk management for COVID-19 was published. In his memory, this article is focused on bankruptcy risk in financial firms. For financial institutions in particular, banks are considered special, given that they perform risk management functions that are unique. Risks in banking arise from both internal and external factors. The GFC underlined the need for comprehensive risk management, and researchers since then have been working towards fulfilling that need. Similarly, the central banks across the world have begun periodic stress-testing of banks' ability to withstand shocks. This paper investigates the machine-learning and statistical techniques used in the literature on bank failure prediction. The study finds that though considerable progress has been made using advanced statistical and computational techniques, given the complex nature of banking risk, the ability of statistical techniques to predict bank failures is limited. Machine-learning-based models are increasingly becoming popular due to their significant predictive ability. The paper also suggests the directions for future research.

Keywords: risk management; bank failure prediction; machine learning; statistical methods

Citation: Liu, Li Xian, Shuangzhe Liu, and Milind Sathye. 2021. Predicting Bank Failures: A Synthesis of Literature and Directions for Future Research. *Journal of Risk and Financial Management* 14: 474. <https://doi.org/10.3390/jrfm14100474>

Academic Editor: Thanasis Stengos

Received: 16 August 2021
Accepted: 24 September 2021
Published: 8 October 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Financial institutions occupy an important position in any economy. Among these, banks in particular perform functions that are unique. The failure of a major bank in any economy would be disastrous for the entire economy due to the risk of contagion, as banks are connected with each other by payment systems. Accepting deposits repayable on demand and making loans and investments are the predominant functions that commercial banks perform, besides a host of other functions. Banks accept deposits of short maturity and make loans that have a long maturity. The unique functions that a bank performs expose it 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. Banks are the major users of technology, and consequently they are exposed to technology risk as well as operational risk. Banks' international lending exposes them to country risk. A combined effect of all these risks could lead to an insolvency risk.

Given the multifarious risks that banks face and the negative externalities they impose on the rest of the economy, banks are subject to strict prudential supervision and periodic stress-testing by regulatory agencies in all countries. The objective is to ensure that banks are prudently run so that their failures and the required bailouts are avoided. A timely prediction of a possible bank failure would considerably help supervisory authorities, as it would help identify areas where the bank is vulnerable to failure risk, and undertake risk-based on-site inspection and an audit. Bank failure prediction models

help in this respect, as they generate a better understanding of a bank's business. Supervisory authorities have also introduced an early-warning system towards this end.

Bank failure prediction has a long history. The CAMELS rating system introduced by the Federal Reserve in the United States in the mid-1990s is still in use, with revisions made from time to time. The five components of the CAMELS system include capital adequacy, asset quality, management administration, earnings, and liquidity (FRBN 1997). A composite rating was produced by CAMELS. However, the difficult dimension was how to measure management quality, since other components could be measured by financial data. The statistical techniques of bank failure prediction that used financial data typically included the use of discriminant analysis and logistic regression function. Some researchers introduced data envelopment analysis, to capture the management efficiency component. However, any model is a prototype of the reality, not the reality per se. The reality in the banking world is quite complex, and as such, predictions must be made in an everchanging dynamic environment. Towards that end, machine-learning techniques (MLTs) are increasingly being used. The major MLTs include the artificial neural network (ANN), support vector machines (SVMs), and k-nearest neighbour algorithm or KNN (Le and Viviani 2018).

Whether these techniques have helped in accurately predicting bank failures is a question that remains to be answered. It is this gap in the literature that the present paper addresses.

The paper is organised as follows: Section 2 presents a literature review of more than 60 key papers in the area; and provides a classification of these papers by methodologies, database used, study period, country and district studied, conclusions drawn, and limitations; Section 3 presents a discussion of findings, and Section 4 concludes the paper.

2. Literature on Bank Failure Prediction

The literature on business failures dates back to the late 1970s, when Beaver (1966) applied a set of financial ratios to assess the likelihood of business failure. Similarly, Altman attempted to assess the corporate bankruptcy issue using a traditional ratio analysis, as well as more rigorous statistical techniques (Altman 1968). Over the years, models of prediction have become more sophisticated.

The review of studies was conducted from two perspectives: a methodological review and a predictive indicators review.

2.1. Review of Methodology

Among statistical techniques, the methods are covered in three categories: (1) logit/probit and discriminant analysis and linear analysis; (2) artificial intelligence methods; and (3) machine-learning methods. Table 1 presents the prior studies in these categories. The papers are reviewed in chronological order.

Table 1. The literature study.

Author(s), Year	Country and District	Number of Banks	Study Period	Methodology	Predicting Indicator	Findings
(Altman 1977)	U.S.	212 savings and loan associations	1966–1973	Quadratic discriminant analysis	32 financial ratios and an additional 24 trends of these ratios are used	The results of the study show that a 12-variable econometric system is both accurate and practical for at least three semiannual periods preceding the serious problem data.
(Martin 1977)	U.S.	5700 banks. with 58 identified failures	1970–1976	Logit regression model	25 financial ratios, classified into four broad groups: asset risk, liquidity, capital	The logit and discriminant models are compared in a discriminant-analysis context by computing classification accuracy for failed and nonfailed banks. The relative merits of logit vs.

				adequacy, and earnings	discriminant analysis, at least in this empirical example, appear to depend on the intended use of the results. If a dichotomous classification into “sound” and “unsound” banks is the goal, then we may be indifferent between discriminant and logit models, since the classification accuracies are similar.	
(Lane et al. 1986)	U.S.	130 failed, 334 matching nonfailed	1978–1984	Cox proportional hazards model	21 CAMELS-related variables are used	Results of the study indicate that total classification accuracy of the Cox model is similar to that of discriminant analysis, although the Cox model produces somewhat lower type I errors. In comparison of actual and predicted times to failure, the Cox model tends to identify bankruptcies prior to the actual failure date.
(Tam 1991)	Texas in U.S.	118 banks, with 59 failed and 59 non-failed	1985–1987	Discriminant analysis model, factor-logistic model, kNN model, decision tree (ID3), and neural network model	19 financial ratios based on CAMELS criterial are used	Results show that neural networks offer better predictive accuracy than the other 4 models adopted in the study.
(Bell 1997)	U.S.	722 failed banks, 928 nonfailed banks	1983–1988	Logistic model and neural network computing	28 financial statement related variables are used.	Research results indicate that both methodologies yield similar predictive accuracy across the range of all possible model cutoff values, with the neural network performing marginally better in the “gray area”, where some failing banks appear to be less financially distressed.
(Olmeda and Fernandez 1997)	Spain	66 banks, with 29 failed and 37 nonfailed	1977–1985	Standard feedforward neural network, discriminant analysis, logit, multivariate adaptive, and C4.5	9 financial and economic ratios are used	Study finds that the ANN’s models could be superior to both classical and recently developed statistical and machine-learning classifiers. The main finding is that when one combines two or more of the methods in a simple manner, the predictions are generally more accurate than the ones obtained by applying any single method.
(Demirgüç-Kunt and Detragiache 1998)	Developed and developing countries	65	1980–1994	Logit model	Macroeconomic, financial, institutional, and past-distress variables are used	The empirical results indicate that systemic banking distress was associated with a macroeconomic environment of low economic growth, high inflation, and high real interest rates.
(Wheelock and Wilson 2000)	U.S.	N/A	1984–1993	Cox proportional-hazard model	Efficiency variables, CAMELS-ratings-related variables based on the categories of capital adequacy, asset quality, earnings,	The study finds that less well-capitalized banks, banks with high ratios of loans to assets or poor-quality loan portfolios, banks with low earnings, and managerially

					liquidity, and miscellaneous factors are used	inefficient banks are subject to greater risk of failure.
(Swicegood and Clark 2001)	U.S.	9117	1993	Multivariate discriminant analysis, neural networks, professional human judgement	23 financial and characteristic variables are used.	When comparing the predictive ability of all three models, the neural network model shows slightly better predictive ability than that of the regulators. Both the neural network model and regulators significantly outperform the benchmark discriminant analysis model's accuracy. These findings suggest that neural networks show promise as an off-site surveillance methodology. Factoring in the relative costs of the different types of misclassifications from each model also indicates that neural network models are better predictors, particularly when weighting Type I errors more heavily.
(Kolari et al. 2002)	U.S.	1079 banks, with 18 failed banks	1989–1992	Logit and trait recognition models	28 financial ratios based on size, profitability, capitalization, credit risk, liquidity, liabilities, and diversification are used	In general, trait recognition outperformed logit in the holdout samples. The prediction accuracy of the logit models was not better than chance. From a supervisory standpoint, the trait recognition model would require less maintenance in terms of updating its parameters than the logit model.
(Molina 2002)	Venezuela	36	1994–1995	Proportional hazard model	13 financial indicators; three indicators that were proxies for three of the CAMELS categories of bank performance are used	The banks with higher ROA and more investments in government bonds were less probable to fail. Yet banks with lower operational costs and higher financial expenses were more probable to fail.
(Canbas et al. 2005)	Turkey	40	1994–2001	Principal component analysis (PCA), discriminant model, the logit and probit models	49 financial ratios based on the CAMELS system are used	Due to different applications of bank regulatory and supervisory actions, CAMELS criteria do not maintain a one-to-one correspondence to the specific financial characteristics of the Turkish banks. A violation of the multivariate normal distribution with different means but equal dispersion matrices associated with models is questioned.
(Lanine and Vennet 2006)	Russia	445 banks, with 89 failed banks	1991–2001	Parametric logit model and a nonparametric trait recognition approach	7 financial ratios are used	Study results indicate that the logit and the modified trait recognition approaches perform well in terms of classification accuracy in the original samples. Both methods show lower predictive power in the holdout samples, but nevertheless they both outperform the naive benchmark forecast. Modified trait recognition outperforms the logit approach in both the original and the holdout samples. Moreover, the interpretation

						of the outcomes of the trait recognition results is theoretically straightforward.
(Ozkan-Gunay and Ozkan 2007)	Turkey	23 failed banks, and 36 unfailed banks	1989–2000	Artificial neural network (ANN)	59 financial ratios are used, grouped into four of five of the CAMELS rating systems	It is found that ANN can be successfully applied as an alternative early warning method for assisting both the banking supervisor and bank managers in emerging economies. When a confidence level of 90% is selected, 76% of the failed banks are correctly indicated, and the nonfailed banks are classified correctly 90% of the time.
(Ravi and Pramodh 2008)	Spain and Turkey	66 Spanish banks, 40 Turkish banks	1977–1985	Neural network architecture for Spanish database; Turkish database; Neural network architecture for Spanish database	12 financial ratios used for Turkish banks and 9 financial ratios used for Spanish banks	In both Spanish and Turkish banks' data, PCNN classifier outperformed all other classifiers. The proposed feature subset selection algorithm is very stable and powerful.
(Schaeck 2008)	U.S.	1000 failures	1984–2003	Quantile regression	21 financial ratios and economic factors are used. A loss rate is calculated as resolution costs divided by total assets, then a breakdown of the dataset is also used.	A quantile regression approach that illustrates the sensitivity of the dollar value of losses in different quantiles to explanatory variables is used in this study. The findings suggest that reliance on standard econometric techniques results in misleading inferences, and that losses are not homogeneously driven by the same factors across the quantiles. It is also found that liability composition affects time to failure.
(Andersen 2008)	Norway	136	2000–2005	Logit analysis	27 financial indicators are used.	The risk index comprising four indicators were not sufficient. A re-estimated of the risk index is proposed. The 6 indicators, which include capital adequacy ratio, ratio of residential mortgages to gross lending, an expected loss measure, a concentration risk measure, the return on assets, and Norges Bank's liquidity indicator, are found to be a better predictor of bank failure in Norway.
(Arena 2008)	East Asia, Latin America	444 banks from East Asia, 307 banks from Latin America	1995–1999	Multivariate logit model	8 financial ratios from the asset quality, solvency, liquidity, and return-on-assets areas are used, along with 4 interest-rate-related variables as proxies for fundamental factors.	Bank-level fundamentals significantly affect the likelihood of collapse for these banks. As shown by the survival time analysis for the Latin American case, the banking system and macroeconomic variables also explain the likelihood of failure.
(Boyacioglu et al. 2009)	Turkey	65 banks	1997–2003	Neural networks such as multilayer perceptron (MLP), competitive	20 financial ratios with six features groups from the CAMELS system are adopted.	After the comparison, MLP and LVQ are considered the most successful models in predicting the financial failure of banks in the sample.

				learning (CL), self-organizing map (SOM), and learning vector quantization (LVQ) are employed; and logit, multivariate discriminant analysis, k-means cluster analysis are employed.		
(Ercan and Evirgen 2009)	Turkey	36 failed banks, 45 nonfailed banks	1997–2006	Principal component analysis, which included multinomial logit model and traditional binary model	8 microeconomic variables based on CAMELS rating categories are used	From the macroeconomic perspective, higher credit growth and real interest rates are associated with a higher probability of banking failures.
(Männasoo and Mayes 2009)	19 Eastern European economies	600	1995–2004	Survival model and panel data analysis	21 macroeconomic, structural and bank-specific factors are used.	Bank-specific variables such as liquidity variables provide a strong signal about approaching failure. Changes in bank earnings, efficiency, and relative size of credit portfolio do not provide an early warning of distress.
(Zhao et al. 2009)	U.S.	480	1991–1992	Datamining methods: logistic regression, decision tree, neural network, and k-nearest neighbor	93 raw accounting variables, and 26 constructed financial ratios are used.	The study empirically demonstrated that constructed high-level features such as financial ratios can significantly improve the performance of classifiers by using different methods. It is important to address the issue of the fusion of data mining and domain knowledge in future studies.
(Cebula 2010)	U.S.	Bank failure rate	1970–2007	Eclectic model	5 economic and financial factors and three federal banking statutes, which include average percentage unemployment rate, average nominal cost of funds, variance of monthly averages of closing prices of the S&P 500 index, the average ratio of net charge-offs to outstanding loans, the average interest rate yield on new 30-year fixed rate mortgages, the FDICIA of 1991; the Riegle–Neal Interstate Banking Act of 1994; and the Gramm–	The bank failure rate was found to be an increasing function of the unemployment rate, the average cost of funds, volatility of the S&P 500 stock index, and charge-offs as percentage of outstanding loans and a decreasing function of the mortgage rate on new 30-year fixed-rate mortgages.

				Leach–Bliley Act of 1999		
(Jordan et al. 2010)	U.S.	225 failed banks	2007–2010	Regression and multiple discriminant analysis	9 market-to-book ratios are used.	The ratio of nonaccrual assets + ORE to total assets, the ratio of interest income to earning assets, Tier 1 capital to total assets ratio, the ratio of real estate loans to total assets, and the savings bank and MSA dummy variables have a strong statistical relationship to bank failure status. The model successfully predicts from 66.0% (4 years prior to failure) to 88.2% (1 year prior to failure) of failed banks, with an overall success rate of 76.8%.
(López-Iturriaga et al. 2010)	U.S.	82 defaulted banks, and 196 nondefaulted banks	2003–2008	Neural networks	41 indicators (explanatory variables for bankruptcy risk) are used.	The study reveals distressed banks were exposed to high credit risks and the loan portfolio was concentrated in real estate loans as a result of careless bank strategies rather than low cost efficiency. Further, the model shows a high discriminant power and is able to differentiate correctly wealthy and distressed banks when the model is used to predict future bankruptcies and test the performance of the model by comparing our predictions with the actual bankruptcies between January and June 2010. Specifically, the model would have been able to predict in December 2009 around 60% of failures that occurred in the first six months of 2010.
(Ravisankar and Ravi 2010)	Spain, Turkey, U.K., and U.S.	150 distressed banks and 145 healthy ones in 4 countries	Different historical periods	Three neural network architectures: group method of data handling (GMDH), counter propagation neural network (CPNN), and fuzzy adaptive resonance theory map (ARTMAP)	12 predictor variables for Turkish banks, 9 for Spanish banks, 5 for U.S. banks, 10 for U.K. banks. Variables are based on CAMELS' 6 functional areas: capital adequacy, asset quality, management expertise, earning strength, liquidity, and sensitivity to market risk.	Results indicate that the GMDH outperformed all the techniques with or without feature selection. Furthermore, the results are much better than those reported in previous studies on the same datasets in terms of average accuracy, average sensitivity, and average specificity.
(Wong et al. 2010)	11 EMEAP economies	Bank	1990–2007	Panel probit model	Macroeconomic fundamentals are used.	The model suggests that slowing GDP growth, rising inflation rate, and an increase in money supply relative to foreign reserves associated with deteriorating creditworthiness of banks and nonfinancial companies and are useful leading indicators of banking distress. Contagion effects are present.

(Tatom and Houston 2011)	U.S.	1470	1988–1994; 2006–2010	Probit, logit, and DEA model	CAMELS-related, local, and national economic variables are used.	The model developed in this study has strong forecasting accuracy in both the in-sample and out-of-sample forecasts.
(Jin et al. 2011)	U.S.	6437	2006–2007	Simple univariate and multivariate analysis	13 accounting and auditing variables are used.	Auditor type, auditor industry specialization, Tier 1 capital ratio, proportion of securitized loans, growth in loans, and loan mix are reliable predictors of bank failure.
(Poghosyan and Čihak 2011)	25 European Union countries	5708	1996–2007	Logistic probability model	6 bank-specific financial ratios are used.	Asset quality and earnings profile of banks are important determinants of bank distress next to leverage. The model correctly classifies 44 out of 79 distress events (55.7%) and 29,706 out of 29,783 nondistress events (99.7%) for the 10% cutoff point. It also failed to correctly classify 35 distress events out of 79 and wrongly classified 77 healthy bank-year observations out of 29,783 as distressed. Overall, the model performs satisfactorily in classifying distressed banks. Further, data points to the presence of contagion effects in the fragility of concentrated banking sectors.
(Cipollini and Fiordelisi 2012)	European	308	1996–2009	Panel probit regression model	Bank level (liquidity and credit risks, asset size, income diversification, and market power), industry level, and macro-level are used.	The empirical findings show that credit risk (measured by the ratio of loan loss provisions to total loans), liquidity risk (measured by the ratio of liquid assets to total assets), and bank market power (measured by the Lerner index) are the most influential determinants of distressed SHVR (small changes in the dependent variable). Moreover, it is found that the pooled probit regression model is the one improving upon a naive predictor in countries such as Portugal, Ireland, Greece, Italy, and Spain during the most recent EMU sovereign debt turmoil period.
(Al-Tamimi 2012)	UAE	23	2007	Modified questionnaire surveys, a linear regression analysis	6 corporate governance practices variables are used.	Results find there is a significant positive relationship between financial distress and CG practices of UAE national banks. However, the results indicate that the role of CG practices is not sufficient in the case of financial distress or financial crisis.
(Cole and White 2012)	U.S.	265	2004–2008	Multivariate logistic regression	15 financial ratios and real estate mortgage and loan variables are used.	Study finds that traditional proxies for the CAMELS ratings are important determinants of bank failures. However, portfolio variables such as real estate construction and development loans, commercial mortgages, and multifamily mortgages are consistently associated

						with a higher likelihood of bank failure.
(DeYoung and Torna 2013)	U.S.	6851	2008–2010	Multiperiod logit model	Income from nontraditional and traditional banking activities are used.	Study suggests that income from pure fee-based nontraditional activities are less likely to contribute distressed bank failure; yet, income with asset-based nontraditional activities such as venture capital, investment banking, and asset securitization likely increase the probability of distressed bank failure.
(Ecer 2013)	Turkey	34 banks with 17 failed	1994–2001	Artificial neural networks, support vector machines	36 financial ratios are used.	This study challenges the superiority of ANNs in classifying problems. However, both ANNs and SVMs are promising prediction models in identifying potentially failing banks.
(Erdogan 2013)	Turkey	42 banks with 21 failed	1997–2003	Support vector machines	19 financial ratios are used based on capital ratios, assets quality, liquidity, profitability, and income-expenditure structure	This study shows that SVMs with the Gaussian kernel are capable of extracting useful information from financial data and can be used as part of an early-warning system.
(Kerstein and Kozberg 2013)	U.S.	7835	2007–2010	Probit model	15 accounting-based proxies that are similar to the 6 categories of the CAMELS rating system are used.	Study finds that six categories of CAMELS—capital adequacy, asset quality, management, earnings, liquidity, and sensitivity to interest rates—are significantly associated with the probability of bank failure when examined individually and nearly all measures maintain their significance when examined collectively.
(Serrano-Cinca and Gutiérrez-Nieto 2013)	US	8293	2008–2011	Partial least-squares discriminant analysis (PLS-DA), linear discriminant analysis, logistic regression, l regression stepwise, multilayer perceptron, k-nearest neighbours, naive Bayes, support vector machine, boosting C4.5, bagging random tree	17 financial ratios were extracted	PLS-DA results are very close to those obtained by Linear Discriminant Analysis and Support Vector Machine.
(Cox and Wang 2014)	U.S.	322	2007–2010	Linear and quadratic discriminant analysis	19 financial variables, including broader categories of types of loan made; asset, liability and equity composition; bank size;	The proportion of illiquid loans in their books and the exposure to the interbank funding markets are the main predictors of bank failures. Quadratic discriminant analysis

					and income statement measures are used.	outperformed LDA models in predicting bank failures.
(Avkiran and Cai 2012)	U.S.	186	2004–2006	CAMELS and CPM regression analysis	Financial ratios based on the CAMELS system are used, and measurement of production efficiency is used.	The results from the CAMELS and CPM models support DEA’s discriminatory and predictive power, suggesting that users can rely on DEA results generated from financial data up to 2 years prior to the crisis. Moreover, the CPM model outperforms the CAMELS model, indicating profitability is a key factor in predicting financial distress in banks.
(Betz et al. 2014)	All EU countries except Cyprus, Estonia, Lithuania, and Romania	546	2000–2013	Logit model	Three categories of indicators: bank-specific indicators, CAMELS rating system indicators, and country-specific macro-financial indicators are used.	The key findings of the paper are that complementing bank-specific vulnerabilities with indicators for macro-financial imbalances and banking sector vulnerabilities improves model performance and yields useful out-of-sample predictions of bank distress during the financial crisis at the time.
(Hong et al. 2014)	U.S.	9349	2001–2011	Time hazard model	NSFR (net stable funding ratio) and LCR (liquidity coverage ratio) based on Basel III requirements	Systemic liquidity risk was a major contributor to bank failures in 2009 and 2010, while the net stable funding ratio (NSFR) and liquidity coverage ratio (LCR) proposed by the Basel Committee in December 2010 had limited effects on bank failures.
(Maghyreh and Awartani 2014)	Gulf cooperation council countries	70	2000–2009	Simple hazard model	A wide set of bank level variables, which include the CAMELS type, non-CAMELS type, and other variables including the influence of bank management, competition, diversification, ownership and regulation are used.	The study finds that good management lowers the likelihood of distress. Moreover, competition and diversification were found to be bad for the health of banks. The institutional development index was a statistically relevant predictor. Finally, by conditioning of the relevant covariates, a simple hazard model has performed fairly well in predicting bank distress in the GCC countries.
(Mayes and Stremmel 2014)	U.S.	16,188	1992–2012	The logit technique and discrete survival time analysis	CAMELS indicators that consider the bank-specific variables and macroeconomic conditions are used.	The study finds that the non-risk-weighted capital measure (the adjusted leverage ratio) explains bank distress and failures best. The logit model is able to distinguish failing from healthy banks with an accuracy of 80%. The corresponding survival time model achieves 98%.
(Chiaramonte et al. 2015)	12 European countries	3242	2001–2011	Probit and complementary log-log models	Z-score, CAMELS variables including capital, asset quality, managerial skills, earnings, liquidity, and sensitivity to market risk are used.	The study finds that the Z-score’s ability to identify distress events, both in the entire period and during the crisis years (2008–2011), is at least as good as the CAMELS variables, but with the advantage of being less data-demanding. Finally, the Z-score

						proves to be more effective when bank business models may be more sophisticated
(Iturriaga and Sanz 2015)	U.S.	386	2012–2013	Neural networks: multilayer perceptron network and self-organizing maps (MLP-SOM) model	32 financial ratios used in the literature that are potentially explanatory for bankruptcy risk are chosen. Additional variables with a criterion adapted to the network to improve the results of the model are chosen as well.	A model combining multilayer perceptrons and self-organizing maps is used. Results show that hybrid MLP-SOM model has a high and stable predictive power over time, reaching a balance between Type I and Type II errors.
(Berger et al. 2016)	U.S.	341	2007–2012	Multivariate logit model	16 corporate governance indicators, 12 accounting indicators, 2 market competition indicators, 2 economic indicators, and 2 primary federal regulator indicators are used.	The study finds that a bank’s ownership structure plays a substantial role in explaining likelihood of failure.
(Chiaramonte et al. 2016)	U.S.	8478	2004–2012	Discrete time proportional hazards model	Z-score estimation, and 9 bank and macro-level factors are used.	The study finds that on average, the Z-score can predict 76% of bank failures, and an additional set of other bank- and macro-level variables do not increase this predictability level. It also was found that the prediction power of the Z-score to predict bank defaults remains stable within the three-year forward window.
(Cleary and Hebb 2016)	U.S.	132	2002–2009	Discriminant analysis	13 financial data, such as retained earnings to total assets, liquidity measure, sustainable profitability measure, operating efficiency measure, leverage measure, reliance on loans, loan quality, capital adequacy, and off-balance-sheet items are used.	Bank capital, loan quality, and cash holdings are associated with bank failure.
(Mompalmer et al. 2016)	Euro zone	155	2006–2012	Machine-learning method, boosted regression trees	25 financial ratios are used.	The findings indicate that the greater the size and the higher the income from nonoperating items and net loans to deposits, the more likely is bank failure; conversely, the higher the Interbank ratio, the lower the chances of bank financial distress. For the sake of their own financial soundness, banks should fund lending activities through clients’ deposits and should avoid relying excessively on nonrecurring sources of income.

(Tanaka et al. 2016)	OECD	18,381	1986–2014	Random forests	48 indicators based on four groups: profitability ratio, capitalization, loan quality, and funding are used.	The results of experiments showed that the random forests EWS outperformed conventional EWSs in terms of prediction accuracy.
(Chiaramonte and Casu 2017)	EU banks	513	2004–2013	Pooled logit model	Structural liquidity and capital ratios as defined in Basel III are used.	Estimates from several versions of the logistic probability model indicate that the likelihood of failure and distress decreases with increased liquidity holdings, while capital ratios are significant only for large banks
(Ekinici and Erdal 2017)	Turkey	37	1997–2001	Three common machine-learning models (logistic, J48, and voted perceptron), random subspaces, bagging, and multiboosting	35 financial ratios, including capital, asset quality, management, earnings, liquidity, and sensitivity ratios (CAMELS) are used.	The models are grouped in the following families of approaches: (i) conventional machine-learning models; (ii) ensemble learning models; and (iii) hybrid ensemble learning models. Experimental results indicate a clear outperformance of hybrid ensemble machine-learning models over conventional base and ensemble models. These results indicate that hybrid ensemble learning models can be used as a reliable predicting model for bank failures.
(Bongini et al. 2018)	20 Central, Eastern, and Southeastern European countries	355	1995–2017	Regression	Z-score and CAMELS-based financial strength indices	The study finds that the predictive power of both types of accounting-based measures is weak.
(Constantin et al. 2018)	European	172 bank distress events	1999–2012	Estimated network linkages based on multivariate extreme value theory	Bank specific vulnerabilities, banking sector and macro-financial indicators, and indicators covering all dimensions in the CAMELS rating system are used.	Beyond standard bank-level risk drivers and macro-financial indicators, a tail-dependence network provides additional information about market’s view on bank interconnectedness in situations of elevated financial stress. It can provide information on potentially vulnerable banks following an early-warning signal or a bank failure, and the potential for financial contagion and a systemic banking crisis.
(Iwanicz-Drozdowska et al. 2018)	Europe	163 distressed banks, 3566 nondistressed banks	1992–2014, 2008–2012	Factor and cluster analysis, logistic regression	12 CAMELS-based variables are used.	It is difficult to predict the distress events with the use of a set of CAMELS-like variables, although they are widely used in academic literature and in practice.
(Jing and Fang 2018)	U.S.	293 failed banks	2002–2010	Logit model, neural networks, support vector machines	16 financial ratios covering CAMELS-related variables, as well as rates of change of the financial ratios are used.	Empirical results indicate that the logit model issues more missed failures and false alarms in-sample, but issues fewer missed failures and false alarms out-of-sample, than the data-mining models. The study suggests that the logit model is a good

						and robust tool to predict bank failures. In addition, the logit model allows a better understanding of the relations between financial variables and bank failures, which enables bank supervisors to assess banks' financial health more efficiently than when using data-mining models. Data-mining models can predict bank failures well when the sample is divided randomly, but this does not hold when the sample is divided by time.
(Gogas et al. 2018)	U.S.	1433 banks, 481 failed	2007–2013	Support vector machine	36 financial ratios are used.	The model exhibits a 99.22% overall forecasting accuracy and outperforms the well-established Ohlson's score.
(Le and Viviani 2018)	U.S.	3000 banks (1438 failures)	various years	Discriminant analysis, logistic regression, artificial neural network, support vector machines, and k-nearest neighbours	31 financial ratios covering 5 main aspects from CAMELS are used.	The empirical result reveals that the artificial neural network and k-nearest neighbour methods are the most accurate.
(Beutel et al. 2019)	15 advanced economies	19 crises	1970–2016	Logit model, random forest, support vector machines, k-nearest neighbours, and decision trees	10 variables based on assets prices and credit developments, macroeconomic environment, external and global imbalances, and time trend are used	The study finds that while machine-learning methods often attain a very high in-sample fit, they are outperformed by the logit approach in recursive out-of-sample evaluations. The study also suggests that further enhancements to machine-learning early-warning models are needed before they are able to offer a substantial added value for predicting systemic banking crises. Conventional logit models appear to use the available information already fairly efficiently, and would, for instance, have been able to predict the 2007/2008 financial crisis out-of-sample for many countries. In line with economic intuition, these models identify credit expansions, asset price booms, and external imbalances as key predictors of systemic banking crises.
(Iwanicz-Drozdowska and Ptak-Chmielewski 2019)	European	3691 banks with 132 distress events	1990–2015	Logistic regression, and k-means clustering	CAMELS-like bank-level variables, and control macroeconomic variables.	The study finds that the probability of distress is connected with macroeconomic conditions via regional grouping (clustering). Bank-level variables that were stable predictors of distress from 1 to 4 years prior to an event are the ratios of equity to total assets (leverage) and loans to funding (liquidity). For macroeconomic factors, the GDP growth is a reasonable variable, but

						with a differentiated impact, which shows the changing role of the macroeconomic environment and indicates the potential impact of favorable macroeconomic conditions on the accumulation of systemic problems in the banking sector.
(Kolari et al. 2019)	European	91	2010, 2011, 2014	AdaBoost Ensemble	21 financial ratios based on 6 groups of the CAMELS rating system are used.	The model is able to identify over 98% of failing and passing banks in the training subsample and predict about 90% of banks in the test validation sample.
(Carmona et al. 2019)	U.S.	156	2001–2015	Gradient boosting approach	30 financial ratios based on performance and condition ratios are used.	The findings indicate that lower values for retained earnings to average equity, pretax return on assets, and total risk-based capital ratio are associated with a higher risk of bank failure. In addition, an exceedingly high yield on earnings assets increases the change of bank financial distress.
(Shrivastava et al. 2020)	India	58	2000–2017	Synthetic minority oversampling technique (SMOTE), lasso regression, random forest, AdaBoost	26 bank-specific, macroeconomic, and market-structure variables are used.	This study offers an analytical approach, including the selection of the most significant bank-failure-specific indicators using lasso regression, converting data from imbalanced to balanced form using SMOTE, and the choice of the appropriate machine-learning techniques, to predict the failure of the bank. AdaBoost was found to have the maximum accuracy.
(de Haan et al. 2020)	147 emerging and developing countries	110 banking crises	1980–2016	Panel logit regression model	Finance balance sheet ratios are used.	The results suggest that low levels of bank liquid assets and domestic financial liabilities, and high levels of foreign liabilities and financial leverage, increase the likelihood of a banking crisis. These results are robust when different dependent variables and control variables are used. Results also show that there is no single optimal lag length for all the indicators. Combining all indicators together, it is found that the indicators have the best predictive power with a lag of 42 months.

2.1.1. Discriminant Analyses

The family of discriminant analyses includes linear discriminant analysis (LDA), multivariate discriminate analysis (MDA), and the quadratic discriminant analysis (QDA). These remained the leading techniques for many years. The first application of a discriminant analysis to explain corporate failure was performed by Altman (1968). Studies related to specific corporate groups such as banking soon followed; for example, the Sinkey (1975) study on commercial banks. Bloch (1969) applied linear discriminant analysis in an exploratory study of savings and loan associations, and the encouraging results helped to initiate Altman’s study in the same area. Altman (1977) adopted a quadratic

discriminant analysis in predicting performance in the savings and loan association industry.

Adopting these methods, researchers used U.S. bank data to identify the main explanatory contributors of bank failure (Cleary and Hebb 2016; Cox and Wang 2014; Jordan et al. 2010). In order to address the classification problem associated with discriminant methods, Lam and Moy (2002) presented a method that combined several discriminant methods to predict the classification of new observations. The simulation experiment proved further enhanced accuracy of classification results. Serrano-Cinca and Gutiérrez-Nieto (2013) performed an empirical study, comparing partial least-squares discriminant analysis (PLS-DA) with other eight techniques widely used for classification tasks. The results showed that PLS-DA performed very well in the presence of multicollinearity, with a satisfactory interpretability. The PLS-DA results resembled the linear discriminant analysis and support vector machine results.

2.1.2. Logit/Probit and Linear Regression Analysis

When independent variables are not normally distributed, maximum likelihood methods such as logit and probit models are used. These were used in many studies on bank failure prediction. A logit model is a nonlinear model with dichotomous outcome variables of failed/nonfailed bank. After Martin's (1977) application of a logit model to predict bank failures in the U.S., various studies adopted this model (univariate or multivariate) to predict bank failures in different countries in different periods. These included, for example, Andersen (2008) in Norway; Arena (2008) in East Asia and Latin America; Ercan and Evirgen (2009) in Turkey; Zhao et al. (2009), Cole and White (2012), DeYoung and Torna (2013), Mayes and Stremmel (2014), and Berger et al. (2016) in the U.S.; Poghosyan and Čihak (2011), Betz et al. (2014), and Chiaramonte and Casu (2017) in most of the European Union countries and banks; Demirgüç-Kunt and Detragiache (1998) in 65 developing and developed economies; and de Haan et al. (2020) in 147 emerging and developing countries.

The probit model is another binary model used in banking failure studies (Chiaramonte et al. 2015; Cipollini and Fiordelisi 2012; Kerstein and Kozberg 2013; Wong et al. 2010). Research in this area found that the accuracy was similar to that of logit models (Barr and Siems 1997).

The hazard model as another statistical model is also applied to predict bank failures; this stream of study includes Lane et al. (1986), Molina (2002), Hong et al. (2014), Maghyreh and Awartani (2014), and Chiaramonte et al. (2016). However, Cole and Wu (2009) compared the out-of-sample forecasting accuracy of the time-varying hazard model and the one-period probit model, using data on U.S. bank failures from 1985–1992, and the study found that from an econometric perspective, the hazard model was more accurate than the probit model in predicting bank failures when more recent information was incorporated in the hazard model.

Although standard discriminant analysis has been a popular technique for bankruptcy studies, it suffers from methodological or statistical problems that have limited the practical usefulness of their erroneous results (Ozkan-Gunay and Ozkan 2007). Violations of the normality assumptions may bias the tests of significance and estimated error rates (Ohlson 1980). However, as an early study of the application of the Cox model in finance literature, empirical results from Lane et al. (1986) indicated that the total classification accuracy of the Cox model was similar to that of discriminant analysis. Lanine and Vennet (2006) and Kolari et al. (2002) both used a logit model and a trait-recognition approach to predict bank failures in Russia and the U.S. Both concluded that a trait-recognition approach outperformed the logit approach.

Prediction can be described as a classification method. In the context of banking failure prediction, we categorized the banks into failed and nonfailed groups, which is exactly what data-mining models focus on. As data-mining models capture the relationships between dependent and independent variables by learning from the data, imposing fewer

constraints than traditional statistical models such as the logit model on the distribution of the data (Jing and Fang 2018). In the next subsection, we will review the studies in this area.

2.1.3. Artificial Intelligence Method

The traditional approach of predicting business distress or failures has been criticized because the validity of its results hinges on restrictive assumptions (Coats and Fant 1993). In order to address the problematic issues brought by linear analysis, researchers began bankruptcy studies through neural network analysis in 1990. Neural networks differ from the classical approach because these models assume a nonlinear relation among variables (De Miguel et al. 1993). Tam (1991) believes a neural network is a learning process when given a collection of failed and nonfailed banks, and a network is trained by using a learning algorithm so that the resultant network not only represents a discriminant function for the sample banks, but also makes generalizations from the training sample. Atiya (2001) argued that there are saturation effects in the relationships between the financial ratios and the prediction of default. The following are the bank failure prediction studies that have applied the neural network approach.

One of the early studies adopting neural network was that of Tam (1991), who examined failed banks in the period of 1985–1987. López-Iturriaga et al. (2010) applied the neural network method, studying U.S. commercial banks during the financial crisis period. The model showed a high discriminant power and was able to differentiate healthy and distressed banks. López Iturriaga and Sanz (2015) developed a hybrid neural network model to study U.S. bank bankruptcies. Based on the data, which spanned between 2002 and 2012, the model detected 96.15% of the failures and outperformed traditional models of bankruptcy prediction. Constantin et al. (2018) studied the European bank network with a distress model that offered information about the external-dependence structure of listed European banks. The model could provide information on potential distress following an early-warning signal, and the potential for financial contagion and a systemic banking crisis.

Similar studies have been applied in emerging markets. Olmeda and Fernandez (1997) examined the bankruptcies of Spanish banks, and found the artificial neural network approach had an 82.4% accuracy, compared with 61.8–79.4% for the competing techniques. Ravisankar and Ravi (2010) adopted three unused neural network architectures for bank distress for four different countries. Ozkan-Gunay and Ozkan (2007) applied the artificial neural network approach for examining bank failures in the Turkish banking sector. A new principal component neural network (PCNN) architecture for commercial bank bankruptcy prediction also was proposed and examined in the Spanish and Turkish banking sectors, and the hybrid models that combined PCNN and several other models of banking bankruptcy prediction outperformed other classifiers used in the literature (Ravi and Pramodh 2008). The superiority of artificial-neural-network-related models was further documented and supported (Bell 1997; Boyacioglu et al. 2009; Swicegood and Clark 2001).

Ecer (2013) compared the ability of an artificial neural network (ANN) and support vector machine (SVM) in predicting bank failures in Turkish banks. Of these two models, neural networks were observed to have a slightly better predictive ability than support vector machines. A similar comparative study was conducted by Jing and Fang (2018); however, the study was in favour of the logit model. Le and Viviani's (2018) comparative study revealed that the artificial neural network and k-nearest neighbour methods are the most accurate models.

2.1.4. Machine-Learning Methods (Including Ensembles, Support Vector Machines, Generalized Boosting, AdaBoost, and Random Forests)

Recent statistical learning techniques such as generalized boosting, AdaBoost, and random forests are used to predict banking failure with the purpose of improving prediction accuracy. Using a comprehensive dataset encompassing systemic banking crises for 15 advanced economies over the past 45 years, Beutel et al. (2019) concluded that machine learning helps us predict banking crises.

Tanaka et al. (2016) adopted a novel random-forests-based approach for predicting bank failures for OECD member countries. The experimental results showed that this method outperformed conventional methods in terms of prediction accuracy. Momparler et al. (2016) found the boosted regression trees method was a better model to identify a set of key leading indicators, and further to anticipate and avert bank financial distress. Ekinci and Erdal (2017) applied three common machine-learning models in analysing bank failure prediction for 37 commercial banks operating in Turkey between 1997 and 2001. The experimental results indicated that hybrid ensemble machine-learning models outperformed conventional base and ensemble models.

Erdogan (2013) found that the support vector machine method with a Gaussian kernel was a good application for bank bankruptcy. Gogas et al. (2018) found that a model trained by a support vector machine had an overall accuracy of 99.22%.

Olson et al. (2012) applied a variety of data-mining tools to bankruptcy data to compare accuracy and number of rules. Decision trees were found to be more accurate than neural networks and support vector machines, albeit with an undesirably high number of rule nodes.

Carmona et al. (2019) adopted an extreme gradient-boosting approach that was not required to be managed like a black box, and found out the predictive power was greater than most conventional methods. Kolari et al. (2019) studied a European bank stress test by using an AdaBoost ensemble approach, and the models' accuracy was found to be 98.4%. A similar result was found by Shrivastava et al. (2020) in the banking sector in India.

Overall, many studies compared the traditional approaches to several machine-learning approaches, as it is well documented that machine-learning methods outperform the traditional models. However, further enhancements to machine learning are needed when we consider the performance metric, crisis or distress event definition, preference parameters, sample length, and regulatory differences among countries.

2.2. Review of Predicting Indicators

In the empirical literature, the prediction of bank failure has been primarily focused on the identification of leading indicators that contribute to generate reliable early warning systems (Chiaramonte et al. 2015). This group of indicators mostly includes financial/accounting-based indicators since Beaver (1966) pioneered the prediction of bankruptcy using financial statement data such as financial leverage, return on assets, and liquidity.

In our particular banking sector, over the years, the Federal Reserve and FDIC developed their own methodology for identifying distress in the banking sector (Kerstein and Kozberg 2013). The initial CAMELS rating comprised five categories: capital adequacy, asset quality, management, earnings, and liquidity, to indicate the condition of a bank. In 1996, the CAMELS system was expanded to include a sixth rating area. Nevertheless, the bank-level fundamentals proxied by CAMELS-related variables has been extensively studied for a particular country or district or at a cross-country level (Arena 2008; Chiaramonte et al. 2015; Iwanicz-Drozdzowska and Ptak-Chmielewska 2019; Kerstein and Kozberg 2013; Kolari et al. 2002; Lane et al. 1986; Maghyereh and Awartani 2014; Mānasoo and Mayes 2009; Molina 2002; Wheelock and Wilson 2000), and most of them were associated with a statistical model such as the logit model.

In the early seminar articles on bankruptcy prediction, Altman (1968), Beaver (1966), and Beaver (1968) used Z-scores that comprise five market- and/or accounting-based ratios to predict business failures. Subsequent articles adapted or expanded the use of Z-score analysis to predict bank failure. Martin (1977) drew a set of 25 financial ratios from the database maintained by the Federal Reserve Bank of New York for research on bank surveillance programs, and used a similar logit analysis to Altman's study to examine bank failures in the period of 1975–1976. Chiaramonte et al. (2016) examined U.S. commercial banks data from 2004 to 2012 and found that on average, the Z-score can predict 76% of bank failures, and an additional set of other bank- and macro-level variables did not increase this predictability level. However, Bongini et al. (2018) found that the predictive power of the Z-score was weak, especially for developing economies.

Although traditional CAMELS indicators are found to be successful in anticipating bank failures in the U.S., Canbas et al. (2005) found that these criteria did not maintain a one-to-one correspondence with the specific financial characteristics for Turkish commercial banks due to different applications of bank regulatory and supervisory actions. Kapinos and Mitnik (2016) proposed a simple method for stress-testing banks using a top-down approach that captured the heterogeneous impact of shocks to macroeconomic variables on banks' capitalization. They performed a principal component analysis on the selected variables and showed how the principal component factors can be used to make projections, conditional on exogenous paths of macroeconomic variables. Ercan and Evirgen (2009) and Canbas et al. (2005) adopted the same approach (principal component analysis). Iwanicz-Drozdowska et al. (2018) also found that it was difficult to predict the distress with a set of CAMELS-like variables in the European setting.

Meanwhile, researchers are attempting to find other explanatory factors to address the distress phenomena. These include macroeconomic and regulation variables (Cebula 2010; Männasoo and Mayes 2009; Schaeck 2008; Wong et al. 2010), accounting and audit quality (Jin et al. 2011), income from nontraditional banking activities (DeYoung and Torna 2013), market and macroeconomic variables (Cole and Wu 2009), commercial real-estate investments (Cole and White 2012), information content of Basel III liquidity risk and capital measures (Chiaramonte and Casu 2017; Hong et al. 2014), corporate governance (Al-Tamimi 2012; Berger et al. 2016).

Data envelopment analysis (DEA) has been widely applied in banking efficiency studies. Although DEA suffers from the usual statistical inefficiency problems found in nonparametric estimation (Kneip et al. 1996), the efficiency variable generated from this method is also used as an indicator to predict banking failure. Wheelock and Wilson (2000) adopted the DEA method by developing an operating efficiency as a measure of management performance, along with other CAMELS-related variables to investigate the determinants of U.S. bank failures. Similar studies were conducted in different banking sectors in different countries (Avkiran and Cai 2012; Barr and Siems 1997; Cipollini and Fiordelisi 2012; Kao and Liu 2004; Tatom and Houston 2011). Barr and Siems (1997) found their model outperformed many previous logistic models in predicting failure when DEA efficiency as the proxy for the management quality and other CAMELS-ratings related variables were used.

3. Discussion

We reviewed 24 papers that in the artificial intelligence and machine-learning research areas, and 41 papers that used regression models and discriminant analyses to assess bank failures—a total of 65 papers. Though regression models formed close to 50% of the papers on bank failures after the global financial crisis (GFC), the recent trend seems to be to use machine-learning techniques for prediction of bank distress. The accuracy rate of machine-learning models as reported above is 95% generally. Almost half of the machine-learning papers used U.S. bank data. The other half was scattered throughout a few European countries. The use of artificial intelligence and machine-learning approaches requires solid skills in these areas, and few banks and regulators may have the necessary

expertise. The statistical techniques, on the other hand, are commonly used, and data are easily available to the banks. In addition, from a cost perspective, data and other associated costs are much higher if artificial intelligence or machine-learning techniques are to be used (Incze 2019). Overall, more research is required using banking data, regulation, macroeconomic conditions, and market structure in non-U.S. countries. Research on Asia Pacific countries is woefully lacking, barring one paper that used Indian bank data.

However, we do not know whether regulatory agencies have adopted these models in practice or whether the banks, in their own interest, use these models to assess their vulnerability periodically. Future studies may consider a survey of banks to find which techniques are being used in practice, and if not, why they are not being used. Similarly, a survey of regulatory agencies could also be conducted. Only a few papers have performed a comparative study of regression models and machine-learning techniques, and these found that machine-learning models performed better in predicting bank distress.

Furthermore, papers are overwhelmingly based on U.S. data. However, the regulatory set-up and banking laws in other countries of the world may not be similar to those in the U.S. Accordingly, there is an inherent bias in the literature. In countries where banks are predominantly under public ownership, such as India or China, the conclusions of prior studies may not be relevant. Similarly, the macroeconomic environment and market structure in these countries would be different, and this fact needs to be taken into consideration.

It is not surprising to see that corporate bankruptcy prediction models have been intensively developed and studied. Researchers found each method had its pros and cons. For example, for the recent trend of the application of neural networks, Olson et al. (2012) argued that decision trees can be just as accurate, and provide the transparency and transportability that neural networks are often criticized for. Further, the breadth and depth of the recent financial crisis indicates that these methods must improve if they are to serve as a useful tool for regulators and managers of financial institutions (Carmona et al. 2019). While research on bankruptcy in the banking sector has been well developed, studies on other financial institutions are rather sparse, such as those on fund management, insurance companies, etc.

The majority stream of research predicting bank failures focuses on the determinant factors or leading indicators, such as accounting and financial ratios, macroeconomic data, and regulation. A small set of studies applied a different dataset, but aligned with banking activities in bank failures during the financial crisis. In light of the ongoing FinTech advancement, it would be beneficial to conduct further studies on the different risks faced by large banks, such as trading risk (off-balance-sheet items), or currency risk or crises (Kaminsky and Reinhart 1999). These authors pointed out that not much attention has been paid to the interaction between banking and currency problems, neither in the older literature nor in the new models of self-fulfilling crises, or technological risk, which would be a logical extension of bank early-warning-sign literature.

4. Conclusions

The paper provided a synthesis of post-GFC studies on bank failures. A total of 39 studies published in reputed journals were compared. The emerging trend was towards the use of machine-learning techniques, although currently, regression-model-based studies dominate. The directions for future research have also been identified.

Author Contributions: Conceptualization, M.S.; methodology, L.X.L.; software, S.L.; validation, L.X.L. and S.L.; formal analysis, M.S. and L.X.L.; investigation, L.X.L., S.L.; resources, M.S., L.X.L., S.L.; data curation, L.X.L.; writing—original draft preparation, M.S., L.X.L.; writing—review and editing, S.L.; visualization, L.X.L.; supervision, M.S.; project administration, M.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The papers cited are available online.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Al-Tamimi, Hussein A. Hassan. 2012. The effects of corporate governance on performance and financial distress: The experience of UAE national banks. *Journal of Financial Regulation and Compliance* 20: 169–81.
- Altman, Edward I. 1968. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance* 23: 589–609.
- Altman, Edward I. 1977. Predicting performance in the savings and loan association industry. *Journal of Monetary Economics* 3: 443–66.
- Andersen, Henrik. 2008. *Failure Prediction of Norwegian Banks: A Logit Approach*. Working Paper, No. 2008/2. Oslo: Norges Bank, ISBN 978-82-7553-424-6. Available online: <http://hdl.handle.net/11250/2497784> (accessed on 10 September 2021).
- Arena, Marco. 2008. Bank failures and bank fundamentals: A comparative analysis of Latin America and East Asia during the nineties using bank-level data. *Journal of Banking & Finance* 32: 299–310.
- Atiya, Amir F. 2001. Bankruptcy prediction for credit risk using neural networks: A survey and new results. *IEEE Transactions on Neural Networks* 12: 929–35.
- Avkiran, Necmi K., and Lin Cynthia Cai. 2012. Predicting Bank Financial Distress Prior to Crises. Paper presented at the New Zealand Finance Colloquium, Auckland, New Zealand, February 8–10.
- Barr, Richard S., and Thomas F. Siems. 1997. Bank failure prediction using DEA to measure management quality. In *Interfaces in Computer Science and Operations Research*. Berlin: Springer, pp. 341–65.
- Beaver, William H. 1966. Financial ratios as predictors of failure. *Journal of Accounting Research* 1: 71–111.
- Beaver, William H. 1968. The information content of annual earnings announcements. *Journal of Accounting Research* 1: 67–92.
- Bell, Timothy B. 1997. Neural nets or the logit model? A comparison of each model's ability to predict commercial bank failures. *Intelligent Systems in Accounting, Finance & Management* 6: 249–64.
- Berger, Allen N., Björn Imbierowicz, and Christian Rauch. 2016. The roles of corporate governance in bank failures during the recent financial crisis. *Journal of Money, Credit and Banking* 48: 729–70.
- Betz, Frank, Silviu Oprică, Tuomas A. Peltonen, and Peter Sarlin. 2014. Predicting distress in European banks. *Journal of Banking & Finance* 45: 225–41.
- Beutel, Johannes, Sophia List, and Gregor von Schweinitz. 2019. Does machine learning help us predict banking crises? *Journal of Financial Stability* 45: 100693.
- Bloch, Ernest. 1969. The setting of standards of supervision for Savings and Loan Associations. *Study of the Savings and Loan Industry* 4: 1619–89.
- Bongini, Paola, Małgorzata Iwanicz-Drozdowska, Paweł Smaga, and Bartosz Witkowski. 2018. In search of a measure of banking sector distress: Empirical study of CESEE banking sectors. *Risk Management* 20: 242–57.
- Boyacioglu, Melek Acar, Yakup Kara, and Ömer Kaan Baykan. 2009. 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 Systems with Applications* 36: 3355–66.
- Canbas, Serpil, Altan Cabuk, and Suleyman Bilgin Kilic. 2005. Prediction of commercial bank failure via multivariate statistical analysis of financial structures: The Turkish case. *European Journal of Operational Research* 166: 528–46.
- Carmona, Pedro, Francisco Climent, and Alexandre Momparler. 2019. Predicting failure in the US banking sector: An extreme gradient boosting approach. *International Review of Economics & Finance* 61: 304–23.
- Cebula, Richard J. 2010. Determinants of bank failures in the US revisited. *Applied Economics Letters* 17: 1313–17.
- Chiaromonte, Laura, and Barbara Casu. 2017. Capital and liquidity ratios and financial distress. Evidence from the European banking industry. *The British Accounting Review* 49: 138–61.
- Chiaromonte, Laura, Ettore Croci, and Federica Poli. 2015. Should we trust the Z-score? Evidence from the European Banking Industry. *Global Finance Journal* 28: 111–31.
- Chiaromonte, Laura, Hong Liu, Federica Poli, and Mingming Zhou. 2016. How Accurately Can Z-score Predict Bank Failure? *Financial Markets, Institutions & Instruments* 25: 333–60.
- Cipollini, Andrea, and Franco Fiordelisi. 2012. Economic value, competition and financial distress in the European banking system. *Journal of Banking & Finance* 36: 3101–9.
- Cleary, Sean, and Greg Hebb. 2016. An efficient and functional model for predicting bank distress: In and out of sample evidence. *Journal of Banking & Finance* 64: 101–11.
- Coats, Pamela K., and L. Franklin Fant. 1993. Recognizing financial distress patterns using a neural network tool. *Financial Management* 1: 142–55.
- Cole, Rebel A., and Lawrence J. White. 2012. Déjà vu all over again: The causes of US commercial bank failures this time around. *Journal of Financial Services Research* 42: 5–29.
- Cole, Rebel A., and Qiongbing Wu. 2009. Is Hazard or Probit More Accurate in Predicting Financial Distress? Evidence from US Bank Failures. Available online: <https://mpa.ub.uni-muenchen.de/id/eprint/24688> (accessed on 10 September 2021).
- Constantin, Andreea, Tuomas A. Peltonen, and Peter Sarlin. 2018. Network linkages to predict bank distress. *Journal of Financial Stability* 35: 226–41.

- Cox, Raymond A. K., and Grace W.-Y. Wang. 2014. Predicting the US bank failure: A discriminant analysis. *Economic Analysis and Policy* 44: 202–11.
- de Haan, Jakob, Yi Fang, and Zhongbo Jing. 2020. Does the risk on banks' balance sheets predict banking crises? New evidence for developing countries. *International Review of Economics & Finance* 68: 254–68.
- De Miguel, L. J., E. Revilla, J. M. Rodríguez, and J. M. Cano. 1993. A comparison between statistical and neural network based methods for predicting bankfailures. Paper presented at the IIIth International Workshop on Artificial Intelligence in Economics and Management, Portland, OR, USA, August 28–September 3.
- Demirgüç-Kunt, Asli, and Enrica Detragiache. 1998. The determinants of banking crises in developing and developed countries. *Staff Papers* 45: 81–109.
- DeYoung, Robert, and Gökhan Torna. 2013. Nontraditional banking activities and bank failures during the financial crisis. *Journal of Financial Intermediation* 22: 397–421.
- Ecer, Fatih. 2013. Comparing the bank failure prediction performance of neural networks and support vector machines: The Turkish case. *Economic Research-Ekonomika Istraživanja* 26: 81–98.
- Ekinci, Aykut, and Halil İbrahim Erdal. 2017. Forecasting bank failure: Base learners, ensembles and hybrid ensembles. *Computational Economics* 49: 677–86.
- Ercan, Hakan, and Özgü Evirgen. 2009. Predicting Bank Failures in Turkey by Discrete Choice Models. *METU Studies in Development*. 35: 95–126.
- Erdogan, Birsan Eygi. 2013. Prediction of bankruptcy using support vector machines: An application to bank bankruptcy. *Journal of Statistical Computation and Simulation* 83: 1543–555.
- Federal Reserve Bank of New York (FRBN). 1997. Revision of CAMEL Rating System Effective January 1, 1997, Federal Reserve Bank of New York. Available online: <https://www.newyorkfed.org/banking/circulars/10905.html> (accessed on 14 August 2021).
- Gogas, Periklis, Theophilos Papadimitriou, and Anna Agrapetidou. 2018. Forecasting bank failures and stress testing: A machine learning approach. *International Journal of Forecasting* 34: 440–55.
- Hong, Han, Jing-Zhi Huang, and Deming Wu. 2014. The information content of Basel III liquidity risk measures. *Journal of Financial Stability* 15: 91–111.
- Incze, R. 2019. The Cost of Machine Learning Projects, Medium. Available online: <https://medium.com/cognifeed/the-cost-of-machine-learning-projects-7ca3aea03a5c> (accessed on 13 September 2021).
- Iturriaga, Félix J. López, and Iván Pastor Sanz. 2015. Bankruptcy visualization and prediction using neural networks: A study of U.S. commercial banks. *Expert Systems with Applications* 42: 2857–69, doi:10.1016/j.eswa.2014.11.025.
- Iwanicz-Drozdowska, Małgorzata, and Aneta Ptak-Chmielewska. 2019. Prediction of Banks Distress—Regional Differences and Macroeconomic Conditions. *Acta Universitatis Lodziensis. Folia Oeconomica* 6: 73–57.
- Iwanicz-Drozdowska, Małgorzata, Erkki K. Laitinen, and Arto Suvas. 2018. Paths of glory or paths of shame? An analysis of distress events in European banking. *Bank i Kredyt* 49: 115–44.
- Jin, Justin Yiqiang, Kiridaran Kanagaretnam, and Gerald J. Lobo. 2011. Ability of accounting and audit quality variables to predict bank failure during the financial crisis. *Journal of Banking & Finance* 35: 2811–19.
- Jing, Zhongbo, and Yi Fang. 2018. Predicting US bank failures: A comparison of logit and data mining models. *Journal of Forecasting* 37: 235–56.
- Jordan, Dan J., Douglas Rice, Jacques Sanchez, Christopher Walker, and Donald H. Wort. 2010. Predicting bank failures: Evidence from 2007 to 2010. Available online: <https://ssrn.com/abstract=1652924> (accessed on 10 September 2021).
- Kaminsky, Graciela L., and Carmen M. Reinhart. 1999. The twin crises: The causes of banking and balance-of-payments problems. *American Economic Review* 89: 473–500.
- Kapinos, Pavel, and Oscar A. Mitnik. 2016. A top-down approach to stress-testing banks. *Journal of Financial Services Research* 49: 229–64.
- Kao, Chiang, and Shiang-Tai Liu. 2004. Predicting bank performance with financial forecasts: A case of Taiwan commercial banks. *Journal of Banking & Finance* 28: 2353–68.
- Kerstein, Joseph, and Anthony Kozberg. 2013. Using accounting proxies of proprietary FDIC ratings to predict bank failures and enforcement actions during the recent financial crisis. *Journal of Accounting, Auditing & Finance* 28: 128–51.
- Kneip, Alois, Byeong U. Park, and Léopold Simar. 1996. *A Note on the Convergence of Nonparametric DEA Efficiency Measures*. Ottignies-Louvain-la-Neuve: Université Catholique de Louvain. Center for Operations Research and and Econometrics [CORE].
- Kolari, James W., Dennis Glennon, Hwan Shin, and Michele Caputo. 2002. Predicting large US commercial bank failures. *Journal of Economics and Business* 54: 361–87.
- Kolari, James W., Félix J. López -Iturriaga, Ivan Pastor Sanz. 2019. Predicting European banks stress tests: Survival of the fittest. *Global Finance Journal* 39: 44–57.
- Lam, Kim Fung, and Jane W. Moy. 2002. Combining discriminant methods in solving classification problems in two-group discriminant analysis. *European Journal of Operational Research* 138: 294–301.
- Lane, William R., Stephen W. Looney, and James W. Wansley. 1986. An application of the Cox proportional hazards model to bank failure. *Journal of Banking & Finance* 10: 511–31.
- Lanine, Gleb, and Rudi Vander Vennet. 2006. Failure prediction in the Russian bank sector with logit and trait recognition models. *Expert Systems with Applications* 30: 463–78.

- Le, Hong Hanh, and Jean-Laurent Viviani. 2018. Predicting bank failure: An improvement by implementing a machine-learning approach to classical financial ratios. *Research in International Business and Finance* 44: 16–25.
- López-Iturriaga, Félix J., Óscar López-de-Foronda, and Iván Pastor-Sanz. 2010. Predicting Bankruptcy Using Neural Networks in the Current Financial Crisis: A Study of US Commercial Banks. Available online: <https://ssrn.com/abstract=1716204> (accessed on 10 September 2021).
- Maghyereh, Aktham I., and Basel Awartani. 2014. Bank distress prediction: Empirical evidence from the Gulf Cooperation Council countries. *Research in International Business and Finance* 30: 126–47.
- Männasoo, Kadri, and David G. Mayes. 2009. Explaining bank distress in Eastern European transition economies. *Journal of Banking & Finance* 33: 244–53.
- Martin, Daniel. 1977. Early warning of bank failure: A logit regression approach. *Journal of Banking & Finance* 1: 249–76.
- Mayes, D., and H. Stremmel. 2014. *The Effectiveness of Capital Adequacy Measures in Predicting Bank Distress: SUERF Study 2014/1*. Brussels: Larcier.
- Molina, Carlos A. 2002. Predicting bank failures using a hazard model: The Venezuelan banking crisis. *Emerging Markets Review* 3: 31–50.
- Momparler, Alexandre, Pedro Carmona, and Francisco Climent. 2016. Banking failure prediction: A boosting classification tree approach. *Spanish Journal of Finance and Accounting/Revista Española De Financiación Y Contabilidad* 45: 63–91.
- Ohlson, James A. 1980. Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research* 1: 109–31.
- Olmeda, Ignacio, and Eugenio Fernandez. 1997. Hybrid classifiers for financial multicriteria decision making: The case of bankruptcy prediction. *Computational Economics* 10: 317–35.
- Olson, David L., Dursun Delen, and Yanyan Meng. 2012. Comparative analysis of data mining methods for bankruptcy prediction. *Decision Support Systems* 52: 464–73.
- Ozkan-Gunay, E. Nur, and Mehmed Ozkan. 2007. Prediction of bank failures in emerging financial markets: An ANN approach. *The Journal of Risk Finance* 8: 465–80.
- Poghosyan, Tigran, and Martin Čihák. 2011. Determinants of bank distress in Europe: Evidence from a new data set. *Journal of Financial Services Research* 40: 163–84.
- Ravi, Vadlamani, and Chelimala Pramodh. 2008. Threshold accepting trained principal component neural network and feature subset selection: Application to bankruptcy prediction in banks. *Applied Soft Computing* 8: 1539–48.
- Ravisankar, Pediredla, and Vadlamani Ravi. 2010. Financial distress prediction in banks using Group Method of Data Handling neural network, counter propagation neural network and fuzzy ARTMAP. *Knowledge-Based Systems* 23: 823–31.
- Schaeck, Klaus. 2008. Bank liability structure, FDIC loss, and time to failure: A quantile regression approach. *Journal of Financial Services Research* 33: 163–79.
- Serrano-Cinca, Carlos, and Begoña Gutiérrez-Nieto. 2013. Partial least square discriminant analysis for bankruptcy prediction. *Decision Support Systems* 54: 1245–55.
- Shrivastava, Santosh, P. Mary Jeyanthi, and Sarbjit Singh. 2020. Failure prediction of Indian Banks using SMOTE, Lasso regression, bagging and boosting. *Cogent Economics & Finance* 8: 1729569.
- Sinkey, Joseph F., Jr. 1975. A multivariate statistical analysis of the characteristics of problem banks. *The Journal of Finance* 30: 21–36.
- Swicegood, Philip, and Jeffrey A. Clark. 2001. Off-site monitoring systems for predicting bank underperformance: A comparison of neural networks, discriminant analysis, and professional human judgment. *Intelligent Systems in Accounting, Finance & Management* 10: 169–86.
- Tam, Kar Yan. 1991. Neural network models and the prediction of bank bankruptcy. *Omega* 19: 429–45.
- Tanaka, Katsuyuki, Takuji Kinkyō, and Shigeyuki Hamori. 2016. Random forests-based early warning system for bank failures. *Economics Letters* 148: 118–21.
- Tatom, John, and Reza Houston. 2011. *Predicting Failure in the Commercial Banking Industry*. Networks Financial Institute Working Paper No. 2011-WP-27. Networks Financial Institute Working Paper. <http://dx.doi.org/10.2139/ssrn.1969091>. Available online: <https://ssrn.com/abstract=1969091> (accessed on 10 September 2021).
- Wheelock, David C., and Paul W. Wilson. 2000. Why do banks disappear? The determinants of US bank failures and acquisitions. *Review of Economics and Statistics* 82: 127–38.
- Wong, Jim, Tak-Chuen Wong, and Phyllis Leung. 2010. Predicting banking distress in the EMEAP economies. *Journal of Financial Stability* 6: 169–79, doi:10.1016/j.jfs.2010.01.001.
- Zhao, Huimin, Atish P. Sinha, and Wei Ge. 2009. Effects of feature construction on classification performance: An empirical study in bank failure prediction. *Expert Systems with Applications* 36: 2633–44.