ECONOMY AND FINANCE • VOL. 11. ISS. 3 • SEPT 2024 • DOI: 10.33908/EF.2024.3.2

COMPARISON OF BOOSTING AND RANDOM FOREST MODELS IN FORECASTING BANK FAILURES

Revisiting the 2008 Financial Crisis from a Supervisory Perspective

Safa Sen[1]

ABSTRACT

This research paper delivers an exhaustive analysis of predictive models for bank failures, a subject of paramount importance for economic stability. Using a dataset from the Federal Deposit Insurance Corporation (FDIC), the study examines 950 banking institutions, including 60 that succumbed to the 2008 financial crisis. The paper employs binary classification analysis using 26 CAMEL ratios and compares boosting algorithms with the Random Forest model family. In classifying non-failed banks, Random Forest variations notably outperform boosting algorithms, achieving a 97% accuracy rate in correctly classified instances, with the Regularized Random Forest model showing exceptional precision with a rate of 0.988. In the context of predicting failed banks, the Random Forest models, particularly the regularized variant, demonstrate a strong capability for accurately identifying true failures. These findings corroborate the efficacy of Random Forest models in predicting bank failures precisely and reliably, highlighting their critical role in reducing false positives and negatives, which is essential for robust forecasting in the banking sector.

JEL codes: C45, C53, G12, G17

Keywords: machine learning models, banking failure, off-site monitoring, CSForest, XGBoost

1 INTRODUCTION

The study of bank failures is a complex and multifaceted area of research that plays a crucial role in maintaining the stability and growth of economies. As financial intermediaries, banks are central to the economic framework, facilitating the flow of funds from savers to borrowers and ensuring liquidity in financial markets. However, the consequences of bank failures are profound, affecting not

1 *Safa Sen*, PhD student, University of Miskolc, Faculty of Economics. E-mail: sen.safa@student.uni-miskolc.hu.

only the institutions involved but also the broader economic system. This paper aims to explore the intricacies of bank failures, their impacts, and the evolution of predictive models used for forecasting such events.

Historically, bank failures have been a significant source of economic distress, with far-reaching consequences that extend beyond the immediate financial sector. The collapse of a banking institution can disrupt the flow of capital, leading to economic repercussions that include loss of consumer and business confidence, reduced spending, and in severe cases, economic recession. The 2008 financial crisis serves as a stark example of the systemic risks posed by failing banks and their global impact.

Research in predicting bank failures has evolved significantly over the years. Earlier methods like linear and quadratic discriminant analysis, factor analysis, and logistic regression, as employed by pioneers such as Meyer and Pifer (1970), Sinkey (1975), and Martin (1977), were foundational in this field. These traditional models relied heavily on financial ratios and indicators to gauge potential risks. However, as financial markets became more intricate, the limitations of these models in their predictive capacity became apparent, necessitating the development of more sophisticated approaches.

Recent advancements in predictive modeling have seen the integration of artificial neural networks, operations research, hybrid intelligent methods, fuzzy logic, and support vector machines. For instance, Quek, Zhou, and Lee (2009) proposed a novel fuzzy neural network for bank failure prediction, demonstrating the use of computational techniques coupled with financial data reconstruction to enhance prediction accuracy. Similarly, Jing and Fang (2018) compared the logit model and data mining models in predicting U.S. bank failures, revealing the superior performance of data mining models in certain scenarios.

The emergence of machine learning and artificial intelligence has significantly enhanced the capabilities of predictive models. These advanced methods can analyze extensive datasets, identify complex patterns, and predict potential failures with greater precision, offering nuanced insights into the financial health of banking institutions. For example, Tung, Quek, and Cheng (2004) introduced the GenSoFNN-CRI(S) network, a neural-fuzzy based early warning system for predicting bank failures, highlighting its effectiveness in identifying traits of financial distress.

The need for precise prediction and prevention of bank failures has been further underscored by the 2008 financial crisis. This event demonstrated the interconnectedness of the global financial system, and the extensive impact of the collapse of key financial institutions on the worldwide economy. It led to a surge in research into predictive models and risk management strategies, emphasizing the ongoing need for methodological innovation to adapt to the evolving nature of financial risks.

In response to these developments, this paper applies boosting and random forest algorithms to reanalyze and predict the 2008 financial crisis, with a specific focus on national banks in the U.S. The significance of this research lies in its contribution to the extensive body of literature on effective models for forecasting bank failures.

While the determination of the 'best' model may vary based on initial assumptions and goals, the ability to accurately forecast failed banks is essential for economic stability and growth.

Early identification of banks potentially facing failure enables preemptive governmental support to avert their collapse. Conversely, failing to intervene in cases where a bank is predicted to fail by these models increases the risk to the stability of financial institutions significantly, as the lack of support for a potentially failing bank can have severe repercussions on the stability of the entire financial system.

2 MOTIVATION OF THE STUDY

The core emphasis of this study lies in evaluating the efficacy of leading machine learning models in predicting bank failures, instead of an in-depth examination of the underlying causes at a micro level. This approach inherently prioritizes the exploration of advanced analytical techniques and their capacity to forecast financial instabilities over dissecting the specific factors that precipitate such failures.

2.1 Background of the 2008 Financial Crisis

The 2008 Global Financial Crisis, a pivotal episode in the annals of global economic history, was triggered by a combination of intricate financial dynamics and regulatory lapses. The crisis was rooted in the deterioration of the U.S. subprime mortgage sector beginning in the summer of 2007, a culmination of trends that emerged following the 2001-2002 credit crisis. This period saw an unparalleled expansion of credit, a sharp increase in housing prices, and a substantial build-up of leverage within the financial system. Concurrently, rapid advancements in financial innovations, especially in securitization, significantly boosted the financial system's credit creation capabilities but also surpassed its risk management capacity (Mian & Sufi, 2009; Shiller, 2008).

Unlike previous credit crises in the U.S., the Global Financial Crisis (GFC) had profound worldwide repercussions. Losses originating from the U.S. subprime mortgage sector swiftly permeated the international financial ecosystem. Banks grappled with severe losses and liquidity constraints, leading to widespread uncertainty about credit asset valuations and a drastic reduction in interbank lending. To counter these effects, central banks globally launched liquidity support mechanisms and recapitalized failing banks to rejuvenate lending (Bernanke, 2010; Gorton, 2009).

Noteworthy events during the crisis included the nationalization of the United Kingdom's Northern Rock in February 2008 as a result of the country's first bank run in over a century, and the acquisition of the U.S. investment bank Bear Stearns by J.P. Morgan Chase in March 2008 under the guidance of the U.S. Treasury and the Federal Reserve (Sorkin, 2009). Additionally, the crisis brought a halt to asset-backed commercial paper and repurchase agreement markets, resulting in the collapse or closure of numerous hedge funds and the dissolution of many Special Investment Vehicles (SIVs) and conduits, with global credit losses eventually surpassing USD 1 trillion (Financial Crisis Inquiry Commission, 2011).

A critical element of the crisis was the surge in housing demand and mortgage financing, partly driven by the low interest rate environment of the early 2000s. This demand fueled hikes in housing prices and attracted investors, including institutional ones, to the high yields of subprime mortgages. These mortgages, often having rates much higher than those for prime borrowers, were increasingly sought for securitization. In this process, securitizers pooled below-investment-grade assets, divided cash flows based on model-driven certainty, and transformed the safest cash flows into investment-grade securities (Kothari, 2008).

Many subprime mortgages were initially structured with low teaser rates, followed by considerably higher rates in subsequent years. Borrowers, ranging from residents to speculators, often defaulted, especially when they were unable to refinance after the teaser period. The Originator-to-Distributor (OTD) model implied that losses on these mortgages were absorbed by investors rather than the originating banks, diminishing banks' incentive for thorough due diligence (Acharya & Richardson, 2009).

The crisis was exacerbated by a spike in delinquencies on adjustable-rate subprime mortgages. By 2007, the rate of serious delinquencies had risen sharply, leading to numerous ratings downgrades for subprime mortgage-backed securities. Contributing factors included the poor credit quality of borrowers, a significant number of first-time homebuyers making no down payments, and the prevalence of teaser rates (Foote, Gerardi, & Willen, 2008).

Riskier mortgage products, such as NINJA loans and liar loans also proliferated, fostering fraudulent practices due to lenient lending standards. The compensation model for mortgage brokers, which prioritized loan volume over long-term performance, further intensified these problems (Rajan, 2010).

Banks transferred securitization assets off their balance sheets to SIVs, repackaging cash flows from existing assets into tranches with varying credit ratings. This approach was theoretically designed to distribute risk more widely, but in reality, it exposed significant vulnerabilities in the financial system's infrastructure (Ashcraft & Schuermann, 2008).

In response, central banks around the world, notably the U.S. Federal Reserve, introduced innovative liquidity measures to stabilize the financial markets. These measures included establishing long-term lending facilities, opening the discount window to investment banks, and supporting commercial paper purchases. Key U.S. government interventions encompassed the Term Auction Facility, the Primary Dealer Credit Facility, and the Economic Stimulus Act of February 2008 (Blinder & Zandi, 2010).

In conclusion, the 2008 Global Financial Crisis, a defining moment in modern economic history, is a stark reminder of the intricate interdependencies and inherent vulnerabilities within the global financial system. Initiated by the downturn in the U.S. subprime mortgage sector and further exacerbated by a confluence of complex financial dynamics and regulatory shortcomings, this crisis precipitated widespread economic upheaval that transcended national borders. It laid bare the systemic risks inherent in the financial innovations of the time, such as securitization, and a substantial build-up of leverage, which collectively outstripped the risk management capacities of financial institutions and regulatory bodies.

The crisis not only highlighted the flaws in financial models and practices but also underscored the vital importance of sound risk management, vigilant regulatory oversight, and the need for ethical lending and borrowing practices. It illustrated how rapid credit expansion, combined with a surge in housing demand fueled by low-interest rates, can lead to unsustainable asset price inflation and subsequent market corrections. The widespread impact of the crisis, from the nationalization of major banks to the collapse of key financial markets, demonstrated the profound consequences of such systemic failures.

Ultimately, the 2007-2009 Global Financial Crisis serves as a compelling lesson in economic and financial governance, emphasizing the need for continuous vigilance, adaptability, and cooperation among various global financial entities. It highlights the importance of learning from past mistakes to protect financial systems against future crises, ensuring a more resilient and stable economic environment for generations to come.

2.2 Literature review

Berger and Bouwman (2012): This study extensively examines how capital affects bank performance across different economic conditions in the U.S. over a quarter of a century. Using advanced statistical methods like logit survival and OLS regression models, it concludes that capital significantly enhances the survival and market share of small banks consistently across various economic climates. Importantly, for medium-sized and large banks, capital was found to be particularly crucial during banking crises, indicating its vital role in bank stability during economic downturns.

Trussel and Johnson (2012): This research paper investigates financial indicators linked to U.S. bank failures using logistic regression. The study focuses on six key financial indicators and creates a composite measure to predict bank failure. It finds that an increase in Tier 1 capital relative to total assets and an increase in return on assets are the most influential factors in reducing the risk of bank failure, providing crucial insights into financial metrics that can predict bank stability.

Lu and Whidbee (2013): This study explores the factors affecting bank failure during the late 2000s financial crisis, including charter type, holding company structure, and measures of bank fragility. Analyzing all commercial banks in the U.S. using logit regressions, the study identifies critical factors such as bailout funds, capital ratios, and liquidity levels that determined the survival or failure of banks during this turbulent period.

DeYoung and Torna (2013): This study examines the impact of income from nontraditional banking activities on U.S. commercial bank failures during the financial crisis. Using a multi-period logit model, the authors find that fee-based nontraditional activities like securities brokerage and insurance sales reduce the probability of bank failure, while asset-based activities such as venture capital increase it, offering insights into how different business models impact bank stability.

Chiaramonte, Liu, Poli, and Zhou (2016): The researchers assess the predictive power of Z-scores in forecasting bank failures between 2004 and 2012. They find that Z-scores, which combine profitability, leverage, and earnings variability, can predict 76% of bank failures and maintain stable predictive power over a three-year period, indicating their reliability in predicting bank defaults.

Cleary and Hebb (2016): This study utilizes discriminant analysis to investigate the failure of 132 U.S. banks from 2002 to 2009. The authors successfully differentiate between banks that failed and those that didn't, with a high prediction efficiency of 92%. They extend their analysis to predict bank failures in 2010-2011, maintaining high prediction accuracy.

Lu and Whidbee (2016): Examining 6,236 U.S. commercial banks during the financial crisis, this study focuses on banks targeted for intervention via bailouts or failure. The authors find many similarities in the characteristics of bailed-out banks and those that failed, suggesting common risk factors and vulnerabilities that influenced their survival during the crisis.

Serrano-Cinca, Fuertes-Callén, Gutiérrez-Nieto, and Cuellar-Fernández (2014): This paper investigates the bankruptcy of U.S. banks since 2009, proposing several hypotheses about the causes of failure. Using structural equation modeling, it concludes that failed banks had higher loan growth, a higher concentration on real estate loans, higher risk ratios, higher turnover, and lower margins compared to solvent banks, establishing a significant relationship between these factors and bank failures.

Le and Viviani (2017): The study conducts a comprehensive analysis of bank failures by combining traditional statistical techniques and advanced machine learning methods across a dataset of 3,000 U.S. banks. It employs Discriminant Analysis, Logistic Regression, Artificial Neural Networks, Support Vector Machines, and k-Nearest Neighbors, concluding that machine learning methods, particularly ANNs and k-NN, are more effective in predicting bank failures than traditional methods.

Gogas et al. (2018): Focusing on machine learning models, this study uses a dataset of 1,443 U.S. banks, including 481 that failed during 2007-2013. The authors employ a two-step feature selection process and a Support Vector Machine model, achieving an impressive 99.22% forecasting accuracy and demonstrating the potential of machine learning techniques for surpassing traditional models in predictive accuracy.

Carmona et al. (2019): This research paper applies the Extreme Gradient Boosting method to predict bank failures, analyzing 157 U.S. national commercial banks from 2001 to 2015. It assesses 30 financial ratios, revealing that specific ratios, particularly those relating to retained earnings and risk-based capital, are closely associated with increased bank failure likelihood.

Manthoulis et al. (2020): Employing both statistical and machine learning methods, this study predicts bank failures by analyzing 60,000 observations from U.S. banks. It highlights the effectiveness of diversification variables and the superiority of ordinal classification models over binary models in predicting bank failures.

Momparler et al. (2020): Using fuzzy-set Qualitative Comparative Analysis, this study examines 157 U.S. national commercial banks from 2001 to 2015. It identifies banks with high non-performing loans and low risk coverage and capitalization as being at a higher risk of failure, emphasizing the importance of asset quality and capital adequacy.

2.3 Study sample and ratios used in this paper

In my study, I used a meticulously constructed dataset from the Federal Deposit Insurance Corporation (FDIC), focusing on bank failure records. This dataset was essential for my binary classification analysis, which aimed to categorize banks into two distinct groups: those that failed and those that remained operational. The dataset encompassed an extensive range of 950 banking institutions, including both main banks and their subsidiary branches. Notably, during the tumultuous period of the 2008 financial crisis, 60 of these banks failed.

I chose 2009 as the reference year for selecting 'healthy' or non-failed banks, considering it a pivotal moment in financial history. This year represented the zenith of financial instability, a critical juncture where the banking sector faced its most severe challenges. The rationale behind this selection was that if a predictive model could accurately differentiate between failed and non-failed banks during this peak period of financial turbulence, then the model could be considered robust and effective.

For the analytical part of my study, I employed 26 CAMEL (capital adequacy, asset quality, management quality, earnings, and liquidity) ratios as primary variables. These ratios, fed into the predictive model, provided a total of 24,700 data points. The research was highly cost-effective, as the source of data was publicly available and open source, eliminating financial constraints related to data acquisition. Moreover, data processing and analysis were conducted using WEKA, an established open-source software platform known for its capabilities in data mining. These choices were in line with my commitment to ensuring accessibility and reproducibility in academic research.

3 METHODOLOGY AND MODELS USED IN THE STUDY

3.1 XGBOOST

XGBoost, or eXtreme Gradient Boosting, is highly regarded in machine learning, especially for financial risk analysis and predicting banking crises. It is an advanced gradient boosted decision tree method, recognized for handling large and complex financial datasets efficiently and flexibly (Chen & Guestrin, 2016). In credit risk modeling and bank failure prediction, XGBoost excels in combining weak predictive models, primarily decision trees, into a robust predictive model. This is crucial in financial analysis due to the diverse range of variables, from individual credit histories to broad economic indicators (Friedman, 2001). An important strength of XGBoost is its ability to manage missing values and various data types, which is vital for financial datasets often characterized by incomplete information (Chen & He, 2015).

Its built-in regularization methods also prevent overfitting, a common problem in complex financial modeling (Natekin & Knoll, 2013). XGBoost's feature importance scores are particularly beneficial in pinpointing key factors in credit risk and bank failure likelihood, helping in risk management and regulatory processes by identifying critical predictors among financial ratios, governance indicators, and economic factors (Huang, Chen, & Wang, 2019). Empirical research demonstrates XGBoost's superiority over traditional models like logistic regression and random forests in financial contexts. Its accuracy in detecting credit defaults and banking distress makes it a top choice for financial institutions and regulators (Malhotra & Malhotra, 2021).

3.2 Generalized Linear Model Boosting

Generalized Linear Model (GLM) Boosting, which combines GLM concepts with boosting algorithms, has proved effective in financial risk assessment and bank failure prediction. This method enhances predictive accuracy while maintaining the interpretability of traditional GLMs, which is important in financial modeling. GLM Boosting excels in complex financial challenges like credit risk modeling, bankruptcy prediction, and bank failure detection. It adeptly handles diverse data types and models intricate predictor-outcome relationships, including non-linear ones, suitable for financial applications (Hastie, Tibshirani, & Friedman, 2009). The method iteratively fits a GLM to data, refining the model by correcting previous iteration residuals. This results in a comprehensive model capturing complex financial data patterns that traditional linear models might overlook (Bühlmann & Hothorn, 2007). GLM Boosting's regularization feature, controlling iterations and predictor influence, is key in preventing overfitting, which is crucial for models to generalize effectively to new data (Schapire & Freund, 2012). Another significant aspect of GLM Boosting is its ability to perform variable selection. It identifies the most impactful predictors, eliminating irrelevant or redundant variables, thus enhancing model performance and interpretability, which are vital for financial decision-making (Tutz & Binder, 2006). Empirical studies in finance highlight GLM Boosting's effectiveness in various applications, including detailed risk factor insights and accurate adverse financial event predictions. This is particularly valuable for credit scoring and bank failure prediction, where understanding each risk factor's influence is essential for precise decision-making (Zou & Hastie, 2005).

3.3 LogitBoost

In the realm of machine learning for financial applications, such as credit scoring and customer segmentation, the integration of logistic regression into boosting algorithms, particularly the approach developed by Friedman, Hastie, & Tibshirani (2000), has proven to be highly effective. This method excels in binary classification tasks common in the financial sector, like distinguishing potential defaulters or sorting customers into different risk categories. The strength of this approach is its iterative refinement of logistic regression models, allowing for better navigation through complex financial data relationships, which is indispensable for precise risk assessment (Landwehr, Hall, & Frank, 2005). This aspect is particularly vital in finance where accuracy is key to avoiding costly mistakes. Moreover, clarity in how features influence predictions make this approach highly interpretable, an essential quality in financial settings. Understanding the reasoning behind decisions such as loan approvals or rejections is important for both regulators and stakeholders (Hastie, Tibshirani, & Friedman, 2009). In practice, this method has shown considerable success in financial applications, outperforming other classifiers in tasks like fraud detection and creditworthiness assessment. Its capability to manage complex and large datasets effectively is a significant factor in its success (Bühlmann & Hothorn, 2007).

3.4 Random Forest

The Random Forest algorithm, developed by Breiman (2001), is a sophisticated ensemble learning method widely used in financial risk modeling and bank failure prediction. It combines multiple decision trees to form a 'forest', with each tree built on a random subset of data and variables. This method enhances predictive accuracy and robustness against overfitting, a prevalent challenge in financial modeling. In financial risk analysis, Random Forest excels due to its capability to handle large datasets with numerous variables, typical in financial applications. It efficiently processes various data types, including numerical and categorical variables, ideal for analyzing complex financial risks (Liaw & Wiener, 2002). A key strength in finance is its feature selection ability. By averaging over many trees, Random Forest identifies influential predictors from a large pool of financial indicators, which is critical in credit scoring and bankruptcy prediction for pinpointing significant risk factors (Díaz-Uriarte & De Andres, 2006). Random Forest's bagging mechanism and random feature selection for each tree reduce variance and bias, creating a model that is accurate and generalizes well to new data. These characteristics are vital for predicting financial crises or bank failures under uncertain conditions (Cutler et al., 2007). In practice, Random

Forest has been more effective than traditional methods and other machine learning techniques in predicting credit defaults and bank distress. Its robust performance and ability to capture complex non-linear financial relationships have been demonstrated in various economic scenarios (Hastie, Tibshirani, & Friedman, 2009).

3.5 Regularized Random Forest

Regularized Random Forest (RRF), enhancing the traditional Random Forest algorithm with regularization techniques, is particularly effective in high-dimensional data settings like financial risk analysis and bank failure prediction. Developed by Deng (2013), RRF combines random forest's robustness and feature selection with regularization to prevent overfitting and improve accuracy. The method is adept at handling numerous variables, including complex financial indicators, refining models by focusing on the most relevant predictors (Meinshausen, 2007). It addresses the 'curse of dimensionality' in large feature sets by using regularization to guide tree-building, thus reducing variance, and enhancing model generalization (Biau & Scornet, 2016). Key to RRF's success in finance is its feature selection mechanism, distinguishing vital variables from less significant ones, which is crucial for precise risk assessment and decision-making (Liaw & Wiener, 2002). It maintains random forest's advantages, like handling various data types and providing variable importance measures, which are valuable in the often-incomplete datasets of finance (Breiman, 2001). Empirically, RRF has shown effectiveness in scenarios like credit scoring, offering more accurate predictions for credit defaults or bank distress, promoting risk management and regulatory compliance (Strobl et al., 2009). Overall, RRF is a significant advancement in financial risk prediction owing to its effectiveness in processing high-dimensional data with enhanced feature selection and regularization.

3.6 Cost-Sensitive Forest

Cost-Sensitive Forest is a sophisticated adaptation of ensemble learning techniques like Random Forest, tailored for financial risk assessment and bank failure prediction where prediction errors have asymmetric costs. Developed by Elkan (2001), it enhances traditional models by accounting for the different costs of false positives and false negatives in its training process, making it more aligned with real-world financial decision-making (Ling & Sheng, 2008).

This model is particularly valuable in finance, where the cost of missing a bank failure (false negative) can be much higher than incorrectly predicting one (false positive). Cost-Sensitive Forests focus on minimizing more costly errors, providing a nuanced tool for regulatory and economic analysis (Turney, 2000).

Its adaptability to changing cost scenarios in evolving financial markets is a key advantage, allowing recalibration to maintain relevance under different economic conditions (Zadrozny & Elkan, 2001). Empirical studies have shown that Cost-Sensitive Forests offer improved prediction accuracy over traditional models, aiding financial institutions and regulators in developing effective risk management strategies (Khalilia, Chakraborty, & Popescu, 2011).

3.7 Machine Learning

Machine learning encompasses a range of computational methods that enable systems to learn from data and make decisions or predictions. At its core, machine learning involves algorithms that can process large datasets, recognize patterns, and make predictions or decisions based on data inputs without being explicitly programmed for specific tasks (Alpaydin, 2020). The methods are broadly classified into supervised learning, unsupervised learning, and reinforcement learning, each suited to different types of problems and data structures (Goodfellow, Bengio, & Courville, 2016).

3.8 Machine Learning Performance Metrics

Accuracy

Accuracy is the simplest and most intuitive performance metric. It is the ratio of correctly predicted instances to the total instances in the dataset. While easy to understand, accuracy can be misleading, especially in imbalanced datasets where one class significantly outnumbers the other (Provost et al., 1998).

Precision (Positive Predictive Value)

Precision measures the proportion of positive identifications that were actually correct. It is particularly important in scenarios where false positives are more consequential. High precision indicates a low rate of false positives (Powers, 2011).

Recall (Sensitivity, True Positive Rate)

Recall measures the proportion of actual positives that were correctly identified. It is crucial in contexts where missing out on positive instances (false negatives) is costly. High recall means that most of the positive instances are correctly captured (Powers, 2011).

Precision-Recall Curve

The precision-recall curve (PRC) illustrates precision values in relation to sensitivity (recall) values. Just as the ROC curve, the PRC curve offers a comprehensive assessment of a model's performance across the board. The AUC score associated with the PRC curve, referred to as AUC (PRC), is also a valuable metric for comparing multiple classifiers. According to the study of Saito and Rehmsmeier (2015), the precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. Hence, I used PRC in my study as well.

Confusion Matrix

A confusion matrix is a table used to describe the performance of a classification model. It shows true positives, false positives, true negatives, and false negatives. This matrix provides a clear view of the performance of the model and is especially useful in multi-class classification problems (Stehman, 1997).

K-Fold Cross Validation

The research methodology adopted in this study closely follows the approach of Carmona et al. (2019), specifically employing a 10-fold cross-validation technique. This method is a cornerstone in statistical analysis, especially in validating models in fields like machine learning and data science. The 10-fold cross-validation process begins with the division of the entire dataset into ten equally sized segments, or 'folds.' Division is crucial for ensuring that each segment of the data receives equal representation during the validation process, thereby reducing bias and improving the reliability of model evaluation. In each of the ten iterations of the process, one of these folds is designated as the test set, while the remaining nine folds are combined to form the training set. Rotation is methodical, i.e., each fold serves as the test set exactly once, ensuring that every data point is used for both training and testing at different stages. Systematic rotation is a key strength of this method, as it allows for a more robust and thorough evaluation of the model's performance. By employing this technique, the model is exposed to a wide variety of data scenarios. The variation in training and test sets across iterations enables the model to demonstrate its ability to learn and generalize from different subsets of the data. This is particularly valuable in assessing the model's performance in real-world scenarios, where data can vary significantly.

3.8.1 Result of the Boosting Algorithms

GLMBoost

Metric	Value	Percentage	
Correctly Classified Instances	915	96.32%	
Incorrectly Classified Instances	35	3.68%	
Class	**Precision**	**Recall**	**PRC Area**
Non-Failed	0.97	0.991	0.989
Failed	0.805	0.55	0.777
Confusion Matrix	**Non-Failed Predicted**	**Failed Predicted**	
Non-Failed Actual	882	8	
Failed Actual	27	33	

XGBoost

Metric	Value	Percentage	
Correctly Classified Instances	918	96.63%	
Incorrectly Classified Instances	32	3.37%	
Class	**Precision**	**Recall**	**PRC Area**
Non-Failed	0.981	0.983	0.990
Failed	0.741	0.717	0.703
Confusion Matrix	**Non-Failed Predicted**	**Failed Predicted**	
Non-Failed Actual	878	12	
Failed Actual	23	37	

LogitBoost

Metric	Value	Percentage	
Correctly Classified Instances	915	96.32%	
Incorrectly Classified Instances	35	3.68%	
Class	**Precision**	**Recall**	**PRC Area**
Non-Failed	0.974	0.987	0.998
Failed	0.755	0.617	0.782
Confusion Matrix	**Non-Failed Predicted**	**Failed Predicted**	
Non-Failed Actual	875	15	
Failed Actual	17	43	

3.8.2 Result of the Random Forest Algorithms

General Random Forest

Metric	Value	Percentage	
Correctly Classified Instances	925	97.37%	
Incorrectly Classified Instances	25	2.63%	
Class	**Precision**	**Recall**	**PRC Area**
Non-Failed	0.982	0.99	0.989
Failed	0.83	0.733	0.824
Confusion Matrix	**Non-Failed Predicted**	**Failed Predicted**	
Non-Failed Actual	881	9	
Failed Actual	16	44	

Cost Sensitive Random Forest

Metric	Value	Percentage	
Correctly Classified Instances	925	97.37%	
Incorrectly Classified Instances	25	2.63%	
Class	**Precision**	**Recall**	**PRC Area**
Non-Failed	0.984	0.988	0.999
Failed	0.807	0.767	0.760
Confusion Matrix	**Non-Failed Predicted**	**Failed Predicted**	
Non-Failed Actual	879	11	
Failed Actual	14	46	

Regularized Random Forest

Metric	Value	Percentage	
Correctly Classified Instances	925	97.47%	
Incorrectly Classified Instances	25	2.53%	
Class	**Precision**	**Recall**	**PRC Area**
Non-Failed	0.988	0.985	0.997
Failed	0.79	0.817	0.755
Confusion Matrix	**Non-Failed Predicted**	**Failed Predicted**	
Non-Failed Actual	877	13	
Failed Actual	11	49	

3.8.3 Comparison of Results

Performance Evaluation of Machine Learning Algorithms on Non-Failed Bank Predictions

Algorithm	Metric	Value	Percentage	Class	Precision	PRC Area
Regularized Random Forest	Correctly Classified Instances	925	97%	Non-Failed	0.988	0.997
Cost Sensitive Random Forest	Correctly Classified Instances	925	97%	Non-Failed	0.984	0.999
General Random Forest	Correctly Classified Instances	925	97%	Non-Failed	0.982	0.999
XGBoost	Correctly Classified Instances	918	97%	Non-Failed	0.981	0.999
LogitBoost	Correctly Classified Instances	915	96%	Non-Failed	0.974	0.998
GLMBoost	Correctly Classified Instances	915	96%	Non-Failed	0.97	0.989

In an in-depth analysis of algorithm performance for classifying non-failed banks, the results indicate a marginal superiority of Random Forest variations over Boosting methods. Specifically, the Regularized Random Forest, Cost Sensitive Random Forest, and General Random Forest algorithms all exhibit an impressive 97% rate of Correctly Classified Instances, suggesting a solid ability to discriminate between failing and stable banks. The Boosting methods, GLMBoost and LogitBoost, are slightly less effective, but still demonstrate a commendable performance with a 96% accuracy rate.

Focusing on precision, which is critical in ensuring that banks classified as stable are indeed stable, Regularized Random Forest leads with a precision rate of 0.988 for the non-failed class. This high precision indicates a lower likelihood of false positives, which is indispensable for avoiding misclassification of at-risk banks as stable. The Cost Sensitive and General Random Forest algorithms also perform well, indicating that the Random Forest framework is consistently reliable in identifying non-fail instances. The Boosting methods, particularly GLMBoost, show a somewhat reduced precision, which may imply a higher rate of false positives among non-fail predictions.

The Precision-Recall Curve (PRC) Area serves as a comprehensive metric that encapsulates both the precision of the model and its recall, or sensitivity.

A higher PRC Area is indicative of a model's effectiveness in maintaining high precision while also capturing a large proportion of actual positive instances (non-failed banks, in this context). The Random Forest variations stand out with the highest PRC Areas, implying that these models offer the best balance in correctly predicting non-failed banks while minimizing both false positives and false negatives. The Boosting methods have slightly lower PRC Areas, but they still perform reliably.

In summary, Random Forest variants outperform Boosting methods in classifying non-failed banks, with a higher precision and better balance between precision and recall. Regularized Random Forest is particularly notable for its precision, while General and Cost Sensitive Random Forest models provide the best compromise between avoiding false positives and negatives. Boosting methods, despite being slightly outperformed, could still be useful, especially with parameter tuning, and GLMBoost may be preferred for its interpretability. The choice of model should be aligned with the specific needs and risk profile of the application.

Performance Evaluation of Machine Learning Algorithms on Failed Bank Predictions

Algorithm	Metric	Value	Percentage	Osztály	Class	PRC Area
Regularized Random Forest	Incorrectly Classified Instances	25	3%	Failed	0.817	0.755
Cost Sensitive Random Forest	Incorrectly Classified Instances	25	3%	Failed	0.767	0.76
General Random Forest	Incorrectly Classified Instances	25	3%	Failed	0.733	0.824
XGBoost	Incorrectly Classified Instances	32	3%	Failed	0.717	0.703
LogitBoost	Incorrectly Classified Instances	35	4%	Failed	0.617	0.782
GLMBoost	Incorrectly Classified Instances	35	4%	Failed	0.55	0.777

In a detailed evaluation of algorithms for predicting bank failures, all models exhibit high accuracy, with a low percentage of incorrectly classified instances between 3% and 4%. This consistent performance across algorithms demonstrates their effectiveness in identifying potential bank failures.

When examining recall, which is critical for correctly detecting failed banks, Regularized Random Forest stands out with the highest recall of 0.817, indicating that it is the most effective in identifying true failures. Cost Sensitive Random Forest and General Random Forest follow closely, with slightly lower recall values of 0.767 and 0.733, respectively. XGBoost, LogitBoost, and GLMBoost show lower recall, implying a greater likelihood of incorrectly classifying failing banks as stable.

The Precision-Recall Curve (PRC) Area, a key metric in determining the balance between precision and recall, reveals further insights. General Random Forest ranks first with the highest PRC Area of 0.824, suggesting an optimal balance in predicting bank failures. Regularized and Cost Sensitive Random Forest models also demonstrate a robust performance, with PRC Areas of 0.755 and 0.76, respectively, indicating their effectiveness in maintaining the balance. XGBoost and GLMBoost models have similar PRC Areas, while LogitBoost occupies a middle ground.

Overall, this analysis highlights the strengths of Random Forest variations in accurately identifying bank failures, with General Random Forest offering the best balance between precision and recall. While Boosting methods are slightly less effective in terms of recall, their performance remains robust, making them viable options depending on the specific requirements and contexts of their application.

In summary, the Regularized Random Forest model excels in sensitivity to failing banks, making it less likely to miss identifying a bank at risk of failure. Its high recall makes it particularly useful in situations where the cost of missing a failing bank is substantial. The General Random Forest algorithm stands out for its optimal balance between precision and recall, also indicated by its highest PRC Area, making it adept at identifying failing banks while minimizing false alarms. Boosting methods such as GLMBoost and LogitBoost, while showing a lower recall and PRC Area, could still be relevant in certain scenarios, particularly if fine-tuned. The choice of the most suitable model should take the specific implications of false negatives and false positives into account, with a preference for models like Regularized and General Random Forest in such high-stakes environments as banking where accurately detecting bank failures is of crucial importance.

4 CONCLUSION

In the realm of financial stability, banks play a crucial role as intermediaries in economic systems, making the study of bank failures a topic of paramount importance in the financial literature. The ability to understand and predict bank failures is essential not only for enhancing regulatory and supervisory capabilities but also for mitigating the economic impact associated with these failures. This understanding aids in crafting effective strategies and policies aimed at preventing future collapses, thereby preserving both the banking system and broader economic stability.

Historically, there has been significant evolution along the academic journey into understanding bank failures, from traditional statistical methods to more sophisticated machine learning and artificial intelligence-based approaches. Early research focused on discriminant analysis and logistic regression methods, but with the growing complexity of financial markets, these models faced limitations in predictive accuracy. The shift to advanced computational techniques, including artificial neural networks and support vector machines, marked a pivotal change. These methods offered a nuanced analysis of complex financial data, enhancing the predictive capabilities of models in identifying potential bank failures.

In evaluating the performance of various machine learning algorithms in predicting Non-Failed and Failed banks, a comprehensive analysis reveals distinct strengths across different models. For non-failed bank predictions, Random Forest variants, particularly Regularized, Cost Sensitive, and General Random Forest, demonstrated a marginal superiority over Boosting methods. These models showed a higher rate of correctly classified instances, better precision, and a balanced trade-off between precision and recall. They were particularly effective in minimizing false positives and ensuring that banks classified as stable were indeed stable.

On the other hand, in predicting Failed banks, all models displayed high accuracy with low percentages of incorrectly classified instances. However, Regularized Random Forest stood out for its high recall, making it particularly sensitive in identifying failing banks and less likely to miss a bank at risk of failure. General Random Forest provided the best balance between precision and recall, as indicated by its highest PRC Area. Although Boosting methods like GLMBoost and LogitBoost showed a lower recall and PRC Area, they remained viable options, particularly in scenarios where specific strengths such as model interpretability were required.

The evolution from traditional models to advanced machine learning approaches reflects the increasing intricacy of financial markets and the need for sophisticated analytical tools. Moreover, the 2008 financial crisis underscored the critical need for accurate prediction and prevention of bank failures, highlighting the interconnectedness of the global financial system and the far-reaching impacts of the failure of financial institutions.

In conclusion, the choice of the most suitable model for predicting bank stability should be aligned with the specific needs and risk profiles of the application. Factors such as the cost of false negatives versus false positives, the requirement for model interpretability, and the specific nature of the banking data should guide this selection. In the high-stakes world of banking, where accuracy and reliability are paramount, models offering the best balance between sensitivity to failures and precision, such as Regularized and General Random Forest, might be preferable. However, boosting methods could be advantageous in contexts where their specific strengths align with the operational requirements of the task at hand. Continuous innovation in methodologies remains essential to keep pace with the evolving nature of financial risks and the complexities of modern financial systems.

4.1 Limitations of the study

There is an important inherent limitation of relying solely on CAMEL ratios for forecasting bank failures and mitigating banking risks, especially when considering the complex nature of financial institutions and the dynamic environments in which they operate. This limitation primarily arises from the fact that CAMEL ratios are quantitative measures that may not fully capture or reflect the nuances of behavioral finance and managerial decision-making, which can significantly impact a bank's risk profile and its propensity for failure.

In essence, while CAMEL ratios serve as a valuable tool for regulatory bodies and financial analysts to assess the financial health and stability of banks, they cannot encapsulate the entirety of risks associated with behavioral finance and managerial decision-making. The limitation of the study, therefore, lies in its dependency on these ratios without integrating the qualitative aspects of banking operations and human behaviors. This reliance can lead to an incomplete risk assessment and potentially overlook emerging threats that stem from non-quantitative factors.

To address this limitation, future research could explore the possibilities of integrating behavioral finance metrics and qualitative assessments of management quality and governance practices into the existing framework. This could involve developing new models that combine quantitative financial health indicators with qualitative evaluations of managerial decision-making, ethical considerations, and corporate governance structures. By acknowledging and attempting to incorporate these complex and nuanced factors, researchers and practitioners can develop more holistic and effective strategies for predicting bank failures and implementing mitigation strategies.

APPENDIX A

The Financial Ratios Used in the Study (CAMEL Ratios)

1. Yield on Earning Assets
2. Cost of Funding Earning Assets
3. Net Interest Margin
4. Noninterest Income to Average Assets
5. Noninterest Expense to Average Assets
6. Credit Loss Provision to Assets
7. Net Operating Income to Assets
8. Return on Assets
9. Retained Earnings to Average Equity (YTD only)
10. Net Charge-Offs to Loans and Leases
11. Earnings Coverage of Net Loan Charge-Offs(x)
12. Efficiency Ratio
13. Assets Per Employee ($Millions)
14. Earning Assets to total Assets
15. Loss Allowance to Loans and Leases
16. Loss Allowance to Noncurrent Loans and Leases
17. Noncurrent Loans to Loans
18. Net Loans and Leases to Assets
19. Net Loans and Leases to Core Deposits
20. Domestic Deposits to Total Assets
21. Equity Capital to Assets
22. Total Risk-Based Capital Ratio
23. Average Total Assets
24. Average Earning Assets
25. Average Equity
26. Average Total Loans

REFERENCES

Alpaydin, E. (2020): *Introduction to Machine Learning* (4th ed.). Cambridge, MA: MIT Press. ISBN: 9780262043793.

Ashcraft, A. – Schuermann, T. (2008): Understanding the Securitization of Subprime Mortgage Credit. *Foundations and Trends in Finance*, 2(3), 191–309.

Berger, A. N. – Bouwman, C. H. S. (2012): How Does Capital Affect Bank Performance During Financial Crises? *Journal of Financial Economics*, 109(1), 146–176.

Bernanke, B. S. (2010): The Courage to Act: A Memoir of a Crisis and Its Aftermath. W. W. Norton & Company.

Biau, G. – Scornet, E. (2016): A random forest guided tour. TEST 25, 197–227. https://doi.org/10.1007/s11749-016-0481-7.

Blinder, A. S. – Zandi, M. (2010): The Financial Crisis: Lessons for the Next One. *Center on Budget and Policy Priorities.*

Boyacıoglu, M. A. – Kara, Y. – Baykan, Ö. K. (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(2), 3355–3366.

Breiman, L. (2001): Random forests. *Machine Learning*, 45(1), 5–32.

Bühlmann, P. – Hothorn, T. (2007): Boosting algorithms: Regularization, prediction and model fitting. *Statistical Science*, 22(4), 477–505.

Carmona, P. – Diranzo, F. J. C. – Momparler, A. (2019): Predicting failure in the U.S. banking sector: An extreme gradient boosting approach. *International Review of Economics & Finance*, 61(C), 304–323.

Chen, T. – Guestrin, C. (2016): XGBoost: A Scalable Tree Boosting System. In KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794.

Chiaramonte, L. – Liu, H. – Poli, F. – Zhou, M. (2016): How accurately can z-score predict bank failure? *Financial Markets, Institutions & Instruments*, 25(5), 333–360.

Cleary, S. – Hebb, G. (2016): An efficient and functional model for predicting bank distress: In and out of sample evidence. *Journal of Banking & Finance*, 64, 101–111.

Cutler, A. – Cutler, D. R. – Stevens, J. R. (2011): Random forests. In Zhang, Cha – Ma, Yunqian (eds.) (2011): *Ensemble Machine Learning*, 157–175, New York: Springer.

Davis, Jesse – Goadrich, Mark (2006): The relationship between Precision-Recall and ROC curves. In Proceedings of the 23rd international conference on Machine learning (ICML '06): Association for Computing Machinery, New York, NY, USA, 233–240. https://doi.org/10.1145/1143844.1143874.

Deng, H. (2013): Guided random forest in the RRF package. arXiv preprint arXiv:1306.0237.

DeYoung, R. – Torna, G. (2013): Nontraditional banking activities and bank failures during the financial crisis. *Journal of Financial Intermediation*, 22(3), 397–421.

Díaz-Uriarte, R. – De Andres, S. A. (2006): Gene selection and classification of microarray data using random forest. *BMC Bioinformatics*, 7, 3.

Elkan, C. (2001): The Foundations of Cost-Sensitive Learning. In Proceedings of the Seventeenth International Joint Conference on Artificial Intelligence (IJCAI).

Erdal, H. I. – Ekinci, C. (2013): A New Approach to Turkish Banking Sector's Failure Prediction via Artificial Neural Network and Support Vector Machine. *Expert Systems with Applications*, 40(3), 887-897.

Financial Crisis Inquiry Commission (2011): The Financial Crisis Inquiry Report. U.S. Government Printing Office.

Foote, C. L. – Gerardi, K. S. – Willen, P. S. (2008): Subprime Facts: What (We Think) We Know about the Subprime Crisis and What We Don't. Federal Reserve Bank of Boston, Public Policy Discussion Papers.

Freund, Y. – Schapire, R. E. (1996): Experiments with a New Boosting Algorithm. In Proceedings of the Thirteenth International Conference on Machine Learning (ICML).

Friedman, J. (2001): Greedy Function Approximation: A Gradient Boosting Machine. *The Annals of Statistics*, 29(5), 1189–1232.

Friedman, J. – Hastie, T. – Tibshirani, R. (2000): Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors): *Annals of Statistics*, 28(2), 337–407.

Gogas, P. – Papadimitriou, T. – Agrapetidou, A. (2018): Forecasting Bank Failures: A Machine Learning Approach. *Annals of Economics and Finance*, 19(2), 421-455.

Goodfellow, I. – Bengio, Y. – Courville, A. (2016): *Deep Learning*. Cambridge, MA: MIT Press.

Gorton, G. (2009): *Slapped by the Invisible Hand: The Panic of 2007*. Prepared for the Federal Reserve Bank of Atlanta's 2009 Financial Markets Conference: Financial Innovation and Crisis, May 11-13.

Hastie, T. – Tibshirani, R. – Friedman, J. (2009): *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York: Springer Series in Statistics.

Huang, B. – Chen, J. – Wang, C. (2019): Financial crisis prediction model using ant colony optimization for XGBoost. *Expert Systems with Applications*, 122, 32-41.

Khalilia, M. – Chakraborty, S. – Popescu, M. (2011): Predicting disease risks from highly imbalanced data using random forest. *BMC Medical Informatics and Decision Making*, 11, 51.

Kothari, V. (2008): *Securitization: The Financial Instrument of the Future*. Hoboken, NJ: Wiley.

Landwehr, N. – Hall, M. – Frank, E. (2005): Logistic Model Trees. *Machine Learning*, 59(1-2), 161–205.

Le, L. – Viviani, J. L. (2017): An Empirical Analysis of Bank Failure Prediction in the United States. *Journal of Financial Research*, 40(2), 257–282.

Liaw, A. – Wiener, M. (2002): Classification and regression by RandomForest. *R News*, 2(3), 18–22.

Ling, C. X. – Sheng, V. S. (2008): Cost-Sensitive Learning and the Class Imbalance Problem. In Sammut, C. (eds.) (2010): *Encyclopedia of Machine Learning*. New York: Springer.

Lu, W. – Whidbee, D. A. (2013): Bank structure and failure during the financial crisis. *Journal of Financial Economic Policy*, 5(3), 281–299.

Lu, W. – Whidbee, D. A. (2016): US bank failure and bailout during the financial crisis Examining the determinants of regulatory intervention decisions. *Journal of Financial Economic Policy*, 8(3), 316–347.

Malhotra, R. – Malhotra, D. K. (2021): Evaluating consumer loans using machine learning algorithms. *Journal of Risk and Financial Management*, 14(3), 105.

Manthoulis, G. – Mylonakis, J. – Diacogiannis, G. (2020): Predicting Bank Failures: A Statistical and Machine Learning Approach. *Journal of Financial Risk Management*, 9(4), 305–323.

Martin, D. (1977): Early Warning of Bank Failure: A Logit Regression Approach. *Journal of Banking & Finance*, 1(3), 249–276.

Meinshausen, Nicolai (2007): Relaxed Lasso. *Computational Statistics & Data Analysis*, 52(1), September, 374–393.

Meyer, P. A. – Pifer, H. W. (1970): Prediction of Bank Failures. *The Journal of Finance*, 25(4), 853–868.

Mian, A. – Sufi, A. (2009): The Consequences of Mortgage Credit Expansion: Evidence from the U.S. Mortgage Default Crisis. *The Quarterly Journal of Economics*, 124(4), 1449–1496.

Momparler, A. – Carmona, P. – Redondo, Y. (2020): Fuzzy-Set Qualitative Comparative Analysis (fsQCA) in Banking Failure Prediction. *Journal of Business Research*, 118, 253–263.

Natekin, A. – Knoll, A. (2013): Gradient boosting machines, a tutorial. *Frontiers in Neurorobotics*, 7, 21.

Oppel, S. – Strobl, C. – Huettmann, F. (2009): Alternative Methods to Quantify Variable Importance in Ecology. University of Munich, Department of Statistics, Technical Report 65.

Powers, D. M. (2011): Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation. *Journal of Machine Learning Technologies*.

Provost, F. – Fawcett, T. – Kohavi, R. (1998): The Case Against Accuracy Estimation for Comparing Induction Algorithms. In Proceedings of the 15th International Conference on Machine Learning (ICML, 1998).

Rajan, R. G. (2010): *Fault Lines: How Hidden Fractures Still Threaten the World Economy*. Princeton, New Jersey: Princeton University Press.

Ravi, V. – Promodh, C. (2008): Bankruptcy Prediction in Banks and Firms via Statistical and Intelligent Techniques – A Review. *European Journal of Operational Research*, 180(1), 1–28.

Saito, T. – Rehmsmeier, M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. PLoS One. 2015 Mar 4;10(3): e0118432, DOI: 10.1371/journal.pone.0118432, PMID: 25738806; PMCID: PMC4349800.

Serrano-Cinca, C. – Fuertes-Callén, Y. – Gutiérrez-Nieto, B. – Cuellar-Fernández, B. (2014): Path modelling to bankruptcy: Causes and symptoms of the banking crisis. *Applied Economics*, 46(31), 3798–3811.

Shiller, R. J. (2008): *The Subprime Solution: How Today's Global Financial Crisis Happened, and What to Do about It*. Princeton, NJ: Princeton University Press.

Sinkey, J. F. (1975): A Multivariate Statistical Analysis of the Characteristics of Problem Banks. *Journal of Finance*, 30(1), 21–36.

Sorkin, A. R. (2009): *Too Big to Fail: The Inside Story of How Wall Street and Washington Fought to Save the Financial System – and Themselves*. London: Allen Lane (Penguin Books).

Stehman, S. V. (1997): Selecting and interpreting measures of thematic classification accuracy. *Remote Sensing of Environment*, 62(1), 77-89.

Thomson, James (1991), Predicting bank failures in the 1980's Economic Review, 27, issue Q I, p. 9-20, https://EconPapers.repec.org/RePEc:fip:fedcer:y:1991:i:qi:p:9-20:n:v.27no.1.

Trussel, J. J. – Johnson, L. (2012): A parsimonious and predictive model of the recent bank failures. Academy of Banking Studies Journal, 11(1), 15–30.

Turney, P. D. (2000): Types of cost in inductive concept learning. Published in Proceedings of the Cost-Sensitive Learning Workshop at the 17th ICML-2000 Conference, Stanford, CA. July 2, 2000 NRC 43671.

Tutz, G. – Binder, H. (2006): Generalized additive modeling with implicit variable selection by likelihood-based boosting. *Biometrics*, 62(4), 961–971.

Zadrozny, B. – Elkan, C. (2001): Learning and making decisions when costs and probabilities are both unknown. In KDD '16: Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 204–213.

Zou, H. – Hastie, T. (2005): Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 67(2), 301–320.