



# **An Improved Approach for Modeling Bank Loan Default in Pursuit of Sustainable Banking**

**Noura Metawa<sup>1\*</sup>, Rania Itani<sup>2</sup>**

<sup>1</sup>University of Sharjah, Sharjah, UAE

<sup>1</sup>University of Murdoch, Dubai, UAE

Emails: [nmetawa@sharjah.ac.ae](mailto:nmetawa@sharjah.ac.ae); [rania.itani@dmu.ac.uk](mailto:rania.itani@dmu.ac.uk)

## **Abstract**

This article presents our research effort to explore the convergence of sustainable banking practices and predictive modeling for bank loan defaults, with a primary emphasis on addressing the pressing need for resilient financial systems. To this end, an applied methodology is presented in this study to model bank loan defaults, emphasizing the incorporation of sustainability criteria into predictive analytics. Given the temporal nature of loan data, our approach leverages Long Short-Term Memory (LSTM) networks as its backbone process for predictive modeling. The empirical results of the public case study underscored the enhanced predictive accuracy completed through this approach, emphasizing the pivotal function of integrating sustainability metrics in predicting mortgage defaults inside the banking area.

**Keywords:** Loan Default; Sustainable Finance; Credit Risk Modeling; Banking Sustainability; Machine Learning; Financial Stability; Default Probability Estimation; Risk Management; Econometric Modeling; Banking Practices.

## **1. Introduction**

The worldwide economic panorama continues to evolve, prompting a vital evaluation of banking structures to foster sustainability and resilience. Within this realm, the predictive modeling of bank loan default stands as a pivotal region influencing financial stability and sustainable practices [1-3]. This paper delves into an imperative pursuit: the refinement of a technique geared toward modeling financial institution mortgage default, looking to reinforce the muse of sustainable banking. In a generation in which economic volatility intertwines with environmental and social considerations, the enhancement of predictive fashions will become imperative to mitigate dangers and nurture a sustainable monetary environment [4-5]. At the heart of sustainable banking lies the project of efficiently managing credit score risks while aligning operations with ethical, social, and environmental principles [6-7]. The inherent complexity of predicting mortgage defaults needs sophisticated modeling strategies that not simplest investigate financial aspects but also integrate broader sustainability metrics. Embracing this project, this paper endeavors to provide a modern framework that amalgamates strong statistical methodologies and contemporary gadget learning algorithms [8-10]. Such an integrative method's objectives are no longer to enhance predictive accuracy in loan default modeling but also to embed sustainability standards into the center of banking decision-making approaches [11].

This research embarks on an exploration of methodologies that go beyond traditional hazard assessment paradigms, acknowledging the multidimensional nature of sustainable banking [12-16]. By bridging the space between quantitative modeling and sustainability imperatives, this paper aspires to make contributions to a nuanced knowledge of the way advanced modeling strategies can bolster the resilience of economic institutions. Furthermore, it seeks to elucidate how the amalgamation of those methodologies fosters a greater holistic approach to danger control, one which fosters the lengthy-term viability of banks within a framework of societal and environmental stewardship [17-19].

In essence, this observation units out to redefine the contours of predictive modeling for financial institution loan defaults by emphasizing the integration of sustainability concerns. By doing so, it aims to not only enhance the predictive energy of fashions but also domesticate a banking ethos rooted in responsible lending practices and a commitment to sustainable development goals [13]. Through an in-depth exploration of this intersection between monetary modeling and sustainability, this paper strives to offer insights so that it will resonate in the banking sector and make contributions to the wider discourse on fostering sustainability inside economic structures [20-21].

This paper is established to provide a complete understanding of the pursuit of sustainable banking via an improved method for modeling financial institution mortgage default. Section 2 navigates thru the prevailing literature, and empirical studies shaping the panorama of mortgage default modeling. Following this, Section Three unveils the innovative framework developed in this study. Section four meticulously info the methodologies used in checking out and validating the proposed approach. Section 5 gives the empirical findings and conducts a rigorous evaluation of the consequences. Finally, Section 6 synthesizes the important thing insights gleaned from this study.

## **2. Related Works**

In terms of sustainable banking and forecasting models for loan defaults, this section examines a comprehensive review of the relevant literature and empirical research underpinning this review Ahmad et al. [22] examined the implications of the global financial crisis in more detail through the lens of Islamic finance, examining its views and responses to economic policy and sustainability in the Islamic framework and that to Pisano and his colleagues. [23] examined the complex relationship between finance and sustainable development, emphasizing the role of the board in promoting or constraining sustainable practices. Betz et al. [24] contributed significantly to the discussion by providing a predictive model of distress in European banks, enhancing the understanding of risk assessment in banking institutions Calomiris and Gorton [25] examined bank panics initially by providing insights into patterns, facts, and legal decisions Providing insights into, and emphasizing the role of institutional legitimacy and market differences in the global economic environment In addition, Nizam and his colleagues [27] conducted a comprehensive global study, which assessed the impact of social and environmental sustainability on the financial performance of the banking sector, providing empirical evidence of the link between growth and finance between the effects of social and environmental sustainability on the financial performance of the banking sector, providing empirical evidence for the interplay between sustainability and financial outcomes.

Sanusi [28] scrutinized the aftermath of the global economic meltdown at the Nigerian banking quarter, focusing on the reforms instituted to mitigate systemic dangers. Poledna and Thurner [29] proposed a systemic hazard transaction tax to take away systemic risk in financial networks, contributing to the discourse on systemic danger management. Furthermore, Gennaioli et al. [30] delivered a version of shadow banking, offering insights into the complexities and implications of this monetary mechanism. Bolton and Freixas [31] delved into the connection between corporate finance and the economic transmission mechanism, dropping mild at the difficult connections between economic selection-making and the broader economy. Weber et al. [32] contributed to the integration of sustainability standards into credit threat management, emphasizing its significance in cutting-edge threat evaluation practices. Bihari and Pradhan [33] explored the relationship between Corporate Social Responsibility (CSR) and performance inside the Indian banking region, highlighting the effect of CSR initiatives on financial institutions' overall performance metrics.

Ozili and Outa [34] carried out a comprehensive evaluation of bank loan loss provisions studies, providing a detailed analysis of the factors influencing those provisions within banking establishments. Moreover, Yap et al. [35] delved into the application of data mining techniques to beautify credit score scoring fashions, emphasizing its importance in assessing creditworthiness. Arafah and Nugroho [36] explored the utility of Maqhashid Sharia in the financing commercial enterprise version for clean water at Islamic banks, supplying insights into the intersection of Islamic standards and sustainable financing. Lastly, Behr et al. [37] analyzed the monetary constraints confronted by way of non-public firms and their implications on bank lending behavior, contributing to the understanding of lending practices concerning limited corporations.

## **3. Methodological Design**

This section carefully presents the methodology used in this study and provides a systematic framework for developing an advanced approach to modeling bank loan defaults in sustainable banking.

Long Short-Term Memory (LSTM) is a kind of recurrent neural community (RNN) architecture designed to method and version sequential information whilst mitigating the constraints related to the vanishing and exploding gradient

troubles in traditional RNNs. LSTM networks are characterized by way of their capability to capture and keep long-term dependencies within sequential information with the aid of incorporating specialized reminiscence cells. These memory cells include gates—along with input, forget about, and output gates—that regulate the flow of statistics, permitting LSTMs to selectively keep or discard statistics over varying time steps. By leveraging these gating mechanisms, LSTMs efficiently manage and hold relevant facts across sequential inputs, making them particularly appropriate for modeling and predicting temporal patterns, developments, and dependencies within dynamic datasets like monetary time collection.

In this study, we used the LSTM network as a sophisticated deep-learning technique to identify and understand the complex structure embedded in the loan database. The inherent ability of LSTM to model temporal dependence made it a suitable candidate to capture the complex sequential relationships of financial data, especially in the forecasting of loan defaults. Particularly, we fit the loan dataset to the LSTM system, the network automatically identified temporal patterns and dependencies embedded in the data and participated in them. Through iterative and optimized training, the LSTM model successfully indexed the sequences, extracted latent factors, and found complex temporal relationships between variables. It allowed us to harness the power of deep learning for regularity patterns, and subtle relationships have been revealed in the loan data set, providing a more accurate and comprehensive model to enable us to predict bank loan defaults in a sustainable banking system.

The following formula articulates the working mechanism for the LSTM cell:

$$f_t = \sigma(W_f [h_{t-1}, X_t] + b_f) \quad (1)$$

$$i_t = \sigma(W_i [h_{t-1}, X_t] + b_i), \quad (2)$$

$$N_t = \tanh(W_n [h_{t-1}, X_t] + b_n), \quad (3)$$

$$C_t = C_{t-1} f_t + N_t i_t. \quad (4)$$

$$O_t = \sigma(W_o [h_{t-1}, X_t] + b_o), \quad (5)$$

$$h_t = O_t \tanh(C_t). \quad (6)$$

the output generated by the Long Short-Term Memory (LSTM) layers progresses to Fully Connected Layers (FCLs) with sigmoid activation, which is expressed as:

$$\text{Output}_{\text{FCL}} = \sigma(Wx + b) \quad (7)$$

Here,  $\text{Output}_{\text{FCL}}$  denotes the output of the FCL,  $W$  represents the weight matrix,  $x$  signifies the input vector from the LSTM output, and  $b$  symbolizes the bias vector.

The cross-entropy loss function,  $\mathcal{L}_{\text{CE}}$ , is employed to evaluate the disparity between the predicted values and the actual labels, and is defined as:

$$\mathcal{L}_{\text{CE}} = -\frac{1}{N} \sum_{i=1}^N (y_i \cdot \log(\hat{y}_i) + (1 - y_i) \cdot \log(1 - \hat{y}_i)) \quad (8)$$

Here,  $N$  signifies the total number of samples in the dataset,  $y_i$  denotes the actual label (either 0 or 1) for the  $i^{\text{th}}$  sample,  $\hat{y}_i$  represents the predicted probability (output from the sigmoid function) for the  $i^{\text{th}}$  sample.

#### 4. Experimental Setup

In this part of our manuscript, the experimental setup hired to validate and test the proposed technique for modeling bank mortgage default within the context of sustainable banking is comprehensively detailed. This section delineates the tricky layout and execution of empirical experiments orchestrated to evaluate the efficacy and performance of the subtle modeling method provided in this section. In Table 1, we present a comprehensive compilation of precise information derived from the mortgage dataset, encapsulating key quantitative measures that provide a descriptive review of the fact's traits. This desk includes vital statistical metrics consisting of mean, well-known deviation, minimum, maximum, and quartile values for applicable variables or capabilities within the mortgage dataset. It is notable that the mean of values of different features lies in different ranges which necessitates normalizing the features'

value to a uniform range before fitting them into the model. In addition, the tabulated count values imply that there is a column that can missing values, which necessitate applying appropriate techniques for imputing the missing values.

Table 1: Summary Statistics of Loan Dataset Variables.

	count	mean	std	min	25%	50%	75%	max
loan_amount	148664. 0	331115. 1	183909. 9	16500. 0	196500. 0	296500. 0	436500. 0	3576500.0
rate_of_interest	112226. 0	4.0	0.6	0.0	3.6	4.0	4.4	8.0
Interest_rate_sprea d	112026. 0	0.4	0.5	-3.6	0.1	0.4	0.8	3.4
Upfront_charges	109023. 0	3225.1	3251.1	0.0	581.9	2596.8	4812.5	60000.0
term	148623. 0	335.1	58.4	96.0	360.0	360.0	360.0	360.0
property_value	133566. 0	497915. 5	359928. 4	18000. 0	268000. 0	418000. 0	628000. 0	16508000. 0
income	139520. 0	6957.3	6496.6	0.0	3720.0	5760.0	8520.0	578580.0
Credit_Score	148664. 0	699.8	115.9	500.0	599.0	699.0	800.0	900.0
LTV	133566. 0	72.5	18.9	1.0	60.5	75.1	86.2	263.5
Status	148664. 0	0.2	0.4	0.0	0.0	0.0	0.0	1.0
dtir1	124549. 0	37.7	10.5	5.0	31.0	39.0	45.0	61.0

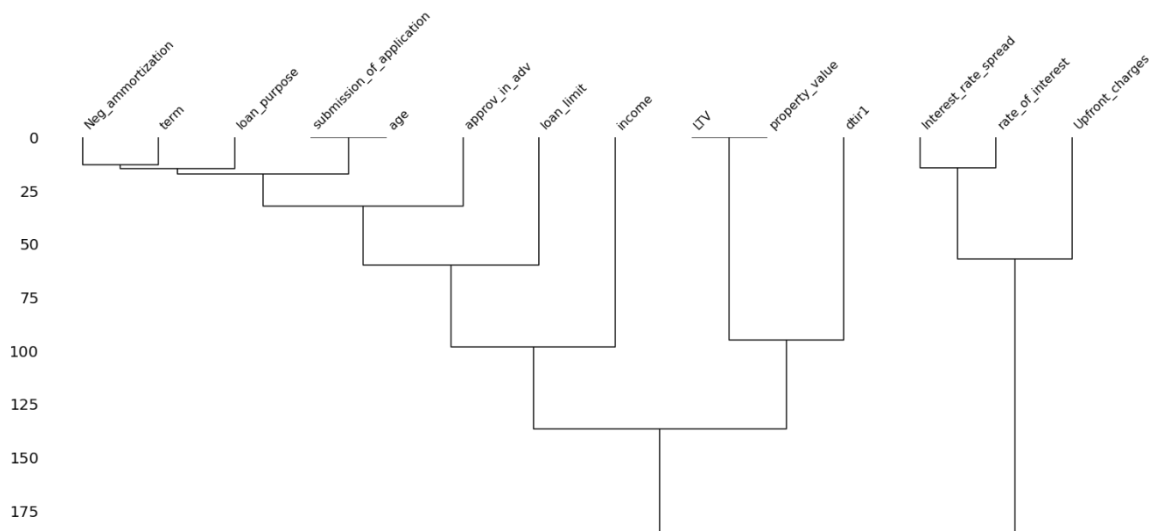


Figure 1: Dendrogram Visualization of Features in Loan Dataset

### 5. Experimental Results

In this phase of our work, we delved into the empirical results and their meticulous evaluation, imparting a detailed exposition of the outcomes derived from the rigorous application of the delicate modeling method in earlier sections. This phase unfolds a complete exploration of the empirical findings acquired via the experimental setups mentioned

right here, aiming to scrutinize the efficacy and sustainability implications of the proposed method for modeling financial institution mortgage default inside sustainable banking frameworks.

In Figure 1, we gift a comprehensive visible representation inside the shape of a lacking fee heatmap, showcasing the distribution and patterns of missing values inside the loaded dataset. This heatmap gives a clean and systematic review of the lacking data throughout distinctive capabilities or variables, permitting a visible assessment of the completeness or sparsity of facts inside the dataset. The various sunshades or colors inside the heatmap correspond to the presence or absence of records, making an allowance for easy identification of missing values and their particular locations inside the dataset. This visualization performs a pivotal position in the knowledge of the quantity and styles of missingness, which is vital for making knowledgeable decisions regarding records imputation or coping with lacking values in subsequent analytical steps. By visually depicting the missing fee styles through this heatmap, we aim to offer treasured insights into the statistics best and facilitate appropriate strategies for dealing with missing facts in our analysis of bank mortgage default modeling within the context of sustainable banking practices.

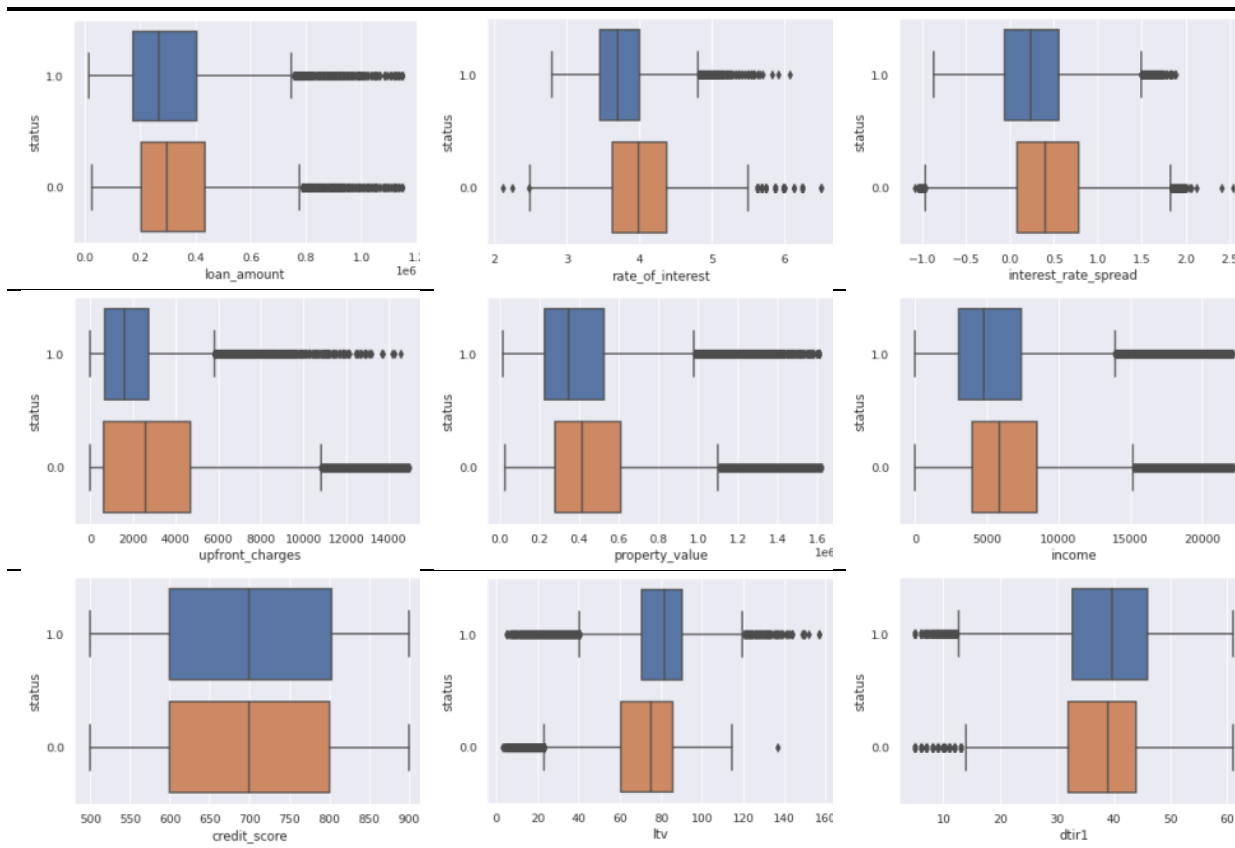


Figure 2: Visualization of Box and Whisker Plots for Loan Data Variables.

In Figure 2 we present a visualization of the dendrogram showing the ranked features in the weighted dataset. This is a dendrogram representation of the relationship and proximity between the various factors used in our study. Through this visualization, hierarchical clustering algorithms sort objects based on their similarities, revealing unique groups or clusters based on shared characteristics or patterns. Examining this dendrogram we gain valuable insight into the underlying structure and relationships between features.

In Figure 3, we present box and whisker plots that offer a comprehensive visible representation of the distribution and variability of variables inside the mortgage dataset. These plots depict the statistical measures inclusive of median, quartiles, and outliers for every variable, offering a clean example of their crucial tendency, unfold, and the potential presence of extreme values. In Figure 4, we gift the confusion matrix derived from the assessment of our version. The

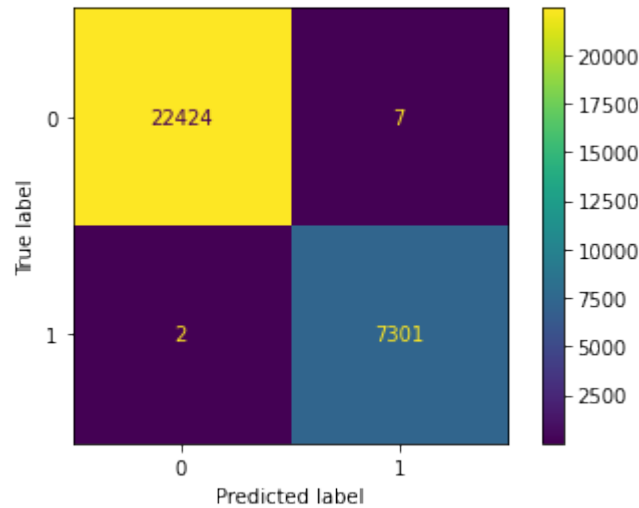


Figure 3: Confusion Matrix Representation of Model Evaluation

confusion matrix illustrates the performance of the version with the aid of summarizing the genuine positives, actual negatives, fake positives, and fake negatives. This visualization serves as a pivotal tool for assessing the version's overall performance, facilitating a deeper knowledge of its strengths, weaknesses, and regions for capacity development in optimizing predictive outcomes.

## 6. Conclusion

This study spans the terrain of sustainable banking by taking a new approach by applying advanced modeling techniques to the forecasting of bank loan defaults. Through the use of long- and short-term memory networks (LSTM) and comprehensive loan data analysis, this study reveals nuanced understandings of the complex relationship between economic volatility and banking behavior it remains permanent. Despite emphasizing the importance of predictive accuracy to identify potential credit failures from its complex models skillfully captured and identified, the study emphasizes the importance of sustainability considerations will accompany predictive modeling efforts, paving the way for responsible lending practices and enabling resilience in monetary policy. Specifically, this study integrates economic and developmental data and advocates a paradigm shift towards a more robust and conscientious banking system. The findings presented here not only highlight the effectiveness of LSTM-based modeling in increasing the accuracy of loan defaults but also highlight the important role of data-driven methods emphasizing sustainable development in the economic environment.

## References

- [1] Yip, Angus W H, and Nancy M P Bocken. 2018. "Sustainable Business Model Archetypes for the Banking Industry." *Journal of Cleaner Production* 174: 150–69.
- [2] Bhardwaj, Broto Rauth, and Aarushi Malhotra. 2013. "Green Banking Strategies: Sustainability through Corporate Entrepreneurship." *Greener Journal of Business and Management Studies* 3 (4): 180–93.
- [3] Bennett, Lynn, and Carlos E Cuevas. 1996. "Sustainable Banking with the Poor." *Journal of International Development*. Wiley Online Library.
- [4] Paulet, Elisabeth, Miia Parnaudeau, and Francesc Relano. 2015. "Banking with Ethics: Strategic Moves and Structural Changes of the Banking Industry in the Aftermath of the Subprime Mortgage Crisis." *Journal of Business Ethics* 131 (1): 199–207.
- [5] Ledgerwood, Joanna. 1999. "Sustainable Banking with the Poor Microfinance Handbook."
- [6] Banga, Josué. 2019. "The Green Bond Market: A Potential Source of Climate Finance for Developing Countries." *Journal of Sustainable Finance & Investment* 9 (1): 17–32.
- [7] Weber, Olaf, and Sven Remer. 2011. *Social Banks and the Future of Sustainable Finance*. Vol. 64. Taylor & Francis.
- [8] Cui, Yujun, Sean Geobey, Olaf Weber, and Haiying Lin. 2018. "The Impact of Green Lending on Credit Risk in China." *Sustainability* 10 (6): 2008.

- [9] Bose, Sudipta, Habib Zaman Khan, Afzalur Rashid, and Shajul Islam. 2018. "What Drives Green Banking Disclosure? An Institutional and Corporate Governance Perspective." *Asia Pacific Journal of Management* 35: 501–27.
- [10] Schoenmaker, Dirk, and Willem Schramade. 2018. *Principles of Sustainable Finance*. Oxford University Press.
- [11] Milano, Riccardo. 2011. "Social Banking: A Brief History." In *Social Banks and the Future of Sustainable Finance*, 15–47. Routledge.
- [12] D'Orazio, Paola, and Lilit Popoyan. 2019. "Fostering Green Investments and Tackling Climate-Related Financial Risks: Which Role for Macroprudential Policies?" *Ecological Economics* 160: 25–37.
- [13] Thomas, Jesu Raju, and Jyothi Kumar. 2016. "Social Performance and Sustainability of Indian Microfinance Institutions: An Interrogation." *Journal of Sustainable Finance & Investment* 6 (1): 38–50.
- [14] Falcone, Pasquale Marcello, and Edgardo Sica. 2019. "Assessing the Opportunities and Challenges of Green Finance in Italy: An Analysis of the Biomass Production Sector." *Sustainability* 11 (2): 517.
- [15] Conley, John M, and Cynthia A Williams. 2011. "Global Banks as Global Sustainability Regulators?: The Equator Principles." *Law & Policy* 33 (4): 542–75.
- [16] Angori, Gabriele, David Aristei, and Manuela Gallo. 2019. "Determinants of Banks' Net Interest Margin: Evidence from the Euro Area during the Crisis and Post-Crisis Period." *Sustainability* 11 (14): 3785.
- [17] Munkhdalai, Lkhagvadorj, Tsenduren Munkhdalai, Oyun-Erdene Namsrai, Jong Yun Lee, and Keun Ho Ryu. 2019. "An Empirical Comparison of Machine-Learning Methods on Bank Client Credit Assessments." *Sustainability* 11 (3): 699.
- [18] Markose, Sheri M, Simone Giansante, Mateusz Gatkowski, and Ali Rais Shaghaghi. 2010. "Too Interconnected to Fail: Financial Contagion and Systemic Risk in Network Model of Cds and Other Credit Enhancement Obligations of Us Banks."
- [19] Gorton, Gary B, and Ping He. 2008. "Bank Credit Cycles." *The Review of Economic Studies* 75 (4): 1181–1214.
- [20] Addo, Peter Martey, Dominique Guegan, and Bertrand Hassani. 2018. "Credit Risk Analysis Using Machine and Deep Learning Models." *Risks* 6 (2): 38.
- [21] Popescu, Cristina Raluca Gh, and Gheorghe N Popescu. 2019. "An Exploratory Study Based on a Questionnaire Concerning Green and Sustainable Finance, Corporate Social Responsibility, and Performance: Evidence from the Romanian Business Environment." *Journal of Risk and Financial Management* 12 (4): 162.
- [22] Ahmed, Adel. 2010. "Global Financial Crisis: An Islamic Finance Perspective." *International Journal of Islamic and Middle Eastern Finance and Management* 3 (4): 306–20.
- [23] Pisano, Umberto, André Martinuzzi, and Bernulf Bruckner. 2012. "Financial Sector and Sustainable Development."
- [24] Betz, Frank, Silviu Opric'ua, Tuomas A Peltonen, and Peter Sarlin. 2014. "Predicting Distress in European Banks." *Journal of Banking & Finance* 45: 225–41.
- [25] Calomiris, Charles W, and Gary Gorton. 1991. "The Origins of Banking Panics: Models, Facts, and Bank Regulation." In *Financial Markets and Financial Crises*, 109–74. University of Chicago Press.
- [26] Ng, Artie W. 2018. "From Sustainability Accounting to a Green Financing System: Institutional Legitimacy and Market Heterogeneity in a Global Financial Centre." *Journal of Cleaner Production* 195: 585–92.
- [27] Nizam, Esma, Adam Ng, Ginanjar Dewandaru, Ruslan Nagayev, and Malik Abdulrahman Nkoba. 2019. "The Impact of Social and Environmental Sustainability on Financial Performance: A Global Analysis of the Banking Sector." *Journal of Multinational Financial Management* 49: 35–53.
- [28] Sanusi, Sanusi L. 2011. "Global Financial Meltdown and the Reforms in the Nigerian Banking Sector." *CBN Journal of Applied Statistics (JAS)* 2 (1): 7.
- [29] Poledna, Sebastian, and Stefan Thurner. 2016. "Elimination of Systemic Risk in Financial Networks by Means of a Systemic Risk Transaction Tax." *Quantitative Finance* 16 (10): 1599–1613.
- [30] Gennaioli, Nicola, Andrei Shleifer, and Robert W Vishny. 2013. "A Model of Shadow Banking." *The Journal of Finance* 68 (4): 1331–63.
- [31] Bolton, Patrick, and Xavier Freixas. 2006. "Corporate Finance and the Monetary Transmission Mechanism." *The Review of Financial Studies* 19 (3): 829–70.
- [32] Weber, Olaf, Roland W Scholz, and Georg Michalik. 2010. "Incorporating Sustainability Criteria into Credit Risk Management." *Business Strategy and the Environment* 19 (1): 39–50.
- [33] Bihari, Suresh Chandra, and Sudepta Pradhan. 2011. "CSR and Performance: The Story of Banks in India." *Journal of Transnational Management* 16 (1): 20–35.
- [34] Ozili, Peterson K, and Erick Outa. 2017. "Bank Loan Loss Provisions Research: A Review." *Borsa Istanbul Review* 17 (3): 144–63.

- [35] Yap, Bee Wah, Seng Huat Ong, and Nor Huselina Mohamed Husain. 2011. "Using Data Mining to Improve Assessment of Credit Worthiness via Credit Scoring Models." *Expert Systems with Applications* 38 (10): 13274–83.
- [36] Arafah, Willy, and Lucky Nugroho. 2016. "Maqhashid Sharia in Clean Water Financing Business Model at Islamic Bank." *International Journal of Business and Management Invention* 5 (2): 22–32.
- [37] Behr, Patrick, Lars Norden, and Felix Noth. 2013. "Financial Constraints of Private Firms and Bank Lending Behavior." *Journal of Banking & Finance* 37 (9): 3472–85.