risks

MDPI

Article

Predicting Credit Risk with ESG Factors Using XGBoost and Structural Learning in Vague Environments (SLAVE) in Commercial Banks

Jamil J. Jaber [1,*], Abdullah A. K. Alkhawaldeh [2], Qusay Ayman Sulaiman Mazahreh [3], Ala'Aldin Al Rowwad [4], Anwar Al-Gasaymeh [5] and Thair A. Kaddumi [1]

1 Department of Banking and Finance, Faculty of Business, Applied Science Private University (ASU), Amman 11937, Jordan; thair_lion@asu.edu.jo
2 Department of Accounting, Faculty of Economics and Administrative Sciences, Hashemite University, Zarqa 13133, Jordan; alkwaldhh@yahoo.com
3 Department of Accounting, School of Business, Applied Science Private University (ASU), Amman 11931, Jordan; qusay.mazahreh@gmail.com
4 Department of Public Administration, Faculty of Business, The University of Jordan, Amman 11942, Jordan; a.alrowwad@ju.edu.jo
5 Department of Accounting and Finance, College of Administrative Sciences, Applied Science University (ASU), Al Eker 5055, Bahrain; anwar.gasaymeh@asu.edu.bh
* Correspondence: j.jaber@ju.edu.jo or jameljaber2011@hotmail.com

Abstract

Predicting credit risk is vital for banks as it safeguards financial stability, minimizes default losses, optimizes capital, and ensures regulatory compliance. This study aims to predict credit risk (High/Low) in commercial banks by integrating machine learning with traditional econometric approaches. The Structural Learning in Vague Environments (SLAVE) fuzzy rule-based model handles ambiguity in financial decisions, while the eXtreme Gradient Boosting (XGBoost) uncovers non-linear patterns among predictors. Input variables—profitability, liquidity risk, ESG (environmental, social, and governance) score, and monetary freedom—were selected via multicollinearity tests and three panel regression models, including ordinary least squares (OLS), fixed effects, and random effects models. The empirical investigation uses a panel dataset of forty commercial banks across seven Middle Eastern countries from 2014 to 2023, yielding 400 observations. Regression results reveal that profitability and ESG score significantly reduce credit risk. Liquidity risk and monetary freedom increase credit risk. XGBoost combined with the SHapley Additive exPlanations (SHAP)-based interpretation identifies ESG Score as the most influential predictor. The SLAVE model was evaluated using three data splits: 70/30, 80/20, and 90/10. The 80/20 split achieved the highest accuracy, with superior performance in identifying low-risk banks. Stronger ESG performance and stable monetary environments contribute to fostering sustainable banking and reducing credit risk, making these indicators valuable for risk management frameworks in the Middle Eastern banking sector.

Keywords: ESG; credit risk; SLAVE model; XGBoost model

Check for updates

Academic Editor: Jian Xu

Received: 27 July 2026
Revised: 25 August 2026
Accepted: 25 August 2026
Published: 27 August 2026

1. Introduction

The most important risk that financial institutions will have to deal with is credit risk, defined as the risk that the borrower or counterparty will not fulfill its obligations as described in the contract (BCBS 2004; Saunders et al. 2026). Furthermore, global regulatory frameworks, such as Basel III, have increased the focus on banks' internal credit risk assessments and the capital buffers required to meet potential losses (BCBS 2017). Credit risk is inherent in financial intermediaries, where banks accept deposits and make loans to generate profits, distinguishing it from market and operational risks. The magnitude of this risk is staggering, as loan portfolios represent the largest asset item on most commercial banks' balance sheets, and credit losses have traditionally been the primary cause of bankruptcy (Bessis 2015). Accurate forecasting of credit risk is essential to the stability and efficiency of financial systems and lies at the heart of the sustainability of lending institutions and the broader economy. The ability to accurately predict a borrower's risk of non-repayment is crucial for allocating capital, managing risk, and meeting regulatory requirements, while also being directly relevant to the profitability of banks and other credit issuers (Dastile et al. 2020).

Effective credit risk management is essential to banks' sustainability, solvency, and profitability. Inadequate credit risk management directly reduces earnings due to provisions for non-performing loans and charge-offs, as well as the capital reserves required to absorb future crises (Hull 2023). On the other hand, a well-managed credit portfolio enables a bank to maximize the risk–return relationship, allocate resources more effectively, and gain a competitive advantage through more precise underwriting and pricing. Therefore, the ability to predict, estimate, and manage credit risk is not just a regulatory obligation but a key factor in the financial strength and reputation of a bank.

Credit risk modeling has been an ongoing journey of finding increasingly predictive and accurate ways of modeling credit risk. The classical statistical methods were largely predominant in the field at the beginning, and logistic regression and discriminant analysis were used as the benchmark methods because of their transparency and ease of interpretation (Dumitrescu et al. 2022). These models are still widely used today, particularly in regulated industries where interpretability is crucial. As the volume, complexity, and dimensionality of the datasets have increased, more complex machine learning algorithms have been used (Dastile et al. 2020). The use of support vector machines (SVM) (Bellotti and Crook 2009), decision trees and ensemble methods such as random forests and boosting algorithms (Marqués et al. 2012) has been widely studied in the last 20 years. The literature also shows that these non-parametric models are often able to outperform traditional statistical methods, especially in the presence of non-linear relationships and complex interactions between borrower characteristics (Runchi et al. 2023; Sohn et al. 2016).

Recently, the focus has shifted towards a new generation of hybrid and deep learning models. To fuse the learning ability of neural networks with the interpretability of fuzzy logic, some architectures have been proposed (such as Adaptive Neuro-Fuzzy Inference Systems (ANFIS) (Akkoç 2012). Despite the potential of deep learning models for credit scoring, their "black-box" nature presents challenges in regulated financial environments (Talaat et al. 2024). This has triggered an important shift in the field, from just predicting to transparency and explainability of the models. The new trend, Explainable AI (XAI), aims to provide explanations for model decisions to lenders, regulators, and customers, as models are increasingly incorporated into credit risk frameworks to improve trust, auditability and fairness, such as by using techniques like SHAP and the Local Interpretable Model-agnostic Explanations—LIME (Machado et al. 2025; Straus et al. 2025; Wilkinson et al. 2026).

The incorporation of ESG (environmental, social, and governance) considerations into credit risk analysis has become a major milestone in the evolution of financial analysis, signaling a shift from focusing on financial data to a comprehensive analysis of corporate resilience and default probability. This increased focus on ESG performance is because non-financial factors can have a tangible impact on a borrower's capacity to meet their debt obligations, including through regulatory compliance, operational efficiency, reputational capital and stakeholder trust (Abdul Razak et al. 2023; Deng et al. 2024). However, a large empirical body of evidence suggests that there is a clear correlation between better ESG outcomes and better credit outcomes, such as lower credit spreads, lower default risk, and higher credit ratings (Anaz and Awad 2025; Okimoto and Takaoka 2024; Zhang 2025).

The results indicate that ESG is not only a signaling tool for reputation management, but also a relevant financial signal of the quality of management, transparency and long-term orientation. A study of the corporate credit default swap (CDS) spreads, for instance, shows that companies with higher ESG ratings have more favorable financing conditions and lower perceived default risk (Abdul Razak et al. 2023; Aslan et al. 2021) and research in global markets has demonstrated that integrating ESG improves corporate creditworthiness through better risk management and stakeholder relations (Anaz and Awad 2025). In addition, ESG engagement has proven to be related to decreased negative news sentiment, decreased information risk, and improved internal control mechanisms, thereby improving the accuracy of credit risk assessment and credit risk pricing (Miao and Nie 2026; Tian and Tian 2022). These channels all highlight the importance of ESG factors as good leading indicators of financial distress and debt servicing.

MENA (Middle East and North Africa) banks have adopted sustainable financing policies, similar to their European counterparts. However, the level of alignment with sustainable development goals differs across countries. According to (UNEP FI 2021, 2023), the UAE leads the region in sustainable finance. Morocco is in an advanced stage of implementation. Egypt has integrated climate finance into select policies and national strategies. Bahrain and Jordan are actively aligning their financial systems with sustainability agendas. Saudi Arabia, meanwhile, prioritizes economic diversification to reduce oil dependence and expand non-oil GDP contributions. Despite these national efforts, a notable gap persists in climate finance maturity and effectiveness. Adaptation finance remains largely untouched by current policy measures. Furthermore, CSR practices in the MENA region are still nascent. Disclosure guidelines for CSR remain voluntary (Farah et al. 2021; Maside-Sanfiz et al. 2024).

The main contribution of this study is the multi-methodological and comprehensive approach to the assessment of credit risk in the banking sector of a strategically important region that has not been studied in sufficient detail. Many studies exist in the field of credit risk modeling, but most of the literature deals with the credit risk problem in developed markets or only uses one analytical method. This study aims to fill the gaps by making four contributions. Firstly, a new and region-specific dataset is created consisting of commercial banks in seven Middle Eastern countries: the United Arab Emirates (UAE), Saudi Arabia, Qatar, Oman, Kuwait, Jordan and Bahrain. These countries are critical because they are a diverse group of oil-reliant economies, rapidly developing financial capital centers and different regulatory environments. This diversity enables a more comprehensive and transferable knowledge of the factors that affect credit risk, beyond the single country or developed market approach. Second, three types of panel regression models (OLS, fixed effects and random effects) are used to rigorously test the significance of the relationships between the input variables and credit risk, and much more careful consideration is given to which variables are included in the models before using more complex models. The selection of four carefully chosen input variables—profitability, liquidity risk, ESG score and monetary freedom—shows a comprehensive approach to bank stability. Specifically, the use of ESG score as a predictor is timely, as sustainability and governance concerns have become more prominent in the financial resilience of financial institutions and are not typically included in credit risk models for the region.

Third, to capture the complex interactions, we use XGBoost, a versatile and robust machine learning classification model that handles non-linear, complex dynamics often seen in financial information. XGBoost can discover hidden interaction effects and patterns in the relationship between profitability, liquidity risk, ESG score, and monetary freedom, unlike traditional linear models. To make this interpretable, we add SHAP values to XGBoost, which rank feature importance in an explainable manner and give us insights into which drivers are most impactful on credit risk in this case. Fourth, and most innovatively, a SLAVE fuzzy rule-based model for credit risk classification is introduced. This is a major methodological leap, since SLAVE is well adapted to deal with the intrinsic uncertainty and fuzziness of financial decision-making. SLAVE classifies banks as high-risk or low-risk by producing interpretable "if-then" rules. Overall, the use of XGBoost and SLAVE presents a strong and hybrid approach, bridging the gap between predictive performance and interpretability, which is crucial for credit risk analytics.

The rest of this paper is organized as follows. Section 2 comprises the detailed literature review and fleshes out the research hypotheses that would lead to the relationships between credit risk and its determinants. The methodology is explained in Section 3, covering regression models (OLS, fixed effects and random effects), the XGBoost machine learning algorithm, as well as the SLAVE fuzzy rule-based classification model. The empirical results and discussion section summarizes all the results of the analyses in Section 4. The proposed SLAVE model is presented in Section 5. Limitations of this study and directions for future research are discussed in Section 6. Finally, Section 7 summarizes the work and provides some suggestions for financial and banking policy and institutions.

2. Literature Review and Hypothesis Development

The aim of this section is to conduct a systematic review of the existing literature on the determinants of credit risk in banking. More specifically, we look at four factors that constitute overall credit risk exposure: profitability, liquidity risk, monetary freedom and ESG scores. Profitability and liquidity risk are traditional financial fundamentals, which have already received much attention in the literature, whereas the roles of monetary freedom and ESG performance are more subtle and are still a subject of scholarly debate.

2.1. Profitability on Credit Risk

Academic interest in the link between risk management and bank profitability has been growing significantly in the wake of the global financial crisis in 2007–2008 and the adoption of the Basel III regulatory frameworks. The relationship between credit risk and bank profitability has been extensively examined in the post-crisis banking literature, and this literature has generally found that increased credit risk, which is reflected in increased non-performing loans (NPLs) and loan loss provisions (LLPs), is consistently negatively associated with bank profit in various geographic and institutional contexts. The theoretical basis for this negative relationship is grounded in the basic principles of banking intermediation, where the losses due to loan defaults directly lower net income due to increased provisioning and write-offs, impact lending power, and reduce future income generation. In all the studies reviewed, whether they focus on Asian developed economies (Abbas et al. 2019), European countries (Raftis et al. 2024), MENA countries (Harb et al. 2023; Saleh and Abu Afifa 2020), or Southeast Asian markets (Sutanto and Ariefianto 2024), the findings remain consistent: higher NPLs and LLPs are associated with lower returns on assets (ROA) and return on equity (ROE).

Comparative evidence from across countries shows that credit risk has a negative impact on profitability in developed and developing economies, but varies by region. The association is found to be stronger in Asia than in the United States (Abbas et al. 2019), which may be a result of differences in regulatory environment, legal framework, and creditor rights. In a study of 903 European banks, Raftis et al. (2024) find that NPLs and LLPs are key drivers of bank profitability, but that LLPs are positively associated with NIM (the net interest margin), which implies that banks might increase interest spreads to mitigate expected credit losses (Rawashdeh 2025; Rawashdeh and Idris 2025; Rawashdeh et al. 2026).

Emerging market studies support the negative credit risk–profitability nexus and highlight some important contextual nuances. The results presented in Saleh and Abu Afifa (2020) indicate that credit risk is statistically significant and negatively related to ROA and NIM for the Jordanian banks, while the relation between credit risk and ROE is not statistically significant, which may indicate that equity buffers absorb some of the losses of the credit risk. Ben-Ahmed et al. (2023) confirm the positive relationship between credit risk and liquidity risk, as a reduction in cash inflow from credit risk further decreases profitability, thus increasing the risk of profitability crises. The non-linear relationship between credit risk effects and NPLs is illustrated in Harb et al. (2023), which shows that, while conservative lending caused market performance to suffer, once a threshold is passed, higher NPLs could lead to better market performance as a result of higher interest income.

Methodological developments have helped to deepen the understanding of dynamic relations. In the study, Sutanto and Ariefianto (2024) use the Dynamic Common Correlated Effects (DCCE) model with an error correction mechanism on banks in Indonesia; it can be seen that there is a significant negative relationship between NPLs and profitability in the short and long terms, with it taking about 2.11–3.73 months for the process of equilibrating. Their analysis also shows that capital adequacy is statistically insignificant in the analysis of the dynamics of credit risk, suggesting that profitability, rather than capital buffers, is the main internal driver for NPL evolution. These relationships have been disrupted by the COVID-19 pandemic; however, profitability shows a positive relationship with credit risk as banks pursued higher margins and fees to sustain profitability (Rahi et al. 2026a, 2026b). The first null hypothesis of this study is:

H_{01}. *Profitability has no statistically significant effect on credit risk.*

2.2. Liquidity Risk on Credit Risk

Liquidity risk is believed to be a drag on profitability, particularly when it limits lending or increases funding pressure. In a study of 13 Jordanian banks over the period 2010–2019, Saleh and Abu Afifa (2020) conclude that liquidity risk (measured through the liquid asset ratios and the loan-to-deposit ratio) negatively affects profitability, though the association varies across bank size and capital levels. In another study in the MENA region, Harb et al. (2023) find that the practices of liquidity risk management have a significant impact on bank performance, where tighter liquidity conditions are related to lower profitability, based on the analysis of 51 banks from the MENA region over the period 2010–2018. Furthermore, a negative relationship between liquidity and profitability in the MENA region is confirmed by hybrid econometric and machine learning approaches, with credit risk mediating this association (Mousa et al. 2025). However, other studies find mixed or even positive results based on the type of bank, market structure or definition of liquidity. However, the relationship between liquidity risk and profitability is not universal, as illustrated in the study of Malaysian banks by Sufian and Habibullah (2010), which shows that in some institutional settings, liquidity risk can boost the profitability of banks. The effects of liquidity risk also vary among US and Asian banks, according to Abbas et al. (2019), with Asian banks exhibiting less negative effects.

There are several studies that highlight the fact that it is not possible to separate credit risk and liquidity risk. Explicitly addressing the interplay between credit and liquidity risks, Ben-Ahmed et al. (2023) focus on how these risks interact with each other, which can markedly influence bank profitability, and highlight how bank credit and liquidity risk exposures can exacerbate losses. Using aggregation of data from the Middle East and North Africa (MENA region), Abdelaziz et al. (2022) show that credit risk and liquidity risk reinforce each other and that their interaction leads to deteriorating bank performance. The most methodologically advanced evidence is that of Mousa et al. (2025), which, using both econometric and machine learning techniques, demonstrated that the relationship between credit risk and profitability is mediated by liquidity risk, and that the relationship between credit risk and liquidity risk is a better predictor of performance than either risk alone. The second null hypothesis of this study is:

H_{02}. *Liquidity risk has no statistically significant effect on credit risk.*

2.3. ESG Score on Credit Risk

The integration of ESG factors into credit risk assessment has become a central concern for investors, regulators, and financial institutions. The Principles for Responsible Investment's "ESG in Credit Risk and Ratings Initiative" has grown from six participating credit rating agencies in 2016 to 28 by 2022, with investor assets under management rising from USD 16 trillion to USD 41 trillion (Hauch 2026). This institutional momentum reflects a prevailing assumption that superior ESG performance signals stronger corporate resilience and, consequently, lower credit risk. A systematic review by Kiss et al. (2025), processing 61 relevant studies published between 2020 and 2024, confirms that "according to the majority of publications, better ESG performance reduces credit risk".

Barth et al. (2022) demonstrate that ESG performance is associated with lower corporate CDS spreads in US and European firms, employing fixed effects and quantile regressions that address unobserved firm heterogeneity. Ma and Hu (2026) extend this evidence to emerging markets, showing that ESG reduces default risk with robustness checks for endogeneity. In the banking sector, Porenta et al. (2026) find a negative association between ESG and credit risk among EU banks using panel fixed-effects models. Multiple studies from China (Wang et al. 2025; Zhang 2025), Japan (Okimoto and Takaoka 2024), Korea (Jang et al. 2020), and Thailand (Sritanee and Chienwattanasook 2026) corroborate these findings across diverse geographic contexts using the GMM method. Moreover, tight regulations and increasing market interest are leading to an ever-expanding number of studies on the ESG and credit risk aspect, with most studies indicating that ESG leads to a reduction in credit risk (Kiss et al. 2025).

According to Frydrych (2026), prior studies have shown that the empirical evidence linking ESG factors to corporate credit risk assessment is either weak or non-existent. There are several specific ways in which this ambiguity comes out. One is that the relationship is not necessarily linear. Truc and Hoa (2025) recognize that there is a U-shaped relationship in SE Asian firms: ESG performance may be positively associated with the default risk below a threshold but negatively associated with the default risk above the threshold. The same holds true for Sritanee and Chienwattanasook (2026), who discover a threshold effect and decreasing returns to scale in Thai corporations in terms of the marginal risk reduction benefit of ESG. Secondly, there are mixed effects on the governance pillar. According to Abdul Razak et al. (2023) and Li et al. (2024), governance performance lowers credit risk; while according to Le et al. (2025) and Yang et al. (2023), governance performance is insignificant or even counterproductive in terms of lowering credit risk. The findings of Wang et al. (2025) reveal that the impact of governance is more mixed for manufacturing firms in China.

Third, context-dependent effects are found in cross-country and region analyses. This result runs counter to that of Porenta et al. (2026), who do not detect any meaningful connection in non-European banks and even notice inverse dynamics in some emerging-market banks. In contrast to the prevailing view, Passas et al. (2025) find that in Europe, higher ESG scores correspond with lower credit ratings when using ordered-probit models. Fourth, the level of disagreement between rating agencies, or ESG rating divergence, causes an extra layer of complexity. The rating similarities between providers are weak, and this difference in the ratings in turn raises corporate credit risk (Hauch 2026). The third null hypothesis of this study is:

H_{03}. *ESG score has no statistically significant effect on credit risk.*

2.4. Monetary Freedom on Credit Risk

The economic freedom literature considers monetary freedom as a measure of the pressure that results from the extent of government intervention on pricing and the stability of the monetary environment (Gwartney and Lawson 2003; Murphy 2024). There is a theoretical linkage that is made through a variety of channels. First, monetary freedom allows for interest rate liberalization, which can lead to lower costs of funding and incentivize banks to offer credit facilities to riskier borrowers, potentially boosting default risk (Bianco and Herrera 2025; Bianco 2021). Second, a combination of stable monetary environments and strong institutions and effective regulation can facilitate credit allocation and financial stability (Apergis et al. 2021; Canh et al. 2021). This dual-channel mechanism implies that the impact is not always positive but crucially depends on the interaction between monetary freedom and banking structure, as well as the policy setting (Djebali 2024; Ghosh 2016).

Evidence from the MENA region and the literature reveals that an increase in monetary freedom is related to an increase in bank risk-taking and a decrease in bank stability, especially in the context of risk incentives for banks (Djebali 2024; Mahrous et al. 2020). During more accommodative monetary conditions, the risk premiums banks charge are likely to decline, banks are more likely to expand lending to financially constrained firms, and banks are more likely to rebalance their portfolios toward riskier but higher-yield assets (Bianco and Herrera 2025; Shen et al. 2021). Smaller, less-capitalized, or less-efficient banks seem to be more sensitive to shifts in monetary conditions (Alzoubi et al. 2022; Kwan and Eisenbeis 1997).

When combined with strong regulation, strong institutions and effective macroprudential supervision, monetary freedom can contribute to stability, but it can exacerbate credit risk in environments with weak institutions and lax lending practices (Alzoubi et al. 2022; Canh et al. 2021; Linh et al. 2022). The last null hypothesis in this study is:

H_{04}. *Monetary freedom has no statistically significant effect on credit risk.*

3. Methodology and Mathematical Models

In this study, the methodology is divided into two parts. First, the preliminary tests and panel regression analysis are performed to identify the significant determinants of credit risk (input variables). The sample includes 40 commercial banks from seven Middle Eastern countries: Qatar, Oman, Kuwait, Jordan, the UAE, Saudi Arabia, and Bahrain. The output variable in the data is credit risk, while the input variables are profitability, liquidity risk, ESG score and monetary freedom. The input variables are selected using descriptive statistics and multicollinearity tests (variance inflation factor (VIF) and tolerance) to make sure that the variables are suitable for regression analysis. After performing these preliminary tests, three regression models, namely ordinary least squares (OLS), fixed effects and random effects, are used to examine the effects of the input variables on the output variable. After that, the Hausman test is performed to determine the efficient model. Furthermore, XGBoost machine learning classification is used to capture non-linear relationships and interactions among the variables, and SHAP values are used to interpret feature importance.

The second part deals with the use of the SLAVE fuzzy rule-based model for classification of banks into high and low credit risk categories. The median value of the credit risk data is 0.35, so banks with a credit risk ≥ 0.35 are considered to be in high credit risk, while banks having a credit risk < 0.35 are considered to be in low credit risk. In banking practice, a credit risk score above 35% of the total loan portfolio is considered a critical threshold. It indicates an elevated default risk.

The SLAVE model is trained with three data splits (70/30, 80/20, 90/10) to examine the stability of the classification results. The four input variables are taken into account by the SLAVE model, and the predicted output value is set to 0 (low-risk) or 1 (high-risk). The two-part approach is used to conduct a comprehensive study of the factors determining credit risk, with powerful and complementary results that allow us to identify the factors that affect credit risk in the banking sector in the Middle East (see Figure 1).

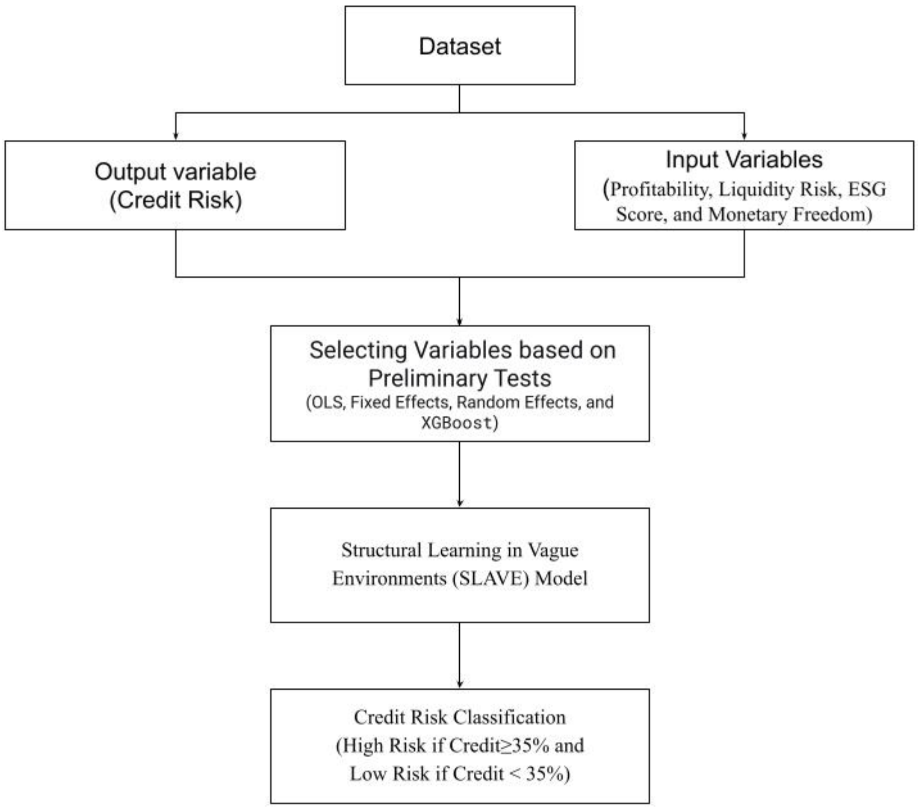


Figure 1. Flowchart based on SLAVE fuzzy model.

3.1. XGBoost Model

An XGBoost Fuzzy model (often referred to as a Fuzzy Gradient Boosting Decision Tree or Fuzzy XGBoost) is a hybrid architecture that combines the robustness of fuzzy logic (handling uncertainty, noise, and imprecise boundaries) with the high performance and scalability of eXtreme Gradient Boosting (XGBoost). While a standard XGBoost model uses hard, binary split conditions, a Fuzzy XGBoost model utilizes fuzzy membership functions to compute a degree of belonging to a branch or leaf node, providing continuous transitions across decision boundaries. In standard XGBoost architecture, predictions are additive. For an ensemble of K trees, the prediction for instance i (Abdelghani et al. 2025; Dhaliwal et al. 2018; Lundberg and Lee 2017)

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i), \quad f_k < \mathcal{F} \tag{1}$$

where $\mathcal{F}$ is the space of CART (Classification and Regression Trees). In Fuzzy XGBoost, the function $f_k(x_i)$ is determined by the combination of leaf values weighted by their fuzzy rule activation strengths (membership degrees). However, the XGBoost model utilizes the adaptive training technique to tune the objective function. Accordingly, every step of the tuning procedure depends on the previous step with respect to the result. The objective function of the XGBoost approach can be mathematically articulated as follows:

$$T^{(t)} = \sum_{i=1}^{n} l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i) + R(f_t) + C \tag{2}$$

where the loss of the i-th round is denoted as l, C is the constant term, and $R(f_t)$ is the regularization parameter of XGBoost, which can be obtained by:

$$R(f_t) = \gamma T_i + \frac{\lambda}{2} \sum_{j-1}^{T} \omega_j^2 \tag{3}$$

In this study, y_i represents the actual credit risk classification (0 for low-risk, 1 for high-risk) for bank I, while $\hat{y}_i^{(t-1)}$ denotes the predicted credit risk probability from the previous boosting round. The function $f_t(x_i)$ represents the contribution of the newly added decision tree at round t, where x_i is the vector of input variables for bank i, including profitability, liquidity risk, ESG score, and monetary freedom. The loss function l measures the discrepancy between the actual credit risk class and the predicted probability, with the binary logistic loss being the appropriate choice for our classification problem. In general, the complexity of the XGBoost tree structure is correlated to the values of γ and λ tuning parameters. The increase in the parameters will result in a simpler tree structure of the model, thereby reducing overfitting and enhancing generalization capability for credit risk prediction. Using a second-order Taylor expansion to approximate the objective function around the previous prediction $\hat{y}_i^{(t-1)}$:

$$T^{(t)} = \sum_{i=1}^{n} \left[l(y_i, \hat{y}_i^{(t-1)}) + g_i f_t(x_i) + \frac{h_i}{2} f_t^2(x_i) \right] + R(f_t) \tag{4}$$

The first and second derivatives of the XGBoost, named g and h, are calculated in the following way:

$$g_i = \frac{\vartheta l(y_i, \hat{y}_i^{(t-1)})}{\vartheta \hat{y}_i^{(t-1)}} h_i = \frac{\vartheta^2 l(y_i, \hat{y}_i^{(t-1)})}{\vartheta (\hat{y}_i^{(t-1)})^2} \tag{5}$$

where g_i represents the first-order gradient (the derivative of the loss function with respect to the predicted credit risk probability), and h_i represents the second-order Hessian (the second derivative), which together enable the algorithm to efficiently optimize the objective function by leveraging second-order approximations, leading to faster convergence and more accurate credit risk predictions. Finally, the solution is determined by applying the following two equations (Alnaser et al. 2026; Liu 2026):

$$\omega_j^* = \frac{\sum g_i}{\sum h_i + \lambda} \tag{6}$$

$$\mathcal{L}_0^* = \frac{1}{2}\sum_{j-1}^{T} \frac{\left(\sum g_i\right)^2}{\sum h_i + \lambda} + \gamma T \tag{7}$$

where the loss function value is denoted by $\mathcal{L}_0^*$, representing the optimized objective function value for credit risk classification, while the solution's weight values are represented by ω_j^*, which correspond to the optimal weights assigned to the leaves of each decision tree. The regularization parameters γ and λ control the complexity of the trees, preventing overfitting to the training data and ensuring that the model generalizes well to unseen banks. The term T represents the number of leaves in the tree, and the summation runs over all leaves, where $\sum g_i$ and $\sum h_i + \lambda$ are aggregated gradient statistics from the data points that fall into each leaf, ultimately determining the optimal weight ω_j^* for each leaf. This XGBoost framework is applied to our dataset of 40 commercial banks across seven Middle Eastern countries, with the goal of accurately classifying banks into high or low credit risk categories based on the four input variables, while the SHAP values derived from the model provide interpretable insights into the relative importance of each predictor variable in determining credit risk outcomes.

3.2. SLAVE Model

The Structural Learning Algorithm on Vague Environments (SLAVE) is a genetic fuzzy rule-based learning system that generates interpretable classification rules through an iterative rule learning approach. SLAVE uses a set-valued model of the rule (González and Pérez 2001):

IF X_1 is A_1 and ... X_n is A_n THEN Y is B

where the value assigned to each antecedent variable, A i, is allowed to be a subset of simple fuzzy values of the domain. For example, let ESG Score be the variable with fuzzy domain D = {*VERY LOW*, *LOW*, *MEDIUM*, *HIGH*, *VERY HIGH*}. An assignment such as *ESG Score* = {*LOW*, *MEDIUM*, *HIGH*} is equivalent to "ESG Score is LOW or MEDIUM or HIGH" and can be interpreted as the convex hull of the fuzzy labels. With this rule model, when the value of a variable coincides with its whole domain, the variable is irrelevant for describing the particular class in the rule and may be eliminated. However, SLAVE follows the genetic iterative approach, learning only one fuzzy rule in each execution of the genetic algorithm (GA). It fixes a class and selects the best antecedent for this class. The complete set of rules is obtained through a sequence of repeated executions by the GA. The rule selection process consists of obtaining the best rule in each execution of the GA based on the notions of completeness and consistency (González and Pérez 2001):

Step 1. Define Positive and Negative examples: The "best rule" in the SLAVE framework is defined by maximizing two primary concepts: **completeness** and **consistency**. These values are derived using the fuzzy cardinalities of "positive examples" ($n^+(R)$) and "negative examples" ($n^-(R)$) for a given rule R.

Step 2. Degree of Completeness: The completeness degree measures how well a rule covers the training examples belonging to its targeted class B:

$$\Lambda(R) = \frac{n^+(R)}{n_B} \tag{8}$$

where n_B is the total number of examples in the training set belonging to class B.

Step 3. Degree of Completeness: To accommodate noise within datasets, SLAVE introduces a soft consistency threshold. First, it defines a set of valid rules Δ^k that do not exceed a fraction k of negative examples relative to positive examples:

$$\Delta^k = \left\{ R \middle| n^-(R) < k \times n^+(R) \right\} \tag{9}$$

Using two fixed parameters, a lower noise boundary (k_1) and an upper noise boundary (k_2) where $k_1 < k_2$, the soft consistency degree $\Gamma_{k_1,k_2}(R)$ is calculated as:

$$\Gamma_{k_1,k_2}(R) = \begin{cases} 1, & if\ R \in \Delta^{k_1} \\ \dfrac{k_2.n^+(R) - \ n^-(R)}{n^+(R)(k_2 - k_1)^2}, & if\ R \notin \Delta^{k_1}\ \wedge\ R \in \Delta^{k_2} \\ 0, & Otherwise \end{cases} \tag{10}$$

Step 4. Optimization Function (Fitness): The GA guides its population optimization task by combining both completeness and soft consistency through a product operator to establish the rule's overall fitness:

$$fitness\ (R) = \Lambda(R) \times \Gamma_{k_1,k_2}(R) \tag{11}$$

The overarching learning objective of the rule extraction module is to maximize these criteria over the fuzzy domain space D for a fixed consequent class B:

$$\max_{A \in D} \{\Lambda(R_B(A)) \times \Gamma_{k_1,k_2}(R)\} \tag{12}$$

where $D = P(D_1) \times P(D_2) \times \ldots \times P(D_n)$ with D_i being the fuzzy domain of X_i variable, $R_B(A)$ represents a rule with antecedent value $A = (A_1, \ldots, A_2) \in D$ and consequent value B, with B being fixed in the optimization problem. The iterative approach will change this consequent value to obtain different values (González and Pérez 2001).

SLAVE model demonstrates competitive performance across the three benchmark datasets, achieving 97.33% accuracy on the Iris dataset, which is the highest or second-highest among all methods, outperforming ksvm (96.00%), random forest (95.33%), and knn (94.67%). On the Pima dataset, SLAVE achieves 72.91%, which is moderate compared to tree and random forest (76.57%), but it outperforms other fuzzy-based methods such as FRBCS.CHI (67.44%), GFS.GCCL (66.54%), and FH.GBML (68.62%), ranking second among fuzzy approaches after fugeR (76.09%). On the Wine dataset, SLAVE achieves 88.17%, outperforming fugeR (89.31%), FRBCS.CHI (92.67%), and tree (92.67%), though ksvm achieves the highest accuracy (98.29%). Overall, SLAVE shows consistent and robust performance across all three datasets, with a notable advantage in interpretability as a fuzzy rule-based model, making it particularly suitable for financial applications where transparent and explainable credit risk assessments are essential (Riza et al. 2015).

3.3. Validation Methodology

To evaluate the performance of our classification models, we employ a comprehensive set of statistical criteria derived from the confusion matrix, which provides a thorough assessment of model effectiveness in distinguishing between high and low credit risk banks. This confusion matrix is structured as a 2 × 2 contingency table that summarizes the classification results—where rows represent the actual classes and columns represent the predicted classes—and is built upon four fundamental components: true positives (TP), which represent the number of high-risk banks correctly classified as high-risk; true negatives (TN), which represent the number of low-risk banks correctly classified as low-risk; false positives (FP), or Type I error, denoting the number of low-risk banks incorrectly classified as high-risk; and false negatives (FN), or Type II error, signifying the number of high-risk banks incorrectly classified as low-risk. Based on these four intersecting components, we compute the subsequent evaluation metrics to rigorously quantify classification performance as defined in (Al Omari et al. 2023).

https://doi.org/10.3390/risks14090194

- The true positive ratio measures the proportion of all banks that are correctly classified as high-risk relative to the total number of banks in the dataset

$$TP\ rate = \frac{TP}{FP + TP + FN + TN} \tag{13}$$

- The true negative ratio measures the proportion of all banks that are correctly classified as low-risk:

$$TN\ rate = \frac{TN}{FP + TP + FN + TN} \tag{14}$$

- The false positive ratio measures the proportion of all banks that are incorrectly classified as high-risk relative to the total number of banks:

$$FP\ rate = \frac{FP}{FP + TP + FN + TN} \tag{15}$$

- The false negative ratio measures the proportion of all banks that are incorrectly classified as low-risk relative to the total number of banks:

$$FN\ rate = \frac{FN}{FP + TP + FN + TN} \tag{16}$$

- Precision measures the proportion of banks classified as high-risk that are actually high-risk:

$$Precision = \frac{TP}{FP + TP} \tag{17}$$

- The F1-score is the harmonic mean of Precision and Recall (TP Rate), providing a single balanced measure of model performance. A higher F1-Score indicates better overall model performance for the correct classification of high- and low-risk banks.

$$F1\text{-}score = \frac{2TP}{2TP + FP + FN} \tag{18}$$

- Accuracy measures the overall proportion of correct classifications made by the model:

$$Accuracy\ (ACC) = \frac{TN + TP}{FP + TP + FN + TN} \tag{19}$$

4. Empirical Results and Discussion

4.1. Data

In this study, an empirical investigation is conducted by using a balanced panel data set of 40 commercial banks from seven countries of the Middle East in the sample period 2014–2023, yielding a total of 400 observations. The sample was restricted to banks listed and reporting their environmental and social scores in the LSEG environmental and social reporting platform, and those banks with missing data were removed, giving rise to the final sample. The sample represents banks from various countries that do business in the UAE, as well as a representative sample of banks from Bahrain, Jordan, Kuwait, Oman, Qatar, Saudi Arabia, and the UAE, reflecting the relative size and diversity of the banking sector in each of these countries. The extent of coverage is in line with the variety and relative size of banking sectors in the region, with a focus on the GCC countries and Jordan as a representative banking market in the MENA region. Jordan's inclusion in the list is especially important because of its unique banking system, with a more advanced regulatory regime; despite the regional geopolitical climate, its financial system has a long history of stability, and its banking system is a link between the Gulf economies and the rest of the Middle East and North Africa region. In addition, Jordan's banking system has features, such as a history of weathering regional crises, which make it a significant comparative case that enriches the cross-country analysis and may help increase the generalizability of our findings to the other oil-rich Gulf economies.

The output variable is the credit risk, which is the ratio of non-performing loans to total gross loans multiplied by 100,000, taken from banks' audited annual financial statements. To ensure the consistency and accuracy of the data across the sample period, the input variables of bank profitability and liquidity are represented by ROE, which is net income divided by total shareholders' equity and the loan-to-deposit ratio, which is total gross loans divided by total deposits, respectively, derived from banks' published financial statements. Furthermore, environmental, social, and governance performance is measured by the composite ESG score derived from the LSEG Data and Analytics database (previously Thomson Reuters' Refinitiv), which includes more than 630 different ESG criteria and represents more than 90% of the global market capitalization, being extensively used in the academic literature and considered to be a reliable source of ESG data (LSEG 2022). Monetary freedom is built in as a macroeconomic control variable using data from Euromoney Country Risk (Euromoney 2025), which includes a measure of the stability and autonomy of monetary policy in each country, which is a measure of economic freedom. R program (4.4.3) was used to run statistical analysis.

4.2. Descriptive Statistics

Table 1 exhibits descriptive statistics for all variables across the full sample and each of the seven countries in our study, giving a preliminary indication of the distributional features and the cross-country differences in the measures of credit risk, profitability, liquidity, ESG and monetary freedom. The credit risk (non-performing loans/total gross loans*100,000) has significant inter-country variation, with the highest mean value being 5.1135 in the UAE, with an unusually high standard deviation value of 11.4033 and maximum value of 57.2966. The dispersion is marked, indicating that there is a high level of heterogeneity between individual UAE banks, possibly meaning that there are strong banks and severely distressed institutions in the UAE banking system, the largest and most diverse in the region.

On the other hand, Saudi Arabia has the lowest mean credit risk of 0.0836 and a standard deviation of 0.0789, indicating the good uniformity of the portfolios of Saudi banks, which could be attributed to the high economic status and conservative lending policies of the kingdom. The mean credit risk levels of Qatar (0.2404) and Kuwait (0.4557) are relatively low, indicating good asset quality management in these oil-rich countries. The statistics for Oman are a mean of 0.5567 and a standard deviation of 0.3551, indicating moderate credit risk exposure, which is a little higher than in the other countries in the Gulf region. It is noteworthy, however, that Jordan has a significantly higher mean credit risk of 2.2222, the second highest among the banks in the Gulf, with a standard deviation of 1.5594, meaning that banks in Jordan are at a much higher risk than their Gulf counterparts of facing asset quality difficulties, probably because of the country's exposure to the geopolitical tensions in the region and its less oil-dependent economy.

In terms of profitability (as reflected in ROE), we see that Saudi Arabia has the highest mean profitability at 0.1062 (10.62%), followed closely by Qatar at ROE of 0.0986 (9.86%). Kuwait and Jordan exhibit mean ROE values of 0.0688 (6.88%) and 0.0794 (7.94%) respectively, while Oman shows a mean of 0.0661 (6.61%). The UAE banking sector, however, recorded the lowest mean profitability of 0.0312 (3.12%), indicating that the earnings performance was relatively less favorable in the UAE banking sector as compared to the other banking sectors in the region during the sample period. The relatively low standard deviations for Saudi Arabia (0.0442), Qatar (0.0470), Oman (0.0273) and Jordan (0.0281) suggest that the profitability performance of the banking sector in these countries has been consistent; importantly, all banks in Qatar, Oman and Jordan had positive ROE values across the whole observation period, indicating that these banking systems have been resilient and their profitability has been stable during the times of economic turmoil in the region.

Table 1. Descriptive statistics for the dataset in different countries.

Variables	Countries	UAE	Saudi Arabia	Qatar	Oman	Kuwait	Bahrain	Jordan
Credit Risk	Mean	5.1135	0.0836	0.2404	0.5567	0.4557	2.1841	2.2222
	St. Dev.	11.4033	0.0789	0.2408	0.3551	0.7473	1.5447	1.5594
	Maximum	57.2966	0.4096	1.0067	1.9208	2.9978	5.4081	6.2991
	Minimum	0.0000	0.0000	0.0000	0.1332	0.0000	0.5959	0.0000
Profitability	Mean	0.0312	0.1062	0.0986	0.0661	0.0688	0.1268	0.0794
	St. Dev.	0.2327	0.0442	0.0470	0.0273	0.0411	0.0185	0.0281
	Maximum	0.3185	0.1920	0.1886	0.1257	0.1684	0.1516	0.1374
	Minimum	−1.6112	−0.0779	0.0156	0.0089	−0.1194	0.0985	0.0204
Liquidity Risk	Mean	0.9111	0.8564	0.9806	0.9914	0.8120	0.6026	0.7174
	St. Dev.	0.1959	0.0889	0.0858	0.0869	0.0886	0.1154	0.0495
	Maximum	1.6708	1.2701	1.1576	1.1513	1.0448	0.7203	0.8216
	Minimum	0.0175	0.6950	0.8242	0.7282	0.6388	0.3653	0.5702
ESG Score	Mean	39.8868	39.8704	49.8091	28.4698	43.0099	52.7443	39.0330
	St. Dev.	19.2453	13.3133	11.2330	15.8499	9.9695	16.2000	12.8214
	Maximum	79.6753	70.0514	73.1037	67.7822	70.7048	87.4243	70.6446
	Minimum	5.9168	17.6680	28.7905	8.9003	28.6086	38.5671	19.0090
Monetary Freedom	Mean	81.4000	86.5000	80.1000	72.6000	68.3000	71.5000	83.6000
	St. Dev.	4.2000	5.7000	1.3000	3.9000	1.9000	3.1000	2.3000
	Maximum	88.0000	90.0000	83.0000	80.0000	70.0000	78.0000	88.0000
	Minimum	76.0000	70.0000	79.0000	68.0000	65.0000	68.0000	80.0000

For liquidity risk (in terms of loan-to-deposit ratio (LDR)), we find that Oman has the highest mean liquidity risk of 0.9914, which shows that Omani banks have a very low liquidity buffer and hence might be impacted by any shock on the funding side or any demand for money withdrawals. With a mean of 0.9806, Qatar comes in very close behind, suggesting similarly aggressive lending practices relative to deposit funding. The UAE has the highest mean LDR of 0.9111, while the UAE banks have the highest standard deviation of 0.1959 and the lowest LDR of 0.0175, indicating that there is a wide spread of liquidity positions across UAE banks and some banks have extremely low LDRs. Saudi Arabia's mean LDR of 0.8564 and Kuwait's of 0.8120 are higher than other Gulf nations, indicating more conservative liquidity management, whereas Jordan's mean LDR is 0.7174 with a remarkably small standard deviation of 0.0495, which might signal that Jordanian banks maintain substantially higher liquidity buffers relative to their Gulf counterparts, which could be considered prudent in light of the country's exposure to regional instability and economic uncertainty.

The ESG scores provide interesting cross-country patterns in sustainability practices of banks in the Middle East. However, Qatar has the highest mean ESG score of 49.8091, indicating that there is stronger performance in the environmental, social and governance sector among Qatar's banks compared to their regional counterparts, which may be due to the strategic investments made in the country in sustainable development and corporate governance reforms. The UAE and Saudi Arabia showed nearly the same ESG mean scores of 39.8868 and 39.8704, respectively, indicating similar ESG integration in these two key banking markets, while Kuwait shows a higher mean score of 43.0099. Jordan has a mean ESG score of 39.0330, which is quite similar to the UAE and Saudi Arabia, while Oman has the lowest score of 28.4698, implying that Omani banks are not as advanced as others in relation to sustainability practices and disclosure. The large standard deviations within countries (especially the UAE 19.2453, Oman 15.8499, Saudi Arabia 13.3133 and Jordan 12.8214) suggest that there is a significant spread in ESG performance within countries, which can be explained by the country-specific size of banks and their commitment to sustainability.

Saudi Arabia had the highest mean monetary freedom (86.5), standard deviation (5.70), and maximum (90.00), suggesting that Saudi Arabia has a fairly stable monetary policy environment and tight inflation control. Jordan has a mean monetary freedom (83.60) with a standard deviation (2.30), highlighting the country's financial stability and good money management, even in the face of regional difficulties. The mean of the UAE is 81.40 with a standard deviation of 4.20, whereas the mean of Qatar is 80.10 with a very small standard deviation of 1.30, which shows that monetary conditions in Qatar are very stable. The least monetary freedom is found in Oman and Kuwait, both with a mean monetary freedom score of 72.60 and with Kuwait showing the lowest standard deviation (1.90), and the lowest range overall, least monetary freedom is found. These cross-country differences in monetary freedom are likely to be associated with varying degrees of monetary policy independence, exchange rate regimes and inflation dynamics, which may have differential impacts on bank credit risk through their impact on interest rates, economic stability and the ability of borrowers to repay.

4.3. Multicollinearity Assessment

This section will discuss the multicollinearity issue, which is caused by a strong correlation between the input variables that make up the regression model. One of the common assumptions that researchers make prior to selecting the variables for a regression model is that they avoid multicollinearity, as it can make the model interpretation difficult and may create overfitting issues. To identify the level of multicollinearity for the multiple regression analysis, VIF and tolerance tests are used. VIF is a statistical term that is a measure of the amount of collinearity. A higher VIF value indicates greater collinearity. This increases the variance of the regression coefficient. The tolerance is computed as 1/VIF, and if the VIF is over four or the tolerance is less than 0.25, the result is generally an indication of multicollinearity. For the correlation analysis, if the correlation coefficient value is less than 0.5, there is a weak correlation; if the value is greater than or equal to 0.5, there is a strong correlation. To conclude that there is no multicollinearity issue, the independent variables should exhibit weak correlations with one another.

The correlation matrix and collinearity statistics for the independent variables used in our analysis are shown in Table 2. The correlation coefficients show that credit risk has a significant negative correlation with profitability at −0.606, which means that the banks with higher profit margins have a lower non-performing loan ratio. Credit risk has a weak positive correlation with liquidity risk at 0.403. Banks with higher loan-to-deposit ratios have higher credit risk exposure. Credit risk also demonstrates a weak negative correlation with ESG Score at −0.341, indicating that banks with superior environmental, social, and governance performance tend to exhibit lower credit risk exposure. Additionally, credit risk displays a weak positive correlation with monetary freedom at 0.132. Among the independent variables, profitability exhibits a weak negative correlation with liquidity risk at −0.320. Profitability shows a weak positive correlation with ESG Score at 0.241, while profitability exhibits a weak negative correlation with monetary freedom at −0.075.

https://doi.org/10.3390/risks14090194

Table 2. Multicollinearity tests.

Variables	Correlations					Collinearity Statistics	
	Credit Risk	Profitability	Liquidity Risk	ESG Score	Monetary Freedom	Tolerance	VIF
Credit Risk	1	−0.606	0.403	−0.341	0.132		
Profitability		1	−0.32	0.241	−0.075	0.867	1.153
Liquidity Risk			1	−0.283	0.033	0.852	1.174
ESG Score				1	0.068	0.887	1.127
Monetary Freedom					1	0.986	1.015

The collinearity statistics give additional confidence that there is no multicollinearity in the model. VIF values are significantly below the threshold of four, and the tolerance values are significantly above 0.25, across all variables. The results showed that multicollinearity is not a threat to the reliability of the multiple regression estimates, because the independent variables in the study have low linear interdependency. The low VIF values mean that correlation among the input variables does not significantly affect the variance of each regression coefficient, which helps to maintain the stability and interpretability of the estimated coefficient values. Therefore, there is no need to worry about multicollinearity influencing the results when using all four input variables (profitability, liquidity risk, ESG Score, monetary freedom) in multiple regression models.

4.4. Multiple Regression Models

This section applies OLS, fixed effects, and random effects to examine the effectiveness of the input variables (profitability, liquidity risk, ESG score, and monetary freedom) on the output variable (credit risk) as shown in Table 3. The estimation results of the OLS and FE models for the determinants of credit risk in Middle Eastern banks are presented in Table 3. While the OLS model assumes homogeneity across countries, we estimate effects based on different countries in the fixed effects and random effects models to capture country-specific effects. This helps us address the unobserved heterogeneity at the country level that may influence credit risk outcomes. We have used country-level effects because our objective is to examine input variables across different countries. These variables are measured at the country level. Future research could extend our analysis by including bank and year fixed effects.

The OLS model shows good explanatory power with an R-squared of 0.4495 and an adjusted R-squared of 0.4439, which means that the independent variables together can explain around 44.5% of the variation in credit risk in the sample. The overall statistical significance of the model is confirmed by the F-statistic (80.63, p-value < 2.2×10^{-16}). The constant term is negative and statistically significant at -7.7188 ($p < 0.05$), which represents the baseline credit risk level when all independent variables are zero. The profitability variable ($\beta = -24.4045$) has a significant negative effect on credit risk at a significance level of less than 1%, which rejects H_{01}, indicating that for every one-unit increase in the profitability ratio, credit risk will decrease by approximately 24.4 units. The liquidity risk variable ($\beta = 7.9020$) has a significant positive effect on credit risk at a considerable level of less than 1%, rejecting H_{02}. The variable of ESG Score ($\beta = -0.0729$) has a significant negative effect on credit risk at a considerable level of less than 1%, which rejects H_{03}, implying that banks with superior environmental, social, and governance performance maintain lower credit risk levels. The variable of monetary freedom ($\beta = 0.0974$) has a significant positive effect on credit risk at a considerable level of less than 1%.

Table 3. OLS, fixed effects, and random effects models.

Models		Intercept	Profitability	Liquidity Risk	ESG Score	Monetary Freedom
OLS	Estimate	−7.7188	−24.4045	7.902	−0.0729	0.0974
	Std. Error	3.3174	1.977	1.6677	0.0166	0.0367
	t-stat	−2.327	−12.344	4.738	−4.401	2.655
	p-value	0.0205	$<2 \times 10^{-16}$	0.000	0.000	0.0083
	R-squared		0.4495		F-stat.	80.63
	Adjusted R-squared		0.4439		*p*-value	$<2.2 \times 10^{-16}$
Fixed effect	Estimate	-	−20.2954	16.1675	−0.0885	0.1295
	Std. Error	-	1.8672	1.9248	0.0161	0.0633
	t-stat	-	−10.8697	8.3995	−5.4945	2.0464
	p-value	-	$<2.2 \times 10^{-16}$	0.000	0.000	0.0414
	R-squared		0.5103		F-stat.	101.336
	Adjusted R-squared		0.4977		*p*-value	$<2.22 \times 10^{-16}$
Random effect: (Wallace–Hussain's transformation)	Estimate	−14.043	−20.8059	14.8177	−0.0863	0.1083
	Std. Error	4.9342	1.8708	1.8872	0.0161	0.0588
	z-value	−2.8461	−11.1214	7.8517	−5.3594	1.8405
	p-value	0.0044	$<2.2 \times 10^{-16}$	0.000	0.000	0.0657
	R-squared		0.4967		Chi (χ^2)	390.517
	Adjusted R-squared		0.4916		*p*-value	$<2.22 \times 10^{-16}$
	Hausman Test with Fixed effect: Chi (χ^2) = 11.991, df = 4, *p*-value = 0.01742					
Random effect: (Amemiya's transformation)	Estimate	−14.5552	−20.6998	15.1018	−0.0867	0.1121
	Std. Error	5.0223	1.867	1.8925	0.0161	0.0597
	z-value	−2.8981	−11.087	7.98	−5.3949	1.8792
	p-value	0.0038	$<2.2 \times 10^{-16}$	0.000	0.000	0.0602
	R-squared		0.4996		Chi (χ^2)	394.797
	Adjusted R-squared		0.4945		*p*-value	$<2.22 \times 10^{-16}$
	Hausman Test with Fixed effect: Chi (χ^2) = 3.735, df = 4, *p*-value = 0.4431					

The fixed effects model, which controls unobserved country-specific characteristics that remain constant over time, provides an R-square of 0.5103 and an adjusted R-square of 0.4977, suggesting that the model accounts for about 51% of the variation in credit risk across countries. The F-statistic of 101.336 with a *p*-value of $<2.2 \times 10^{-16}$ confirms the overall significance of the model. Notably, the fixed effects model reveals several differences in coefficient magnitudes compared to the OLS estimates, highlighting the importance of controlling country-level heterogeneity. The profitability term is negative and highly significant at −20.2954 ($p < 0.01$), which rejects H_{01}. The variable of liquidity risk (β = 16.1675) has a significant positive effect on credit risk at a considerable level of less than 1%, which rejects H_{02}. The variable of ESG Score (β = −0.0885) has a significant negative effect on credit risk at a considerable level of less than 1%, which rejects H_{03}. The variable of Monetary freedom (β = 0.1295) has a significant positive effect on credit risk at a substantial level of less than 5%, which rejects H_{04}.

In order to control for country-specific random effects and the robustness of our results, we estimate random effects models using two alternative transformations besides the fixed effects model, namely Wallace–Hussain's transformation and Amemiya's transformation. The adjusted R-square is 0.4916, indicating that the model explains 49.7% of the variation in credit risk. The overall significance of the model is supported by the chi-square statistic of 390.517 with a *p*-value of $<2.2 \times 10^{-16}$. The intercept term is negative and statistically significant at −14.0430 ($p < 0.01$), indicating the baseline credit risk level after accounting for country-specific random effects. The variable of profitability (β = −20.8059) has a significant negative effect at a considerable level (less than 1%), which rejects H_{01}. This result is consistent with the OLS and supports the theory that higher profitability is associated with lower credit risk levels of banks. The variable of liquidity risk (β = 14.8177) has a significant positive impact on credit risk at a very significant level (less than 1%), which rejects H_{02}, meaning that the higher the ratio of loans to deposits, the more the bank is exposed to credit risk. The variable of ESG Score ($\beta = -0.0863$) has a significant negative effect on credit risk at a considerable level of less than 1%, which rejects H_{03}, supporting the notion that superior sustainability performance contributes to lower credit risk. The variable of monetary freedom ($\beta = 0.1083$) is not statistically significant at conventional levels ($p = 0.0657$), indicating that we fail to reject H_{04} for this specification at the 1% and 5% significance levels; however, at the 10% significance level, we reject H_{04}, suggesting that monetary freedom does not have a significant effect on credit risk when using the Wallace–Hussain transformation.

The Amemiya random effects model yields an R-squared of 0.4996 and an adjusted R-squared of 0.4945, indicating that the model explains approximately 49.96% of the variation in credit risk. The chi-square statistic of 394.797 with a p-value of $<2.2 \times 10^{-16}$ confirms the overall significance of the model. The intercept term is negative and statistically significant at -14.5552 ($p < 0.01$), representing the baseline credit risk level after accounting for country-specific random effects using Amemiya's transformation. The variable of profitability ($\beta = -20.6998$) has a significant negative effect on credit risk at a considerable level of less than 1%, which rejects H_{01}, consistent with the Wallace–Hussain and OLS findings. The variable of liquidity risk ($\beta = 15.1018$) has a significant positive effect on credit risk at a considerable level of less than 1%, which rejects H_{02}, confirming the robust positive relationship between liquidity risk and credit risk across all model specifications. The variable of ESG Score ($\beta = -0.0867$) has a significant negative effect on credit risk at a considerable level of less than 1%, which rejects H_{03}, providing further evidence that ESG performance is negatively associated with credit risk. The variable of monetary freedom ($\beta = 0.1121$) is not statistically significant at conventional levels ($p = 0.0602$), indicating that we fail to reject H_{04} for this specification at the 1% and 5% significance levels; however, at the 10% significance level, we reject H_{04}, suggesting that monetary freedom has a marginally significant positive effect on credit risk when employing the Amemiya random effects estimator.

To determine the appropriate specification between fixed effects and random effects models, we conduct the Hausman test, which evaluates whether the unique errors are correlated with the regressors. The Hausman test comparing the fixed effects model with the Wallace–Hussain random effects model yields a chi-square statistic of 11.991 with four degrees of freedom and a p-value of 0.01742. Since the p-value is less than 0.05, we reject the null hypothesis that random effects are consistent, indicating that the fixed effects model is preferred over the Wallace–Hussain random effects specification. This suggests that there is a systematic difference between the coefficients estimated by the fixed effects and random effects models, and that the fixed effects model provides more reliable estimates by controlling for unobserved country-level heterogeneity that is correlated with the independent variables. However, the Hausman test comparing the fixed effects model with the Amemiya random effects model yields a chi-square statistic of 3.735 with four degrees of freedom and a p-value of 0.4431. Since the p-value is greater than 0.05, we fail to reject the null hypothesis, suggesting that the Amemiya random effects model is consistent and may be more efficient than the fixed effects model.

4.5. The XGBoost Model Using SHAP Values

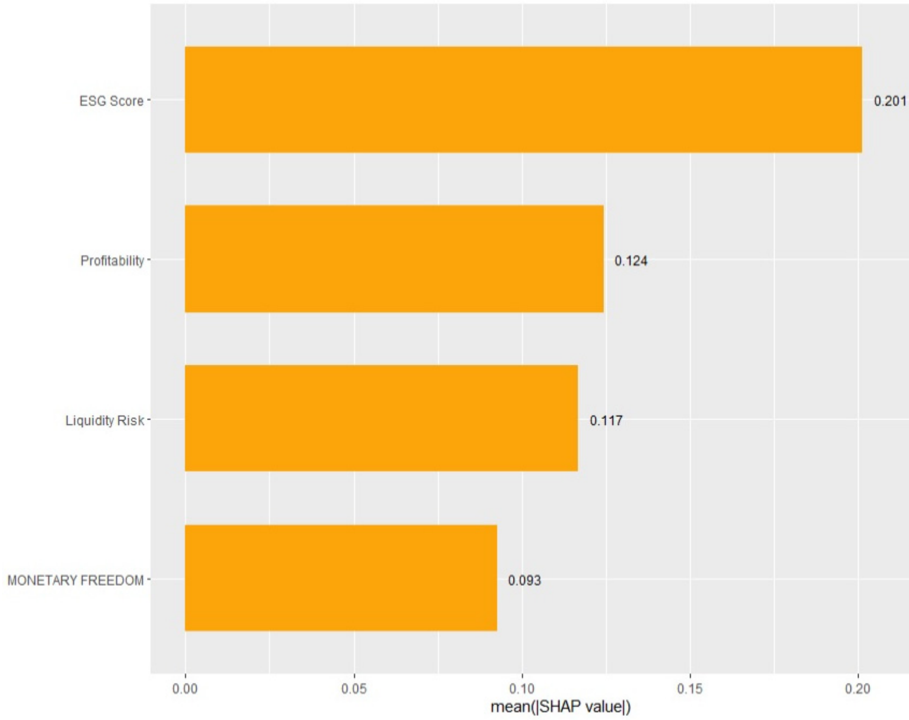
This section uses eXtreme Gradient Boosting (XGBoost). It is a non-parametric and tree-based machine learning approach. It is used alongside panel regression analysis because it can account for non-linear relationships and interactions between the input variables. The credit risk output variable is a continuous variable and is converted to a binary and categorical variable for ease of classification. The median value of credit risk in our sample is 0.35; consequently, banks with credit risk values equal to or greater than 0.35 are classified as exhibiting high credit risk and assigned a value of one, while banks with credit risk values below 0.35 are classified as exhibiting low credit risk and assigned a value of zero. For the XGBoost implementation, we used the following tuning procedures and parameters in our R code: learning rate (eta) = 0.1, maximum tree depth (max_depth) = 6, subsample ratio (subsample) = 0.8, and number of boosting rounds (nrounds) = 100, with the objective function set to "reg:squarederror" for regression.

The use of the median as the threshold for classification is based on the following basic statistical concepts: the central limit theorem and the characteristics of the median as a robust measure of central tendency. When the credit risk distribution is skewed and/or contains influential observations, the median is a more representative central value than the mean, since it is not affected by extreme values and outliers. The median is a reliable and solid measure because it is not as sensitive to extreme values of credit risk as the mean is, and because it is representative of the typical credit risk level in our sample of banks. We chose the median as our classification threshold so that the high-risk and low-risk groups are balanced in size to prevent class-imbalance problems that can affect the performance of a machine learning classifier and result in skewed predictions for the majority class.

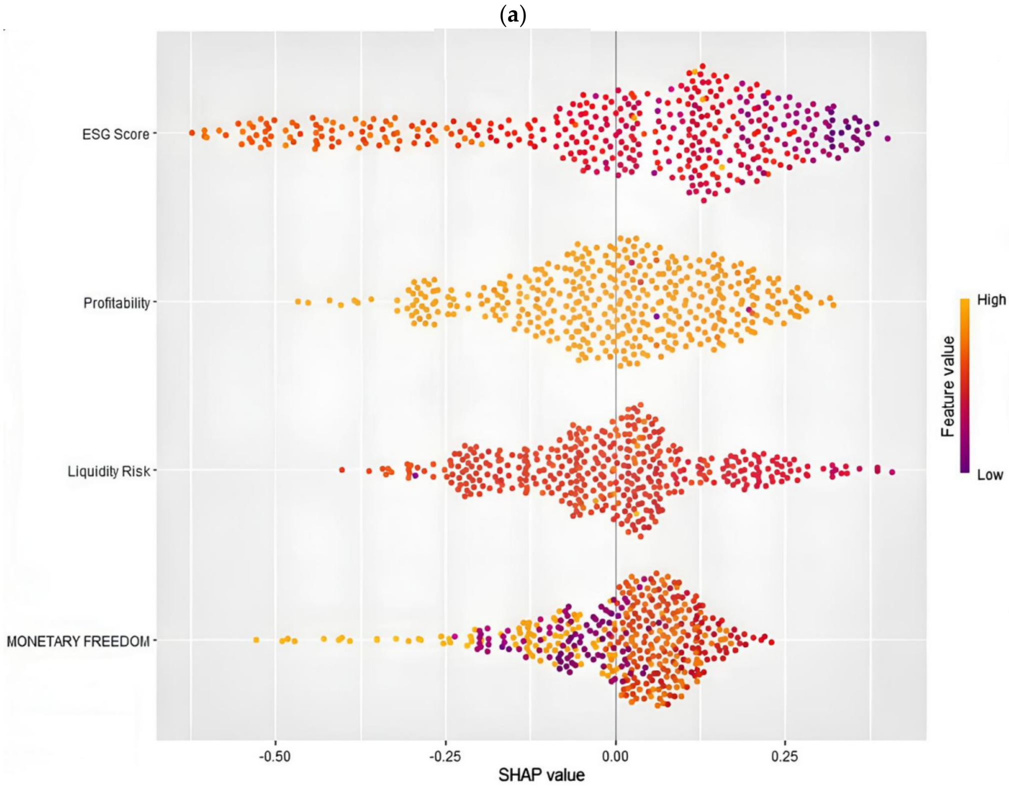
XGBoost provides interpretable measures of feature importance, which can be used to determine which of our input variables (profitability, liquidity risk, ESG score, and monetary freedom) is most relevant and significant in determining whether a bank is high or low credit risk. This also adds to our regression analysis, as it will reveal non-linearities and interaction effects that linear panel models may fail to explain. Using XGBoost on our binary credit risk outcome provides a way to confirm or challenge our results of the panel regressions and gives an extra perspective on the non-linear relationship of the credit risk determination process of Middle Eastern banks, which can be easily seen in the credit risk output. When applied to our binary credit risk outcome, our credit risk conclusions from our panel regressions can be validated, and an added perspective is gained by the complex and non-linear aspects of credit risk determination in Middle Eastern banks shown in our credit risk output (Lundberg and Lee 2017).

Figure 2 shows the SHAP interpretation of the XGBoost credit risk prediction model: In Figure 2a, the feature importance analysis provides the list of features that are most significant in the prediction model. In Figure 2b, a histogram of the feature contributions reveals the distribution and the direction of the feature contributions. The predicted credit risk is predominantly driven by ESG Score (mean SHAP = 0.193), as revealed in Figure 2a, which indicates that this feature has the greatest share of explanatory power among all other features. In line with recent empirical studies on the incorporation of sustainability metrics into financial decision-making (Dorfleitner et al. 2015; Henisz and McGlinch 2019), these findings point to the increasing materiality of environmental, social and governance aspects in the assessment of credit risk today.

Liquidity Risk (mean SHAP value = 0.120) and Profitability (mean SHAP value = 0.116) are second and third most important, respectively, and have similar explanatory power, about 62% and 60% of the ESG Score's importance, respectively, and confirm the longstanding relevance of traditional financial fundamentals in credit risk assessment, as documented in the existing literature (Altman 1968, Brunnermeier and Pedersen 2009). Monetary Freedom (mean SHAP = 0.109) is fourth, contributing a non-trivial but still relatively small amount that shows the impact of the general macroeconomic environment on credit risk, and the mean SHAP for the three variables differ by only a small amount, indicating the three are about equal contributors to the model.



(a)



(b)

Figure 2. SHAP features for the XGBoost credit risk prediction model. (**a**) Bar chart; (**b**) Dot plot.

Moving on to Figure 2b, the summary plot shows the magnitude and direction of all the predictor variables' contributions, and ESG Score has the largest range of SHAP values between about −0.6 and +0.6, and it has a large concentration of positive contributions around +0.6, suggesting that ESG Score not only has the largest overall influence but it is also very variable across the observations. The Profitability variable is bounded between about −0.6 and +0.6, with a clear clustering of values above +1.0 for some very high profitability observations, indicating that a very high profitability performance may lead to disproportionately high impact on credit risk reduction, and highlighting the non-linear relationship between profitability and creditworthiness where above-average performance has more than double the positive impact on credit risk assessments. Both Liquidity Risk and Monetary Freedom show spread in SHAP values similar to the one observed in the credit risk predictions and have a significant concentration of negative SHAP values, with Liquidity Risk having a wider distribution of SHAP values, ranging from about −0.6 to +0.6, but with a clear concentration below 1.0, and Monetary Freedom having a narrower distribution of SHAP values, ranging from about −0.6 to +0.6, with a concentration near 1.0, suggesting that the magnitude and direction of the effects of changes in the liquidity and/or monetary environment, respectively, on the credit risk predictions are relatively stable and moderate, and are consistent with their respective mean SHAP values of 0.120 and 0.109. The overall spread of SHAP values of ESG Score and Profitability indicates that there is significant variation in the credit risk dimension between different banks.

Figure 3 shows a country-wise comparison of the importance of the SHAP values for the XGBoost credit risk prediction model for seven different countries, namely Bahrain, Jordan, Kuwait, Oman, Qatar, Saudi Arabia, and the UAE, and is shown in stacked bars to allow for cross-country comparisons of the contributions of the predictors. What is apparent in the visualization is the almost complete overlap among the countries regarding the feature importance, with mean SHAP values of ~0.33 for Liquidity Risk, ESG Score, Profitability, and Monetary Freedom for most of the countries (Bahrain, Kuwait, Qatar, Saudi Arabia, and the United Arab Emirates). The uniformity implies that the factors driving credit risk are of similar relative importance in these closely intertwined markets in the Gulf, consistent with the coordinated monetary policy and the similarities of the financial systems in the Gulf regions.

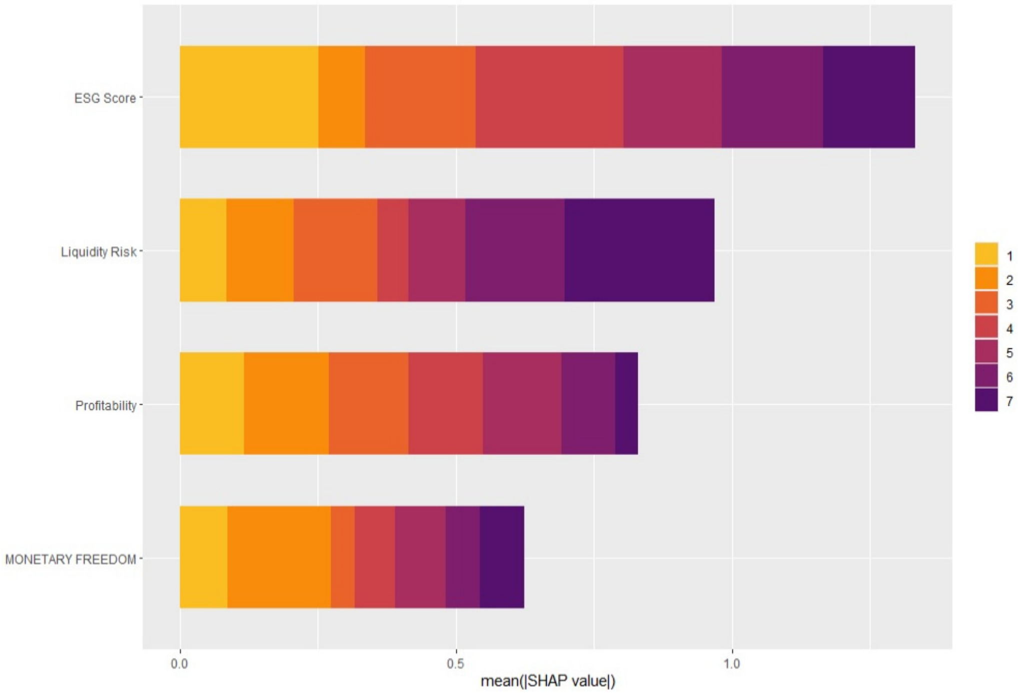


Figure 3. SHAP feature of countries for the XGBoost credit risk prediction model.

Figure 3 also shows significant variations between the mean SHAP values in the predictions of the ESG Score and Monetary Freedom for Oman and Jordan, which stand at around −0.10, suggesting that these act as much less important drivers of credit risk predictions in these countries than their peers in the region. Oman's finding might be linked to the country's economic structure, which is more heavily reliant on hydrocarbon revenues, and a more concentrated financial structure, which may dampen the impact of ESG factors on credit assessment. Likewise, Jordan's deviation could be explained by its distinct geopolitical situation, as well as by the fact that limited natural resources might be more oriented towards liquidity and profitability considerations rather than sustainability indicators. The above figure further illustrates that Liquidity Risk and Profitability are relatively stable in importance across all countries, with no country deviating notably from the regional average, indicating that the fundamentals of credit risk evaluation—Liquidity Risk and Profitability—remain relevant across all countries irrespective of the national context (Altman 1968; Brunnermeier and Pedersen 2009).

Collectively, the results underscore the diversity of ESG integration among the countries in the GCC region, with most countries showing consistency between sustainability performance and creditworthiness, and Oman and Jordan showing significant outliers that require further research on country-specific institutional and economic factors that influence the materiality of ESG factors.

5. Structural Learning in Vague Environments Model

In this section, we use the Structural Learning in Vague Environments (SLAVE) model to classify banks according to their credit risk levels. This section complements our previous analyses with a fuzzy rule-based learning approach that captures uncertainty and vagueness inherent in credit risk assessment. The SLAVE model is used to classify whether the data value of the credit series is high-risk or low-risk with the help of a fuzzy rule extraction methodology in order to extract understandable rules that explain the relationship between the input variables and the classification of the credit risk. The four input variables (ESG score, profitability, liquidity risk and monetary freedom) are assumed to be predictors, and the output credit risk value is binary and equal to zero if its value is below the median 0.35 and one otherwise. The data were split into three possible training and testing split for the model robustness testing, namely 70:30 (280:120), 80:20 (320:80) and 90:10 (360:40). The number of classes was two (low-risk, high-risk), number of linguistic labels for each input variable was two, cross validation percentage was 0.9, maximum number of iterations was 10, maximum number of generations was set as 10, mutation percentage was set as 0.3 and the epsilon threshold was set as 0.8, the simulation was designated as "sim-0".

Table 4 presents the classification results of the SLAVE model for the three different sample splits, where the true positive (TP) and true negative (TN) rates, false positive (FP) and false negative (FN) rates, F1-score, precision and accuracy are reported for each matrix. As shown in the matrices, the model is over-performing on the high-risk banks in all splits compared to the low-risk banks, which is a significant advantage for the SLAVE approach, as it is able to reflect the features of financially strong banks. For the 70/30 split (280 training, 120 testing), the confusion matrix shows 42 true low-risk, eight true high-risk, 18 false high-risk, and 52 false low-risk classifications. For the 80/20 split (320 training, 80 testing), the matrix shows 29 true low-risk, five true high-risk, 11 false high-risk, and 35 false low-risk classifications. For the 90/10 split (360 training, 40 testing), the matrix shows 15 true low-risk, zero true high-risk, seven false high-risk, and 18 false low-risk classifications.

The analysis of the three splits shows significant trends in model performance in terms of the criteria used for the classification. For the 70/30 split (280 training, 120 testing), the model achieved a TP Rate of 0.3500, a TN Rate of 0.0667, and an F1-score of 0.1860. For the 80/20 split (320 training, 80 testing), the model yielded a TP Rate of 0.3625, a TN Rate of 0.0625, and an F1-score of 0.1786. For the 90/10 split (360 training, 40 testing), the model produced a TP Rate of 0.3750, a TN Rate of 0.0000, and an F1-score of 0.0000. As more data is used for training (from 70% to 90%), the TP Rate improves slightly from 0.3500 to 0.3750, and thus a larger training sample will help the model to learn how to recognize the high-risk banks. The TN Rate drops significantly, however, from 0.0667 to 0.0000. The F1-score reduces from 0.1860 to 0.1786 and finally to 0.0000, which corresponds to a decrease in the mean of precision and recall, especially if the test data is small (40 observations) and thus not sufficient for performance evaluation. The results in this work provide evidence that a larger training set is important for enhancing the identification of high-risk instances, but that the number of test points must be kept sufficient for a good assessment of the model.

Table 4. Predictive classification statistics for credit risk.

Sample Size	**Train (70%)/Test (30%)**		**Train (80%)/Test (20%)**		**Train (90%)/Test (10%)**	
	280	**120**	**320**	**80**	**360**	**40**
Matrix	**Low**	**High**	**Low**	**High**	**Low**	**High**
Low	42	18	29	11	15	7
High	52	8	35	5	18	0
TN Rate	0.35		0.3625		0.375	
TP Rate	0.0667		0.0625		0	
FP Rate	0.4333		0.4375		0.45	
FN Rate	0.15		0.1375		0.175	
F1-score	0.186		0.1786		0	
Precision	0.3077		0.3125		0	
Accuracy	0.4167		0.425		0.375	

6. Limitations and Directions for Future Research

This study employed XGBoost classification and SLAVE fuzzy rule-based modeling to detect high and low credit risk levels of commercial banks across seven Middle Eastern countries. Although our analysis was as complete as it could be, some limitations exist that could stimulate further research. First, we used profitability (ROE), liquidity risk (loan-to-deposit ratio), ESG score, and monetary freedom as input variables in our models. Future research could expand the model to include additional indicators and bank-specific variables such as capital adequacy ratios, asset quality indicators, bank size, and operational efficiency metrics. Researchers could apply feature selection techniques to identify the most predictive combinations of these variables and improve classification model capability.

Second, we focused on commercial banks from seven Middle Eastern countries (UAE, Saudi Arabia, Qatar, Oman, Kuwait, Jordan and Bahrain). Future experiments could be extended to include banks in other developed areas like the United States, the United Kingdom and Europe. Cross-regional comparative studies would reveal differences in ESG–credit risk relationships across diverse regulatory and economic environments, thereby enhancing the generalizability of our results. Third, year-to-year financial information was used, though it spans the COVID-19 pandemic years, which could also be considered in the future to be a longer time horizon and/or more macroeconomic variables added to the dataset, such as GDP growth, inflation rates, and market risks, to capture the dynamic nature of how credit risk evolves. Rolling-window analysis and time-varying coefficient models could better capture the dynamic nature of how credit risk evolves over time in response to changing economic conditions.

Fourth, the classification accuracy obtained by the SLAVE model shows that the current set of parameters gives an accuracy of up to 42.5%, so additional studies may involve considering optimizing model parameters, other fuzzy logic methods and developing hybrid models that integrate fuzzy systems with neural networks and/or evolutionary algorithms. Specifically, hybrid architectures combining fuzzy systems with convolutional neural networks could enhance pattern recognition capabilities and improve prediction accuracy. Fifth, this study uses ESG scores as a new measure of credit risk, but ESG data come from LSEG and might not cover unique sustainability efforts specific to the Middle East, while future research could create region-specific ESG metrics that better identify the unique environmental, social, and governance challenges in the Middle East among banks. Weighted ESG indices incorporating regional sustainability priorities would offer more contextually relevant assessments.

Finally, a dichotomization of credit risk was made according to the median value, which is statistically sound, but may not reflect the whole spectrum of credit risk exposure. Future research could investigate whether multi-class classification approaches or continuous risk scoring methods give a more fine-grained and accurate classification of risk levels.

7. Conclusions and Recommendations

This study achieved the prediction of the credit risk of commercial banks in seven countries of the Middle East using XGBoost classification and SLAVE fuzzy rule-based modeling. The multicollinearity tests revealed that the independent variables—profitability, liquidity risk, ESG score and monetary freedom—are appropriate for inclusion, with all VIF values under four and all tolerance values over 0.25. The OLS results reveal that profitability ($\beta = -24.4045$, $p < 0.01$) and ESG score ($\beta = -0.0729$, $p < 0.01$) have significant negative effects on credit risk, while liquidity risk ($\beta = 7.9020$, $p < 0.01$) and monetary freedom ($\beta = 0.0974$, $p < 0.01$) have significant positive effects. The fixed effects model reveals that profitability ($\beta = -20.2954$, $p < 0.01$) and ESG score ($\beta = -0.0885$, $p < 0.01$) have negative effects on credit risk, whereas liquidity risk ($\beta = 16.1675$, $p < 0.01$) and monetary freedom ($\beta = 0.1295$, $p < 0.05$) have positive effects. The Hausman test favors the fixed effects model over Wallace–Hussain ($\chi^2 = 11.991$, $p = 0.01742$) but not over Amemiya ($\chi^2 = 3.735$, $p = 0.4431$).

The output variable is derived from 400 observations in the UAE, Saudi Arabia, Qatar, Oman, Kuwait, Jordan and Bahrain over a period of 10-years from 2014 to 2023. The XGBoost model was implemented, and SHAP interpretation showed that credit risk and ESG (mean SHAP = 0.193) have the largest influence on the model, with Profitability (mean SHAP = 0.126), Liquidity Risk (mean SHAP = 0.120), and Monetary Freedom (mean SHAP = 0.100) following. ESG Score and Profitability have the highest dispersion of SHAP values, reflecting significant heterogeneity in the credit risk profile of banks. The SLAVE model was applied for three data splits (70/30, 80/20, and 90/10), with credit risk values ≥ 0.35 (the median) considered as high-risk (1) and <0.35 as low-risk (0). The accuracy obtained by the SLAVE model ranged from 0.3750 to 0.4250, with the best accuracy of 0.4250 achieved by the 80/20 split. As more training data was added, the TP rate went up from 0.3500 to 0.3750 while the F1-score went down from 0.1860 to 0.0000.

The results suggest that sustainability matters in banking risk assessment as the ESG score remains a consistent and significant predictor of credit risk. A higher ESG rating is correlated with lower credit risk, as are more stable financial conditions. We recommend that banks apply these indicators to bolster risk management and regulators integrate ESG factors into the supervisory framework to encourage sustainable banking and reduce regional credit risk.

Author Contributions: Conceptualization, J.J.J. and Q.A.S.M.; methodology, J.J.J. and A.A.R.; software, J.J.J.; validation, J.J.J. and A.A.R.; formal analysis, J.J.J. and A.A.-G.; investigation, A.A.K.A., Q.A.S.M., A.A.-G. and T.A.K.; resources, Q.A.S.M.; data curation, A.A.K.A., A.A.R. and A.A.-G.; writing—original draft, J.J.J., A.A.K.A., A.A.R. and A.A.-G.; writing—review and editing, Q.A.S.M.; visualization, Q.A.S.M. and A.A.-G.; supervision, A.A.R., A.A.-G. and T.A.K.; project administration, A.A.K.A. and T.A.K.; funding acquisition, A.A.K.A., A.A.R. and T.A.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no additional funding.

Data Availability Statement: The data presented in this study are available in the Euromoney Institutional Investor PLC Annual Report & Accounts (Euromoney 2025) [https://www.euromoney.com]. Furthermore, LSEG is the source of the ESG scores [https://www.lseg.com/en] (accessed on 10 June 2026).

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

References

Abbas, Faisal, Shahid Iqbal, and Bilal Aziz. 2019. The impact of bank capital, bank liquidity and credit risk on profitability in postcrisis period: A comparative study of US and Asia. *Cogent Economics & Finance* 7: 1605683. [CrossRef]

Abdelaziz, Hakimi, Boussaada Rim, and Hamdi Helmi. 2022. The interactional relationships between credit risk, liquidity risk and bank profitability in MENA region. *Global Business Review* 23: 561–83. [CrossRef]

Abdelghani, Rana M., Iman Morsi, Abd-Elhady B. Kashyout, Walid El-Shafie, and Taha El-Sayed. 2025. Comparative Analysis of XGBoost, Fuzzy Logic, and Hybrid Models for Acetone, Ethanol, and Methanol Gas Concentration Prediction. Paper presented at the 2025 International Telecommunications Conference (ITC-Egypt), Cairo, Egypt, July 28–31.

Abdul Razak, Lutfi, Mansor H. Ibrahim, and Adam Ng. 2023. Environment, social and governance (ESG) performance and CDS spreads: The role of country sustainability. *The Journal of Risk Finance* 24: 585–613. [CrossRef]

Akkoç, Soner. 2012. An empirical comparison of conventional techniques, neural networks and the three stage hybrid Adaptive Neuro Fuzzy Inference System (ANFIS) model for credit scoring analysis: The case of Turkish credit card data. *European Journal of Operational Research* 222: 168–78. [CrossRef]

Alnaser, Firas, Samar Rahi, Mahmoud Alghizzawi, and Abdul Hafaz Ngah. 2026. An integrative research framework to investigate factors influencing citizen's intention to adopt e-health applications: Post-COVID-19 perspective. *Global Knowledge, Memory and Communication* 75: 1263–84. [CrossRef]

Al Omari, Rania, Rami S. Alkhawaldeh, and Jamil J. Jaber. 2023. Artificial neural network for classifying financial performance in Jordanian insurance sector. *Economies* 4: 106. [CrossRef]

Altman, Edward I. 1968. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance* 4: 589–609. [CrossRef]

Alzoubi, Marwan, Alaa Alkhatib, Ayman Abdalmajeed Alsmadi, and Hamad Kasasbeh. 2022. Bank size and capital: A trade-off between risk-taking incentives and diversification. *Banks and Bank Systems* 17: 1–11. [CrossRef]

Anaz, Ameera Fatima, and Antoine B. Awad. 2025. ESG and corporate creditworthiness: An ordinal-logistic exploration of global markets. *Journal of Financial Reporting and Accounting*, 1–49. [CrossRef]

Apergis, Nicholas, Ahmet F. Aysan, and Yassine Bakkar. 2021. How do institutional settings condition the effect of macroprudential policies on bank systemic risk? *Economics Letters* 209: 110123. [CrossRef]

Aslan, Aydin, Lars Poppe, and Peter Posch. 2021. Are sustainable companies more likely to default? Evidence from the dynamics between credit and ESG ratings. *Sustainability* 13: 8568. [CrossRef]

Barth, Florian, Benjamin Hübel, and Hendrik Scholz. 2022. ESG and corporate credit spreads. *The Journal of Risk Finance* 23: 169–90. [CrossRef]

BCBS. 2004. *International Convergence of Capital Measurement and Capital Standards: A Revised Framework*. Basel: Bank for International Settlements.

BCBS. 2017. *Basel III: Finalising Post-Crisis Reforms*. Basel: Bank for International Settlements.

Bellotti, Tony, and Jonathan Crook. 2009. Support vector machines for credit scoring and discovery of significant features. *Expert Systems with Applications* 36: 3302–08. [CrossRef]

Ben-Ahmed, Kais, Naziha Kasraoui, and Asaad Soulami. 2023. Exploring the interplay of credit and liquidity risks: Impacts on banks' profitability. *International Journal of Advanced and Applied Sciences* 10: 64–70. [CrossRef]

Bessis, Joel. 2015. *Risk Management in Banking*, 4th ed. Hoboken: John Wiley & Sons.

Bianco, Timothy. 2021. Monetary policy and credit flows. *Journal of Macroeconomics* 70: 103362. [CrossRef]

Bianco, Timothy, and Ana María Herrera. 2025. Monetary policy and credit flows: A tale of two effective lower bounds. *Journal of Economic Dynamics and Control* 175: 105084. [CrossRef]

Brunnermeier, Markus K., and Lasse Heje Pedersen. 2009. Market liquidity and funding liquidity. *The Review of Financial Studies* 6: 2201–38. [CrossRef]

Canh, Nguyen Phuc, Christophe Schinckus, Thanh Dinh Su, and Felicia Hui Ling Chong. 2021. Institutional quality and risk in the banking system. *Journal of Economics, Finance and Administrative Science* 26: 22–40. [CrossRef]

Dastile, Xolani, Turgay Celik, and Moshe Potsane. 2020. Statistical and machine learning models in credit scoring: A systematic literature survey. *Applied Soft Computing* 91: 106263. [CrossRef]

Deng, Guoying, Shibo Ma, Jingzhou Yan, Can Shuai, and Hanying Liu. 2024. Dissecting the impact of the three E, S, G pillars on credit risk. *Economic Analysis and Policy* 83: 301–13. [CrossRef]

Dhaliwal, Sukhpreet Singh, Abdullah-Al Nahid, and Robert Abbas. 2018. Effective intrusion detection system using XGBoost. *Information* 9: 149. [CrossRef]

Djebali, Nesrine. 2024. Assessing the determinants of banking stability in the MENA region: What role for economic and financial freedom? *Journal of Banking Regulation* 25: 127–44. [CrossRef]

Dorfleitner, Gregor, Gerhard Halbritter, and Mai Nguyen. 2015. Measuring the level and risk of corporate responsibility–An empirical comparison of different ESG rating approaches. *Journal of Asset Management* 16: 450–66. [CrossRef]

Dumitrescu, Elena, Sullivan Hué, Christophe Hurlin, and Sessi Tokpavi. 2022. Machine learning for credit scoring: Improving logistic regression with non-linear decision-tree effects. *European Journal of Operational Research* 297: 1178–92. [CrossRef]

Euromoney. 2025. Annual Report & Accounts 2025 [Euromoney Database]. Available online: https://www.euromoney.com/ (accessed on 10 June 2026).

Farah, Bassam, Rida Elias, Ruth Aguilera, and Elie Abi Saad. 2021. Corporate governance in the Middle East and North Africa: A systematic review of current trends and opportunities for future research. *Corporate Governance: An International Review* 29: 630–60. [CrossRef]

Frydrych, Sylwia. 2026. Transition Risk and ESG Materiality in Credit Rating Downgrades: Evidence from Explainable Machine Learning. *Finance Research Letters* 107: 110408. [CrossRef]

Ghosh, Saibal. 2016. Does economic freedom matter for risk-taking? Evidence from MENA banks. *Review of Behavioral Finance* 8: 114–36. [CrossRef]

González, Antonio, and Raúl Pérez. 2001. Selection of relevant features in a fuzzy genetic learning algorithm. *IEEE Transactions on Systems, Man, and Cybernetics, Part B* 31: 417–25. [CrossRef]

Gwartney, James, and Robert Lawson. 2003. The concept and measurement of economic freedom. *European Journal of Political Economy* 19: 405–30. [CrossRef]

Harb, Etienne, Rim El Khoury, Nadia Mansour, and Rima Daou. 2023. Risk management and bank performance: Evidence from the MENA region. *Journal of Financial Reporting and Accounting* 21: 974–98. [CrossRef]

Hauch, Selina. 2026. ESG rating divergence and corporate credit risk. *The Journal of Risk Finance* 27: 132–57. [CrossRef]

Henisz, Witold J., and James McGlinch. 2019. ESG, material credit events, and credit risk. *Journal of Applied Corporate Finance* 31: 105–17. [CrossRef]

Hull, John C. 2023. *Risk Management and Financial Institutions*. Hoboken: John Wiley & Sons.

Jang, Ga-Young, Hyoung-Goo Kang, Ju-Yeong Lee, and Kyounghun Bae. 2020. ESG scores and the credit market. *Sustainability* 12: 3456. [CrossRef]

Kiss, Boglarka, Dániel Homolya, and György Walter. 2025. The Relationship between ESG Factors and Corporate Credit Risk: Research Trends from a Bibliometric Perspective. *Financial and Economic Review* 24: 96–117. [CrossRef]

Kwan, Simon, and Robert A. Eisenbeis. 1997. Bank risk, capitalization, and operating efficiency. *Journal of Financial Services Research* 12: 117–31. [CrossRef]

Le, Thuong Thi, Tanvir Bhuiyan, Thi Le, and Ariful Hoque. 2025. Exploring Governance Failures in Australia: ESG Pillar-Level Analysis of Default Risk Mediated by Trade Credit Financing. *Journal of Risk and Financial Management* 18: 464. [CrossRef]

Li, Quan, Tianshu Li, and Yuan Zhang. 2024. ESG performance, heterogeneous creditors, and bond financing costs: Firm-level evidence. *Finance Research Letters* 66: 105527. [CrossRef]

Linh, Nguyen Tran Xuan, Nguyen Ngoc Thach, and Nguyen Ngoc Tan. 2022. Monetary Policy, Macroprudential Policy, Institutional Quality and Bank Risk: Evidence from Eagle Group. In *Prediction and Causality in Econometrics and Related Topics. ECONVN 2021.*. Cham: Springer.

Liu, Weida. 2026. Price prediction model of cultural products based on improved XGBoost. In *Second International Conference on Communication, Information, and Digital Technologies (CIDT 2025)*. Bellingham: SPIE.

LSEG. 2022. London Stock Exchange Group ESG Scores Database. Available online: https://www.lseg.com/en (accessed on 10 June 2026).

Lundberg, Scott M., and Su-In Lee. 2017. A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems 30*. Red Hook: Curran Associates, Inc.

Ma, Yujie, and Wenxiu Hu. 2026. Does ESG performance mitigate corporate debt default risk? Evidence from emerging markets. *Finance Research Letters* 102: 110079. [CrossRef]

Machado, Marcos R., Daniel Tianfu Chen, and Joerg R. Osterrieder. 2025. An analytical approach to credit risk assessment using machine learning models. *Decision Analytics Journal* 16: 100605. [CrossRef]

Mahrous, Sherif Nabil, Nagwa Samak, and Mamdouh Abdelmoula M. Abdelsalam. 2020. The effect of monetary policy on credit risk: Evidence from the MENA region countries. *Review of Economics and Political Science* 5: 289–304. [CrossRef]

Marqués, Ana I, Vicente García, and José Salvador Sánchez. 2012. Exploring the behaviour of base classifiers in credit scoring ensembles. *Expert Systems with Applications* 39: 10244–50. [CrossRef]

Maside-Sanfiz, José Manuel, Ana Iglesias-Casal, Qusay Ayman Sulayman Mazahreh, and Mª Celia López-Penabad. 2024. The impact of competition on environmental and social performance in the MENA banking sector. *Corporate Social Responsibility and Environmental Management* 31: 6290–317. [CrossRef]

Miao, Xia, and Zi Nie. 2026. Credit risk early warning for listed companies based on ESG information: Evidence from China. *Applied Economics Letters* 33: 637–40. [CrossRef]

Mousa, Rojer, Julia Nabil, Ahmed Safty, Ibrahim Hassan, and Yara Ibrahim. 2025. Liquidity–credit risk dynamics and bank profitability: Hybrid econometric and machine learning evidence from MENA. *Journal of Financial Reporting and Accounting*, 1–27. [CrossRef]

Murphy, Ryan H. 2024. Mapping inflation to economic freedom in the post-COVID era. *Journal of Institutional Economics* 20: e16. [CrossRef]

Okimoto, Tatsuyoshi, and Sumiko Takaoka. 2024. Sustainability and credit spreads in Japan. *International Review of Financial Analysis* 91: 103052. [CrossRef]

Passas, Ioannis, Dimitrios I Vortelinos, Christos Lemonakis, Voicu D Dragomir, and Stavros Garefalakis. 2025. Impact of Environmental, Social, and Governance (ESG) Scores on International Credit Ratings: A Sectoral and Geographical Analysis. *Sustainability* 17: 9755. [CrossRef]

Porenta, Jan, Iza Matko, and Riste Ichev. 2026. Sustainability and stability: ESG–credit risk dynamics in EU banks. *Economic and Business Review* 28: 35–58. [CrossRef]

Raftis, Achilleas, Christos Karpetis, Stephanos Papadamou, and Eleftherios Spyromitros. 2024. Monetary winds of change: Exploring the link between policy shifts and bank profitability in developed and emerging European markets. *Global Finance Journal* 59: 100932. [CrossRef]

Rahi, Samar, Mazuri Abd Ghani, Firas Mohammed Alnasr, Mahmoud Alghizzawi, and Aamir Rashid. 2026a. What influences consumer's intention to adopt Islamic digital wallet? Decoding the role of diffusion of innovation and Hofstede's cultural model. *Journal of Islamic Marketing* 17: 37–53. [CrossRef]

Rahi, Samar, Mazuri Abd Ghani, Mahmoud Alghizzawi, Aamir Rashid, and Rizwana Rasheed. 2026b. Understanding the electric vehicle consumers' AWE experience and intention to recommend electric vehicles with the moderating effect of trust in sustainable manufacturers. *Asia-Pacific Journal of Business Administration* 18: 314–33. [CrossRef]

Rawashdeh, Awni. 2025. Catalysts of transformation: Unravelling the nexus between regulatory changes, perceived efficiency gains, and business performance in blockchain implementation. *International Journal of Productivity and Performance Management* 74: 3577–600. [CrossRef]

Rawashdeh, Awni, and Mohammed Idris. 2025. Synchronized curriculum adaptation: A strategy for enhancing skill development in accounting education. *Journal of International Education in Business* 18: 269–88. [CrossRef]

Rawashdeh, Awni, Esam Shehadeh, Anwar Salameh Al-Gasaymeh, and Amani Ben Tareef. 2026. Student perceptions of ChatGPT's educational value: Exploring motivational correlates in accounting education. *Asian Education and Development Studies*, 1–23. [CrossRef]

Riza, Lala Septem, Christoph Bergmeir, Francisco Herrera, and José M. Benítez. 2015. frbs: Fuzzy rule-based systems for classification and regression in R. *Journal of Statistical Software* 65: 1–30. [CrossRef]

Runchi, Zhang, Xue Liguo, and Wang Qin. 2023. An ensemble credit scoring model based on logistic regression with heterogeneous balancing and weighting effects. *Expert Systems with Applications* 212: 118732. [CrossRef]

Saleh, Isam, and Malik Abu Afifa. 2020. The effect of credit risk, liquidity risk and bank capital on bank profitability: Evidence from an emerging market. *Cogent Economics and Finance* 8: 1814509. [CrossRef]

Saunders, Anthony, Marcia Millon Cornett, and Otgo Erhemjamts. 2026. *Financial Institutions Management: A Risk Management Approach*, 11th ed. New York: McGraw Hill.

Shen, Qingyuan, Binbin Chang, Ruifeng Ma, and Xiaoyan Tao. 2021. How does interest rate liberalisation influence the profits of commercial banks? Evidence from China. *Transformations in Business and Economics* 20: 278.

Sohn, So Young, Dong Ha Kim, and Jin Hee Yoon. 2016. Technology credit scoring model with fuzzy logistic regression. *Applied Soft Computing* 43: 150–58. [CrossRef]

Sritanee, Norrasate, and Krisada Chienwattanasook. 2026. Diminishing returns: The non-linear impact of ESG ratings on credit risk in Thai corporations–evidence from the Stock Exchange of Thailand. *Cogent Economics and Finance* 14: 2665539. [CrossRef]

Straus, Justus V. D., Valeri K. Andreev, Tijn Koeleman, Thomas Sepanosian, and Marcos R. Machado. 2025. Explaining to defend: SHAP-based defense mechanism against adversarial attacks in P2P credit risk. *Expert Systems with Applications* 301: 130432. [CrossRef]

Sufian, Fadzlan, and Muzafar Shah Habibullah. 2010. Does economic freedom fosters banks' performance? Panel evidence from Malaysia. *Journal of Contemporary Accounting and Economics* 6: 77–91. [CrossRef]

Sutanto, Hendra, and Moch Doddy Ariefianto. 2024. The dynamic relationships of credit risk, profitability, and capital: Evidence from Indonesia. *Asian Economic and Financial Review* 14: 191–207. [CrossRef]

Talaat, Fatma M., Abdussalam Aljadani, Mahmoud Badawy, and Mostafa Elhosseini. 2024. Toward interpretable credit scoring: Integrating explainable artificial intelligence with deep learning for credit card default prediction. *Neural Computing and Applications* 36: 4847–65. [CrossRef]

Tian, Haowen, and Gaoliang Tian. 2022. Corporate sustainability and trade credit financing: Evidence from environmental, social, and governance ratings. *Corporate Social Responsibility and Environmental Management* 29: 1896–908. [CrossRef]

Truc, Anh Nguyen Thi, and Le Thanh Hoa. 2025. The nonlinear impact of ESG on firm default risk: Evidence from Southeast Asia's firms. *International Journal of Sustainable Economy* 17: 423–41. [CrossRef]

UNEP FI. 2021. *Promoting Sustainable Finance and Climate Finance in the Arab Region*. Geneva: UNEP FI.

UNEP FI. 2023. *Adapting to a New Climate in the MENA Region*. Geneva: UNEP FI.

Wang, Yanan, Xiao Zhang, Michal Wojewodzki, Yuxin Jian, and Fadey AbiDaoud. 2025. ESG Performance and Credit Risk: Evidence from Chinese Manufacturing Companies. *International Journal of Finance and Economics* 31: 3720–46. [CrossRef]

Wilkinson, Connor, John Yawney, and Stephen Andrew Gadsden. 2026. Explaining explainability: A comprehensive survey on explainable artificial intelligence and relevant industry applications. *Intelligent Systems with Applications* 30: 200647. [CrossRef]

Yang, Xin, Ahmad Fahmi Sheikh Hassan, Wei Theng Lau, and Nazrul Hisyam Ab Razak. 2023. The discordance of governance performance from environmental and social performance on idiosyncratic risk: The effect of board composition. *Cogent Economics and Finance* 11: 2276556. [CrossRef]

Zhang, Meng-Yao. 2025. The Impact of Corporate ESG Performance on Debt Default Risk: Evidence from China. *Journal of Mathematics* 2025: 8873471. [CrossRef]