

Article

Not peer-reviewed version

Machine Learning-Based Early Warning System For Banking Crises: A Case Study Of Nigeria

[Ntanganedzeni Mandiwana](#), [Thakhani Ravele](#)*, [Caston Sigauke](#), [Rendani Netshikweta](#)

Posted Date: 27 May 2026

doi: 10.20944/preprints202605.1839.v1

Keywords: banking crises; crisis prediction; early warning systems; emerging markets; random forest; XGBoost







Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC, OpenAlex.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Machine Learning-Based Early Warning System for Banking Crises: A Case Study of Nigeria

Ntanganedzeni Mandiwana ¹, Thakhani Ravele ^{1,*}, Caston Sigauke ¹
and Rendani Netshikweta ²

¹ Department of Mathematical and Computational Sciences, University of Venda,
Private Bag X5050, Thohoyandou 0950, Limpopo, South Africa

² Department of Mathematics and Applied Mathematics, University of Limpopo,
Private Bag X1106, Sovenga 0727, Limpopo, South Africa

* Correspondence: thakhani.ravele@univen.ac.za (T.R.); Tel.: +27-15-962-8098

Abstract

Banking crises pose a constant threat to macroeconomic stability in emerging markets, where standard econometric Early Warning Systems (EWS) often fail to model nonlinear macro-financial relationships. This paper examines whether machine learning algorithms, rather than standard logistic regression, can improve forecasts of banking crisis risk in Nigeria. We compare the performance of Random Forests, Support Vector Machines (SVMs), and Extreme Gradient Boosting (XGBoost) to logistic regression on the African Financial Crises dataset (1954-2014) with annual data. Resampling is restricted to the training set to compensate for the rarity of crisis instances. In a strict out-of-time validation setting, the model's accuracy is assessed by accuracy, precision, recall, F1-score, and the area under the receiver operating characteristic curve (AUC). Our findings show that tree-based ensemble models outperformed logistic regression on the test set: XGBoost generalises better (AUC = 1.0; F1 = 0.95 for non-crisis, 0.80 for crisis) instances, although Random Forest yields the highest cross-validated F1-score on the training set. Exchange rate volatility, inflation, systemic crisis variables, and defaults on external sovereign debt are identified as key predictors through feature importance analysis. Crisis years exhibit the strongest predictive signals, suggesting that annual data have limited early-warning capacity. Due to the small sample size and lack of crisis observations during the test period, results should be interpreted cautiously. All things considered, the findings provide strong early evidence that, although not yet ready as fully functional policy tools, machine learning models can support conventional tools for tracking banking crises in Nigeria.

Keywords: banking crises; crisis prediction; early warning systems; emerging markets; random forest; XGBoost

1. Introduction

1.1. Overview

In emerging economies like Nigeria, where the banking industry plays a crucial role in both financial intermediation and economic development, predicting banking crises is a vital component of financial stability analysis. The early 1990s, late 1990s, and the global financial crisis of 2009–2014 revealed structural flaws in Nigeria's banking system, such as high non-performing loan levels, lax regulatory oversight, and susceptibility to macroeconomic and external shocks [1,2]. These recurrent crises highlight the crucial need to build trustworthy early warning systems to spot financial vulnerabilities before they become systemic disruptions.

Traditionally, econometric models such as logit and probit regression models have been employed in Early Warning Systems (EWS) for banking crises. These models are the most favoured models by financial regulators and policymakers due to their simplicity and clarity [3]. However, their usefulness is constrained by the high degree of linearity and the models' inability to capture complex relationships

among macro-financial variables. In poorly structured nations like Nigeria, with nonlinear patterns and high limitations, this can lead to high false-positive rates and reduced accuracy, limiting their use in applications and, therefore, should be reformed to achieve good results in these countries [3].

The latest developments in machine learning (ML) have rendered alternative methods for crisis prediction more suitable for identifying intricate interactions and nonlinear associations in high-dimensional financial information [4]. Empirical research has found that machine learning models, including Random Forest, Support Vector Machines (SVM), and Extreme Gradient Boosting (XGBoost), are more precise than traditional econometric models in predicting the occurrence of financial crisis in multiple empirical studies of financial crisis predictions in developed countries and emerging markets alike, such as China and India, respectively. Their respective financial reporting systems [5–7]. Ensemble learning-based methods have demonstrated enormous potential, even in the turbulent economic environment, to identify early warning signs concealed in macro-financial variables [6].

Although research on this subject worldwide has increased, there is a gap in knowledge of empirical studies on the use and effectiveness of machine learning (ML)-based early warning systems in the banking sector in Nigeria. This is particularly relevant because Nigeria has been prone to banking instability in the past and requires precise crisis-prediction tools [2]. The present study will fill this gap by examining the predictive performance of various machine learning methods in developing an early warning system for a financial crisis in Nigeria. In particular, the research enhances the random forests, support vector machines (SVMs), and extreme gradient boosting (XGBoost) models using the African Financial Crises dataset created by the author [8]. These models are then compared to the predictive performance of a baseline Logistic Regression model. The effectiveness and robustness of the model are evaluated using various measures, including accuracy, precision, recall, F1-score, and the area under the Receiver Operating Characteristic curve (ROC-AUC).

First, the study is important because it will add to the current body of knowledge in both the academic field and policy-making. The research in the Nigerian case studies compares machine learning models with the more traditional logistic regression model to provide empirical data on the trade-offs between the accuracy of predicted results and the interpretability of early warning systems. The findings of this study will allow practitioners, financial regulators, and policymakers to adopt more information-based approaches to prevent crises, thereby strengthening the financial stability base in Nigeria and the banking sector's stability globally to mitigate cybernetic economic shocks.

1.2. Literature Review

1.2.1. Traditional Early Warning Systems for Banking Crises

Early Warning Systems (EWS) are created to identify weaknesses in the banking system so that immediate policy measures can be taken. In econometric models such as probit and logit regression, macro-financial variables are mostly employed in conventional EWS models, including debt service ratios, exchange rate volatility, and credit-to-GDP ratios, among others [9]. Financial regulators and policymakers find these models very attractive due to their simplicity and ease of interpretation.

However, several notable flaws in traditional EWS models exist. These models have a linear structure, which complicates the determination of the nonlinear characteristics of financial regimes. This leads to a high false-positive rate, rendering these models useless. The Nigerian structure's reliance on oil payables, its vulnerability to external shocks, and the inaccessibility of data further limit the predictability of conventional EWS models [2].

Recent research has shown the importance of having a more flexible early warning system. Reference [10] emphasises the importance of early warning indicators being consistent with policymakers' decision-making cycles, and [11] illustrates that sophisticated statistical and computational techniques can be used to improve predictive accuracy. The Nigerian banking crisis of 2009, precipitated by risk-taking, commodity price shocks, and stock market exposure, illustrates the importance of an EWS that can identify both country-specific vulnerabilities and global financial linkages.

1.2.2. Machine Learning Methods for Banking Crisis Prediction

Machine learning (ML) has been identified as a promising new approach to predicting banking crises, distinct from traditional econometric models. Unlike linear regression, machine learning methods can capture complex, nonlinear relationships and interactions in high-dimensional data [12]. Machine learning is particularly useful for early warning systems in unstable, fundamentally vulnerable financial systems because of its flexibility.

Random Forests, Support Vector Machines (SVMs), and Extreme Gradient Boosting (XGBoost) have shown improved predictive accuracy over traditional methods [13,14]. Although logistic regression is a good baseline model because of its interpretability, machine learning methods based on ensembles and kernels provide significant improvements in accuracy and resistance, especially during times of increased economic stress.

In merging economies, an early-warning system is particularly significant, and an ML-driven system can deliver significant benefits [1]. Conversely, another article [15] argues that exogenous shocks, including increases or decreases in oil prices, may cause financial instability in oil-exporting nations, such as Nigeria. Compared to classical models, those based on ensemble techniques such as Random Forests and XGBoost are more likely to identify nonlinear correlations among macro-financial variables, enabling stronger predictions about a branch of financial matters of interest [16].

Despite the common outperformance of machine learning models over conventional econometric models in the prediction domain, their black box properties may make them difficult to accept by policymakers in need and desirous of understanding the effects of input variables on forecasts, one variable at a time, to guide policy formulation and execution [17]. To overcome this weakness, explainable artificial intelligence (XAI) models, including Shapley Additive Explanations (SHAP), have been proposed to improve model transparency and estimate the contribution of individual features, thereby aiding in determining the essential causes of financial instability.

Although SHAP and similar XAI methods are effective in other settings, this study aims to assess the ability of machine learning models to detect early warning signs of financial crises in Nigeria. Even though this falls out of the research, the use of explainable AI tools is also essential to discuss. This weakness provides future studies with an opportunity to enhance predictive power and interpretation.

Although there is growing evidence of machine learning's effectiveness worldwide, its application in Nigeria has not been widely used for predicting crises. A majority of studies with a Nigerian focus use conventional econometric methods, which often yield high false-positive rates and cannot capture the dynamics of nonlinearity. Moreover, the comparison of logistic regression and machine learning models in the Nigerian case is not done properly. Consequently, policymakers lack empirical evidence on which modelling method offers the best trade-off between interpretability and forecast accuracy. This research fills these gaps by comparing the performance of Random Forest, SVM, and XGBoost with a baseline logistic regression model. This work has value for researchers and policymakers because it analyses projected performance and clearly states that the implementation of explainable AI could only be pursued in future research.

Table 1 in the companion source (literature comparison) gives a summary of the literature on how banking crises are predicted in the literature and how the methodology has been changing over time through the transition of traditional econometric models to machine learning algorithms, and how it may be applied to new markets, like Nigeria.

This paper models the Nigerian banking crisis using logistic regression, Random Forests, Support Vector Machines, and XGBoost.

Table 1. Overview of banking crisis prediction strategies.

Author/Year	Methodology Type	Key Features	Performance Metrics	Limitations
[9]	Traditional EWS: Logit/Probit	Signal extraction approach	False alarm rates	Linear assumptions; high false positives
[1]	Historical analysis	Crisis incidence analysis	Frequency analysis	Descriptive only; not predictive
[2,15]	Systemic crisis database	Crisis dating	Systemic risk indicators	Retrospective; limited prediction
[10]	Credit-to-GDP gap	Single-indicator EWS	Signal extraction	Country-specific limitations
[11]	Multivariate logit	Stress testing	AUROC, Calibration	Data requirements for emerging markets
[4]	XGBoost	Ensemble learning	Accuracy	Black-box nature
[16]	SVM, Random Forests	Kernel methods	AUROC	Limited emerging economy focus
[13]	ML vs Traditional	Comparative analysis	AUROC	Data constraints
[14]	Ensemble methods	Random Forests	F1-score, Accuracy	Interpretability challenges
[6]	XGBoost, Random Forests	Ensemble learning	AUROC, F1-score	European focus only
[17]	XAI with SHAP	Explainable AI	SHAP values	Accuracy-interpretability trade-off
[3]	Hybrid EWS	Bayesian methods	Predictive accuracy	Implementation complexity

1.3. Contributions and Research Highlights

By systematically assessing the predictive accuracy of machine learning models such as Random Forest, XGBoost, and Support Vector Machines against the traditional logistic regression model in the event of a banking crisis in Nigeria, this study, therefore, adds to the existing literature on financial stability and early warning systems. The main contribution of this research is to show that ML-based methods, especially XGBoost, achieve better predictive performance and validity in designing effective early warning systems in developing countries. The highlights of this study are:

- Three machine learning models, Random Forest, XGBoost, and SVM, were assessed and compared to logistic regression to forecast financial crises in Nigeria.
- XGBoost came in second to Random Forest, which performed the best during cross-validation on the training data with high F1-scores, recall, and ROC-AUC values.
- XGBoost demonstrated its durability on unseen data by achieving the best overall performance on the test set, with the greatest accuracy and crisis-class F1-score.
- Exchange rate volatility, inflation, systemic crises, and sovereign debt defaults are important macro-financial indicators that underscore the role of economic fundamentals at the onset of crises.
- Leveraged estimated crisis probability to build a workable Early Warning System (EWS) that correctly recognised every past crisis inside the Red alert zone, demonstrating the policy's applicability.
- Provides policymakers with advice on model selection and risk management, and empirical evidence to justify the use of machine learning models for crisis forecasting in Nigeria.

This research aims to determine whether advanced machine learning techniques outperform traditional logistic regression in forecasting banking crises in Nigeria. As the dataset provides only one observation per year, the study uses an exploratory benchmarking method. It acknowledges the limitations of the available data, as well as the small number of crisis events. The procedure of this study will be the selection of the features and detailed data preprocessing. Multiple machine learning algorithms are run to capture a potential nonlinear relationship in the system, including Support Vector Machines (SVMs), Extreme Gradient Boosting (XGBoost), Random Forests, and Logistic Regression.

Since this dataset is very imbalanced, the metrics used to evaluate model performance are the F1 score and the Receiver Operating Characteristic Area Under the Curve (ROC AUC), as these are considered relevant performance measures in practice.

The rest of the paper is structured as follows. Section 2 outlines the modelling framework used in this study. The empirical findings are presented in Section 3, and the performance of the models is compared in Section 4. Finally, Section 5 provides the concluding remarks.

2. Methods

This section explains the methodological framework that was used to test the hypothesis that machine learning-based Early Warning Systems (EWS) are useful in predicting banking crises in Nigeria. The aim is to ensure it is transparent, replicable, and robust, as outlined in the early warning systems literature. The methodology will include data description and preprocessing, gradient-boosted feature selection, model training with traditional econometric and machine learning methods, and model testing and explainability through cross-validation and rigorous out-of-time assessments. Figure 1 summarises the workflow.

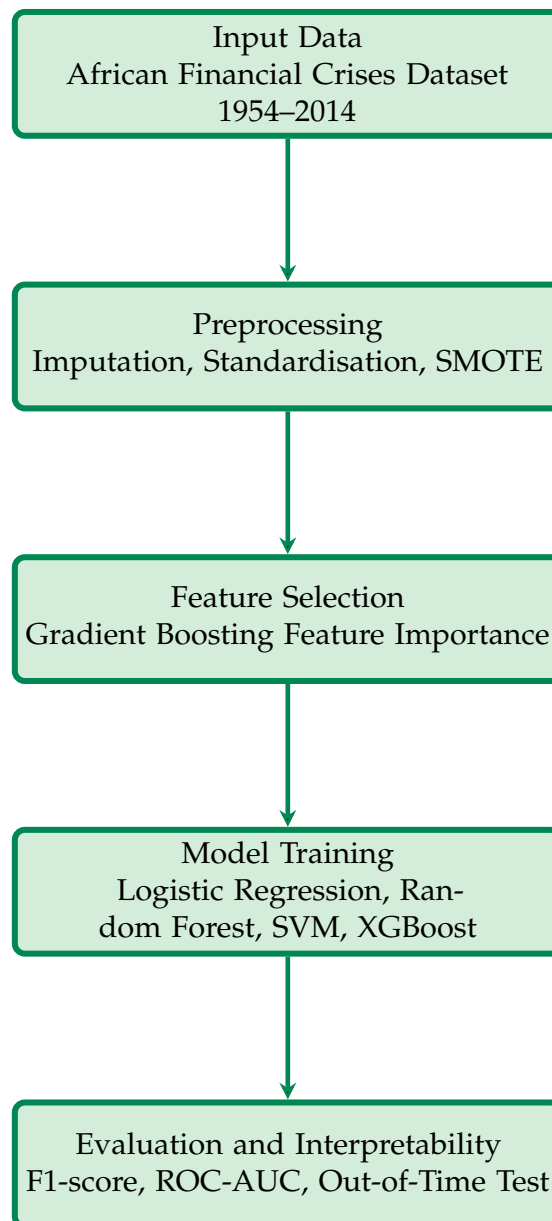


Figure 1. Methodological framework for banking crisis prediction.

2.1. Data Description and Preprocessing

Annual macro-financial data for Nigeria are analysed using the African Financial Crises Dataset [8], which covers 1954-2014 and contains approximately 60 annual observations.

The data set includes the exchange rate volatility (calculated as standard deviation of the annual change of the exchange rates), inflation (year-over-year CPI), credit-to-GDP ratio, sovereign external debt default variables, a systemic crisis dummy variable, which is used to capture simultaneous crises in other industries, debt service ratios, variables of oil price exposure and variables of current account balance.

Banking crises are very rare in the dataset, and there are only approximately 11 years of crises. As a result, emphasis was placed on data preparation to prevent overfitting and information leakage. Less than 5% of the data was missing, and this was addressed by replacing continuous data with the mean and binary data with the most common value (mode). Furthermore, all continuous variables were standardised to have a mean of 0 and a variance of 1, thereby enhancing numerical stability and facilitating comparisons across models.

Because banking crises occur much less frequently than non-crisis periods, the data is highly imbalanced. To correct this problem, the Synthetic Minority Over-sampling Technique (SMOTE) was used only on the training set, while the test set remained unchanged. The data were split temporally into a training sample (1954-2000) and a distinct out-of-time test sample (2001-2014). This arrangement mirrors actual forecasting practice and avoids look-ahead bias. In addition, it ensures that synthetic data does not influence the evaluation phase and prevents information leakage from the training data to the test period. The minority class corresponds to crisis years, and resampling was applied to help the models better identify these crisis instances without changing the original class distribution during the evaluation phase.

2.2. Feature Selection

Because of the small number of observations and the relatively large number of macro-financial variables, feature selection was used to improve the models' generalisability and robustness. A feature selection method based on Gradient Boosting was used because it can identify nonlinear dependencies and interaction effects, which often occur in macro-financial variables. The Extreme Gradient Boosting (XGBoost) algorithm was applied to the preprocessed training data, and feature importance was measured using gain-based criteria. This quantifies each feature's contribution to the model's predictive power. Features with low importance were eliminated, and economically significant predictors were retained.

Starting with a set of around 15-20 indicators, the final set of features was narrowed down to 8-10 variables. These features included exchange rate volatility, inflation, systemic crisis variables, sovereign external debt default, credit-to-GDP gap, oil price volatility, debt service ratio, and current account balance. This method of feature selection is important because it balances statistical significance and economic theory, which is especially relevant in the context of policy-oriented early warning systems.

2.3. Model Development

Four classification models are employed in this study, which represent various modelling strategies that can be used in early warning models. These are a linear probabilistic model (logistic regression), an ensemble bagging approach (Random Forest), a margin-based model (Support Vector Machine), and an ensemble model based on boosting (Extreme Gradient Boosting). By employing these models, it is possible to compare the results of traditional econometric approaches with those of machine learning models.

Given the small dataset and limited number of crisis event observations, measures were taken to control model complexity and avoid overfitting. The hyperparameters of the machine learning models were set to standard values recommended in previous studies, and the models' performance was evaluated using cross-validation on the training data.

The models were trained on the preprocessed training data (1954-2000) and tested on a separate out-of-time period (2001-2014). Hyperparameter optimisation was carried out using five-fold stratified cross-validation on the training data to ensure that the very low probability of banking crisis events was maintained.

2.3.1. Logistic Regression

Logistic regression is used as a benchmark model because it is widely applied in early warning system research and is easy to interpret. Policymakers and regulators often prefer this model because its estimated coefficients can be clearly linked to economic theory, facilitating transparent communication and decision-making.

The model estimates the probability of a banking crisis ($y = 1$) as:

$$P(y = 1 | X) = \frac{1}{1 + \exp(-(\beta_0 + \beta^\top X))}, \quad (1)$$

where X represents the set of macro-financial predictors and β denotes the coefficients estimated using maximum likelihood. Logistic regression assumes a linear relationship between the predictors and the log-odds of a crisis, which may not fully capture the complex nonlinear patterns that often precede banking crises.

No regularisation was applied ($C = 1.0$) to ensure consistency with traditional econometric early warning models. The model, therefore, serves as a baseline for evaluating whether machine learning methods provide additional predictive improvements.

2.3.2. Random Forest

Random Forest is an ensemble learning algorithm that grows many decision trees on different samples of the data and uses random subsets of predictors. This makes Random Forest a good choice for small macro-financial data sets.

The algorithm is very useful for EWS because it can automatically model nonlinear relationships, threshold effects, and interactions without specifying functional forms. This is particularly important for predicting banking crises, where risks tend to accumulate nonlinearly.

For a forest with K trees, the predictions are made by majority voting:

$$\hat{y} = \arg \max_c \sum_{k=1}^K \mathbb{I}(\hat{y}_k = c), \quad (2)$$

where $\mathbb{I}(\cdot)$ is the indicator function. Hyperparameters were chosen via grid search, and the final parameters were set to $n_{\text{estimators}} = 100$, $\text{max_depth} = 10$, and $\text{min_samples_split} = 5$. These decisions are flexible, but can be generalised, as there are only a few crisis observations.

2.3.3. Support Vector Machines

Support Vector Machines (SVMs) are margin-based classifiers that aim to determine the optimal boundary between crisis and non-crisis observations. SVMs are particularly well-suited to high-dimensional data with small sample sizes, both of which are characteristic of macro-financial early warning systems.

Optimising the margin between classes improves generalisation and resistance to noisy predictors. It is especially significant in cases where the number of observations of a crisis is small and is potentially subject to measurement error.

The optimisation problem is as follows:

$$\min \frac{1}{2} \|w\|^2 + C \sum_i \xi_i, \quad (3)$$

subject to $y_i(w^\top x_i + b) \geq 1 - \xi_i$ and $\xi_i \geq 0$. A radial basis function (RBF) kernel is employed to model nonlinear patterns:

$$k(x, x') = \exp(-\gamma \|x - x'\|^2). \quad (4)$$

Hyperparameters $C = 1.0$ and $\gamma = 0.1$ were chosen via stratified cross-validation to balance the model's flexibility and robustness on an imbalanced dataset.

2.3.4. Extreme Gradient Boosting

Extreme Gradient Boosting (XGBoost) is a sophisticated ensemble learning algorithm that builds decision trees iteratively, with each new tree attempting to correct errors made by previous trees. This adaptive learning method enables XGBoost to effectively identify early warning signals of a severe, though uncommon, condition, such as a banking crisis.

XGBoost proves effective in early warning models, as it can learn complex nonlinear relationships and handle imbalanced data, and employs regularisation to prevent overfitting. The features are useful for modelling macro-financial systems that are affected by structural changes and regime shifts.

The objective function that the model attempts to minimise is as follows:

$$L = - \sum_i l(y_i, f_k) - \sum_k \Omega(f_k), \quad (5)$$

in which $l(\cdot)$ is the logistic loss, and $\Omega(\cdot)$ is the cost of model complexity. The algorithm leverages second-order gradient information to improve convergence rate and efficiency.

The last XGBoost model was trained with a learning rate of 0.1, $n_{\text{estimators}} = 100$, $\text{max_depth} = 5$, and a subsample ratio of 0.8.

2.4. Model Evaluation and Interpretability

The model's performance is assessed using accuracy, precision, recall, F1-score, and the area under the receiver operating characteristic (ROC) curve (ROC-AUC). Due to the class imbalance in the dataset and the need to correctly classify crisis periods for policy-making purposes, higher priority is given to recall and F1-score for the crisis class. Accuracy can be highly deceptive in this problem, where a model can be highly accurate by simply predicting the non-crisis periods. ROC-AUC is also used to assess the model's overall performance in correctly classifying crisis and non-crisis periods across different thresholds.

The models were validated using five-fold stratified cross-validation on the training set and a strict out-of-sample validation on the test period from 2001 to 2014. The feature importance scores of the XGBoost models (gain-based) were used to enhance the proxies for model interpretability. Probabilities above 0.5 were considered "red alerts" in the context of the early warning system approach.

The approach adopted is guided by the need for transparency and replicability, especially given the limitations of annual macro-financial data and the low incidence of banking crises. While the approach allows for a systematic comparison of machine learning models under strict validation criteria, the results are to be considered indicative rather than conclusive of real-time forecasting ability.

3. Results

3.1. Exploratory Data Analysis

Before embarking on model estimation, this section provides a brief description of the Nigerian banking crisis data set.

The data set contains 60 annual observations for Nigeria from 1954 to 2014; 11 of these years (18.3%) are classified as banking crisis years, while the remaining 49 years are non-crisis years. A common characteristic of financial crisis data is the imbalance between crisis and non-crisis observations, which encourages the use of resampling techniques in model training.

Table 2 provides descriptive statistics for the entire set of macroeconomic and crisis-related variables. The summary statistics reveal significant variation between key indicators. In particular, there is noticeable dispersion in exchange rate movements, reflecting periods of sharp currency

depreciation driven by external shocks and volatility in oil prices. Additionally, annual inflation (CPI) shows significant variability, with inflation exceeding 70 per cent during periods of macroeconomic instability, indicating substantial economic volatility throughout the sample period.

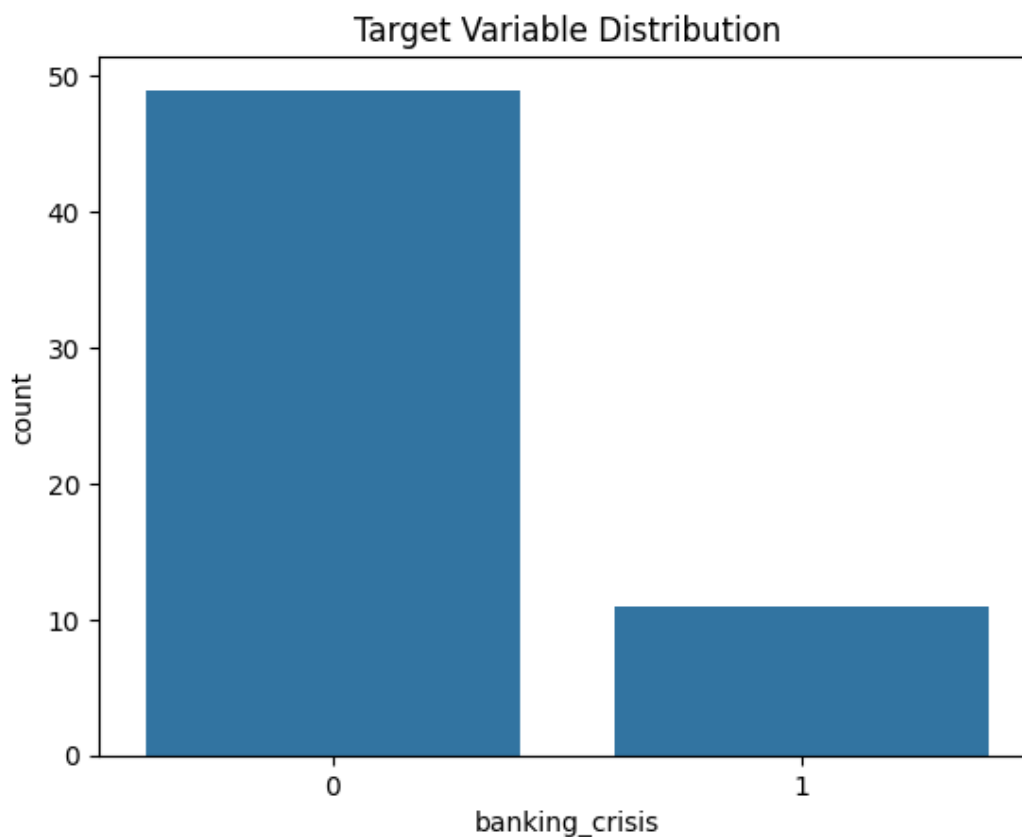


Figure 2. Nigeria's distribution of crisis and non-crisis years.

Table 2. Key variable summary statistics.

Variable	Count	Mean	Std	Min	25%	50%	75%	Max
case	60	45.0	0.0	45.0	45.0	45.0	45.0	45.0
year	60	1983.88	17.88	1954.0	1968.75	1983.5	1999.25	2014.0
systemic_crisis	60	0.167	0.376	0.0	0.0	0.0	0.0	1.0
exch_usd	60	38.95	59.08	0.0	0.0	0.78	100.85	158.27
domestic_debt_in_default	60	0.0	0.0	0.0	0.0	0.0	0.0	0.0
sovereign_external_debt_default	60	0.15	0.36	0.0	0.0	0.0	0.0	1.0
gdp_weighted_default	60	0.0	0.0	0.0	0.0	0.0	0.0	0.0
inflation_annual_cpi	60	14.77	15.65	-4.55	6.10	10.76	16.30	72.73
independence	60	0.9	0.30	0.0	1.0	1.0	1.0	1.0
currency_crises	60	0.167	0.376	0.0	0.0	0.0	0.0	1.0
inflation_crises	60	0.2	0.403	0.0	0.0	0.0	0.0	1.0
banking_crisis	60	0.183	0.390	0.0	0.0	0.0	0.0	1.0

A selection of Nigeria's macroeconomic indicators over the sample period is shown in Figure 3.

In Figure 4, a correlation heatmap is shown to facilitate further analysis of the variables' correlations. The correlation between exchange rate volatility and inflation is positive, and periods of high inflationary pressures and exchange rate volatility are associated with sovereign external debt defaults. Early findings from the correlations indicate that macroeconomic volatility is associated with the banking crisis in Nigeria.

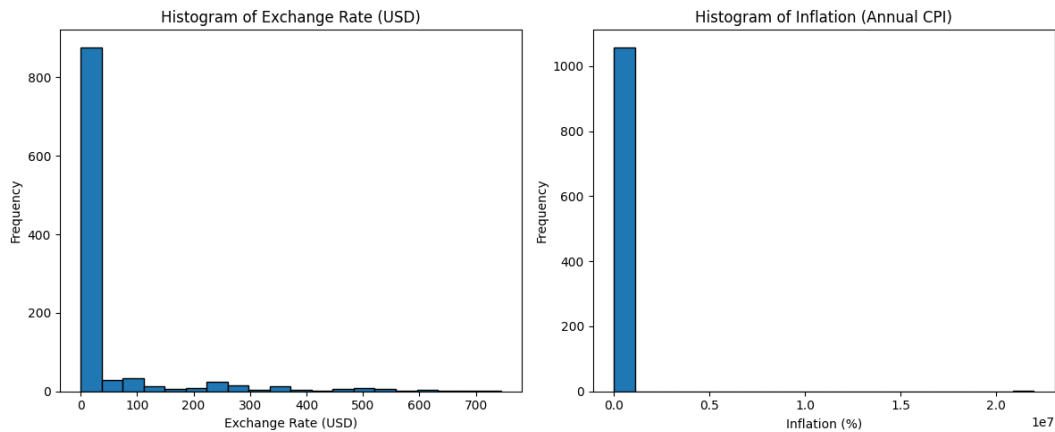


Figure 3. Distribution of a few chosen Nigerian macroeconomic indicators.

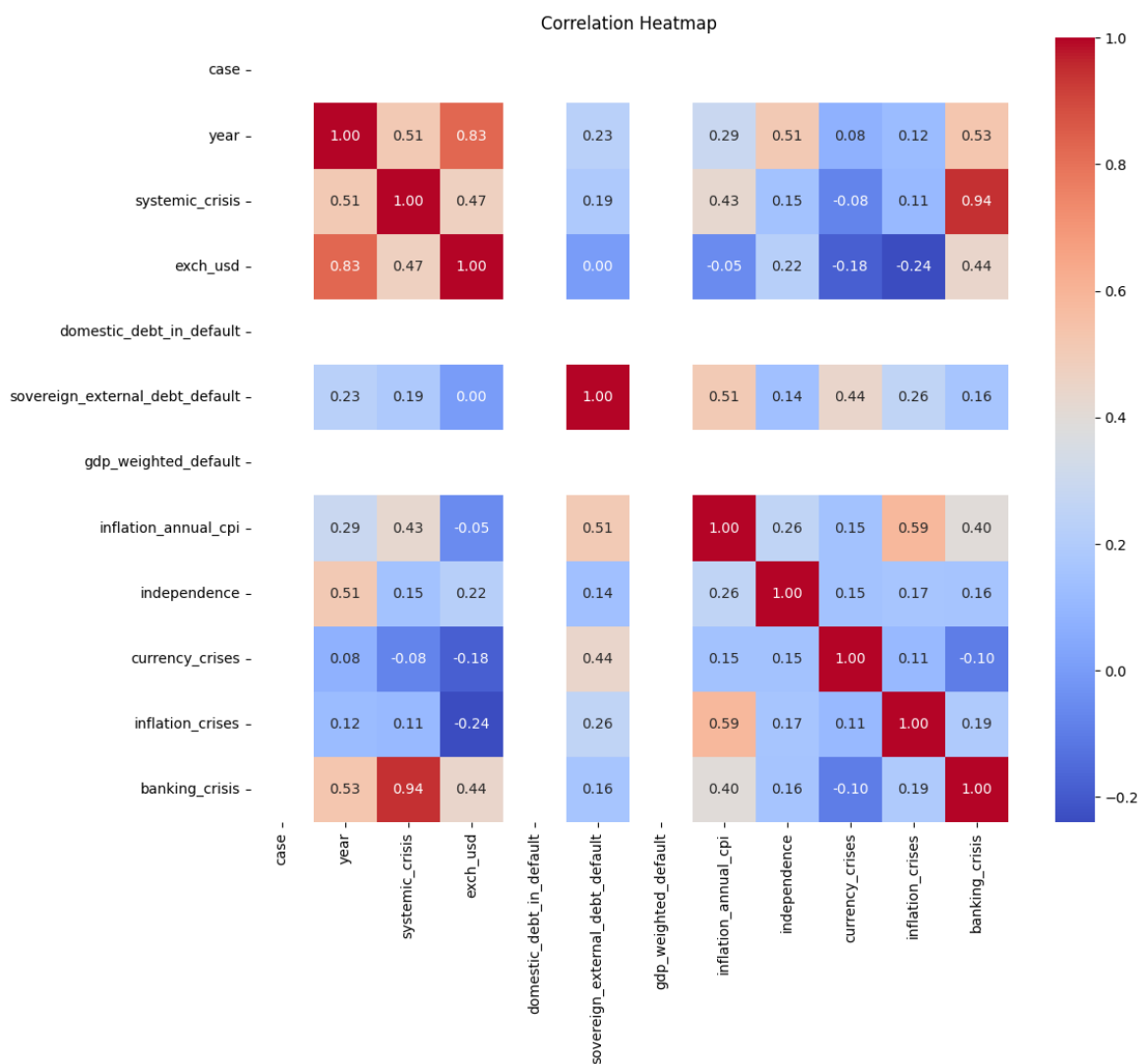


Figure 4. Correlation heatmap of macroeconomic indicators.

Generally, the exploratory analysis confirms that the dataset captures economically significant variation and crisis-related information, supporting its appropriateness for early warning system modelling.

3.2. Feature Importance

Feature importance was analysed using Gradient Boosting and Random Forest algorithms to identify the most important predictors of banking crises in Nigeria. The results are presented in Figure 5.

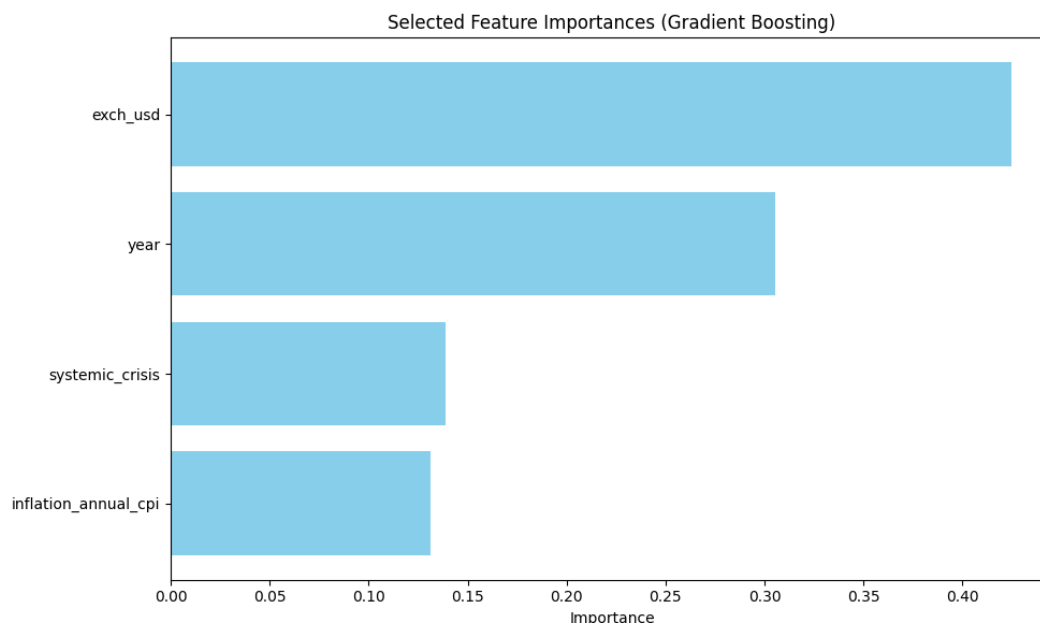


Figure 5. Feature importance for predicting banking crises in Nigeria.

The most significant predictors are the Exchange rate, Inflation, systemic crisis indicators, and time effects (year). Sovereign external debt defaults are also significant predictors of crises, underscoring the importance of macroeconomic volatility and systemic financial distress.

3.3. Model Performance

3.3.1. Cross-Validation

Model performance was first evaluated using five-fold cross-validation on the SMOTE-resampled training dataset. Average performance metrics are reported in Table 3.

Table 3. Cross-validation performance metrics on the SMOTE-resampled training set (AUC refers to ROC-AUC).

Model	Accuracy	F1-score	Precision	Recall	ROC-AUC
Logistic Regression	0.7808	0.7569	0.8731	0.7393	0.9638
Random Forest	0.9233	0.9270	0.9200	0.9500	0.9953
SVM	0.5742	0.4682	0.4981	0.5000	0.8821
XGBoost	0.8717	0.8687	0.8950	0.8679	0.9828

Random Forest delivered the strongest overall results, achieving an F1 Score of 0.93, a recall of 0.95, and an ROC AUC of 0.99. XGBoost also demonstrated strong performance, whereas Logistic Regression exhibited moderate predictive capacity. In contrast, the Support Vector Machine (SVM) model performed relatively poorly.

Note: Acc. = Accuracy; Prec. = Precision; Rec. = Recall.

3.3.2. Test Set Evaluation

While Random Forest achieved the strongest performance during cross-validation on the training data, XGBoost outperformed all other models on the strictly out-of-time test set. Test set performance metrics are reported in Table 4. XGBoost achieved the highest accuracy (0.92) and the strongest crisis-class F1-score (0.80). Random Forest and Logistic Regression also provided reliable results. While

test-set AUC and recall are high, these results should be interpreted cautiously, given the small sample size and the limited number of crisis observations.

Because the test set represents a strictly out-of-time evaluation, performance metrics reflect a more realistic assessment of predictive ability under real-world forecasting conditions.

Table 4. Evaluation metrics on the test set. Class 1 corresponds to banking crises (AUC refers to ROC-AUC).

Model	Acc.	F1 (0)	F1 (1)	Prec. (1)	Rec. (1)	AUC
Logistic Regression	0.83	0.89	0.67	0.50	1.00	1.0
Random Forest	0.75	0.82	0.57	0.40	1.00	1.0
SVM	0.83	0.89	0.67	0.50	1.00	0.0
XGBoost	0.92	0.95	0.80	0.67	1.00	1.0

Test set performance metrics should be interpreted with caution, as they are based on a small out-of-time sample (14 observations) with very few crisis events, which can lead to extreme values in threshold-dependent metrics such as ROC-AUC.

Note: Acc. = Accuracy; Prec. = Precision; Rec. = Recall.

3.4. ROC and Precision–Recall Analysis

Figures 6 and 7 show the model's discrimination ability using Precision-Recall curves and ROC curves. XGBoost and Random Forest perform better than other models, as they can almost perfectly distinguish crisis and non-crisis data points.

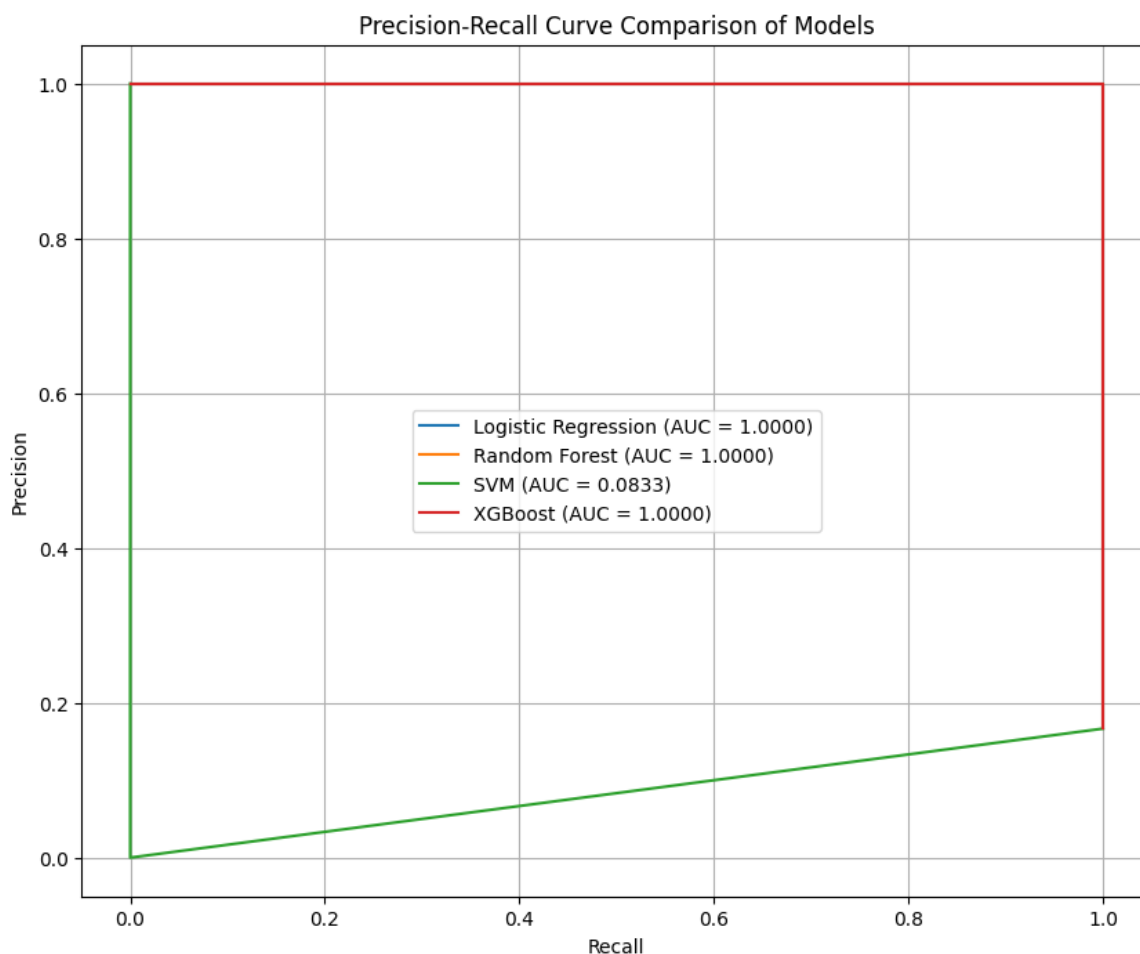


Figure 6. Precision–Recall curves for all models.

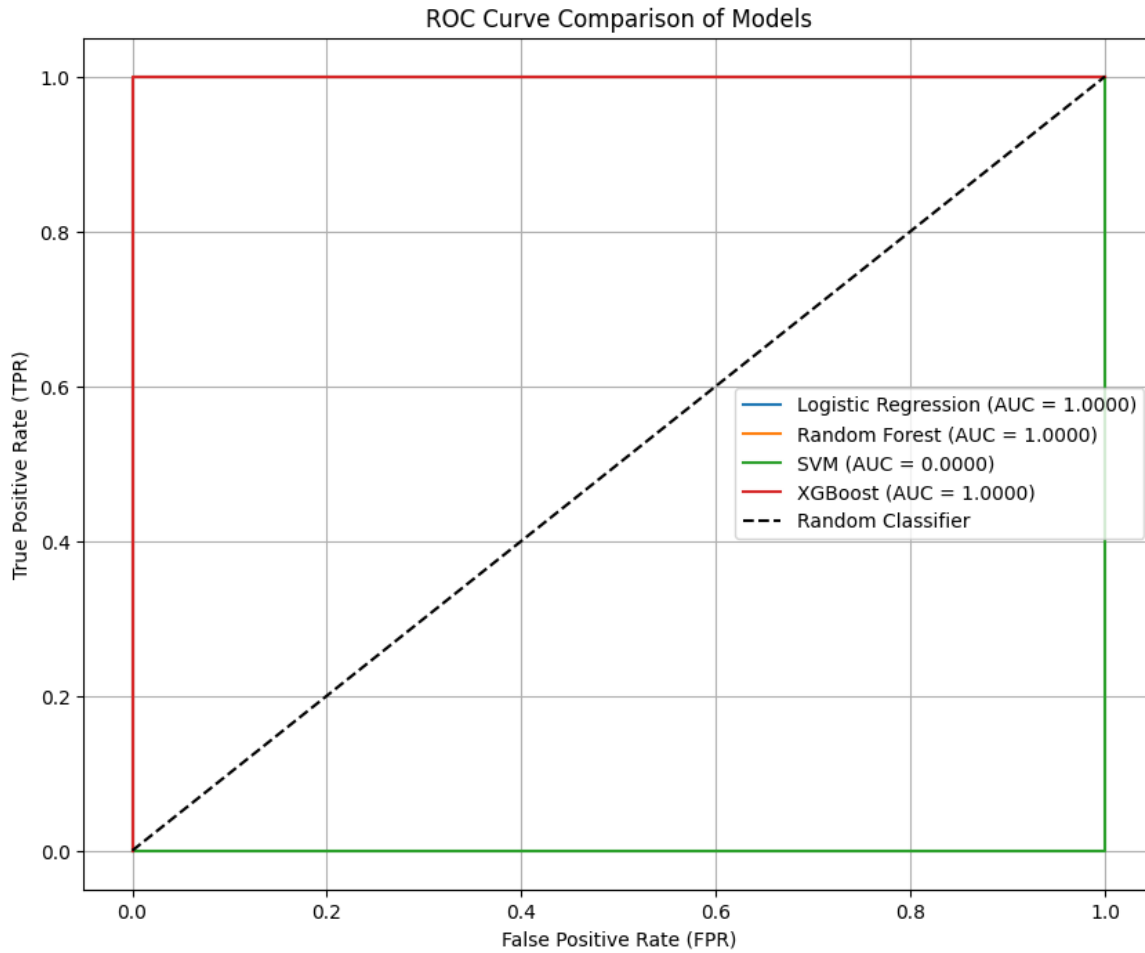


Figure 7. ROC curves for all models.

3.5. Early Warning System Outputs

The Early Warning System (EWS) was built using the predicted crisis probabilities from the XGBoost model, which achieved the best generalisation performance on the out-of-time test set. The risk levels were categorised into Green (low risk), Yellow (moderate risk), and Red (high risk).

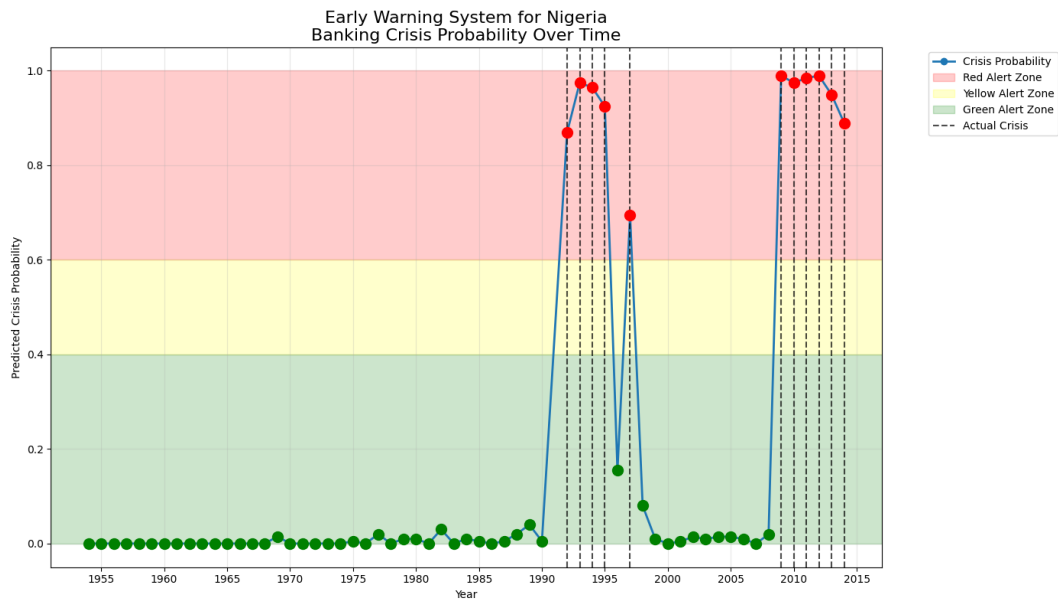


Figure 8. Early Warning System timeline: predicted crisis probabilities and observed banking crises.

All 11 historical crises are in the Red alert zone, indicating a strong crisis-detection capability. The probability of predicted values tends to increase significantly in the crisis year itself, indicating poor predictive power.

3.6. Summary of Key Findings

The exploratory study offers evidence on the relative performance of machine learning models on banking crisis risk identification in Nigeria under severe data constraints. The key findings are summarised as follows:

1. Tree-based ensemble models, especially random forest and XGBoost, always perform better in classifying historical banking crisis episodes than logistic regression and Support Vector Machines, which suggests that they can discern nonlinear macro-financial relationships.
2. Factors found to be significant predictors across a variety of model specifications include exchange rate volatility, inflation, systemic crisis variables, and sovereign external debt default events, highlighting the role of macroeconomic instability in the Nigerian banking sector crisis.
3. As the data are measured annually, the model is most appropriate for detecting crisis periods simultaneously. The probability of a crisis is projected to be high in a crisis year, indicating no predictive ability.
4. The results should be viewed as indicative of above-average performance, instead of clear evidence of a functioning early warning system, although the devised framework is capable of defining preceding crisis cases accurately.
5. Overall, the results suggest that machine learning approaches can be combined with traditional econometric approaches to provide effective early warning of banking crisis risk in Nigeria. However, their application in practice depends on the availability of more extensive data, a clearer definition of lead-time performance, and improved interpretability.

It is necessary to consider the limitations of the data and the study's evaluation when interpreting the results. However, they indicate good classification accuracy for tree-based ensemble methods, namely Random Forest and XGBoost. The application of resampling methods, the utilisation of annual data, and the limited number of crisis instances indicate that the performance results do not guarantee accurate real-time forecasting but rather reflect favourable learning conditions. Thus, the next section will discuss the empirical results within the context of the literature, their interpretation, and their limitations for the development of early warning systems in emerging markets.

4. Discussion

This study demonstrates that ensemble machine learning methods, particularly Random Forest and XGBoost, outperform conventional logistic regression in identifying historical banking crises in Nigeria [18,20]. Random Forest showed exceptional performance on the in-sample dataset, achieving a recall of 0.95 and an F1-score of 0.93, as it effectively captured patterns from past crisis events through bagging and deep decision trees. On the other hand, XGBoost generalised better on the out-of-time test dataset, achieving an accuracy of 0.92. The reason for this improved performance is its sequential boosting process and regularisation techniques, which enhance the model's ability to handle temporal uncertainty.

Also, the analysis of feature importance revealed that exchange rate volatility, inflation, systemic crisis indicators, and sovereign external debt defaults are some of the most powerful predictors. These results confirm the argument that macroeconomic instability is a decisive factor contributing to financial vulnerability in emerging economies. The impact of the year variable is likely due to structural adjustments, including global financial cycles and policy changes.

The characteristics of the dataset, frequency, and the number of crisis events are influencing these outcomes. This particular historical dataset exhibits nonlinear patterns that can be exploited by ensemble models, which is why they perform better than alternative methods for predicting on this dataset [18,20]. However, its general approach, such as data processing, feature extraction, and

ensemble design, can be readily generalised to other contexts with only moderate modification, since its counterparts have been proven efficient across different emerging markets. Hyperparameter tuning and temporal dependency modelling are important factors to consider when applying to other datasets.

The mode of temporal validation (out-of-time testing) in this case provides an easily available template for assessing model robustness, which can be directly replicated to other issues. The main challenge in making this approach applicable to other datasets with very different properties, such as higher-frequency data, cross-country panels, or more frequent crisis events, is doing justice to the minority class of crisis observations. In other crisis forecasting problems, methods like SMOTE have been successfully combined with temporal cross-validation to overcome this problem.

As anticipated in the literature, tree-based ensemble models tend to perform better than linear econometric models in predicting emerging-market crises [18,20]. Nonetheless, other studies have observed that a standard logistic regression model can outperform a machine learning model in recursive out-of-sample tests [13]. Nevertheless, the weak signal of a long-lead warning in annual data strongly supports the notion that the model's performance is still best suited for the task of crisis prediction. The model is more of a crisis-tracking device than a comprehensive early warning system at present. Future research should focus on higher-frequency data, the inclusion of global and institutional factors, and the use of time-series approaches such as recurrent neural networks or LSTM models. It is proposed to integrate SHAP/LIME explainable AI methods with stress testing.

5. Conclusions

This study demonstrates that machine learning algorithms, particularly tree-based ensemble models, can be more effective at detecting banking crises in Nigeria than traditional logistic regression. The findings indicate that the most important factor in determining the banking system's distress is macroeconomic uncertainty, as measured by exchange rate volatility, inflation rates, and sovereign debt default. However, the results are constrained by the limited number of banking crises, and the models are unable to predict crises with much advance notice. Instead, they identify increasing risk as crises materialise, and they can be used for monitoring purposes.

This research demonstrates that machine learning can be used alongside traditional approaches in data-scarce settings. It offers a clear assessment framework that emphasises both the potential and the limitations of these approaches. The models need to be interpreted with care and should not be taken as functional early warning systems in their current form. Future research will focus on the use of high-frequency data and panel models to improve robustness and policy relevance.

Author Contributions: Conceptualisation, N.M., T.R., and R.N.; methodology, N.M.; software, N.M.; validation, N.M., R.N., T.R., and C.S.; formal analysis, N.M.; investigation, N.M., R.N., and T.R.; data curation, N.M.; writing—original draft preparation, N.M.; writing—review and editing, N.M., R.N., T.R., and C.S.; visualisation, N.M.; supervision, T.R., R.N., and C.S.; project administration, T.R. and R.N.; funding acquisition, N.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the DST-CSIR National e-Science Postgraduate Teaching and Training Platform (NEPTTP) <http://www.escience.ac.za/> (accessed on 1 January 2025).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data were obtained from the Kaggle website <https://www.kaggle.com/jaimetricz/african-crisis> (accessed on 16 May 2025). The analytic dataset used in this study was derived by filtering and processing the original Nigerian data and is available from the corresponding author upon reasonable request.

Acknowledgments: The authors gratefully thank the support given to this research by the DST-CSIR National e-Science Postgraduate Teaching and Training Platform (NEPTTP). All opinions, findings, and conclusions expressed are the sole responsibility of the authors and are not necessarily those of NEPTTP. The authors also thank the anonymous reviewers for their constructive and insightful feedback.

Conflicts of Interest: In this manuscript, the authors stated that there is no conflict of interest. The funding bodies were not involved in the design of the study, data collection, analysis, interpretation, manuscript writing, or the decision to publish findings.

Abbreviations

The abbreviations listed in this manuscript are utilised:

EWS	Early Warning System
ML	Machine Learning
SVM	Support Vector Machine
RF	Random Forest
XGBoost	Extreme Gradient Boosting
GDP	Gross Domestic Product
CPI	Consumer Price Index
ROC	Receiver Operating Characteristic
AUC	Area Under the Curve
SMOTE	Synthetic Minority Oversampling Technique
KNN	K-Nearest Neighbours
IMF	International Monetary Fund

References

1. Reinhart, C.M. *This Time Is Different: Eight Centuries of Financial Folly*; Princeton University Press: Princeton, NJ, USA, 2009. [<https://doi.org/10.2307/j.ctvc4m4gqx>]
2. Laeven, L.; Valencia, F. *Systemic Banking Crises Revisited*; International Monetary Fund: Washington, DC, USA, 2018. [<https://doi.org/10.5089/9781484376379.001>]
3. Peychev, A. *Reforming the European Stability Mechanism: Too Much but Never Enough*; Wilfried Martens Centre for European Studies, 2021. [<https://doi.org/10.1177/17816858211063978>]
4. Chen, T.; Guestrin, C. XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 785–794. [<https://dl.acm.org/doi/10.1145/2939672.2939785>]
5. Claessens, S.; Kose, A.; Laeven, L.; Valencia, F. *Understanding Financial Crises: Causes, Consequences, and Policy Responses*; International Monetary Fund: Washington, DC, USA, 2014. [<https://doi.org/10.2139/ssrn.2295199>]
6. Reimann, C. Predicting financial crises: An evaluation of machine learning algorithms and model explainability for early warning systems. *Rev. Evol. Political Econ.* **2024**, *5*, 51–83. [[CrossRef](#)]
7. Voskamp, J. *Machine Learning for Financial Crisis Prediction*; Ph.D. Thesis, 2024.
8. Reinhart, C.M.; Rogoff, K.S.; Trebesch, C.; Reinhart, V.R. *Global Crises Data by Country*; Harvard Business School: Boston, MA, USA, 2019.
9. Kaminsky, G.; Lizondo, S.; Reinhart, C.M. Leading indicators of currency crises. *Staff Papers* **1998**, *45*, 1–48. [[CrossRef](#)]
10. Drehmann, M.; Juselius, M. Evaluating early warning indicators of banking crises: Satisfying policy requirements. *Int. J. Forecast.* **2014**, *30*, 759–780. [[CrossRef](#)]
11. Betz, F.; Oprică, S.; Peltonen, T. A.; Sarlin, P. Predicting distress in European banks. *J. Bank. Financ.* **2014**, *45*, 225–241. [[CrossRef](#)]
12. Langley, P. *Elements of Machine Learning*; Morgan Kaufmann: San Francisco, CA, USA, 1996.
13. Beutel, J.; List, S.; von Schweinitz, G. Does machine learning help us predict banking crises? *J. Financ. Stab.* **2019**, *45*, 100693. [[CrossRef](#)]
14. El Halabi, L. *Predicting Banking Crises Using Machine Learning: The Case of Lebanon*; Ph.D. Thesis, 2023.
15. Laeven, L. Banking crises: A review. *Annu. Rev. Financ. Econ.* **2011**, *3*, 17–40. [[CrossRef](#)]
16. Gogas, P.; Papadimitriou, T.; Agrapetidou, A. Forecasting bank failures and stress testing: A machine learning approach. *Int. J. Forecast.* **2018**, *34*, 440–455. [[CrossRef](#)]
17. Bussmann, N.; Giudici, P.; Marinelli, D.; Papenbrock, J. Explainable machine learning in credit risk management. *Comput. Econ.* **2021**, *57*, 203–216. [[CrossRef](#)]
18. Liu, L.; Chen, C.; Wang, B. *Predicting Financial Crises with Machine Learning Methods*; Journal of Forecasting, *41*, 871–910, 2022. [<https://doi.org/10.1002/for.2840>]

19. Trickz, J. African Crises Dataset. Available online: <https://www.kaggle.com/jaimetricz/african-crises> (accessed on 16 May 2025).
20. Yin, S. *Machine Learning Algorithms for Early Warning Systems: Predicting Systemic Financial Crises through Non-Linear Econometric Models*; International Journal of Economics and Finance Studies, 16, 286–311, 2024. [<https://doi.org/10.34109/ijefs.202416414>.]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.