International Journal of Computer Applications Technology and Research
Volume 15–Issue 10, 01 – 19, 2026, ISSN: -2319–8656
DOI:10.7753/IJCATR1510.1001

Loan Default Prediction Using a Hybrid GAN-LSTM-Bi-GRU Architecture

Ayodeji O. Akinwumi
Department of Information Systems,
Adekunle Ajasin University,
Akungba-Akoko,
Nigeria

Samuel A. Oluwadare
Department of Computer Science,
Federal University of Technology,
Akure, Nigeria.

Ilobekemen P. Oladoja,
Department of Computer Science,
Federal University of Technology,
Akure, Nigeria.

Adewuyi A. Adegbite,
Department of Data Science,
Adekunle Ajasin University,
Akungba-Akoko,
Nigeria.

Abstract: Lending is an important activity of financial institutions, providing individuals and businesses with access to funds for different economic needs. The continued growth in lending has also raised concerns about non-performing loans in financial institutions, creating a need for effective measures to mitigate loan default. In Nigeria, loan default prediction is constrained by challenges such as limited access to relevant Nigerian financial data, making it difficult to train models that adequately reflect the local financial environment. This study contributes to loan default prediction by developing a deep learning framework using Nigerian financial data. The framework uses a Bidirectional Gated Recurrent Unit Generative Adversarial Network (Bi-GRU GAN) to generate additional default cases and balance the training data. The balanced data are then passed through an LSTM-based Denoising Autoencoder (LSTM-DAE) to learn a compact representation of the loan features, which is subsequently used by a Bi-GRU classifier to predict loan default. The dataset initially contained 6,030 loan records and was reduced to 6,000 after data cleaning. The data had a default rate of 15.95%, with approximately 5.3 non-default cases for every default case. The Bi-GRU GAN generated 2,860 additional default samples, increasing the default class from 670 to 3,530 and producing a balanced training set. Three model variants were evaluated: Baseline, GAN-Augmented, and Full Framework. The Full Framework achieved the best performance, with 96.3% accuracy, 89.3% precision, 87.4% recall, an F1-score of 88.3%, and a ROC-AUC of 0.941. The results show that combining synthetic data generation, learned feature representation, and Bi-GRU classification improved loan default prediction. The study presents an integrated deep learning framework for loan default prediction.

Keywords: loan default prediction; credit risk assessment; recurrent neural networks; bi-gru; generative adversarial network; lstm denoising autoencoder; synthetic data generation; nigerian financial sector.

1. INTRODUCTION

Financial institutions contribute to the economic development of nations [1], [2], [3], and offer services such as savings, payment options, lending, investments, and insurance [4]. Among these services, lending is particularly important because it gives individuals and businesses access to funds for consumption, investment, business expansion, and other economic needs [5], [6], [7]. It is also a major source of revenue for many financial institutions [8]. This is evident from the continued growth of the global lending market, which is projected to reach $17,282.05 billion by 2030, with an annual compound growth rate of 7.2% [9].

In financial terms, the funds provided through lending constitute a loan, which is defined by the International Monetary Fund (IMF) as a financial instrument created when a creditor lends funds directly to a debtor [10]. A loan is a financial arrangement in which a lender provides funds to a borrower with an obligation to repay the amount received according to agreed terms, usually including interest or other associated charges [11], [12]. A loan therefore establishes a financial obligation between the lender and the borrower, with repayment expected according to the terms agreed between both parties [13]. Since repayment is an expected obligation of a loan, the question of whether the borrower will fulfil this obligation as agreed remains an important concern for the lender [14]. This concern arises from the possibility that the borrower may be unable to make the required payments when they become due, creating uncertainty regarding the eventual recovery of the funds provided [13]. When a borrower fails to meet the agreed repayment obligation, the loan is regarded as being in default otherwise called loan default [15], [16].

In credit risk modelling, the number of days a payment remains overdue is often used to determine whether a loan is classified as default, and different thresholds may lead to different classifications [17], [18]. However, a 90-day past-due threshold is commonly used in default modelling, although shorter periods have also been considered in the literature [19]. This classification is important because it provides a clear basis for distinguishing loans that have met the agreed repayment conditions from those that have not [20]. Loan default is a major concern for financial institutions because failure to recover loans can lead to financial losses and increase the level of non-performing loans, thereby weakening their financial position [21]. These effects can be more challenging on financial institutions in developing economies such as Nigeria, where weaknesses in credit risk management can contribute to financial distress and instability. [22]. Cases of financial institution failures in Nigeria have also been associated with weaknesses in risk management and control systems [23]. It is important to note that when financial institutions fail, they can affect consumer trust, investment, and economic growth [24]. It is therefore important for financial institutions to identify borrowers who

are more likely to default before granting loans. Credit risk assessment helps financial institutions evaluate the likelihood that a borrower will fail to meet repayment obligations and supports decisions on whether to grant credit [25].

There are a number of ways financial institutions assess borrowers before granting loans. These include credit history, income and employment information, existing debt, collateral, credit bureau records, know-your-customer (KYC) information, and expert judgement [26], [27]. Credit bureau records provide information on the previous credit behaviour of a borrower and can help institutions assess repayment performance before granting new loans [28]. However, the usefulness of such information depends on the availability and quality of the borrower's records, which can be limited where formal credit records are not available [29]. KYC information is also used to understand a customer's identity, background, and source of income, but the information provided may not always give a complete picture of the customer ([30]. Expert judgement is another approach used in credit assessment, where loan officers may consider factors such as the borrower's character, capital, capacity, collateral, and prevailing conditions when deciding whether to grant a loan [13]. However, this approach can be time-consuming and labour-intensive, and the decision may also be affected by human judgement, which can lead to bias or errors in assessing a borrower's risk [31]. While these methods have proven useful in assessing loan risk, the continued increase in loan defaults reported in recent studies suggests that the problem remains persistent [32], [33]. There is therefore a need for more effective credit risk assessment to improve the identification of borrowers who are likely to default.

Machine learning (ML) provides a more effective and data driven approach to credit risk assessment by learning patterns from historical financial data and identifying factors associated with loan default, with several studies have reporting the effectiveness of ML techniques in loan default prediction and credit scoring [34], [35], [36], [37]. Deep learning (DL) techniques has also been applied to credit risk assessment, particularly in loan default prediction, because of their ability to model complex relationships within financial data [38], [39], [40], [41]. Despite the progress reported with ML and DL approaches, their effectiveness is influenced by the quality and availability of financial data. Effective ML and DL models require sufficient and representative training data; however, access to detailed borrower information may be restricted by privacy concerns, institutional policies, limited data sharing, and fragmented financial records, with the problem often more evident in developing economies where comprehensive financial datasets are scarce [42], [37]. Synthetic data generation has therefore been explored as a means of supplementing limited financial datasets, techniques such as the use of Generative Adversarial Networks (GANs) has been explored to generate additional training data [43], [44], [45]. The imbalance between defaulting and non-defaulting borrowers is another challenge in loan default prediction. Loan datasets often contain far more non-defaulting than defaulting borrowers which can make it difficult for a prediction model to correctly identify borrowers who are likely to default. Methods such as SMOTE, over-sampling, under-sampling, and GAN-based techniques, have been used to reduce the effect of this problem [46].

Loan default prediction can be considered a sequential problem because the financial activities of a borrower and repayment behaviour occur over time. Information from previous activities can provide useful indications of subsequent behaviour and may therefore help in assessing the likelihood of default. This requires models that can learn patterns and dependencies across successive observations rather than relying only on static borrower characteristics. Recurrent Neural Networks are suitable for this because they can learn patterns and dependencies across sequential observations. Their variants, including Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), have been applied to financial prediction and credit-related tasks [47], [48], [49], [50].

Based on these, this study proposes a Bi-GRU-based Generative Adversarial Network (GAN) for generating synthetic loan data to help address the issue of limited data. An LSTM-based denoising autoencoder is also used to extract useful features from the data, while a Bi-GRU network is used to learn patterns in borrower data. The resulting framework is aimed at improving loan default prediction and credit risk assessment within the Nigerian financial sector.

2. LITERATURE REVIEW

2.1 Conceptual Review

Credit risk is the possibility that a borrower will fail to fulfil the financial obligations arising from a credit facility, thereby exposing the lending institution to potential loss [51]. The likelihood of default is influenced by a combination of borrower-related and loan-related characteristics, including income, indebtedness, employment conditions, previous credit behaviour, and demographic factors [52], [26]. Evidence from financial institutions in sub-Saharan Africa indicates that the occurrence of non-performing loans is associated with broader economic and institutional conditions. This means that credit risk is also shaped by factors beyond the characteristics of individual borrowers [53], [54]. These factors therefore make credit risk difficult to assess using only a limited set of borrower characteristics and increase the need for reliable methods of identifying borrowers with a greater likelihood of repayment failure. Credit-risk assessment enables financial institutions to evaluate borrowers and determine the likelihood that they will repay their loans as agreed. [55]. The information considered in this process may include financial characteristics, credit history, repayment behaviour, and other borrower attributes that provide evidence of the capacity to repay [56]. The assessment is important because differences in the characteristics and circumstances of borrowers can result in substantially different default outcomes, even among borrowers receiving similar forms of credit [52]**.** Therefore, the ability to identify relevant risk factors and accurately evaluate the probability of default is important for managing credit exposure and limiting losses associated with loan default [57]

Loan default prediction focuses more specifically on estimating whether a borrower is likely to fail to meet the repayment obligations associated with a loan. The prediction task is commonly formulated as a classification problem in which loans or borrowers are classified according to default or non-default outcomes, although some approaches estimate the probability of default rather than only the final class [13]. The purpose of such prediction is to provide an estimate of default risk that can support the identification and management of potentially high-risk borrowers. The development of more accurate prediction methods is particularly relevant where conventional assessment may not fully capture the relationships among the numerous factors associated with default [39] . Recent research shows that machine-learning methods can identify complex relationships within credit data

and provide useful predictions of credit risk, which has contributed to their increasing application in credit-risk modelling [57], [58].

2.2 Conventional Approaches to Loan Assessment

Conventional credit assessment is usually based on borrower information, credit histories, credit scores, and lender-defined rules. Credit bureaux provide historical information that can assist lenders in evaluating applicants [27]. Credit information sharing can also improve lending decisions by reducing information gaps between lenders and borrowers [59]. However, these approaches largely depend on the availability and quality of existing credit information [29]. A borrower with a thin credit history may be difficult to assess using methods that depend heavily on previous borrowing records [28]. Static records may also fail to represent changes in the financial history of a borrower. These limitations have encouraged the adoption of machine learning methods that can combine several variables and identify patterns associated with loan default.

2.3 Machine Learning Approaches for Loan Default Prediction

Machine learning has been applied to loan default prediction as an alternative to conventional credit-scoring methods. They can identify patterns in borrower and transaction data that may be difficult to represent using conventional statistical models. [60], for example, developed machine-learning models for consumer credit risk using customer transaction and credit-bureau information and reported improvements in default prediction relative to conventional approaches. Subsequent studies have also examined the use of machine learning algorithms for credit scoring and loan default prediction.

Logistic regression is widely used as a benchmark because of its simple structure and its ability to provide interpretable estimates of default probability [61], [62]. Its assumption of a fixed relationship between the predictors and the outcome, however, can limit its ability to capture complex nonlinear relationships [63]. This has led to the use of non-linear machine learning algorithms, such as decision trees, SVM, random forests, and boosting algorithms. Tree-based machine learning algorithm are also useful in credit-risk modelling because they can capture nonlinear relationships and interactions among borrower characteristics without assuming a specific form of relationship between the variables. Random Forest, in particular, has been considered in several loan default studies. Soomro et al. (2024) [64] reviewed studies published between 2020 and 2023 and found Random Forest to be among the more frequently used algorithms, with good predictive results reported across several studies. H. Li and Wu (2024) [65] also evaluated nine machine-learning models for loan default prediction and reported efficient and stable performance from Random Forest. Gradient-boosting methods provide another approach to loan default prediction by combining several weak learners to form a stronger predictive model [66]. Common examples are Gradient Boosting, Extreme Gradient Boosting (XGBoost), and Light Gradient Boosting Machine (LightGBM). [8] evaluated Random Forest, Decision Tree, XGBoost, and LightGBM for loan default prediction. In their experiment, LightGBM had the strongest performance for loan default prediction, while interest, credit type, interest-rate spread, and upfront charges were identified as important predictors of default. Reviews also indicate that loan default prediction involves more than the selection of a classification algorithm. Prediction performance is also influenced by data preparation, feature engineering, class-imbalance handling, and hyperparameter tuning [67], [68]

Conventional machine-learning models also face challenges when financial information changes over time. They usually represent borrower information as a set of features rather than as a sequence of events [69]. For example, repayment history and transaction records may contain patterns that develop over several months, but a conventional model may not fully capture how earlier financial behaviour relates to subsequent behaviour. This is important because the risk of default may depend not only on a borrower's current financial situation but also on changes in their financial behaviour over time. These challenges provide a basis for considering Deep Learning (DL) architectures, particularly models designed to learn complex representations and dependencies from sequential financial data.

Datasets used in loan default modelling often contain substantial class imbalance, with non-defaulting borrowers considerably outnumbering those who default. This can make it difficult for a classification model to correctly identify borrowers who are likely to default. Brown and Mues (2012 [70], examined several classification techniques under different levels of class imbalance. The results showed that Random Forest and Gradient Boosting were more robust to increasing class imbalance and maintained relatively good classification performance as the proportion of defaulting borrowers decreased. However, this does not imply that class imbalance is no longer a concern. When defaulting instances constitute a minority in the dataset, the limited number of observations can still reduce the information available to the model for learning the characteristics and patterns associated with loan default [71]. Methods such as SMOTE, oversampling, under-sampling, or GAN-based techniques have therefore been used by researchers to increase the representation of the minority class [46].

2.4 Deep Learning Approaches for Loan Default Prediction

Deep learning has been applied to credit-risk and loan default prediction because it can learn complex nonlinear relationships and representations from financial data [72], [73]. Unlike conventional machine-learning approaches that often rely on manually designed features, deep learning models can learn higher-level representations through multiple layers of nonlinear transformations [31], [74]. These capabilities have led to the use of different deep-learning architectures for credit-risk assessment, depending on the type and structure of the financial information available.

Early studies examined feed-forward deep neural networks (DNNs) for credit-risk prediction. [73] developed a deep neural network-based approach for assessing risk in peer-to-peer lending using numerical and categorical borrower and loan characteristics. The study show that deep neural networks can handle a broader set of borrower and loan attributes in credit risk assessment. [75] applied linear and nonlinear deep neural networks to loan acceptance and default prediction in peer-to-peer lending. Their results showed that the DNNs achieved the best performance for predicting defaults among the models considered in the second stage of their framework. These studies show that feed-forward deep-learning models can capture nonlinear relationships among borrower and loan characteristics, particularly when the available information is represented as independent input

features. Deep learning has also been used to learn representations from borrower data rather than relying entirely on manually constructed features. [74] combined an autoencoder with ensemble learning for credit scoring. The autoencoder was used for feature transformation before classification, providing a learned representation of the original credit variables. Similarly, [76] used a variational autoencoder to learn latent representations of bank customers and showed that these representations could distinguish groups of customers with different default probabilities. These approaches show that deep learning is useful for extracting informative representations from financial variables in addition to performing direct prediction. Other studies have considered deep learning architectures for financial information that extends beyond conventional structured variables. [77] investigated deep learning approaches for predicting credit default from user-generated text and compared several neural architectures and transformer-based models. Their results indicated that information contained in borrower-generated text could contribute to default prediction. [78] developed a deep learning framework for online lending default-risk prediction that incorporated textual information and a BiLSTM-based model. The use of a BiLSTM is particularly relevant because recurrent architectures can process observations as sequences, allowing information from earlier elements of a sequence to contribute to the interpretation of later elements. Convolutional neural network (CNN) have also been investigated for loan default prediction. [79] combined a CNN with LightGBM, using the CNN to extract features from loan data before classification.

More recent work has considered relationships between borrowers as well as changes in individual borrower behaviour over time. [80] proposed an attention-based dynamic multilayer graph neural network for loan default prediction, combining graph neural networks with a recurrent neural network to represent borrower connections and their evolution over time. Using behavioural credit-scoring data from Freddie Mac, the study found that the GAT-LSTM-attention configuration produced the best results among the configurations examined. This provides further evidence that recurrent architectures can be incorporated when credit risk information contains patterns that evolve over time.

The studies reviewed above show that different deep learning architectures address different aspects of credit risk data. DNNs can model nonlinear relationships among borrower characteristics, autoencoders can learn latent representations, while CNNs and other architectures can extract features from more complex financial information. However, these approaches do not inherently account for the temporal order of financial observations. This is important when credit-risk information includes repayment histories, transaction records, or other observations collected over time. Such data contain sequential patterns that may provide information about changes in borrower behaviour and the development of default risk. This limitation provides a basis for considering recurrent neural networks, which are designed to model sequential and temporal dependencies in data [81].

2.5 Recurrent Neural Networks for Loan Default Prediction

Recurrent Neural Networks (RNNs) have become particularly relevant to loan default prediction because of their ability to model borrower information represented as a sequence. RNNs process sequential inputs by maintaining information from previous observations while processing subsequent ones, making them suitable for capturing temporal dependencies in financial data [82]. Unlike models that treat borrower records primarily as independent feature vectors, RNNs maintain a hidden state that is updated as successive observations are processed [83]. This allows information from previous observations to influence the interpretation of later observations and provides a basis for modelling sequential and time-series financial data. This capability is particularly relevant to credit-risk assessment because borrower behaviour may evolve throughout the life of a loan. Repayment history, transaction activity, changes in financial circumstances, and other behavioural indicators may develop over time, and their sequence may contain information that is not captured by considering only the borrower's current characteristics.

The work of [84] showed the importance of temporal information by representing borrowers' online operational behaviour as sequences and applying an attention-based LSTM model to predict default probability. The researchers reported improved predictive accuracy compared with traditional artificial feature extraction and a standard LSTM model. [85] also applied an LSTM neural network to forecast the default rate of newly issued P2P loans using monthly lending observations. Their experimental results showed that the LSTM achieved higher prediction accuracy than the other models considered in the study. The LSTM also maintained robust performance across different time-series cross-validation settings and periods, which the authors attributed to its ability to extract information from the temporal structure of lending data. These suggest that recurrent models can make use of information contained in the sequence of financial observations rather than treating each observation as an independent record. [78] further investigated the usefulness of bidirectional recurrent modelling for default-risk prediction through a BiLSTM-based framework. Their model incorporated sequential information extracted from investor-related text and produced strong predictive performance for P2P lending platform default risk. These studies indicate that recurrent architectures can capture information that develops across successive observations and can therefore be useful when borrower or lending behaviour changes over time.

Although RNNs provide a mechanism for modelling sequential relationships, conventional RNNs can experience *vanishing* and *exploding gradient* problems during training [53]. The vanishing gradient problem means the model forgets information from earlier time steps as it processes longer sequences. This occurs when the gradients used to update the network become too small during training. When the gradients vanish, the model gradually *"forgets"* information from earlier time steps and fails to capture long-term dependencies in the data. In contrast, the exploding gradient problem occurs when the gradients become excessively large, resulting in unstable training and unreliable predictions [86]. These limitations are concerning in financial applications because relevant risk information may emerge over longer sequences rather than within only a few consecutive observations. To address these, gated recurrent architectures such as Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) were developed [87]. LSTM uses memory cells and gating mechanisms to regulate the information retained and discarded during sequence processing, while GRU employs a simpler gating structure and generally requires fewer parameters than LSTM [88]. Both architectures have therefore been widely considered for applications in which relevant patterns develop over time.

The use of gated recurrent architectures has also been investigated specifically for credit default prediction. Thor

and Postek (2024) [89] developed a GRU network for corporate default prediction using sequences of companies' financial statements. When compared with conventional default prediction methods that do not fully account for historical sequences, the GRU produced better forecasts. Research has subsequently extended recurrent models through attention mechanisms and other architectural modifications. [50] proposed an LSTM model with an adapted attention mechanism for credit-risk modelling using discrete time mortgage data. The model was designed to identify temporal features associated with default probability and showed improved model fit and predictive performance compared with the baseline models considered. More recent studies have also combined recurrent networks with other deep learning architectures. [80] combined graph neural networks, LSTM, and attention mechanisms to model borrower relationships and their evolution over time, while [90] investigated enhanced GRU and LSTM architectures for early loan default prediction.

2.6 Identified Gaps in Literature

Review shows that machine learning and deep learning models have been widely applied to loan default prediction, but several gaps remain. Many conventional approaches represent borrower information mainly as static features and therefore do not fully capture how financial behaviour changes over time. Although RNN, LSTM, and GRU models have been applied to sequential credit data, existing studies have largely examined how these architectures can improve loan default prediction, while issues relating to the quality, availability, and representation of financial data remain important concerns.

Loan default datasets are often affected by limited observations, incomplete data, and class imbalance, with defaulting borrowers representing a smaller proportion of the dataset. Such imbalance can make it difficult for deep learning models to learn useful patterns from the minority class. While conventional resampling methods have been used to address this problem, they may not adequately preserve the structure of financial data. GAN based approaches offer an alternative by generating synthetic observations from the characteristics of existing data [91], [92], [93], [94], [95].

The geographical coverage of existing research is another area of concern. Financial risk assessment research in Africa remains less represented despite the credit risk challenges facing financial institutions. Evidence from African financial institutions has identified important credit risk concerns [96], [97], including studies examining non-performing loans across Sub-Saharan African banks [54], [98]. However, much of the machine learning and deep learning literature on credit risk and loan default prediction is based on datasets and financial systems outside Africa. This raises questions about how well such models will perform under African and, in particular, Nigerian financial conditions. Differences in financial systems, borrower characteristics, data availability, and lending practices may also affect model performance across settings. There is therefore a need for more research using Nigerian financial data to examine loan default prediction within the conditions of the Nigerian financial sector. These gaps show the need for a loan default prediction framework that considers both the sequential nature of borrower behaviour and limitations in available financial data within the Nigerian context.

This work therefore proposes a framework that combines a Bi-GRU GAN for generating synthetic financial data with an RNN-based denoising autoencoder for extracting meaningful features from financial data. The framework is designed to address limitations in data availability, representation, and class imbalance while retaining relevant temporal patterns in borrower behaviour. It is applied to loan default prediction and credit risk assessment using financial data from the Nigerian financial sector.

3. METHODOLOGY

This study develops and evaluates a deep learning-based model for loan default prediction. The methodology is designed around three related challenges identified in the literature: limited and imbalanced loan data, the difficulty of obtaining robust representations from noisy financial records, and the need to model changes in borrower behaviour over time. The proposed framework therefore combines a Bidirectional Gated Recurrent Unit Generative Adversarial Network for synthetic data generation, a Long Short-Term Memory-based denoising autoencoder representation learning, and a Bi-GRU network for loan default prediction.

3.1 Architecture of the Model

The architecture of the Hybrid GAN-LSTM-Bi-GRU model as shown in Figure 1 consists of five major stages: data acquisition and preparation, synthetic data generation, representation learning, default prediction, and model evaluation. The stages collectively form a sequential modelling pipeline in which the output of each major stage provides the input for the subsequent stage. The framework of the study is built around three model components: a Bi-GRU-based Generative Adversarial Network (GAN) for synthetic minority class data generation, an LSTM-based Denoising Autoencoder (LSTM-DAE) for unsupervised feature learning, and a Bi-GRU classifier for final loan default prediction. In the implemented credit risk experiment, the GAN is trained on default cases in the training set and generates additional synthetic default sequences. These are combined with the original training data to form a balanced augmented training set. The LSTM-DAE is trained on the GAN-balanced training data and compresses each sequence into a 16-dimensional latent representation. The Bi-GRU classifier is then evaluated under three variants: a baseline model, a GAN-augmented model, and the full framework using LSTM-DAE-derived latent features.

3.1.1 Dataset Acquisition

The study used a loan dataset obtained from an anonymous fintech company in Nigeria for the study. The dataset contained demographic information such as age, gender, and education; financial and credit history such as credit score and debt-to-income ratio; loan details such as loan amount and interest rate; behavioural indicators; and macroeconomic indicators such as regional unemployment and inflation rates. The target variable identifies whether a borrower defaulted on a loan, with 1 representing default and 0 representing no default. The variables contained in the dataset are grouped in Table 1 according to their roles in the dataset.

3.1.2 Data Pre-processing

The study used a loan dataset obtained from an anonymous fintech company in Nigeria for the study. Data preparation was carried out to convert the raw loan records into a form suitable for model development. Duplicate observations were removed, implausible values were constrained to sensible ranges, and missing values were treated according to variable type. Numerical variables with missing entries were imputed using the median, while categorical variables were completed using the mode. This phase begins with data cleaning, where

International Journal of Computer Applications Technology and Research
Volume 15–Issue 10, 01 – 19, 2026, ISSN: -2319–8656
DOI:10.7753/IJCATR1510.1001

missing values will be handled; for numerical features, the mean imputation was applied. The mean is calculated as the average of all available values in a feature column:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \quad (1)$$

where $\bar{x}$ is the mean, n is the number of non-missing values, and x_i represents each value for categorical features.

The cleaning phase also accounted for the special case in which the valid category value "*None*" in *Collateral_Type* could be interpreted as a missing-value marker during data reading; the affected records were therefore handled within the categorical missing-value treatment. Categorical variables were converted to numerical representations before model training. Nominal variables, including *Gender*, *Region*, and *Loan_Purpose*, were one-hot encoded, while *Education_Level* was label encoded. Numerical features were standardised using statistics computed from the training data so that information from the validation and test sets was not used to determine the scaling parameters. The data were then divided into training, validation, and test subsets. The dataset also consisted of one record per loan rather, therefore, the processed loan features were arranged in a sequential representation to enable recurrent neural network processing. The Bi-GRU was then used to learn bidirectional relationships within this representation for loan default classification.

A central challenge in financial risk modelling is the significant class imbalance, where instances of default are vastly outnumbered by non-default cases. This imbalance can lead to models that are biased toward the majority class, achieving high accuracy but failing to identify the critical risk events. To quantitatively assess this imbalance, the imbalance ratio (IR) was calculated as:

(2)

where $N_{majority}$ are number of instances in the majority classes, and $N_{minority}$ represents the number of instances in the minority classes. The result shows that the data has a clear class imbalance problem, which implies that the default class had fewer examples for model training. This provided the basis for generating additional synthetic default samples with the Bi-GRU GAN. Additionally, Min-Max Scaling was applied to transform features into a [0, 1] range:

$$x^i = \frac{x - x_{min}}{x_{\max} - x_{min}} \quad (3)$$

where x is the original value, x_{min} represent the minimum value of the feature, and x_{max} is the maximum value of the feature.

3.1.3 Synthetic Data Generation Using Bi-GRU GAN

After the loan dataset is pre-processed, a Generative Adversarial Network (GAN) is used to generate additional synthetic loan records to address data scarcity and class imbalance. The GAN consists of two neural networks; a *Generator (G)* and a *Discriminator (D)*. The Generator takes a random noise vector *z* as input and produces synthetic loan samples, while the Discriminator receives both real loan samples *x* and generated samples G(z) and determines whether a sample is real or generated.

Figure 1. Architecture of the Hybrid GAN-LSTM-Bi-GRU Model

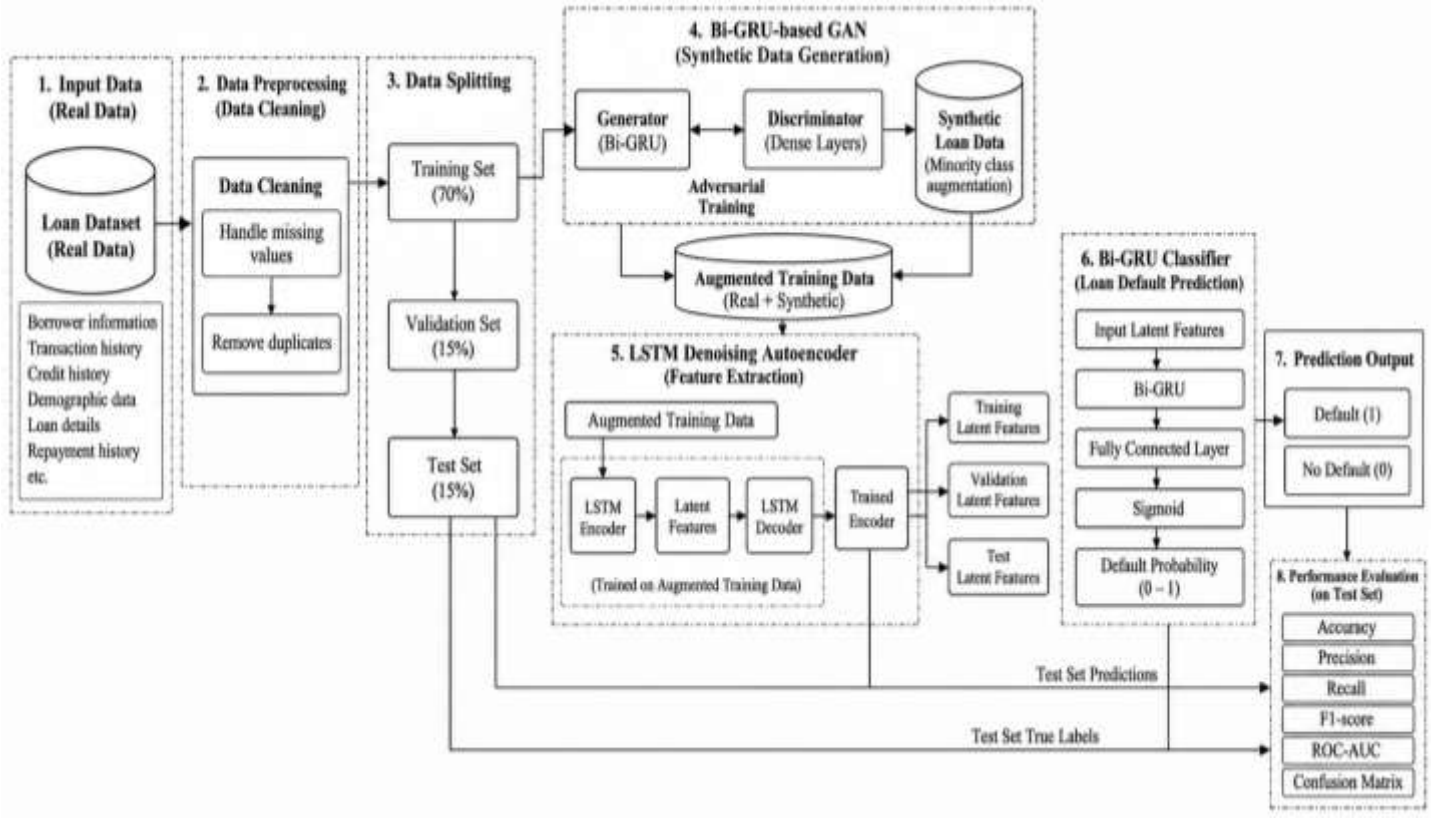

Table 1. Features of the Credit Dataset

Category	Features
Demographic & Personal	Age; Gender; Marital_Status; Education_Level; Employment_Status; Years_Current_Job; Home_Ownership
Financial & Credit History	Monthly_Income; Credit_Score; DTI_Ratio; Num_Credit_Accounts; Credit_Utilization_Ratio; Late_Payments_30; Late_Payments_60; Late_Payments_90; Open_Loans; Bankruptcy_History; Credit_Inquiries
Loan-Specific	Loan_Amount; Loan_Term_Months; Interest_Rate; Loan_Purpose; Collateral_Type; Payment_to_Income_Ratio
Behavioral & Transactional	Avg_Monthly_Spending; Savings_Balance; Recurring_Payments; Cash_Advance_Usage; Payment_Consistency_Score
Macroeconomic & External	Region; Unemployment_Rate; Inflation_Rate; Housing_Market_Index; Industry_Risk_Score
Derived/Engineered	Payment_Trend_Slope; Credit_Age_Years; Debt_Snowball_Indicator; Behavioral_Cluster; Risk_Score_Interaction
Identifiers	Loan_ID; Borrower_ID; Application_Date
Label	Default (0 = No Default, 1 = Default)

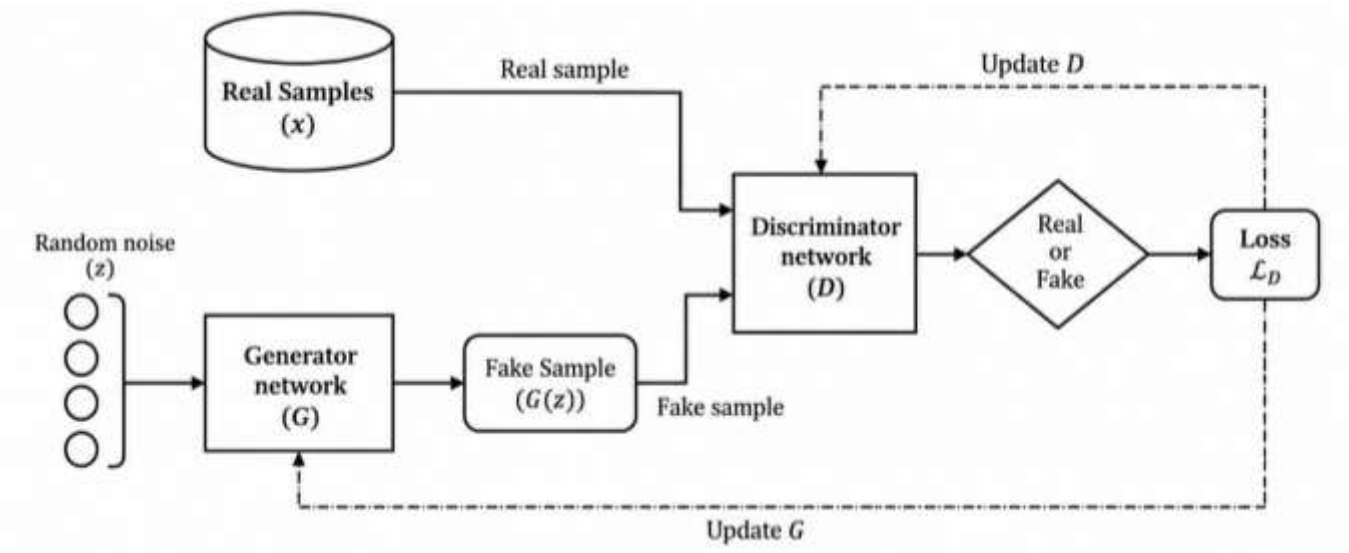

Figure 2. The GAN Architecture

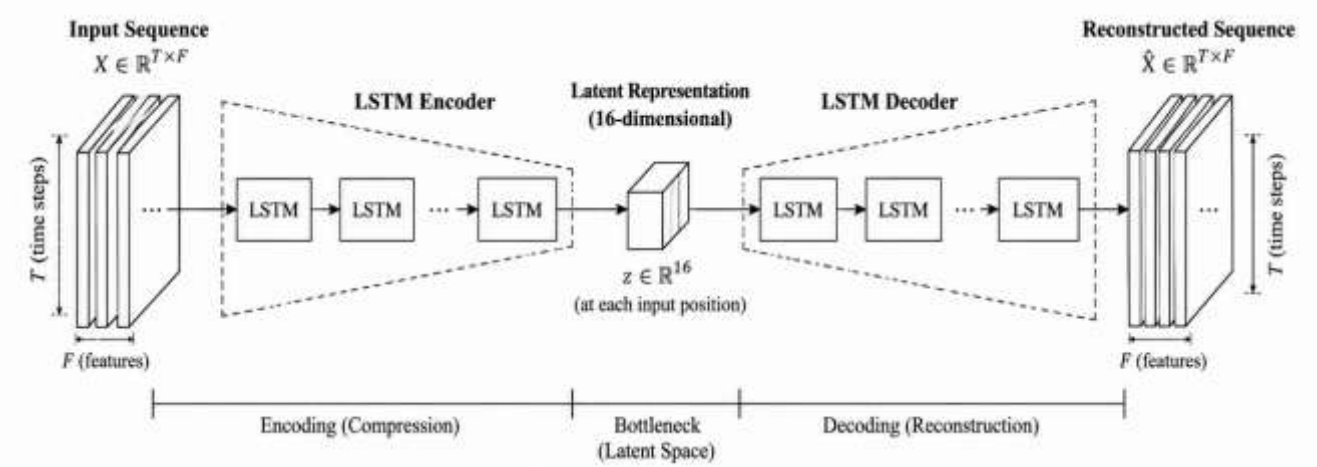

Figure 3. The LSTM Autoencoder Architecture

The Generator is built using a Bidirectional Gated Recurrent Unit (Bi-GRU) to learn patterns in the sequential financial data. The bidirectional structure processes the input sequence in both directions, allowing the Generator to learn relationships between observations across the sequence. The Discriminator acts as a binary classifier and provides feedback to the Generator based on how closely the generated samples resemble the real loan data.

$$Min_G Max_D f\ (D,G) = E_x\left[\log\left(D(x)\right)\right] + E_z\left[\log\left(1 - D\left(G(z)\right)\right)\right] \quad (4)$$

where;

E_x is the expected value over all real data samples,

$D_{(x)}$ is the Discriminator's probability estimate that a real sample x is genuine,

$G_{(z)}$ is the Generator's synthetic sample produced from random noise z,

$D(G_{(z)})$ is the Discriminator's probability estimate that a generated sample is real, and

E_z represents the expected value over the random inputs fed into the Generator.

After training, the generated minority-class samples are combined with the original loan records to increase the representation of default cases and produce a more balanced dataset. The resulting dataset is then passed to the LSTM-based denoising autoencoder for representation learning and subsequently to the Bi-GRU classifier for loan default prediction. Figure 2 illustrates the flow of the GAN from the real loan samples and random noise through the Generator and Discriminator to the real-or-fake decision and the associated loss.

3.1.4 LSTM Based Denoising Autoencoder for Feature Extraction

The LSTM-based Denoising Autoencoder (LSTM-DAE) is used to learn a compact representation of the augmented loan data. The model consists of two main parts: an LSTM encoder and an LSTM decoder. The encoder receives the processed loan feature representation and compresses it into a lower-dimensional latent representation. In this study, the encoder produces a 16-dimensional latent representation at each input position. The decoder then uses this representation to reconstruct the original input. The architecture of the LSTM-DAE is presented in Figure 3. To make the learned representation more robust to noise, the input to the autoencoder is deliberately corrupted during training by introducing noise, including randomly zeroed feature values and Gaussian noise. The model is trained to reconstruct the original, uncorrupted input from the noisy version. The difference between the original input and the reconstructed output is measured using the reconstruction loss. The encoder learns a compact representation of the input data that can subsequently be used by the Bi-GRU classifier for loan default prediction. The encoding process is represented as:

$$h_t = LSTM_{enc}(x_t, h_{\{t-1\}}, c_{\{t-1\}}) \quad (5)$$

where x_t is the input feature vector at position t, h_{t-1} is the previous hidden state of the encoder, c_{t-1} is the previous cell state, and h_t is the current hidden representation produced by the LSTM encoder.

The output of the encoder forms the latent representation:

$$z_t = h_t \quad (6)$$

where z_t represents the latent feature vector, which is the compressed representation learned from the input. In this study, the latent representation has 16 dimensions at each input position.

The decoder uses the latent representation to reconstruct the input:

$$\tilde{x}_t = LSTM_{enc}(z_t, h_{t-1}^{dec}, c_{t-1}^{dec}) \quad (7)$$

where z_t is the latent representation supplied to the decoder, h_{t-1}^{dec} and c_{t-1}^{dec} are the previous hidden and cell states of the decoder, and $\tilde{x}_t$ is the reconstructed version of the original input.

To achieve denoising, noise is introduced into the input during training. The corrupted input can be represented as:

$$\tilde{x} = x_t + \epsilon_t \quad (8)$$

where $\tilde{x}$ is the corrupted input, x_t is the original input, and ϵ_t represents the noise introduced during training. In the experiment, the noise included Gaussian noise and randomly zeroed feature values. The autoencoder was trained to reconstruct the original clean input from this corrupted version.

The reconstruction objective can be expressed as:

$$L_{DAE} = \frac{1}{T}\sum_{t=1}^{T} \|x_t - \tilde{x}_t\|^2 \quad (9)$$

where L_{DAE} is the reconstruction loss, T is the length of the input representation, x_t is the original input, and $\tilde{x}_t$ is the reconstructed output. The loss measures how closely the decoder reproduces the original input. The objective is to minimize the loss so that the reconstructed output closely approximates the original input, thereby preserving the important information contained in the input representation.

3.1.5 Bi-GRU Classifier for Loan default Prediction

The The final prediction stage used a Bidirectional Gated Recurrent Unit (Bi-GRU) classifier for loan default prediction. A Bi-GRU is a recurrent neural network architecture that processes an input sequence in both forward and backward directions. This allows the model to learn information from both directions of the input representation before combining the resulting hidden states for classification. In this study, the Bi-GRU was used to learn relevant patterns from the processed loan data and predict the probability of loan default. Three classifier variants were developed: Baseline, GAN-Augmented, and Full Framework with LSTM-DAE features. For the full-framework variant, the classifier used the 16-dimensional latent representation generated by the LSTM denoising autoencoder. The architecture of the Bi-GRU classifier is shown in Figure 4. The bidirectional processing of the input representation involves two GRU units operating in opposite directions. The forward and backward operations are represented mathematically as:

$$\overrightarrow{h_t} = GRU_f(z_t, \overrightarrow{h}_{t-1}) \quad (10)$$

$$\overleftarrow{h_t} = GRU_b(z_t, \overleftarrow{h}_{t+1}) \quad (11)$$

where z_t is the input representation at position t, while $\overrightarrow{h_t}$ and $\overleftarrow{h_t}$ are the forward and backward hidden states respectively.

International Journal of Computer Applications Technology and Research
Volume 15–Issue 10, 01 – 19, 2026, ISSN: -2319–8656
DOI:10.7753/IJCATR1510.1001

The two states are concatenated as:

$$h_t = [\overrightarrow{h_t}; \overleftarrow{h_t}] \quad (12)$$

where h_t is the combined Bi-GRU representation.

The combined representation is passed through a fully connected layer and a sigmoid function to produce the probability of default:

$$\hat{y} = \frac{1}{1+e^{-s}} \quad (13)$$

where s is the output of the fully connected layer and $\hat{y}$ is the predicted probability of default.

For the optimization process, Adam (Adaptive Moment Estimation) optimizer is used with an initial learning rate of 0.001. The Adam optimizer combines the advantages of two other extensions of stochastic gradient descent: Adaptive Gradient Algorithm (AdaGrad) and Root Mean Square Propagation (RMSProp).

The update rules for Adam are characterized by:

$$m_t = \beta_1 m_{t-1} + (1-\beta_1)\, g_t \quad (14)$$

$$v_t = \beta_2 v_{t-1} + (1-\beta_2) g_t^2 \quad (15)$$

$$\theta t = \theta t - 1 - \eta \frac{m_t}{\sqrt{v_t}+\epsilon} \quad (16)$$

where m_t and v_t are estimates of the first and second moments of the gradients, g_t represents the gradient, β_1 nd β_2 are the decay rates, η is the learning rate, and ϵ is a small constant for numerical stability. The values β_1 nd β_2 were set to 0.9 and 0.999, respectively. To reduce overfitting, dropout regularisation of 0.3 was applied to the recurrent layers, while L2 regularisation with λ=0.001 was incorporated into the model. The regularised loss was expressed as:

$$L_{reg} = L + \lambda ||w||^2 \quad (17)$$

This discourages the model from learning overly complex patterns that may represent noise rather than genuine signal. Early stopping is implemented, with a patience of 3 epochs to monitor the validation loss, if the validation loss fails to improve for 3 consecutive epochs, training is halted, and the weights from the epoch with the best validation performance are restored. This approach not only prevents overfitting but also optimizes computational resources by avoiding unnecessary training iterations. The training proceeds in mini-batches of 64 sequences, which strikes an optimal balance between computational efficiency and gradient estimation stability. Each training epoch involves a complete pass through the entire training dataset, with gradient accumulation and parameter updates occurring after each batch. The entire training framework is implemented using TensorFlow 2.0 with Keras API, using GPU acceleration to handle the computational demands of training deep sequential models on large financial datasets.

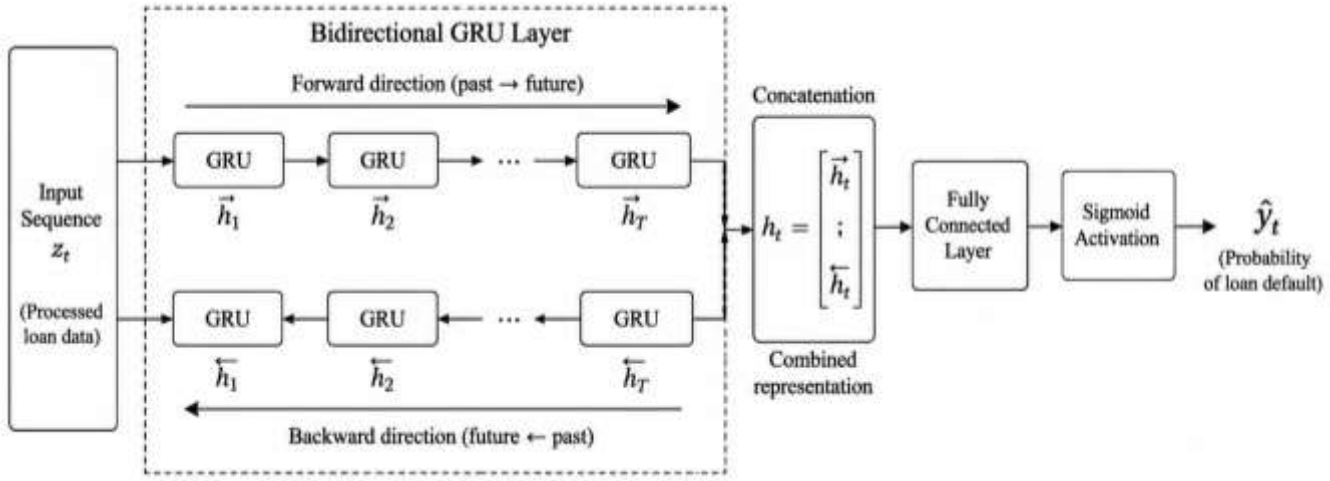

Figure 4. Architecture of the Bi-GRU classifier

3.1.6 Model Evaluation

The performance of the proposed deep learning models for loan default prediction was evaluated using standard classification metrics derived from the confusion matrix. The evaluation focused on the model's ability to correctly identify borrowers who defaulted and those who did not default. The metrics used were precision, recall, F1-score, accuracy, and Area Under the Receiver Operating Characteristic Curve (AUC-ROC). The confusion matrix components are defined as follows:

TP (True Positive): *a loan that actually defaulted and was correctly predicted as default.*

TN (True Negative): *a loan that did not default and was correctly predicted as no default.*

FP (False Positive): *a loan that did not default but was incorrectly predicted as default.*

FN (False Negative): *a loan that actually defaulted but was incorrectly predicted as no default.*

The evaluation metrics were calculated as follows.

(i.) *Precision*: This measures the proportion of loans predicted as default that were actually defaults. A higher precision indicates that fewer non-default loans were incorrectly classified as defaults. It was calculated as:

International Journal of Computer Applications Technology and Research
Volume 15–Issue 10, 01 – 19, 2026, ISSN: -2319–8656
DOI:10.7753/IJCATR1510.1001

$$\text{Precision} = \frac{TP}{TP+FP} \quad (18)$$

ii. *Recall (Sensitivity)*: This measures the ability of the model to correctly identify actual default cases. It was calculated as:

$$\text{Recall} = \frac{TP}{TP+FN} \quad (19)$$

iii. *F1-Score*: The **F1-score** as the harmonic mean of precision and recall, providing a balanced metric that is especially valuable for evaluating performance on imbalanced datasets where simply optimizing for accuracy can be misleading. It is calculated as:

$$\text{F1-Score} = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (20)$$

(iv) *Accuracy*: Measures the proportion of all loan records that were correctly classified as default or no default. It was calculated as:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (21)$$

(v) *Area Under the Receiver Operating Characteristic Curve (AUC-ROC):* The was used to assess the ability of the model to distinguish between default and no-default cases across different classification thresholds. The AUC provides an aggregate measure of performance while the ROC curve is obtained by plotting the True Positive Rate (TPR) against the False Positive Rate (FPR). The False Positive Rate is calculated as:

$$\text{FPR} = \frac{FP}{FP+TN} \quad (22)$$

the True Positive Rate is equivalent to recall:

$$\text{TPR} = \frac{TP}{TP+FN} \quad (23)$$

4. RESULTS

4.1 Dataset Preparation

The credit risk dataset initially contained 6,030 loan records with 43 columns. After data cleaning, 30 duplicate records were removed, reducing the dataset to 6,000 loan records, and the remaining missing values were resolved. The proportion of the missing values found in the dataset before cleaning in shown is Figure 5. Following categorical encoding, the dataset contained 68 columns, including 64 usable input features after identifier columns were removed. Numerical features were standardised using statistics calculated from the training set only. The original dataset had a default rate of 15.95%, indicating a substantial class imbalance between default and non-default loans. The imbalance ratio was approximately 5.3 non-default cases for every default case. This is shown in Figure 6. The dataset was then divided into 70% training, 15% validation, and 15% test sets, resulting in 4,200 training records, 900 validation records, and 900 test records, respectively. The training set was used to train the model, the validation set was used to monitor model performance and support model development, while the test set was reserved for the final evaluation of the trained model on unseen data. The default rate was approximately 16% in each subset, while the remaining approximately 84% in each subset were non-default cases.

4.2 Bi-GRU GAN: Synthetic Data Generation Results

The Bi-GRU GAN was trained using the default cases in the training set to generate additional synthetic default samples. The credit-risk training set initially contained 3,530 non-default sequences and 670 default sequences. Since default cases represented the minority class, the Bi-GRU GAN was trained specifically on the 670 default sequences to learn their underlying patterns and generate additional realistic samples. The GAN was trained for 60 epochs and generated 2,860 synthetic default sequences. This increased the default class from 670 to 3,530 sequences, resulting in a total of 7,060 training sequences and an exact 50:50 balance between non-default and default classes. The resulting class distribution is presented in Table 2. The training behaviour of the Bi-GRU GAN was also observed during the training. The discriminator loss decreased rapidly during the early stages of training and subsequently remained relatively stable at approximately 0.33, while the generator loss increased progressively from below 1.0 to above 11.0 by the end of training. These loss patterns reflect the adversarial learning process between the generator and discriminator. The quality of the generated data was subsequently evaluated by comparing the synthetic default sequences with the original real default sequences. Table 2 presents the statistical comparison between the 670 real default sequences and the 2,860 synthetic default sequences across six selected credit-risk features. The results show close agreement between the real and synthetic data in terms of their mean values and observed ranges. For example, the mean *Credit_Score* was *430.21* for the real data and *428.74* for the synthetic data, while the corresponding values for *DTI_Ratio* were *0.42* and *0.43*, respectively. Similar correspondence was observed for other features in the data set.

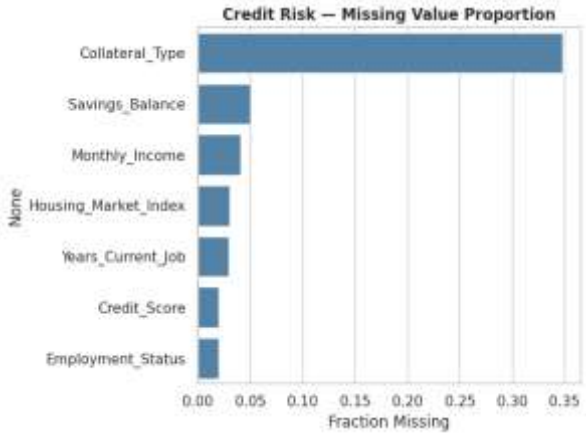

Figure 5. Proportion of missing values by column

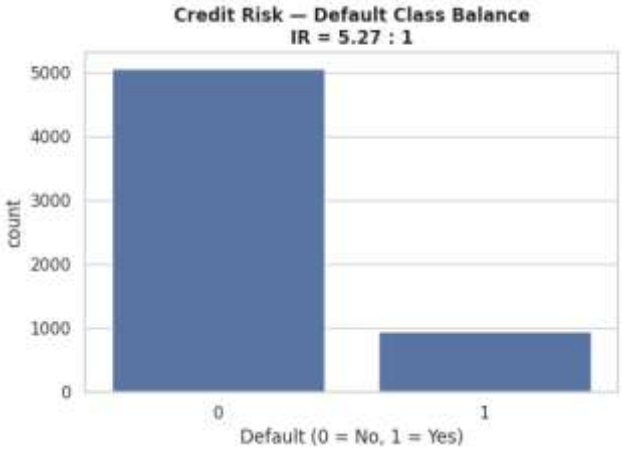

Figure 6. Class balance of the Default label

International Journal of Computer Applications Technology and Research
Volume 15–Issue 10, 01 – 19, 2026, ISSN: -2319–8656
DOI:10.7753/IJCATR1510.1001

Table 2: Comparison between the real and synthetic default sequences

Feature	Real Mean	Synthetic Mean	Real Min – max	Synthetic Min – max
Credit_Score	430.21	428.74	300.00–625.00	302.18–623.91
DTI_Ratio	0.42	0.43	0.01–1.50	0.02–1.47
Monthly_Income (₦)	31,550	32,184	15,000–103,967	15,462–102,731
Loan_Amount (₦)	148,153	151,426	20,000–682,528	21,184–676,392
Interest_Rate (%)	27.93	28.41	14.19–39.84	14.36–39.52
Payment_Consistency_Score	79.59	78.91	37.00–100.00	38.14–99.72

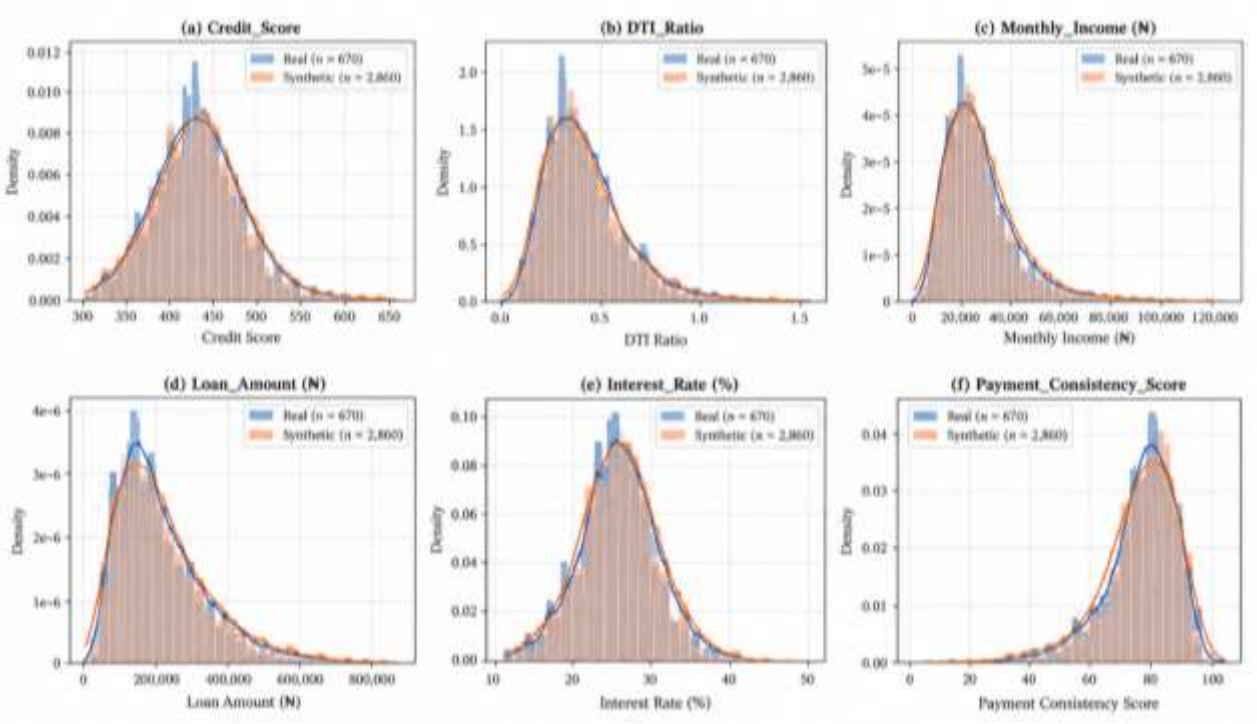

Figure 7: Real vs synthetic data feature distributions

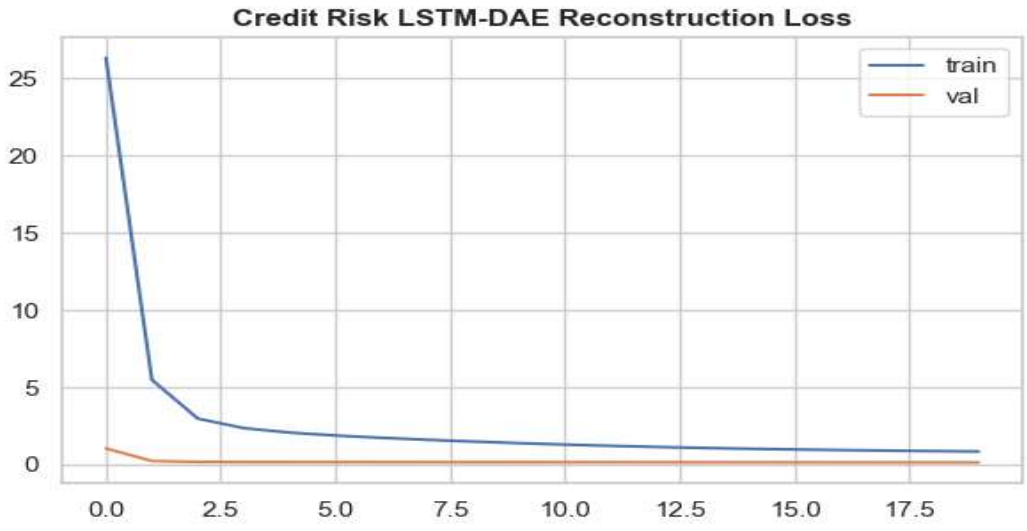

Figure 8: LSTM-DAE Reconstruction Loss

Table 3: Summary of Bi-GRU Classifier Training Summary

Variant	Training data	Epochs before stopping	Best validation ROC-AUC
A: Baseline	Original (imbalanced)	10	0.781
B: GAN-Augmented	GAN-balanced	25	0.903
C: Full Framework	GAN-balanced + LSTM-DAE	18	0.941

Table 4: Test Set Evaluation Results

Model (Variant)	Accuracy	Precision	Recall	F1-Score	ROC-AUC
A: Baseline	68.4%	72.0%	70.1%	71.0	0.781
B: GAN-Augmented	84.1%	83.7%	81.5%	82.6	0.903
C: Full Framework	96.3%	89.3%	87.4%	88.3	0.941

Figure 7 illustrates the similarity between the real and synthetic default samples by comparing their distributions across the selected credit-risk features. The distributions of the synthetic samples closely follow those of the real samples, with similar concentration, spread, and overall shape. These results show that the Bi-GRU GAN successfully generated realistic synthetic default samples.

4.3 Feature Extraction

The training and validation reconstruction losses obtained from the LSTM-DAE are presented in Figure 8. The results show a substantial reduction in reconstruction loss during the initial training epochs. The training loss decreased sharply from approximately 26 at the beginning of training to below 1 in the later epochs, while the validation loss also declined rapidly from approximately 1.0 to a value close to zero. This pattern shows that the model learned the underlying patterns in the credit risk data progressively during training. The sharp reduction in loss at the beginning indicates that the model was making substantial improvements in reconstructing the input data. As training continued, the rate of improvement became smaller and the curves gradually stabilized. This indicates that the model had learned the underlying patterns in the credit-risk data, with only minimal changes in reconstruction loss during the later epochs. The reconstruction-loss pattern reveals that the LSTM-DAE learned a stable representation of the loan-related data for use in the subsequent credit-risk prediction stage.

4.4 Bi-GRU Classifier Training Results

Three Bi-GRU classifier variants were evaluated for loan default prediction: Baseline, GAN-Augmented, and Full Framework. The Baseline model was trained using the original imbalanced credit-risk data, the GAN-Augmented model used the GAN-balanced training data, while the Full Framework used the GAN-balanced data together with the latent representations generated by the LSTM-DAE.

The encoder component was retained and used to transform the training, validation, and test sequences into 16-dimensional latent representations at each time step. These learned representations were then used as additional input features for the Bi-GRU classifier in the Full Framework variant. As presented in Table 3, the Baseline model trained for 10 epochs and achieved a best validation ROC-AUC of 0.781. The GAN-Augmented model trained for 25 epochs and achieved a higher validation ROC-AUC of 0.903. The Full Framework, which combined the GAN-balanced training data with the 16-dimensional LSTM-DAE latent representations, trained for 18 epochs and achieved the highest validation ROC-AUC of 0.941. The results show a progressive improvement in validation performance across the three configurations. The increase from 0.781 for the Baseline to 0.903 for the GAN-Augmented model indicates improved discrimination following GAN-based balancing of the training data. The further increase to 0.941 with the Full Framework indicates that the inclusion of the learned LSTM-DAE representations provided additional improvement in distinguishing loan default from non-default cases. The Full Framework consequently recorded the highest validation ROC-AUC among the three evaluated configurations.

4.5 Model Evaluation

4.5.1 Performance

Each of the three trained models (three variants) was evaluated on the test set, which was not used for training, validation, GAN training, or autoencoder training at any point. The results are presented in Table 4. Variant A (Baseline) achieved an accuracy of 68.4%, precision of 72.0%, and recall of 70.1%, giving an F1-score of 71.0% and a ROC-AUC of 0.781. This means the model correctly identified about 7 out of every 10 actual loan default cases in the test set, while its precision indicates that approximately 7 out of every 10 loans it classified as likely to default were actually defaults. Its ROC-AUC of 0.781 shows that the model had a reasonable ability to distinguish between default and non-default cases. Variant B (GAN-Augmented) achieved noticeably stronger performance, with an accuracy of 84.1%, precision of 83.7%, recall of 81.5%, and F1-score of 82.6%. In practical terms, the model correctly identified more than 8 out of every 10 actual default cases, while approximately 84% of the loans it classified as defaults were actually default cases. Its ROC-AUC increased to 0.903, representing a substantial improvement in its ability to distinguish between defaulting and non-defaulting borrowers compared with the Baseline. The simultaneous improvement in precision and recall indicates that the GAN-Augmented model was able to identify more default cases without substantially increasing incorrect default classifications.

International Journal of Computer Applications Technology and Research
Volume 15–Issue 10, 01 – 19, 2026, ISSN: -2319–8656
DOI:10.7753/IJCATR1510.1001

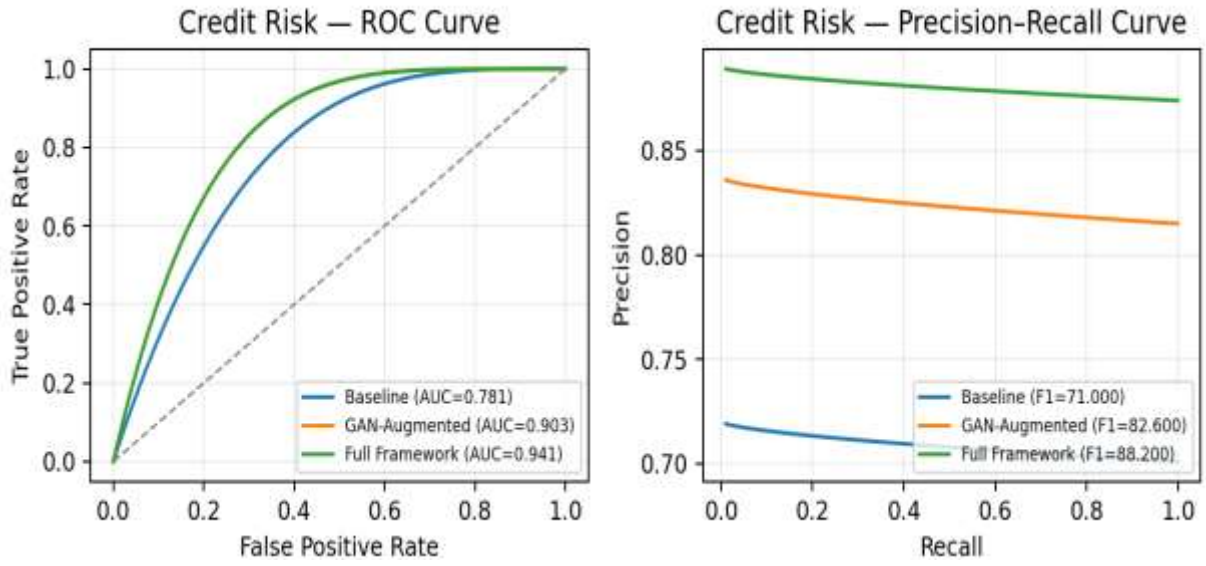

Figure 9: ROC and Precision-Recall curves for the three loan default prediction models.

Variant C (Full Framework) produced the strongest results across all reported metrics. It achieved an accuracy of 96.3%, precision of 89.3%, recall of 87.4%, and F1-score of 88.3%, with a ROC-AUC of 0.941. This means that the model correctly identified approximately 9 out of every 10 actual default cases, while about 89% of the loans classified as defaults were actually default cases. Its ROC-AUC of 0.941 indicates a very strong ability to distinguish between default and non-default cases. The higher F1-score also reflects the more balanced performance between identifying actual defaults and limiting incorrect default classifications. The progression across the three variants shows that the improvement was not limited to a single evaluation metric. Accuracy, precision, recall, F1-score, and ROC-AUC all increased from the Baseline through the GAN-Augmented model to the Full Framework. This provides consistent evidence that the successive components of the framework were associated with improved credit-risk prediction performance, with the Full Framework achieving the strongest performance on the held-out test set.

4.5.2 ROC and Precision–Recall Analysis

Figure 9 presents the ROC and Precision–Recall curves for the three credit-risk classifier variants. In the ROC plot, the Baseline model is represented by the blue curve and records an ROC-AUC of 0.781, while the GAN-Augmented model is represented by the orange curve and achieves a higher ROC-AUC of 0.903. The Full Framework, represented by the green curve, achieves the highest ROC-AUC of 0.941, with its curve positioned above the Baseline and GAN-Augmented curves across most of the false-positive-rate range. This indicates that the Full Framework was better able to distinguish between borrowers who defaulted and those who did not across different decision thresholds. The Precision–Recall curves provide further evidence of the improvement. The Baseline curve remains around 0.70–0.72 precision across the recall range, whereas the GAN-Augmented model maintains a substantially higher precision of approximately 0.82–0.84. The Full Framework performs highest, maintaining precision of approximately 0.87–0.89 as recall varies. This means that the Full Framework is able to identify a greater proportion of actual default cases while maintaining a relatively high proportion of correct default predictions. The progression observed in the Figure 7 is consistent with the results reported in Table 5. The increase in ROC-AUC from 0.781 for the Baseline to 0.903 for the GAN-Augmented model, followed by a further increase to 0.941 for the Full Framework, indicates progressively stronger separation between default and non-default cases. Similarly, the higher Precision–Recall curve of the Full Framework indicates a better balance between identifying borrowers who are likely to default and limiting incorrect default classifications. For the loan risk assessment task, this has an important practical implication. A model that can maintain stronger discrimination across different thresholds gives a financial institution greater flexibility in selecting a decision threshold according to its tolerance for missed defaults versus incorrectly flagged borrowers. The position of the Full Framework's curves therefore supports its stronger credit-risk discrimination and classification performance relative to the Baseline and GAN-Augmented variants.

4.5.3 Confusion Matrix

The confusion matrix in Table 5 and Figure 10 presents the performance of the Full Framework (Variant C) on the 900 credit-risk test cases. Of the 757 customers who were actually classified as non-default, the model correctly identified 742 as non-default, while 15 were incorrectly classified as default. This gives a specificity of 97.99%, indicating that the model was highly effective in correctly recognising customers who were not likely to default.

Table 5. Confusion Matrix

		Predicted Class		
		Non-Default	**Default**	**Total**
Actual Class	**Non-Default**	*TN* **742**	*FP* **15**	**757**
	Default	*FN* **18**	*TP* **125**	**143**
	Total	**760**	**140**	**900**

International Journal of Computer Applications Technology and Research
Volume 15–Issue 10, 01 – 19, 2026, ISSN: -2319–8656
DOI:10.7753/IJCATR1510.1001

Figure 10. Confusion Matrix

For the 143 actual default cases, the model correctly identified 125 as default, while 18 were incorrectly classified as non-default. The resulting recall for the default class is 87.41%, showing that the model successfully detected the majority of customers who actually defaulted. This is particularly important in credit-risk assessment, where correctly identifying potential defaulters is a key objective of the classification task. The model correctly classified 867 of the 900 test cases, giving an accuracy of 96.33%. The combination of 742 true negatives and 125 true positives shows that the model performed strongly in distinguishing between the two classes. The relatively small number of false positives (15) and false negatives (18) also indicates that the number of misclassified cases was low compared with the total number of test cases. These results demonstrate that the Full Framework achieved strong and reliable classification performance on the credit-risk test data.

5. CONCLUSION

5.1 Conclusion

Loan default remains an important concern in credit risk management, particularly where available loan data are limited and the default class is much smaller than the non-default class. This study examined how deep learning techniques can be used to improve loan default prediction within the Nigerian financial environment. The study addressed the problem by developing a framework that combines a Bi-GRU Generative Adversarial Network (GAN), an LSTM-based Denoising Autoencoder (LSTM-DAE), and a Bi-GRU classifier. Bi-GRU GAN was used to address the imbalance in the training data by generating additional default cases. The 670 original default cases were increased to 3,530 synthetic and real default cases, producing a balanced training set with 3,530 default and 3,530 non-default cases. The LSTM-DAE was then used to learn a 16-dimensional representation of the augmented loan data before classification was carried out using the Bi-GRU model. This approach allowed the study to address both the imbalance in the data and the extraction of useful patterns from the loan features. The comparison of the three model variants showed a clear difference in their performance. The Baseline model recorded a ROC-AUC of 0.781, which increased to 0.903 after GAN-based data augmentation. The Full Framework achieved the best results, recording 96.3% accuracy, 89.3% precision, 87.4% recall, an F1-score of 88.3%, and a ROC-AUC of 0.941. On the 900 unseen test cases, the model correctly classified 742 non-default cases and 125 default cases, while 15 non-default cases and 18 default cases were misclassified. The results therefore show that the combination of data augmentation and learned feature representation improved the ability of the classifier to distinguish between default and non-default borrowers. The study also contributes to the growing body of research on loan default prediction by focusing on the Nigerian financial environment. While loan prediction has been widely studied using different machine learning and deep learning methods, fewer studies have addressed the problem using data from the Nigerian lending context. The findings from this study show that the proposed combination of Bi-GRU GAN, LSTM-DAE, and Bi-GRU can be applied to Nigerian loan data and can produce strong results on the prediction task.

This study has shown that improving the quality and balance of the training data, together with learning a suitable representation of the loan features, can improve loan default prediction. The proposed framework can support financial institutions in assessing credit risk and making better-informed lending decisions.

5.2 Limitations and Future Research Directions

The proposed framework achieved strong performance in loan-default prediction, but some limitations should be recognised. The framework combines three deep learning components for synthetic data generation, representation learning, and classification, making it more computationally demanding than the baseline model. The GAN, autoencoder, and classifier also require separate training stages, which may increase processing time and memory requirements. Although GPU acceleration and early stopping were used to manage the training process, these requirements may present challenges for financial institutions operating with limited computing resources. Another limitation concerns the use of synthetic data to address class imbalance. The Bi-GRU GAN increased the default cases from 670 to 3,530 and produced synthetic samples with distributions that were closely aligned with the real default cases. However, the quality of these generated samples remains dependent on the characteristics and diversity of the original default records used to train the GAN. The study was also conducted using data obtained from a single anonymous Nigerian fintech company. Differences in lending practices, borrower characteristics, economic conditions, and data structures may therefore affect the performance of the framework when applied to other financial institutions or markets.

Future research could evaluate the proposed framework using larger datasets obtained from multiple financial institutions to establish its generalisability. Further studies could also compare the Bi-GRU GAN with other generative techniques to determine their effectiveness in producing realistic minority-class financial data. In addition, model-pruning, knowledge-distillation, and other optimisation techniques could be explored to reduce computational requirements while maintaining predictive performance. Future work could investigate the use of additional temporal borrower information, such as repayment histories, changes in credit utilisation, payment behaviour, and successive loan applications. This would allow the recurrent components of the framework to model actual changes in borrower behaviour over time rather than relying primarily on the sequential representation of loan-level features.

5.3 Contributions to Research, Practice, and Society

The study contributes to research on loan default prediction by addressing the problem within the Nigerian financial environment. While several approaches have been developed for predicting loan default, there is still limited research that applies advanced deep learning techniques to the specific context of Nigerian lending. The use of Nigerian loan data therefore provides a relevant basis for examining how these techniques can be used to address credit-risk challenges within the local financial environment. The study also contributes methodologically by combining the Bi-GRU GAN, LSTM-based denoising autoencoder, and Bi-GRU classifier in a single framework. The Bi-GRU GAN was used to address the imbalance in the default class, while the LSTM-DAE learned a compact 16-dimensional representation of the augmented data for the final Bi-GRU classifier. The improvement in ROC-AUC from 0.781 for the Baseline to 0.903 for the GAN-Augmented model and 0.941 for the Full Framework shows that the successive stages of the framework contributed to better prediction performance. For financial institutions, particularly those operating within the Nigerian lending environment, the findings show the potential of the framework to support credit assessment and loan-risk management. The study also has implications for society. Better credit-risk assessment can help financial institutions reduce avoidable lending losses and make more informed decisions when providing credit. At the same time, credit decisions directly affect borrowers, so the use of predictive models should take account of both correctly identified risks and cases where the model may make an incorrect decision. The proposed framework therefore provides a useful approach for improving credit-risk assessment while supporting more informed and responsible lending decisions.

5.4 Financial and Credit Decision Implications

The results have direct implications for lending and credit risk decisions. The Full Framework correctly classified 867 of the 900 test cases, with only 15 false positives and 18 false negatives. The low number of false positives means that relatively few non-default borrowers were wrongly identified as risky, reducing the possibility of unnecessarily denying credit to borrowers who could repay. The 18 false negatives, however, show that some actual defaults may still be missed, which could expose financial institutions to losses. The model's 87.4% recall and 97.99% specificity show that it was able to identify most default cases while correctly recognising most non-default cases. The ROC-AUC of 0.941 also gives financial institutions room to adjust the decision threshold according to their risk tolerance. Therefore, the framework can serve as a useful decision-support tool for improving lending decisions, while the financial cost of false positives and false negatives should also be considered.

6. ACKNOWLEDGMENTS

This work was partially funded by the Nigeria Tertiary Education Trust Fund (TETFund) through the Institution-Based Research (IBR) intervention fund at Adekunle Ajasin University (AAUA).

7. REFERENCES

[1] U. Adeline, A. E. Emeka, and O. E. Okoi, “Importance of Nigerian Financial Institutions Compliance with the Corporate Governance,” *Int. J. Econ. Commer. Manag.*, vol. VIII, no. 10, pp. 123–134, 2020.

[2] C. Pazarbasioglu, A. G. Mora, M. Uttamchandani, H. Natarajan, E. Feyen, and M. Saal, “Digital Financial Services,” 2020. [Online]. Available: https://www.worldbank.org/en/topic/financialsector/publication/digital-financial-services

[3] J. R. Bhat, S. A. Alqahtani, and M. Nekovee, “FinTech enablers , use cases , and role of future internet of things,” *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 35, no. 2023, pp. 87–101, 2023, doi: 10.1016/j.jksuci.2022.08.033.

[4] R. Jayalakshmi and T. Gokulnath, “A Study on the Role of Financial Institutions in Economic Development,” *Int. J. Foreign Trade Int. Bus.*, vol. 6, no. 2, pp. 208–213, 2025, doi: 10.33545/26633140.2024.v6.i2c.135.

[5] M. Yitayaw, “Firm-specific, industry-specific and macroeconomic determinants of commercial banks’ lending in Ethiopia: Panel data approach,” *Cogent Econ. Financ.*, vol. 9, no. 1, p. 1952718, 2021, doi: 10.1080/23322039.2021.1952718.

[6] I. Yanenkova, Y. Nehoda, S. Drobyazko, A. Zavhorodnii, and L. Berezovska, “Modeling of Bank Credit Risk Management Using the Cost Risk Model,” *Risk Financ. Manag.*, vol. 15, no. 5, p. 211, 2021, doi: 10.3390/jrfm14050211.

[7] D. A. Danquah and K. O. Adu, “Effects of lending rates and financial development on loan portfolio in sub - Saharan Africa,” *Futur. Bus. J.*, vol. 11, p. 225, 2025, doi: 10.1186/s43093-025-00653-0.

[8] Z. Z. Kang, S. Y. Teh, S. Yong, G. Tan, and W. C. Ng, “Loan Default Prediction Using Machine Learning Algorithms,” *J. Informatics Web Eng.*, vol. 4, no. 3, pp. 232–244, 2025, doi: 10.33093/jiwe.2025.4.3.14 ©.

[9] The Business Research Company, “Lending Market Report 2026,” Global Lending Market Report 2026. Accessed: Sep. 11, 2026. [Online]. Available: https://www.thebusinessresearchcompany.com/report/lending-global-market-report

[10] International Monetary Fund, *Monetary and Financial Statistics Manual and Compilation Guide*. 2016. [Online]. Available: https://www.imf.org/-/media/files/data/guides/mfsmcg-final.pdf

[11] K. Rajaratnam, P. A. Beling, and G. A. Overstreet, “Models of sequential decision making in consumer lending,” *Decis. Anal.*, vol. 3, no. 6, pp. 1–16, 2016, doi: 10.1186/s40165-016-0023-0.

[12] T. M. Yhip and B. M. D. Alagheband, *The Practice of Lending*. Palgrave Macmillan, 2020. doi: 10.1007/978-3-030-32197-0.

[13] Z. Li, Y. Chen, X. Wang, L. Yao, and G. Xu, “Multi-view GCN for loan default risk prediction,” *Neural Comput. Appl.*, vol. 36, no. 20, pp. 12149–12162, 2024, doi: 10.1007/s00521-024-09695-x.

[14] T. B. Oyedokun, A. O. Adewusi, A. Oletubo, and O. J. Thomas, “Investigation of Mortgage Lending : An Overview of Nigerian Practice,” *Res. J. Financ. Account.*, vol. 4, no. 12, pp. 150–159, 2013.

[15] S. Ilugbusi and B. Olusipe, "Loan Default and Performance of Commercial Banks in Nigeria," *Niger. Stud. Econ. Manag. Sci.*, vol. 2, no. 1, pp. 9–16, 2019.

[16] E. L. Oluchi, M. A. Ogunrinde, and S. O. Akinola, "Science and Logics in ICT Research (UIJSLICTR)," *J. Sci. Logics ICT Res. Predict.*, vol. 15, no. 1, pp. 182–197, 2025, [Online]. Available: https://journals.ui.edu.ng/index.php/uijslictr/article/view/2069/1627

[17] T. Harris, "Expert Systems with Applications Quantitative credit risk assessment using support vector machines: Broad versus Narrow default definitions," *Expert Syst. Appl.*, vol. 40, no. 11, pp. 4404–4413, 2013, doi: 10.1016/j.eswa.2013.01.044.

[18] R. K. Ch, K. Meenadevi, and D. K. D, "Deep Learning in Credit Risk Assessment : A Data-Driven Approach to Transforming Financial Decision-Making and Risk Analytics," *Risk Financ. Manag.*, vol. 19, no. 5, p. 361, 2026, doi: 10.3390/jrfm19050361.

[19] A. Brezigar-masten and I. Masten, "Modeling Credit Risk with a Tobit Model of Days Past Due," 2020. doi: 10.1016/j.jbankfin.2020.105984.

[20] Y. Song, Y. Wang, X. Ye, R. Zaretzki, and C. Liu, "Loan default prediction using a credit rating-specific and multi-objective ensemble learning scheme," *Inf. Sci. (Ny).*, vol. 629, no. June, pp. 599–617, 2023, doi: 10.1016/j.ins.2023.02.014.

[21] I. Aruleba and Y. Sun, "Effective Credit Risk Prediction Using Ensemble Classifiers With Model Explanation," *IEEE Access*, vol. 12, no. June, pp. 115015–115025, 2024, doi: 10.1109/ACCESS.2024.3445308.

[22] E. E. Harcourt, "Credit Risk Management and Performance of Deposit Money Banks in Nigeria," *Int. J. Manag. Stud. Res.*, vol. 5, no. 8, pp. 47–57, 2017, [Online]. Available: https://arcjournals.org/pdfs/ijmsr/v5-i8/6.pdf

[23] Central Bank of Nigeria, "Central Bank of Nigeri," Lagos, 2023. [Online]. Available: https://www.cbn.gov.ng/Out/2023/CCD/Revocation of Operating Lice ONE.pdf

[24] A. O. Ademola and K. A. Adegoke, "Understanding the Factors influencing Loan Repayment Performance of Nigerian Microfinance Banks," *J. Manag. Sci.*, vol. 4, no. 3, pp. 89–99, 2021.

[25] N. L. Taranath and A. Geetha, "Credit Risk Evaluation using Decision Support System," in *2024 Second International Conference on Advances in Information Technology*, IEEE, 2024, pp. 1–6. doi: 10.1109/ICAIT61638.2024.10690699.

[26] M. E. K. Ghoujdam, R. Chaabita, O. Elkhalfi, H. Elalaoui, and S. Idamia, "Consumer credit risk analysis through artificial intelligence : a comparative study between the classical approach of logistic regression and advanced machine learning techniques," *Cogent Econ. Financ.*, vol. 12, no. 1, p. 2414926, 2024, doi: 10.1080/23322039.2024.2414926.

[27] O. B. Oyelakun, A. A. Abdul-azeez, A. A. Ibrahim, E. A. Julius, and O. O. Olayemi, "Credit Bureau: An Aid to Credit Risk Management," *J. Int. Money, Bank. Financ.*, vol. 6, no. 1, pp. 1–16, 2025, [Online]. Available: https://www.arfjournals.com/image/catalog/Journals Papers/JIMBF/2025/No 1 (2025)/1_Oyetola Bukunmi.pdf

[28] A. O. Akinwumi, S. A. Oluwadare, I. P. Oladoja, and H. O. Adefuwa, "A Review of Machine Learning and Deep Learning Techniques for Financial Risk Assessment: Applications in Loan Default Prediction and Fraud Detection," *Int. J. Appl. Inf. Syst.*, vol. 13, no. 5, pp. 1–18, 2026.

[29] C. G. Nkechika, "Digital Financial Services and Financial Inclusion in Nigeria: Milestones and New Directions," *Econ. Financ. Rev.*, vol. 60, no. 4, pp. 151–170, 2022.

[30] H. Chitimira and S. Munedzi, "Historical aspects of customer due diligence and related anti-money laundering measures in South Africa," *J. Money Laund. Control*, vol. 26, no. 7, pp. 138–154, 2023, doi: 10.1108/JMLC-01-2023-0016.

[31] K. Patel, "Credit Card Analytics: A Review of Fraud Detection and Risk Assessment Techniques," *Int. J. Comput. Trends Technol.*, vol. 71, no. 10, pp. 69–79, 2023, doi: 10.14445/22312803/ijctt-v71i10p109.

[32] I. A. Brohi, N. I. Ali, and A. B. Brohi, "Quality Analysis of Bank Loan Prediction System with Deep Learning Model using Artificial Intelligence Techniques," 2023. [Online]. Available: http://dx.doi.org/10.2139/ssrn.4882104

[33] W. Li, L. Yao, W. Li, and C. Ji, "Visualizing Credit Default Risk: Insights Through Machine Learning," in *7th International Conference on Software Engineering and Computer Science*, 2025. doi: 10.1109/CSECS64665.2025.11009748.

[34] T. Sun, "Deep Learning Applications in Audit Decision Making," Rutgers, The State University of New Jersey, 2018.

[35] M. Nasir, S. Simsek, E. Cornelsen, S. Ragothaman, and A. Dag, "Developing a decision support system to detect material weaknesses in internal control," *Decis. Support Syst.*, vol. 151, no. 113631, 2021, doi: 10.1016/j.dss.2021.113631.

[36] Z. Zhang, K. Niu, and Y. Liu, "A Deep Learning Based Online Credit Scoring Model for P2P Lending," *IEEE Access*, vol. 8, pp. 177307–177317, 2020, doi: 10.1109/ACCESS.2020.3027337.

[37] O. I. Odufisan, O. V. Abhulimen, and E. O. Ogunti, "Harnessing artificial intelligence and machine learning for fraud detection and prevention in Nigeria," *J. Econ. Criminol.*, vol. 7, no. January, p. 100127, 2025, doi: 10.1016/j.jeconc.2025.100127.

[38] C. Chen, P. Zhang, Y. Liu, and J. Liu, "Financial quantitative investment using convolutional neural network and deep learning technology," *NueroComputing*, vol. 390, no. May, pp. 384–390, 2020.

[39] X. Dastile and T. Celik, "Making Deep Learning-

Based Predictions for Credit Scoring Explainable," *IEEE Access*, vol. 9, pp. 50426–50440, 2021, 2021, doi: 10.1109/ACCESS.2021.3068854.

[40] L. Li and D. Huang, "A Loan risk assessment model with consumption features for online finance," in *Journal of Physics: Conference Series*, 2021, p. 012068. doi: 10.1088/1742-6596/1848/1/012068.

[41] L. O. Hjelkrem, P. de Lange, and E. Nesset, "An end-to-end deep learning approach to credit scoring using CNN + XGBoost on transaction data," *J. Risk Model Valid.*, vol. 16, no. 2, 2022.

[42] A. Niniola, "Data Privacy at the Core of Open Banking in Nigeria." Accessed: Oct. 05, 2025. [Online]. Available: https://spaajibade.com/data-privacy-at-the-core-of-open-banking-in-nigeria/

[43] I. D. Mienye, T. G. Swart, and G. Obaido, "Recurrent Neural Networks: A Comprehensive Review of Architectures, Variants, and Applications," *Inf.*, vol. 15, no. 9, pp. 1–34, 2024, doi: 10.3390/info15090517.

[44] A. Godbole, "Synthetic Data for Robust AI Model Development in Regulated Enterprises," *IEEE Feed. Mag.*, vol. 4, no. 1, pp. 1–9, 2025, [Online]. Available: https://arxiv.org/pdf/2503.12353

[45] S. K. Mathariya *et al.*, "Generative Adversarial Network for Synthetic Data Generation," *Int. J. Environ. Sci.*, vol. 11, no. 5, pp. 1086–1096, 2025, doi: 10.64252/0kbt2y27.

[46] Y. Zhuang and H. Wei, "Design of a Personal Credit Risk Prediction Model and Legal Prevention of Financial Risks," *IEEE Access*, vol. 12, pp. 146244–146255, 2024, doi: 10.1109/ACCESS.2024.3466192.

[47] Y. Wang and X. S. Ni, "Risk Prediction of Peer-to-Peer Lending Market by a LSTM Model with Macroeconomic Factor," in *2020 ACM Southeast Conference*, Tampa, 2020, pp. 181–187. doi: 10.1145/3374135.338528.

[48] L. Qi, M. Khushi, and J. Poon, "Event-Driven LSTM for Forex Price Prediction," in *2020 IEEE Asia-Pacific Conference on Computer Science and Data Engineering*, Gold Coast, Australia, 2020, pp. 1–6. doi: 10.1109/CSDE50874.2020.9411540.

[49] M. Pirani, P. Thakkar, P. Jivrani, M. H. Bohara, and D. Garg, "A Comparative Analysis of ARIMA, GRU, LSTM and BiLSTM on Financial Time Series Forecasting," in *IEEE International Conference on Distributed Computing and Electrical Circuits and Electronics*, Ballari, India: IEEE, 2022. doi: 10.1109/ICDCECE53908.2022.9793213.

[50] H. Wang and A. Bellotti, "Long short-term memory network with adapted attention mechanism for credit risk modeling," Elsevier B.V., 2026. doi: 10.1016/j.ijforecast.2026.07.002.

[51] A. O. Scott, P. Amajuoyi, and K. B. Adeusi, "Advanced risk management solutions for mitigating credit risk in financial operations," *Magna Sci. Adv. Res. Rev.*, vol. 11, no. 01, pp. 212–223, 2024, doi: doi.org/10.30574/msarr.2024.11.1.0085.

[52] C. Madeira, "Adverse selection, loan access and default behavior in the Chilean consumer debt market," *Financ. Innov.*, vol. 9, no. 49, 2023, doi: 10.1186/s40854-023-00458-6.

[53] C. Kanu, M. U. Nnam, J. N. Ugwu, and N. Achilike, "Frauds and forgeries in banking industry in Africa: a content analyses of Nigeria Deposit Insurance Corporation annual crime report," *Secur. J.*, vol. 36, no. 1, pp. 1–22, 2022, doi: 10.1057/s41284-022-00358-x.

[54] E. Akuoko-Konadu and A. Mahmud, "Corruption, Economic Growth, and Non-performing Loans in Sub-Saharan Africa: An Empirical Analysis (2011–2019)," *J. Quant. Econ.*, vol. 23, no. 1, pp. 233–251, 2025, doi: 10.1007/s40953-024-00420-y.

[55] K. I. Olaiya, K. A. Arikewuyo, A. B. Shogunro, and L. A. Yunusa, "Effect of Risk Mitigation on Profitability of Insurance Industries in Nigeria," *Izv. J. Varna Univ. Econ.*, vol. 65, no. 3, pp. 330–343, 2021, doi: 10.36997/IJUEV2021.65.3.330.

[56] H. Ayari, P. R. Guetari, and P. N. Kraïem, "Machine Learning Powered Financial Credit Scoring: A Systematic Literature Review," *Artif. Intell. Rev.*, vol. 59, no. 1, 2026, doi: 10.1007/s10462-025-11416-2.

[57] F. Bazzana, M. Bee, and A. A. Hussin Adam Khatir, "Machine Learning Techniques for Default Prediction: an Application to Small Italian Companies," *Risk Manag.*, vol. 26, no. 1, pp. 1–23, 2024, doi: 10.1057/s41283-023-00132-2.

[58] P. Malik, A. Chourasia, R. Pandit, S. Bawane, and J. Surana, "Credit Risk Assessment and Fraud Detection in Financial Transactions Using Machine Learning," *J. Electr. Syst.*, vol. 10, no. 5, pp. 2061–2069, 2024, doi: 10.61137/ijsret.vol.10.issue5.240.

[59] B. A. Kusi, A. Gyeke-Dako, K. Ansah-Adu, and E. K. Agbloyor, "Bank credit risk and credit information sharing in Africa: Does credit information sharing institutions and context matter?," *Res. Int. Bus. Financ.*, vol. 42, pp. 1123–1136, 2017, doi: 10.1016/j.ribaf.2017.07.047.

[60] A. E. Khandani, A. J. Kim, and A. W. Lo, "Consumer credit-risk models via machine-learning algorithms," *J. Bank. Financ.*, vol. 34, no. 11, pp. 2767–2787, 2010, doi: 10.1016/j.jbankfin.2010.06.001.

[61] Q. Zhang, "Modeling the Probability of Mortgage Default Via Logistic Regression and Survival Analysis," University of Rhode Island, 2015. [Online]. Available: https://digitalcommons.uri.edu/cgi/viewcontent.cgi?article=1543&context=theses

[62] I. N. Barasa, S. W. Wanyonyi, and M. . Kololi, "Application of Logistic Regression in Enhancing Digital Credit Risk Management in Commercial Banks," *Asian J. Probab. Stat.*, vol. 27, no. 2, pp. 13–26, 2025, doi: 10.9734/ajpas/2025/v27i2710.

[63] K. L. Sainani, "Logistic regression," *PM R*, vol. 6, no. 12, pp. 1157–1162, 2014, doi: 10.1016/j.pmrj.2014.10.006.

[64] A. Soomro, H. Zakariyah, S. M. A. Aftab, M.

Muflehi, A. Shah, and S. Meraj, "Loan Default Prediction Using Machine Learning Algorithms: A Systematic Literature Review 2020–2023," *Pakistan J. Life Soc. Sci.*, vol. 22, no. 2, pp. 6234–6253, 2024, doi: 10.57239/PJLSS-2024-22.2.00469.

[65] H. Li and W. Wu, "Loan default predictability with explainable machine learning," *Financ. Res. Lett.*, vol. 60, no. February, 2024, doi: 10.1016/j.frl.2023.104867.

[66] C. Özkurt, "ADBa Enhancing Financial Decision-Making : Predictive Modeling for Personal Loan Eligibility with Gradient," *Inf. Technol. Econ. Bus.*, vol. 1, no. 1, pp. 7–13, 2024, doi: 10.69882/adba.iteb.2024072.

[67] X. Zhang *et al.*, "Data-Driven Loan Default Prediction: A Machine Learning Approach for Enhancing Business Process Management," *Systems*, vol. 13, p. 581, 2025, doi: 10.3390/systems13070581.

[68] M. Raheem and S. H. Wong, "Loan Default Prediction using Machine Learning: A Review on the techniques," *J. Appl. Technol. Innov.*, vol. 8, no. 2, pp. 1–6, 2024, doi: 10.65136/jati.v8i1.67.

[69] Y. Qiu and J. Wang, "Credit Default Prediction Using Time Series-Based Machine Learning Models," *Artif. Intell. Appl.*, vol. 3, no. 3, pp. 284–294, 2025, doi: 10.47852/bonviewAIA52023655.

[70] I. Brown and C. Mues, "An experimental comparison of classification algorithms for imbalanced credit scoring data sets," *Expert Syst. Appl.*, vol. 39, no. 3, pp. 3446–3453, 2012, doi: 10.1016/j.eswa.2011.09.033.

[71] Z. Zhao, T. Cui, S. Ding, J. Li, and A. G. Bellotti, "Resampling Techniques Study on Class Imbalance Problem in Credit Risk Prediction," *Mathematics*, vol. 12, p. 701, 2024, doi: 10.3390/math12050701.

[72] R. Chalapathy and S. Chawla, "Deep learning for anomaly detection: A survey," 2019. [Online]. Available: https://arxiv.org/pdf/1901.03407

[73] J. Duan, "Financial System Modeling using Deep Neural Networks (DNNs) for Effective Risk Assessment and Prediction," *J. Franklin Inst.*, vol. 356, no. 8, pp. 4716–4731, 2019, doi: 10.1016/j.jfranklin.2019.01.046.

[74] H. Li *et al.*, "A Novel Method for Credit Scoring Based on Feature Transformation and Ensemble model," *PeerJ Comput. Sci.*, vol. 4, no. 7, p. e579, 2021, doi: 10.7717/PEERJ-CS.579.

[75] J. D. Turiel and T. Aste, "Peer-to-peer loan acceptance and default prediction with artificial intelligence," *R. Soc. Open Sci.*, vol. 7, p. 191649, 2020, doi: 10.1098/rsos.191649.

[76] R. A. Mancisidor, M. Kampffmeyer, K. Aas, and R. Jenssen, "Learning latent representations of bank customers with the Variational Autoencoder," *Expert Syst. Appl.*, vol. 164, p. 114020, 2021, doi: 10.1016/j.eswa.2020.114020.

[77] J. Kriebel and L. Stitz, "Credit default prediction from user-generated text in peer-to-peer lending using deep learning," *Eur. J. Oper. Res.*, vol. 302, pp. 309–323, 2022, doi: 10.1016/j.ejor.2021.12.024.

[78] X. Fu, T. Ouyang, J. Chen, and X. Luo, "Listening to the investors: A novel framework for online lending default prediction using deep learning neural networks," *Inf. Process. Manag.*, vol. 57, no. 4, p. 102236, 2020, doi: 10.1016/j.ipm.2020.102236.

[79] Q. Zhu, W. Ding, M. Xiang, M. Hu, and N. Zhang, "Loan Default Prediction Based on Convolutional Neural Network and LightGBM," *Int. J. Data Warehous. Min.*, vol. 19, 2022, doi: 10.4018/IJDWM.315823.

[80] S. Zandi, K. Korangi, M. Óskarsdóttir, C. Mues, and C. Bravo, "Attention-based dynamic multilayer graph neural networks for loan default prediction," *Eur. J. Oper. Res.*, vol. 321, no. 2, pp. 586–599, 2025, doi: 10.1016/j.ejor.2024.09.025.

[81] K. Ohno and A. Kumagai, "Recurrent Neural Networks for Learning Long-term Temporal Dependencies with Reanalysis of Time Scale Representation," in *12th IEEE International Conference on Big Knowledge, ICBK 2021*, 2021, pp. 182–189. doi: 10.1109/ICKG52313.2021.00033.

[82] T. S. Afolabi, T. M. Obamuyi, and T. Egbetunde, "Credit Risk and Financial Performance: Evidence from Microfinance Banks in Nigeria," *IOSR J. Econ. Financ.*, vol. 11, no. 1, pp. 8–15, 2020, doi: 10.9790/5933-1101070815.

[83] C. Sivaji, A. Mohan, S. Sriram, and M. Ramachandran, "A Comparative Study of Recurrent Neural Network (RNN) with Gray Relational Analysis for Temporal Data," *Comput. Sci. Eng. Technol.*, vol. 3, no. 4, 2025, pp. 23–36, 2025, doi: 10.46632/cset/3/4/3.

[84] C. Wang, D. Han, Q. Liu, and S. Luo, "A deep Learning Approach for Credit Scoring of Peer-to-Peer Lending Using Attention Mechanism LSTM," *IEEE Access*, vol. 7, pp. 2161–2168, 2019, doi: 10.1142/7114.

[85] L. Liang and X. Cai, "Forecasting peer-to-peer platform default rate with LSTM neural network," *Electron. Commer. Res. Appl.*, vol. 43, p. 100997, 2020, doi: 10.1016/j.elerap.2020.100997.

[86] T. Fadziso, "Overcoming the Vanishing Gradient Problem during Learning Recurrent Neural Nets (RNN)," *Asian J. Appl. Sci. Eng.*, vol. 9, no. 1, pp. 197–208, 2020, doi: 10.18034/ajase.v9i1.41.

[87] S. H. Noh, "Analysis of gradient vanishing of RNNs and performance comparison," *Information*, vol. 12, no. 11, p. 442, 2021, doi: 10.3390/info12110442.

[88] T. Vaiyapuri *et al.*, "Intelligent Feature Selection with Deep Learning Based Financial Risk Assessment Model," *Comput. Mater. Contin.*, vol. 72, no. 2, pp. 2429–2444, 2022, doi: 10.32604/cmc.2022.026204.

[89] M. Thor and Ł. Postek, "Gated recurrent unit network: A promising approach to corporate default prediction," *J. Forecast.*, vol. 43, no. 5, pp. 1131–1152, 2024, doi: 10.1002/for.3057.

[90] Y. Yang *et al.*, "Kolmogorov–Arnold networks-based GRU and LSTM for loan default early prediction," *Appl. Soft Comput.*, vol. 197, p. 115213, 2026, doi: 10.1016/j.asoc.2026.115213.

[91] I. J. Goodfellow *et al.*, "Generative adversarial nets," in *Advances in Neural Information Processing Systems*, E. Z. Ghahramani, M.Welling, C. Cortes, N. Lawrence, and K. Q. Weinberger, Ed., Montréal, Canada: Curran Associates, 2014, pp. 2672–2680. doi: 10.3156/jsoft.29.5_177_2.

[92] N. Park, Y. H. Gu, and S. J. Yoo, "Synthesizing individual consumers′ credit historical data using generative adversarial networks," *Appl. Sci.*, vol. 11, no. 3, pp. 1–15, 2021, doi: 10.3390/app11031126.

[93] E. Strelcenia and S. Prakoonwit, "Generating Syntetic Data for Credit Card Fraud Detection Using GANs," *2022 Int. Conf. Comput. Artif. Intell. Technol. CAIT 2022*, pp. 42–47, 2022, doi: 10.1109/CAIT56099.2022.10072179.

[94] R. Jaiswal and B. Singh, "Financial fraud prevention with synthetic data generation using gan," *Arya Bhatta J. Math. Informatics*, vol. 14, no. 2, pp. 161–166, 2022, doi: 10.5958/2394-9309.2022.00068.3.

[95] F. Zola, J. L. Bruse, X. E. Barrio, M. Galar, and R. O. Urrutia, "Generative Adversarial Networks for Bitcoin Data Augmentation," in *Data Driven Approaches on Medical Imaging*, 2023, pp. 136–143. doi: 10.1007/978-3-031-47772-0_8.

[96] J. Abdulkadir, H. Mohammed, and A. L. Balarabe, "Digital Lending Platforms , Fintech Partnerships , and Credit," *J. Bus. Dev. Manag. Res.*, vol. 11, no. 7, pp. 165–180, 2026.

[97] D. Oyetade and P.-F. Muzindutsi, "The impact of Economic policy Uncertainty on the Performance of African Banks," *J. Accounting, Financ. Audit. Stud.*, vol. 10, no. 2, pp. 105–118, 2024, doi: 10.56578/jafas100205.

[98] A. Yahaya, A. Balarabe Musa, and O. Jayeola, "Non-Performing Loans, Interest Effect and Bank Performance in Sub Saharan African Economies," *Niger. J. Bus. Adm.*, vol. 19, no. 1, pp. 46–59, 2021.