ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VI, June 2026

An Integrated Explainable Deep Learning Framework for Loan Default Prediction and Credit Risk Decision Support System

Ponsak .S. Bande, Chimuzuoroke E. Ugwuja*, Blessing .C. Uzo, Aminat .O. Atanda

Department of Computer Science, University of Nigeria, Nsukka

***Corresponding author**

DOI: https://doi.org/10.51583/IJLTEMAS.2026.150600205

Received: 08 July 2026; Accepted: 13 July 2026; Published: 22 July 2026

ABSTRACT

The credit risk assessment remains a critical challenge for financial institutions as manual and semi-automated loan evaluation processes often produce inconsistent decisions, high default rates, and operational inefficiencies. This paper discusses an Integrated Data-Driven Loan Management Framework that comprises an Artificial Neural Network (ANN) for credit default prediction, Structured Query Language (SQL) to systematically extract and transform data, and Interactive Visual Analytics Dashboards to aid in providing transparency within the decision-making process. For the study, a public loan dataset was used to pre-process the target dataset containing a total of 38,577 records and 24 attributes with pre-processing techniques such as feature engineering, one-hot encoding and class re-balancing with the use of Synthetic Minority Oversampling Technique (SMOTE) yielding 43 model-ready inputs.

The final ANN architecture was developed with 2 hidden layers (43 and 21 neurons with ReLU activation) and a sigmoid output neuron for binary classification. SHapley Additive exPlanations (SHAP) were computed to provide interpretability for each prediction made by the model. The overall accuracy of the model is 87% and the weighted precision, recall, and F1-score are 0.87 for the held-out test set (n = 12,858). The framework is implemented as a web-based decision support tool using Flask that enables users to receive risk scores in real-time along with explainable outputs and visual dashboards. The results from this experimentation indicate that an integrated pipeline (including data querying, predictive modeling, interpretability, and visualization) provides better decision-making and stakeholder transparency than using a model in isolation; therefore, it serves as a proof-of-concept prototype for intelligent loan management in banks and other financial institutions.

Keywords: Credit risk assessment, artificial neural network, loan default prediction, SMOTE, explainable AI.

INTRODUCTION

The assessment of credit risk is an essential component of establishing sustainable lending practices. Financial institutions must assess the likelihood that a borrower will meet their obligations to repay loans, as opposed to assessing those borrowers who will default on their loans. Non-performing loans (NPLs) continue to represent a substantial financial drain for banks and a major threat to the stability of the international financial system (Altman, 1968). Many developing countries continue to rely on the manual assessment process performed by loan officers when establishing a borrower's creditworthiness. The result is that decisions are made based on subjective factors, which creates continued inconsistencies as well as a lack of appropriate documentation (Bakpo & Kabari, 2009).

The inefficiency of this process results in higher rates of default and prevents creditworthy borrowers from having access to funding. For many years, traditional statistical modeling techniques used for credit scoring have been logistic regression and linear discriminant analysis (Altman, 1968). While effective, these models rely on rigid assumptions that are linear in nature and assume independence of the features being evaluated as well as failing to account for the fact that there are complex and nonlinear relationships between the various

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VI, June 2026

characteristics of the borrower that affect their likelihood of repaying a loan (West, 2000). Recently, machine learning (ML) techniques, specifically Artificial Neural Networks (ANNs), have demonstrated a significantly improved level of predictive accuracy over traditional statistical modeling methods for the classification of credit risk (Bakpo & Kabari, 2009; Chuang & Huang, 2011; El Khair Ghoujdam et al., 2024; Hsieh, 2005; Peng & Tu, 2005). ANNs can learn hierarchical feature representations and nonlinear decision boundaries from data.

Despite these advances, several gaps remain in the literature and practice. The most notable gap is that a large majority of studies have evaluated machine learning (ML) models in an isolated manner, simply presenting classification metrics without providing any contextual information (e.g., complete decision support workflows) within which the model resides (Rudin, 2019).

In addition, the established “black box” nature of neural networks presents an obstacle to adoption in regulated financial environments, due primarily to requirements for explainability (James et al., 2013). Finally, there exists a serious lack of research that combines structured data gathering/processing (e.g., SQL), predictive modeling, and interactive data visualization into a single, cohesive framework—each of which is necessary when implementing a practical application.

This paper presents an integrated solution to these gaps by providing an end-to-end Framework for Credit Risk Assessment using ML, with the following contributions: (i) Data pipeline integration: SQL-based extraction is combined with Python-based preprocessing, including encoding, scaling, and SMOTE rebalancing, to create a reproducible and auditable data pipeline; (ii) Using an Artificial Neural Network (ANN) to create, train, and evaluate a feed-forward neural network model for the purpose of predicting loan defaults, yielding a prediction accuracy of 87% on the held-out test set; (iii) Providing explanatory support to end-users through the generation of per-prediction SHAP explanations, thereby alleviating the interpretability problem associated with using neural networks to create a predictive credit model; (iv) Providing aggregate insight into a portfolio through the use of Tableau dashboards (including approval rate, funded amount, default distribution by loan grade, purpose, and geography); these dashboards complement the SHAP-generated per-instance explanations; (v) Delivering a complete decision-support system by packaging the framework as a Flask web application with a user-friendly HTML interface, thereby demonstrating deployment feasibility. The rest of the paper is organized into sections. Related work is reviewed in Section II; the proposed framework and methodology are described in Section III; experimental results and discussion are detailed in Section IV; and the paper's conclusions and future research opportunities are presented in Section V.

Related Work

A. Traditional Credit Scoring Models

The Z-score model designed by Altman (1968) is known as one of the earliest mathematical methods to assess corporate credit risk and bankruptcy likelihood. This early attempt was based on the use of several financial ratios through a type of statistical technique known as multiple discriminant analysis (MDA).

With continued research, logistic regression has become the most commonly used method for creating credit scorecards because of its ease of interpretation and because it is widely accepted by regulatory agencies (Chawla et al., 2002). However, due to the linear nature of MDA and logistic regression models, they are not always able to appropriately explain nonlinear interactions among the different features in a borrower's behavior data (West, 2000).

B. Neural Networks for Credit Risk

Another popular method of evaluating creditworthiness is the use of neural networks. A study performed by West (West, 2000) evaluated five different neural network types to help in developing a scoring system for credit applications, and it was determined that neural networks consistently demonstrated better performance when it came to predicting how likely or unlikely an applicant would repay their loans (i.e., lower credit risk).

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VI, June 2026

Hsieh (2005) further explored this area and developed a combined method of using self-organizing maps (SOMs) and a second method by using K-means clustering for preprocessing the data before that data were fed into an ANN, reporting improved accuracy after outlier reduction. Peng and Tu (2005) integrated a self-organizing map with a probabilistic neural network in a two-stage credit scoring system.

Chuang and Huang (2011) proposed a hybrid system in which an ANN classifier was followed by case-based reasoning to reassess borderline applicants. The study by Bakpo and Kabari (2009) used an ANN-based method to assess credit worthiness in Nigeria's banking sector and determined that using neural networks resulted in improved credit evaluations compared with traditional, statistical models, but without incorporating explainability or visualization components.

C. Addressing Class Imbalance

A typical loan data set is an imbalanced data set where the default representation is only a small number of instances. Chawla et al. . (2002) proposed the synthetic minority oversampling technique (SMOTE) as an interpolation method for generating additional synthetic data points in the minority class. The use of SMOTE and its variants has become a standard preprocessing technique in credit risk modeling (García et al., 2015; Rudin, 2019).

D. Explainable AI in Finance

The lack of clarity in neural networks has been highlighted as one of the major obstacles limiting their adoption within the finance sector (James et al., 2013).

Lundberg and Lee (2017) introduced SHAP, which provides theoretically grounded local feature attributions based on Shapley values from cooperative game theory. While SHAP has been utilized in credit scoring as a method of explaining the prediction for each customer and assessing model behavior (Liu, 2022),, there have been very few papers that link SHAP-derived explanations directly to a decision support interface for end-users.

E. Integrated Decision Support Systems

El Khair Ghoujdam et al. (2024) reviewed ML-based credit scoring studies and identified persistent challenges in class imbalance handling, model interpretability, and practical deployment. Noriega et al. (2023) developed a backpropagation neural network for early credit risk prediction in commercial banks and reported better performance than logistic regression, but the study did not address full end-to-end system integration.

F. Research Gap

The literature reviewed suggests that credit risk research has predominantly focused on the accuracy of models, with little regard for a) the entire data collection/processing/decision cycle, b) the accessibility of per-prediction explanation to users who are not familiar with the technical aspects, and c) visualisation tools for reviewing credit portfolio performance. This research addresses all three of these gaps in a single framework.

PROPOSED FRAMEWORK AND METHODOLOGY

Overview of the proposed framework

Figure 1 illustrates the overall architecture of the proposed explainable credit risk decision-support framework. The framework consists of six interconnected stages: data acquisition, data preprocessing, feature engineering, predictive modelling, explainability, and decision support. These stages collectively transform raw borrower information into transparent and actionable lending recommendations.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VI, June 2026

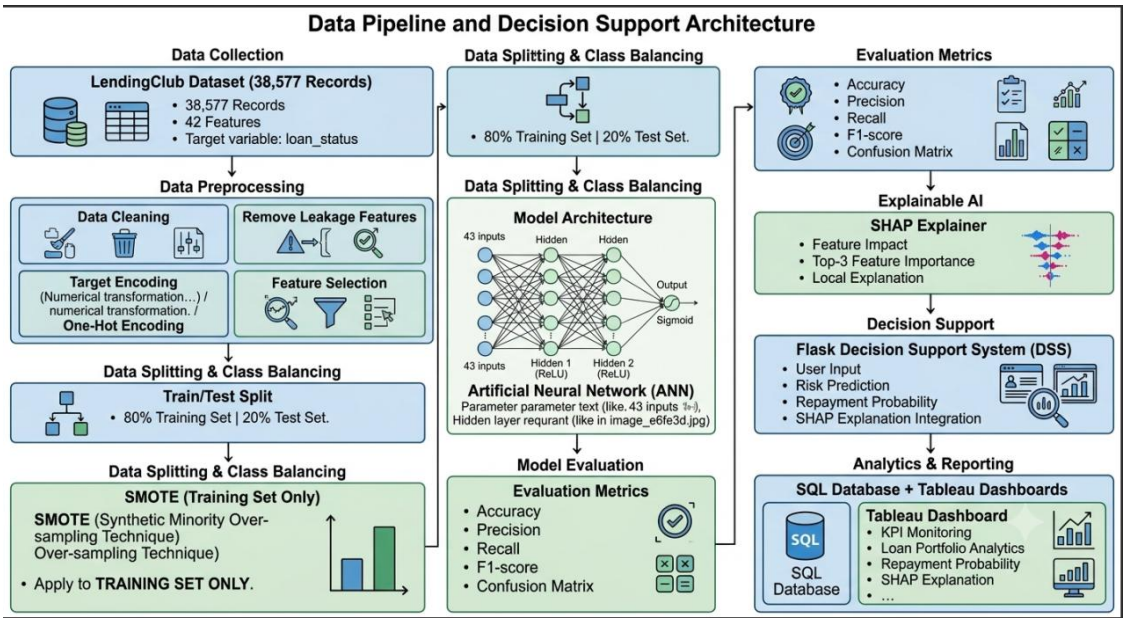

Figure 1: presents the overall workflow of the proposed framework. The architecture follows a modular pipeline beginning with data acquisition and preprocessing, followed by feature engineering and class balancing using SMOTE. The processed data are supplied to the Artificial Neural Network for loan default prediction. To improve transparency, SHAP generates feature-level explanations for every prediction. Finally, the trained model is deployed through a Flask web application, while SQL queries and Tableau dashboards provide portfolio-level analytics that support strategic lending decisions. This modular design enables each component to operate independently while collectively forming an integrated explainable decision-support framework.

Dataset Description

A publicly available dataset from LendingClub, obtained via Kaggle (https://www.kaggle.com/datasets/datawitharyan/financial-loan-dataset), is used. There are 38,577 records in the raw loan dataset, and these records contain 24 attributes, including the demographic information on borrowers, loan terms, credit ratings and status of loan repayments. One of the attributes (target variable) is `loan_status`, which has three classifications: Fully Paid, Charged Off and Current. To address class imbalance, records with `Current` status were first removed, leaving 37,478 records. An 80/20 stratified train-test split was then performed, after which SMOTE was applied only to the training partition. This generated synthetic minority-class samples for the `Charged Off` class, while the test set was left unchanged to provide an unbiased evaluation. The SMOTE output is shown in Figure 2.

```
C:\Users\OWNER\anaconda3\envs\tf_env\lib\site-packages\sklearn\base.py:474: FutureWarning
r._validate_data` is deprecated in 1.6 and will be removed in 1.7. Use `sklearn.utils.val
_data` instead. This function becomes public and is part of the scikit-learn developer AP
  warnings.warn(
Out[141… loan_status
         0    32145
         1    32145
         Name: count, dtype: int64
```

Figure 2: SMOTE output

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VI, June 2026

Class Balancing Using SMOTE

Loan default prediction suffers from severe class imbalance because defaulted loans constitute a relatively small proportion of all observations. To address this problem, the Synthetic Minority Oversampling Technique (SMOTE) was applied exclusively to the training dataset using the equation 1 (Chawla et al. 2002). Given a minority instance x_i a synthetic sample is generated as:

$$x_{new} = x_i + \lambda(x_{nn} - x_i) \quad (1)$$

where

x_i denotes the minority instance,

x_{nn} denotes one of its nearest neighbours,

$\lambda \in [0, 1]$ is a random interpolation factor.

Data Preprocessing

There are five steps involved in preprocessing data:

I.Removal of leakage and irrelevant features. Fifteen columns were removed, including post-issuance fields (e.g. `int_rate`, `total_payment`, `last_payment_date`), unique identifiers (e.g. `id`, `member_id`), noisy free-text (e.g. `emp_title`), and metadata (e.g. `issue_date`, `application_type`, `address_state`, `verification_status`, `sub_grade`) to prevent information leakage and reduce noise.

II.Target variable encoding. All records that have a `loan_status = 'Current'` were removed, as their outcome is unknown. The remaining records were encoded as follows: Charged Off = 1, representing default, and Fully Paid = 0, representing non-default.

III.Feature selection. Eight features were retained based on their relevance to the domain of interest and empirical discriminative power as follows:

- Numerical: `annual_income`, `loan_amount`, `installment`
- Categorical: `emp_length`, `grade`, `home_ownership`, `purpose`, `term`

IV. All categorical features underwent one-hot encoding, producing 43 dimensions for the dataset.

V. Class rebalancing was performed because there was an imbalance of classes (around 86% were Fully Paid and 14% Charged Off) after removing Current records, and the training set only (not the test set) was augmented with synthetic minority class samples (using SMOTE Chawla et al., 2002)) so that the two classes would be equal in number (i.e., 1:1).

The dataset after preprocessing consisted of 37,478 records (i.e., before SMOTE) and 43 features. An 80/20 stratified train-test split was applied, and SMOTE was then applied only to the training partition, resulting in 51,432 training samples. The test set of 12,858 records was left unchanged.

The proposed ANN Model Architecture

The feed-forward neural network was implemented in TensorFlow/Keras, and its architecture is summarized in Table 1. The model is compiled to use an Adam optimizer with a binary cross-entropy loss function and trained for 50 epochs while monitoring training loss and accuracy during training. The sigmoid output produces a probability (p) between 0 and 1, representing the predicted probability of default. A threshold value of 0.5 is used to classify the customers into binary classes. Each of the hidden layer's dimensions follows a tapering or

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VI, June 2026

narrowing scheme, as the first hidden layer is equal to the input layer (43 dimensions) so that when the data is transformed through this layer there is sufficient rank (full rank), which provides maximum possible transformation of the feature space. The second hidden layer contains 21 neurons (21 ≈ 43/2), approximately half of the input dimension, allowing the network to compress the learned representation before classification.

Table 1: Summary of the proposed ANN Architecture

Layer	Neurons	Activation	Parameters
Input	43	—	—
Hidden 1	43	ReLU	1,892
Hidden 2	21	ReLU	924
Output	1	Sigmoid	22
Total			**2,838**

Model Interpretability via SHAP

To address the black-box limitation, Shapley Additive Explanations (SHAP) were integrated as part of the prediction pipeline. A Kernel Explainer was initialized using a neutral baseline (all-zero vector of size 43). For every prediction, SHAP values were calculated. The top three features (in terms of absolute SHAP value magnitude) were logged and displayed to the user with their respective predictions, indicating whether the feature increased or decreased the predicted probability of default for the applicant. From Figure 3, the most influential feature was Loan Grade (mean |SHAP| = 0.260), followed closely by Employment Length (mean |SHAP| = 0.210), Home Ownership (mean |SHAP| = 0.120), Loan Amount (mean |SHAP| = 0.052), and Annual Income (mean |SHAP| = 0.051).

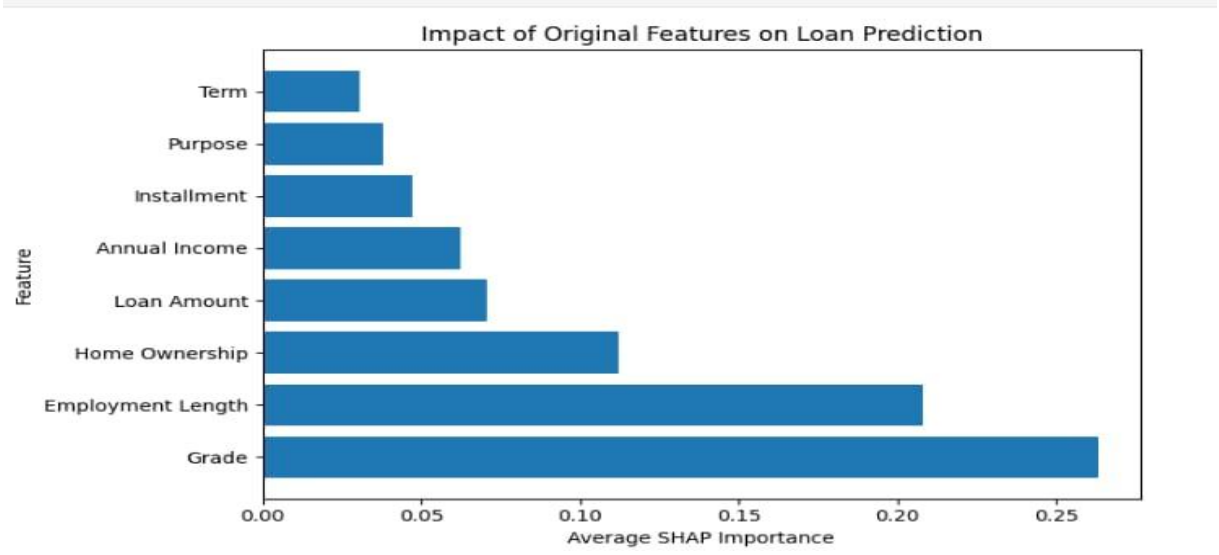

Figure 3: SHAP summary plot for Loan decision model

Visual Analytics

The calculation of key performance indicators (KPIs) using SQL Queries includes:

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VI, June 2026

- The sum of loan applications and comparison to previous month and month to date.
- The total amount funded versus the amount received.
- The average interest rate and the average DTI.
- The percentage of good (Fully Paid/Current) versus bad (Charged Off) loans.

These KPIs are displayed as Tableau dashboards across three separate perspectives (Summary, Overview with geographic and time data, and Detail at the loan level). These dashboards provide portfolio-level monitoring that complements the instance-level ANN predictions.

Deployment

The deployment process of the end-to-end system consists of the following:

I. The trained model was serialized as `loan_model.keras`.

II. Saving feature column metadata to `data_columns.json`.

III. The Flask application (`app.py`) loads the model and SHAP explainer at application startup.

IV. User inputs from the HTML form are submitted through a POST request, pre-processed (one-hot encoded and numerically formatted) and delivered to the model for inference.

V. The prediction result (non-default/default), repayment probability and top-3 SHAP explanations are returned to the client as JSON data that can then be rendered in the browser.

Algorithm 1: End-to-End Explainable Loan Prediction Framework

Input:

LendingClub loan dataset D

Output:

Loan default prediction

Default probability

SHAP explanation

Interactive dashboard visualization

1: Load raw loan dataset D

2: Remove incomplete and irrelevant records

3: Remove all records with loan_status = "Current"

4: Encode target variable:

Charged Off → 1

Fully Paid → 0

5: Select predictive features

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VI, June 2026

6: Apply one-hot encoding to categorical variables

7: Normalize numerical features

8: Split dataset into Training Set (80%) and Test Set (20%)

9: Apply SMOTE only on the Training Set

10: Construct ANN architecture

Input Layer (43)

Hidden Layer 1 (43, ReLU)

Hidden Layer 2 (21, ReLU)

Output Layer (1, Sigmoid)

11: Train ANN using Adam optimizer

12: Minimize Binary Cross-Entropy loss

13: Evaluate model using

Accuracy

Precision

Recall

F1-score

Confusion Matrix

14: Initialize SHAP Kernel Explainer

15: For each applicant do

Predict default probability

Compute SHAP values

Identify Top-3 influential features

Generate explanation

End For

16: Store trained model (.keras)

17: Store feature metadata (.json)

18: Deploy model using Flask

19: Accept applicant information

20: Preprocess user input

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VI, June 2026

21: Predict loan status

22: Generate SHAP explanation

23: Display prediction, explanation and repayment probability

24: Generate Tableau dashboards using SQL queries

25: Display portfolio analytics

End Algorithm

EXPERIMENTAL RESULTS AND DISCUSSION

Training Convergence

Figure 4 depicts both the Loss Function and the training Accuracy curves for 50 epochs of training. The Loss Function's value decreased from roughly 35 at epoch 1 to below 1 at epoch 20, suggesting initial rapid convergence behaviour. The Accuracy metric improved from 60% to 85% over the first 20 epochs, and ultimately reached 87% by epoch 50. Convergence is consistent and shows no evidence of catastrophic overfitting, though the absence of a separate validation curve indicates that future work should incorporate early stopping.

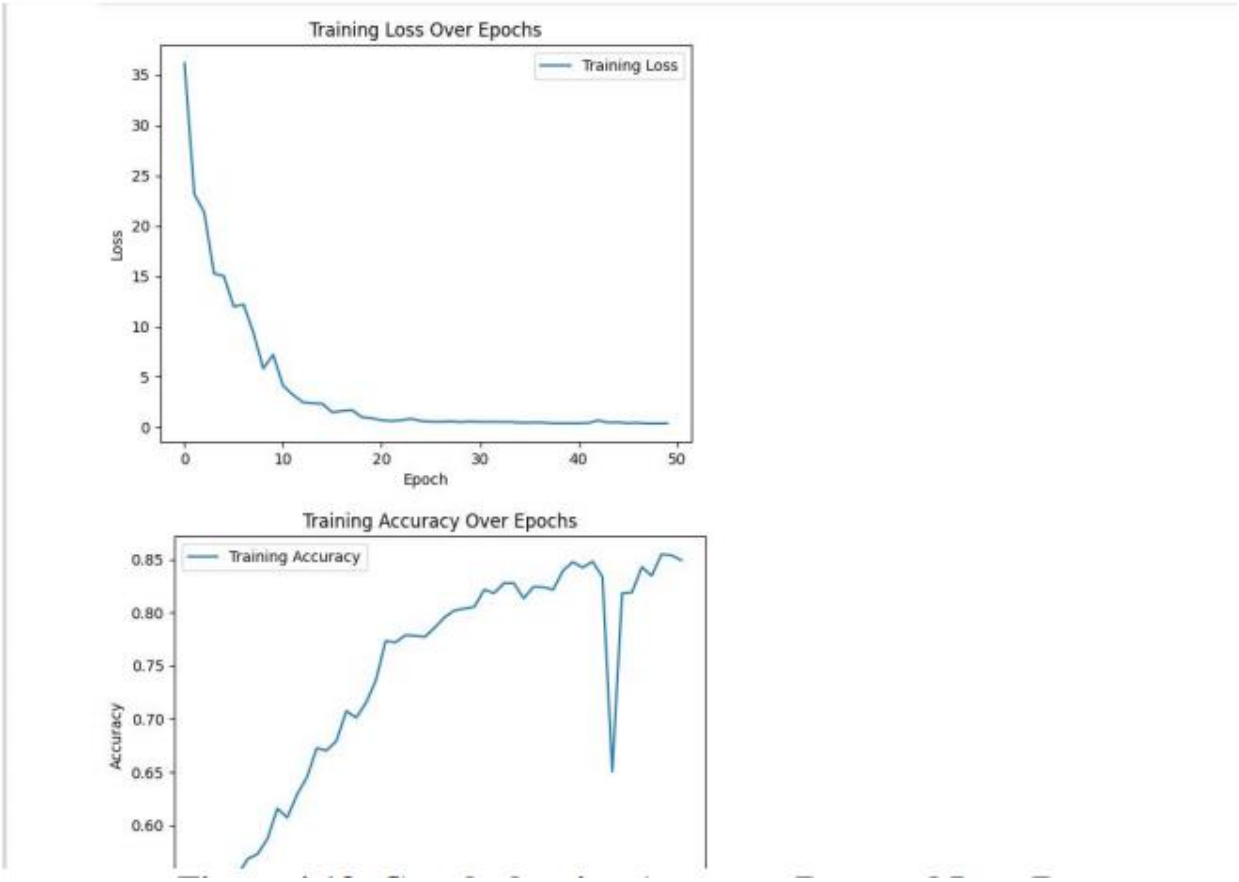

Figure 4: Accuracy and loss rates across training epochs

Classification Performance

The model was evaluated on a held-out test set of 12,858 samples, and the classification metrics are reported in Table 2. The overall accuracy was 87%, with a near-symmetric precision-recall trade-off (0.86–0.88) across both classes, indicating that the SMOTE technique successfully mitigated the original class imbalance.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VI, June 2026

Table 2: Classification Report on Test Set

Class	Precision	Recall	F1-Score	Support
0 (Fully Paid)	0.86	0.88	0.87	6,413
1 (Charged Off)	0.88	0.86	0.87	6,445
Weighted Avg	**0.87**	**0.87**	**0.87**	**12,858**

Confusion Matrix Analysis

The confusion matrix, shown in Figure 5, summarizes the classification results. The confusion matrix has the following information:

- True negatives (TN): 5,668 - Fully Paid loans correctly classified as non-default.
- False positives (FP): 745 - Fully Paid loans incorrectly classified as default.
- False negatives (FN): 891 - Charged Off loans incorrectly classified as non-default.
- True positives (TP): 5,554 - Charged Off loans correctly classified as default.

The false negative rate FN/(TP + FN) = 891/6,445 ≈ 13.8% is somewhat greater than the false positive rate FP/(TN + FP) = 745/6,413 ≈ 11.6%. In credit risk applications, false negatives (approved defaulters) have a greater financial impact than false positives (good borrowers rejected). This asymmetry in rates indicates that threshold optimization or cost-sensitive learning may improve real-world performance of the model.

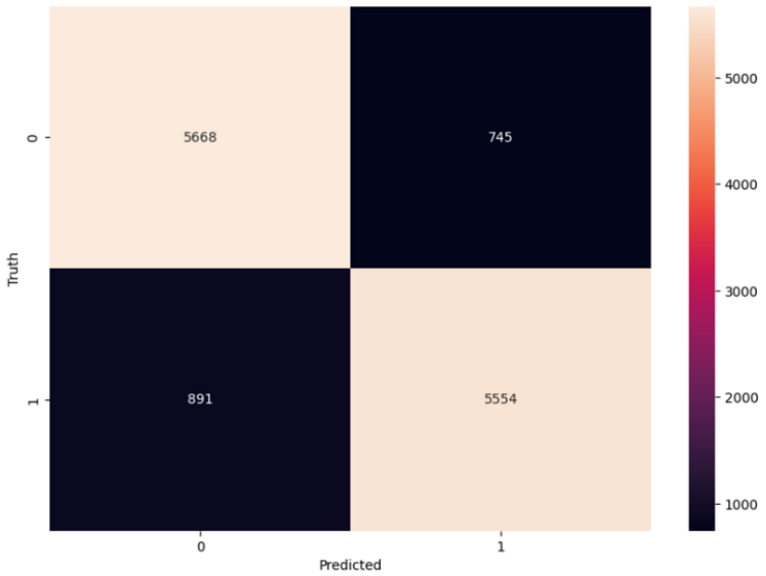

Figure 5: Confusion matrix

SHAP Interpretability

The insights gained from SHAP analysis indicate that the most influential features across the test set are the following (see Figure 3):

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VI, June 2026

I. `annual_income` - A higher annual income reduces the predicted probability of default, which supports the pre-existing knowledge that annual income is one of the key predictors of repayment capacity.

II. `grade` (particularly grades A and B) - A higher credit grade also reduces the predicted probability of default, which supports the credit bureau's own model of assessing risk.

III. `emp_length` - Longer employment tenure reduces the predicted probability of default and serves as a proxy for income stability.

IV. `loan_amount` and `installment` - The larger the requested loan amount and the larger the required monthly installments, the higher the predicted probability of default because of the increased level of exposure to risk.

These SHAP outputs are supplied to the loan officer at the time of making their decision and provide the loan officer with both a risk score and an actionable explanation of what influenced that risk score.

Visual Analytics Insights

The Tableau dashboard (shown in Figure 5) indicates various trends at the portfolio level that include:

- Loan volume: 38,600 applications for a total of $435,800,000 funded and $473,100,000 received.
- Good versus bad loans: 86.2% (as depicted in Figure 6a) of loans were classified as good loans (Fully Paid or Current), while 13.8% were classified as bad loans (Charged Off).
- Distribution of purpose: Debt consolidation represented the majority of the purpose for loans (funded for a total of $232,500,000), followed by credit card refinancing ($58,900,000).
- Length of employment: Borrowers with 10 years or more of employment received a total of $136,100,000 in funded loans, which is greater than any other borrower group.
- Term distribution: 62.66% of funded amounts were for loans with terms of 36 months and 37.34% for loans with terms of 60 months.

These insights provide the ability for portfolio managers to recognize concentrations of risk, track repayment patterns, and develop policies for lending that are based on empirical data rather than intuition.

Figure 6a: Summary of the dataset using Tableau

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VI, June 2026

Figure 6b: Overview of the bank loan report from the dataset

Comparison with Related Work

The proposed framework achieves competitive or superior accuracy compared to related studies reviewed. Among the studies reviewed in this paper, the proposed framework uniquely combines SHAP-based explainability, Tableau-based visual analytics, and a deployable Flask web interface.

Table 3: Comparison of the proposed framework with key related studies

Study	Model	Accuracy	Explainability	Visualization	End-to-End System
Hsieh (2005)	SOM + ANN	80.0%	No	No	**No**
Chuang & Huang (2011)	ANN + CBR	82.5%	Partial (CBR)	No	**No**
Bakpo & Kabari [2009]	ANN	83.2%	No	No	**No**
Liu (2022)	BPNN	>85.0%	No	No	**No**
Proposed	**ANN + SHAP**	**87.0%**	**Yes (SHAP)**	**Yes (Tableau)**	**Yes (Flask)**

Limitations

Several limitations have been noted:

- The model was built from data of US-based LendingClub borrowers, and therefore, deployment in other lending markets (e.g., banks in Nigeria) would necessitate adapting the model to the new domain and retraining.
- The number of input features used to train the model is limited to eight. The incorporation of additional input features, such as credit bureau scores and transaction history, may produce better results.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VI, June 2026

- The model architecture (ANN) is relatively shallow. If applying deeper networks or advanced architectures (e.g., attention mechanisms), the algorithm may capture more complex patterns, subject to empirical validation.
- The model was tested in a prototype manner, using localhost as a testing environment. The deployment of the model into production will require security hardening, scalability testing, and an assessment of compliance with regulations.
- A validation set was not formally used for early stopping, which may allow mild overfitting.

CONCLUSION AND FUTURE WORK

This paper presented an integrated framework for assessing credit risk that includes (i) using an Artificial Neural Network (ANN) to predict loan defaults; (ii) using SHAP-based explanation of individual predictions to better understand ANN predictions; (iii) managing loan data using SQL; and (iv) Tableau dashboards for portfolio-level visualization. Using the proposed ANN method for predicting default, the loan dataset produced an accuracy of 87%, with balanced precision and recall across both classes. Providing SHAP explanations alongside predictions via a web-based interface helps address the interpretability issue that has hindered the adoption of neural networks by regulated financial institutions. Tableau dashboards allow for the visualization of data at the portfolio level using a variety of different metrics related to lending patterns, default distributions, and demographics of borrowers.

The overall direction of future work will be in five areas: (i) comparing the implemented ANN model with other ensemble methods (e.g. XGBoost, Random Forest, LightGBM) and deep learning architectures; (ii) incorporating temporal features and recurrent architectures to develop time-aware credit scoring models; (iii) deploying this solution in the cloud using containerization (e.g. Docker/Kubernetes) to enable scalability; (iv) integrating with credit bureau APIs to provide real-time feature enrichment; and (v) validating the proposed framework on datasets from multiple geographic areas (e.g. developing economies) to assess cross-domain generalizability.

ACKNOWLEDGEMENT

The authors express their sincere appreciation to the Department of Computer Science, University of Nigeria, Nsukka, for providing the academic environment and support that facilitated this research. The authors also acknowledge the developers and maintainers of the publicly available LendingClub loan dataset made accessible through Kaggle, which served as the basis for the experimental evaluation conducted in this study. Furthermore, the authors appreciate the open-source developer communities behind TensorFlow/Keras, SHAP, Flask, Tableau Public, and other Python libraries used in implementing and evaluating the proposed framework.

Declaration of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The study was conducted using publicly available datasets and open-source software resources.

Data Availability

The dataset used in this study is publicly available through the LendingClub loan dataset hosted on Kaggle. As the research utilized publicly accessible secondary data, new datasets were generated during this study using SMOTE technique. The processed data and implementation details are available from the corresponding author upon reasonable request.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VI, June 2026

Ethical Statement

This study used only publicly available, anonymized secondary data and did not involve human participants, animals, or personally identifiable information. Therefore, ethical approval and informed consent were not required for this research.

Author Contributions

Ponsak S. Bande: Conceptualization, methodology, software development, data curation, formal analysis, model development, visualization, validation, writing—original draft preparation.

Ugwuja E. Chimuzuoroke: Conceptualization, methodology refinement, critical review and editing of the manuscript, project administration, validation.

Blessing C. Uzo: Literature review, data preprocessing support, results interpretation, manuscript review and editing, validation.

Aminat O. Atanda: Software testing, visualization, documentation, manuscript review and editing, quality assurance, validation.

All authors reviewed and approved the final manuscript and agreed to its submission for publication.

REFERENCES

1. Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, *23*(4), 589–609.
2. Bakpo, F. S., & Kabari, L. G. (2009). Credit risk evaluation system: An artificial neural network approach. *Nigerian Journal of Technology*, *28*(1), 36–45.
3. West, D. (2000). Neural network credit scoring models. *Computers & Operations Research*, *27*(11–12), 1131–1152.
4. Peng, Y., & Tu, J. (2005). A hybrid approach for credit scoring using self-organizing maps and probabilistic neural networks. In *Proceedings of the International Conference on Computational Intelligence for Finance and Economics* (pp. 67–72). IEEE.
5. Hsieh, N.-C. (2005). Hybrid mining approach in the design of credit scoring models. *Expert Systems with Applications*, *28*(4), 655–665.
6. Chuang, C.-L., & Huang, S.-T. (2011). A hybrid neural network approach for credit scoring. *Expert Systems*, *28*(2), 185–196.
7. El Khair Ghoujdam, M., Chaabita, R., & Idamia, S. (2024). Machine learning in credit risk assessment: A systematic review. *International Journal of Applied Management and Economics*, *2*(8), 255–274.
8. Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, *1*(5), 206–215.
9. James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning*. Springer.
10. Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, *16*, 321–357.
11. García, S., Luengo, J., & Herrera, F. (2015). *Data preprocessing in data mining*. Springer.
12. Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems* (pp. 4765–4774).
13. Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). *IEEE Access*, *6*, 52138–52160.
14. Liu, L. (2022). A self-learning BP neural network assessment algorithm for credit risk of commercial bank. *Wireless Communications and Mobile Computing*, *2022*, Article 9650934. https://doi.org/10.1155/2022/9650934
15. Noriega, J. P., Rivera, L. A., & Herrera, J. A. (2023). Machine learning for credit risk prediction: A systematic literature review. *Data*, *8*(11), Article 169. https://doi.org/10.3390/data8110169