INTERNATIONAL JOURNAL OF LATEST TECHNOLOGY IN ENGINEERING,
MANAGEMENT & APPLIED SCIENCE (IJLTEMAS)

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

Credit Worthiness Prediction Model Using Artificial Neural Network

Echezona Stephenson C., Uzo Blessing Chimezie, Ukekwe Emmanuel C.

University of Nigeria, Nsukka

DOI: https://doi.org/10.51583/IJLTEMAS.2026.150700178

Received: 12 August 2026; Accepted: 17 August 2026; Published: 27 August 2026

ABSTRACT

The study focused on developing a creditworthiness prediction model utilizing artificial neural network. Credit risk evaluation has a relevant role for financial institutions, as lending could result in real and immediate losses. In particular, default prediction was one of the most challenging activities in managing credit risk. The objective was to enhance the accuracy and reliability of credit risk assessments by leveraging the computational power and learning capabilities of artificial neural networks. The parameters of the dataset include the customer's place of work, loan history, monthly salary, loan amount, transaction history, and credit history, all stored in the trained database. When a customer comes to apply for a loan, the system checks if the user is qualified based on these parameters and then approve or disapprove the loan accordingly. Rigorous testing and validation were conducted to ensure the model's robustness and generalizability. The results demonstrated that the neural network-based model significantly outperformed traditional statistical methods, providing more precise predictions of creditworthiness. An Object-Oriented Analysis and Design Methodology (OOADM) approach, which incorporated Unified Modeling Language for analysis and design, was used. The development stage was completed using a set of software tools, including Python and the MySQL database system.

Keywords: Credit worthiness, Prediction, Credit risk assessment, Artificial neural network, Loan

INTRODUCTION

Based on the situations some average Nigerian's found themselves, loan default has become the order of the day. Many of them find themselves borrowing money from different loan apps and microfinance banks thereby making the scenario of loan default increasing on daily basis.

Hence base on the rapid loan default, it's important to know how worthy a borrower is, which is the creditworthiness of a borrower, Schaeck, Klaus, et al., (2009) wrote that with the excessive competition between different financial institutions to attract more consumers, it became important not to deny the loan or simply provide the loan to consumer. And that to provide loan, it is also important to know whether the consumer will be able to pay back the provided loan or not. Although, financial institutes provide "Consumer Loan" with extreme caution and after various affirmations, but still there are some cases in which consumer is not able to pay back the loan. These types of cases are what we call loan default cases and as such creditworthiness has to be considered.

These cases can vary from extreme minor cases to major cases resulting in scams. Some of these cases may even lead the financial institution on the verge of bankruptcy.

Bulus Y. D., et al, (2024), wrote that, the credit transmission channel has been damaged as regards the quantity, price, and distribution of credit in Nigeria. Moreover, several banks have personnel that give grants to people that may need to get a loan. Most a time the discussion on how to grant loan is based on their intellect ideology and also, their exigencies over time will determine if they should give the loan or not. With the economic situation in Nigeria, several institutions like, Help pay, O pay, Mid Land Microfinance banks, and small-scale businesses have emerged to help young Nigerians with soft loan in other to enable them start up business or execute one project or the other without necessarily having any tangible thing to use as

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

collateral. This is a major problem for the small and medium size firms in Nigeria, which has also suffered from bank regulatory concerns of capital adequacy, heightened emphasis on default risk of bank counterparts and the general malfunctioning of credit extension and private sector growth, Ajayi O. Bolaji, et al, (2016). One of the major reasons for establishing banks is to advance loans to customers. But in order to stay in business, banks advance these loans to people who have the ability to pay back the money, thereby minimizing the risk of bad debt. However, risk management; knowing who is credit worthy is still an on-going challenge within the banking sector. The ability to identify a risk score of a customer base on some features such as occupation, age, marital status, salary range/amount of equity, credit history, etc. is an important step that banks go through before giving credit to customers. Credit risk score helps the banks to decide on how much interest to charge on the loan, etc. However, these risk factors sometimes do not give an informed decision on the credit worthiness of customers. Moreover, many banks lack a central well integrated, automated finance and risk management system due to the inability to develop a robust and scalable risk management system to forecast risk score of customers, written by, Olokoyo O. F., et al, (2016).

What then is Credit Worthiness? The easiest way to consider that is to think of one's situation.

Djeundje, A. et al, (2021), take the case where an acquaintance, turns to you and asks you to lend them some money. Not a trivial amount to pay the bus fare home, but a sufficient amount that, if they do not repay you as promised, you are left significantly out of pocket. What do you do? Do you lend the individual the money? They may not repay you. Therefore, it is better to refuse. But again, you may lose out on a possible profitable opportunity. The crux of the decision is whether the individual honors the promise to repay or defaults. The desirable result occurs if the loan is repaid (with interest). The undesirable outcome that you wish to avoid is when the individual fails to repay the loan or in the parlance of credit, "defaults."

Creditworthiness is a lender's willingness to trust you to pay your debts. A borrower deemed creditworthy is one a lender considers willing, able and responsible enough to make loan payments as agreed until a loan is repaid. Lenders evaluate creditworthiness in a variety of ways, typically by reviewing your past handling of credit and debt, and, in many cases, by assessing your ability to afford the payments required to repay the debt, Zhong G. et al., (2018).

Financial institutions use credit ratings to quantify and decide whether an applicant is eligible for credit. Credit ratings are also used to fix the interest rates and credit limits for existing borrowers Ampountolas A. C., et al, (2021). A higher credit rating signifies a lower risk premium for the lender, which then corresponds to lower borrowing costs for the borrower. Across the board, the higher one's credit rating, the better. A credit report provides a comprehensive account of the borrower's total debt, current balances, credit limits, and history of defaults and bankruptcies if any. Due to high levels of asymmetries of information in the market, lenders rely on financial intermediaries to compile and assign credit ratings to borrowers and help filter out bad debtors or "lemons." The independent third parties are called credit rating agencies. The rating agencies access potential customers' credit data and use sophisticated credit scoring systems to quantify a borrower's likelihood of repaying debt. Lenders usually pay for the services, but borrowers may also request their credit score to gauge their worthiness in the market.

Li O., et al, (2019) conducted research on using attributes of customers to assess credit risk by using a weighted-selected attribute bagging method. They bench-marked their result experimentally by using two credit databases and reported outstanding performance both in term of prediction accuracy and stability as compared with another state-of-the-art methods. A data mining approach is also proposed by Moro B., et al., (2021) to predict the success or otherwise of a Portuguese retail bank in telemarketing. They applied various data mining models on the bank telemarketing data and reported that the neural network data mining method was the best for analyzing the data. The work of Shao V. T., et al., (2025), described the strengths and weaknesses of various machine learning techniques within the context of business data mining approach.

Their analysis revealed that Rule Induction Technique was the best approach in mining business data, followed by that of the neural network approach. Therefore, with this growth, the numbers of credit card applications with various credit products are increasing day by day because many people want to avail these

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

services for their personal interest. Due to the large number of potential borrowers, it is necessary to use models and algorithms that avoid human failures in the analysis of credit application in consumer lending. In the proposed system neural networks is use for the Prediction credit worthiness.

Neural networks, usually simply called neural networks or neural nets, are computing systems inspired by the biological neural networks (NN) that constitute animal brains. An NN is based on a collection of connected units or nodes called artificial neurons, which loosely uses the metaphor of the neurons in a biological brain according to Wikiversity, (retrieved, 2025). The neural network will be trained to be able to predict if an individual is able to pay back the loan within the agree period of time.

The aim of the work is to develop a Credit Worthiness Prediction Model using Artificial Neural Networks (ANNs). The objectives are to: design a Prediction Model that can predict user credit worthiness, implement the model using Python Programming Language, evaluate the model using metrics and accuracy.

Theoretical Background

Lender- Based Theory of Collateral, Mueller R., et al, (2021) developed the lender-based theory of collateral to provide a guide on loan management in an imperfectly competitive loan market. The theory indicates that there are mainly two types of lenders: local lenders and transaction distant lenders in which the former can have access to specific and private information about borrowers while the latter can only provide credit basing on publicly available information on borrowers. For example, a local lender tends to have past information about borrowers and therefore can make decision on credit more easily owing to this familiarity. To leverage the lack of soft information, lenders require placement of pledge-able assets by borrowers whose prior relationship is not available or cast a doubt on ability to pay. Moreover, collateral requirement leads to optimal credit decision by reducing the inefficiency occasioned by lack of information. Arguably, collateral creates an optimal credit contract for both borrower and lender in that the borrower manages project efficiently in order not to lose the asset charged as collateral while the lender is able to manage credit risk. Collateral increases the chance of obtaining credit. In this model customers who can raise more collateral are likely to obtain more credit and at credit terms.

substitutes. In addition, proponents of this theory view that informal lending is more prevalent in developing

Loanable Funds Theory

The loanable funds theory of interest is an expansion of the classical savings and investment theory of interest rate. It integrates monetary and non-monetary factors of savings and investment. The theory was developed by Dennis R. et al, (1930). The theory states that the level of interest rate in the financial market is determined by the supply and demand of loanable funds. Saunders C. et al, (2020), suggested that the determination of interest rate could be likened to how the demand and supply of goods determine the price of goods. Holding all other factors constant, the supply of loanable funds increases as interest rate rises, while the demand for loanable funds grows as interest rate falls, all other factors held constant. Saunders C. et al. (2020) further asserts that economic conditions and monetary expansion cause demand curve for loanable funds to shift. The theory also explains that the sum of money offered for lending and demanded by consumers and investors in a given period, as well as, the interest rate is determined by the interaction between potential borrowers and savers. According to the theory, economic agents seek to make the best use of the resources available to them over their lifetime and borrow funds to take advantage of investment opportunities in the economy in their bid to increase real income. This would only work if the rate of return available from the investment is higher than the cost of borrowing. Borrowers would be unwilling to pay a higher real rate of interest than the rate of return available to capital. On the other hand, savers would be willing to save and lend.

Artificial Neural Networks (ANNs)

Suleiman S., et al (2018) Artificial Neural Networks are mathematical representations inspired by the human brain nerve cells and their communication and processing information techniques. The Neural Network is

INTERNATIONAL JOURNAL OF LATEST TECHNOLOGY IN ENGINEERING, MANAGEMENT & APPLIED SCIENCE (IJLTEMAS)
ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

ideally composed of three layers, the input layer, the hidden layer, and the output layer. The input layer consists of input nodes which represent the system's variable. The hidden layer consists of nodes which facilitate the flow of information from the input to the output layers. The flow is controlled by weight factors associated with each connector. The output layer consists of nodes which represent the system's classification decision. The value of the output nodes is compared with cut-offs to determine the output and classify each case. The weight adjustment is known as training. The training process consists of running input values over the network with predefined classification output nodes. This process runs until the weight values are minimized to an error function. Testing samples are used to verify the performance of the trained network. In the context of credit scoring, numerous studies have proven that Neural Network perform remarkably better than any other statistical approach, such as logistic regression or discriminant analysis Yu H. et al, (2007).

The underlying structure of an ANN is a directed graph, i.e. it consists of vertices and directed edges, in this context called neurons and synapses. The neurons are organized in layers, which are usually fully connected by synapses. In ANN, a synapse can only connect to subsequent layers. The input layer consists of all covariates in separate neurons and the output layer consists of the response variables. The layers in between are referred to as hidden layers, as they are not directly observable. Input layer and hidden layers include a constant neuron relating to intercept synapses, i.e. synapses that are not directly influenced by any covariate. To each of the synapses, a weight is attached indicating the effect of the corresponding neuron, and all data pass the neural network as signals. The signals are processed first by the so-called integration function combining all incoming signals and second by the so-called activation function transforming the output of the neuron. The simplest multi-layer perceptron (also known as perceptron) consists of an input layer with n covariates and an output layer with one output neuron. It calculates the function (2.1)

$$O(X) = f(\omega_0 + \sum_{i=1}^{n} \omega_i x_i) = f(\omega_0 + \omega^T X) \qquad 2.1$$

Where ω_0 denotes the intercept $\omega = (\omega_1, \cdots \omega_n)$, the vector consisting of all synaptic weights without the intercept, and

$X = (x_1, x_2, \ldots, x_n)$ the vector of all covariates. The function is mathematically equivalent to that of Generalized Linear Model (GLM) with link function f^{-} . Therefore, all calculated weights are in this case equivalent to the regression parameters of the GLM.

To increase the modeling flexibility, more hidden layers can be included. However, Hornik T. et al. (1989) showed that one hidden layer is sufficient to model any piecewise continuous function. Such an MLP with a hidden layer consisting of hidden neurons calculate the following function (2.3):

$$O(X) = f(\omega_0 + \sum_{j=1}^{J} w_j . h_j) = f(\omega_0 + \sum_{j=1}^{J} \omega_j . f(\omega_{0j} + \sum_{i=1}^{n} \omega_{ij} x_i)) \qquad 2.2$$

$$O(X) = f(\omega_0 + \sum_{j=1}^{J} \omega_j . f(\omega_{oj} + \omega_j^T X)) \qquad 2.3$$

Where $h_j = f(w_{0j} + \sum_{i=1}^{n} w_{ij} x_i), j = 1,2 \ldots J$

In the function (2.2), where ω_0 denotes the intercept of the output neuron and the intercept of the j_{th} hidden neuron. Additionally, ω_j denotes the synaptic weight corresponding to the synapse starting at the j_{th} hidden neuron and leading to the output neuron, $\omega_j = (\omega_{1j}, \ldots, \omega_{nj}$) the vector of all synaptic weights corresponding to the synapses leading to the j_{th} hidden neuron, and $X = (x_1, \ldots x_n$), the vector of all covariates. This shows that the neural networks are direct extensions of GLMs. However, the parameters i.e. the weights cannot be interpreted in the same way anymore. The figure 2.1 shows the architecture of a multi-layer neural network wit n covariates.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

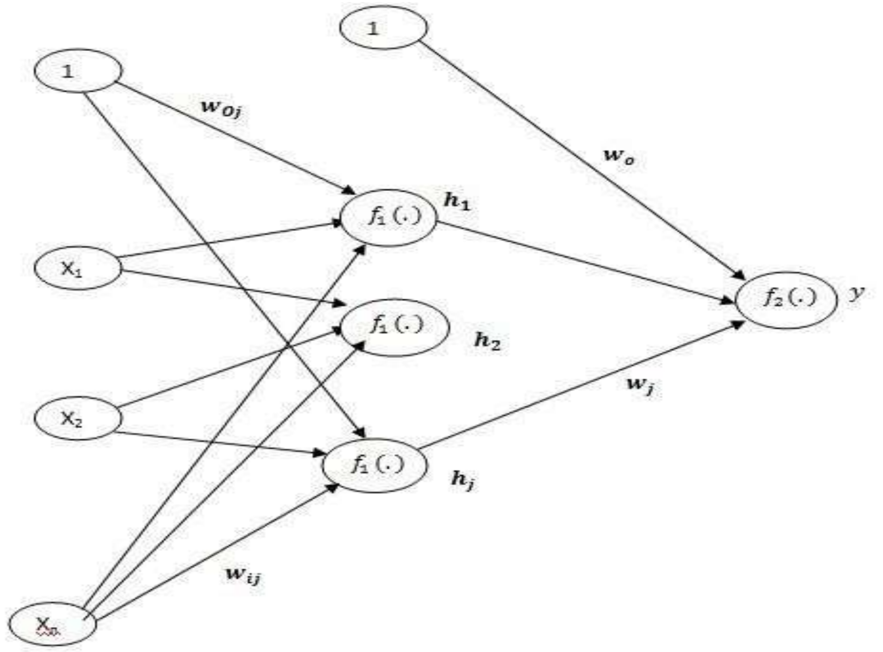

Figure 2.1: the Architecture of multi-layer feed forward neural network with n covariates

REVIEW OF RELATED LITERATURE

Pang, S. V. et al (2022), proposed Credit scoring systems are based on Operational Research and statistical models which seek to identify who of previous borrowers did or did not default on loans. This study looks at the question, when will borrowers default, not if they will default. It suggests that some of the reliability modelling approaches may be useful in this context and may help identify who will default as well as when they may default.

Laitinen E. K., et al (2022), this research dealt with the challenge of reducing banks' credit risks associated with the insolvency of borrowing individuals. To solve this challenge, we propose a new approach, methodology and models for assessing individual creditworthiness, with additional data about borrowers' digital footprints to implement comprehensive analysis and prediction of a borrower's credit profile. We suggest a model for borrowers' clustering based on the method of hierarchical clustering and the k-means method, which groups actual borrowers having similar creditworthiness and similar credit risks into homogeneous clusters.

We also design the model for borrowers' classification based on the stochastic gradient boosting (SGB) method, which reliably determines the cluster number and therefore the risk level for a new borrower. The developed models are the basis for decision making regarding the decision about lending value, interest rates and lending terms for each risk homogeneous borrower's group. The modified version of the methodology for assessing individual creditworthiness is presented, which is to reduce the credit risks and to increase the stability and profitability of financial organizations.

Edelman S. K., et al, (2021), grounded some principles for variables selection. The practical concepts of credit scoring stress to include all the variables related with applicant or applicant's environment that will aid in predicting the default risk. Some attributes offer information about the stability of the applicant (for example, time at present address, time at present employment), some characteristics present information about the financial capacity of the applicant (for example, having a current or checking account, having credit cards, time with current bank), some variables provide information about the applicant's resources (for example, residential status, employment status, spouse's employment) and the rest of the variables offer information on the possible outgoings (for example, number of children, number of dependents). Moreover, it is illegal to use

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

some characteristics like race, religion and gender, etc. Furthermore, there are some specific types of variables those that are not illegal but culturally unacceptable (for example, poor health records or lots of driving convictions). And sometimes, creditors look for insurance protection against the credit card debts during the change of the employment status (suppose, unemployment). So, all variables are subjective.

Jentzsch S., et al, (2021), reviewed some important considerations regarding the selection of the explanatory variables. The author briefs that commercial models possess 30 variables on an average and those models are good at estimating the probability of debt repayment. Additionally, some specific variables are prohibited in some countries, as presented in the table

Here, in the table, "0" signifies "No Legal Restriction" and "1" represents "Legal Restrictions". So, it differs from country to country which variables are restricted. For an example, age is allowed in the European countries, but it is not authorized in the USA unless credit scoring models assign a positive value. Similarly, gender is allowed in the European countries but it is not permitted in the USA. On the other hand, credit behaviour of Americans differs from the behavior of the Australians and there is more credit information available in the USA than European countries. So, credit scoring models differ from each other critically. Edelman S. K., et al. (2021), states that there should be a clear, rational, explanatory relationship between each variable and credit performance. The variables those possess a causal or explanatory relationship with credit performance, are genuine variables and the examples for these types of variables may look like income, debt and living expenses etc. On the other hand, designers have a tendency to use only statistically significant variables like rent (or own), debt to a finance company, age of automobile owned (or financed) and occupation, and the author has a confusion about the validity of these types of variables and discouraged this practice. Those variables should be included that have visible relationships. Moreover, the authors state that some "Prohibited Variables" are discriminatory in nature like race, color, sex, marital status and so on, and should be excluded from analysis. The legislators of the United States and the United Kingdom don't permit to use these variables.

Burgers D. H. et al, (2022), found out that the performance of SVM is either identical or significantly better than other competing techniques in the applications. Considering credit scoring problem, used eight credit scoring datasets to apply various classification algorithms. It finds that SVM achieves the highest classification accuracy rate among the methods tests. Moreover, Wu G., et al (2020), compared the performance in credit risk evaluation between SVM method and back propagation neural networks, concluding that SVM achieves only slight improvement over back propagation neural networks. In later research, they compare with other three techniques (neural networks, genetic programming and decision tree, suggesting that SVM approach could provide similar classification accuracy, even though there are relatively few input variables. Support vector machine is applied for credit rating with good performance.

Kondeti D. B., et al, (2022), proposed a prediction of loan approval using machine learning algorithms, aiming to streamline the decision-making process in financial institutions. Despite the promising results and advancements in predictive accuracy demonstrated by the study, several knowledge gaps were identified that warranted further investigation. Firstly, the dataset used in the research primarily consisted of historical loan approval data from a single financial institution. This limitation posed a challenge in terms of generalizability. The model's performance might vary significantly when applied to data from other banks or credit unions with different customer demographics, lending policies, and economic conditions. Future studies needed to incorporate diverse datasets from multiple sources to validate the model's robustness and adaptability across various financial landscapes.

Secondly, the study focused on a limited number of features for predicting loan approval, such as applicant income, credit score, loan amount, and employment status. However, real-world loan approval decisions often involve a broader range of factors, including the applicant's debt-to-income ratio, existing liabilities, loan purpose, and even macroeconomic indicators. Expanding the feature set to include these additional variables could enhance the model's predictive power and provide a more comprehensive assessment of an applicant's creditworthiness. The Object Oriented Analysis and Design Methodology (OOADM) was adopted in the study. The methodology promises to reduce development time.

INTERNATIONAL JOURNAL OF LATEST TECHNOLOGY IN ENGINEERING, MANAGEMENT & APPLIED SCIENCE (IJLTEMAS)

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

Analysis of the Existing System

Kondeti and Veerabadhra (2022) proposed a Prediction of Loan Approval Using Machine Learning Algorithm.

The prediction of loan approval using machine learning focused on developing a predictive model to evaluate an applicant's eligibility based on financial and demographic factors. The model utilized historical data, incorporating features such as income, loan amount, credit history, employment status, education, and debt-to-income ratio. Its primary objective was to automate the decision-making process for financial institutions, reduce errors in manual evaluations, and improve overall efficiency. As seen in figure 3.1, it consists of Upload data, Display output to user and train dataset sub systems.

Model Explanation

A classification algorithm, such as Logistic Regression, Decision Trees, or Random Forest, were used to predict loan approval. The model development process included several key steps:

1. Data Preprocessing: Missing values in the dataset were filled using statistical techniques, ensuring no loss of critical information. Categorical variables, like education level or marital status, were transformed into numerical formats using encoding techniques such as one-hot encoding or label encoding. Finally, the data was normalized to maintain uniformity across all features.
2. Feature Selection: Significant factors influencing loan approval, including credit history, income, loan amount, and debt-to-income ratio, were identified through statistical analyses and expert knowledge. This step ensured the model focused on the most relevant inputs.
3. Model Training: The dataset was divided into training and testing sets. The training set was used to teach the algorithm patterns that linked input features to loan approval outcomes. The model learned to generalize these patterns to predict eligibility.
4. Evaluation: The model's effectiveness was assessed using metrics such as accuracy, precision, recall, and F1-score on the test set. To enhance its robustness and minimize over fitting, cross-validation was employed during training.
5. Deployment: After optimization, the model was integrated into a loan decision-making system. This system allowed loan officers to receive real-time predictions, expediting the application review process.

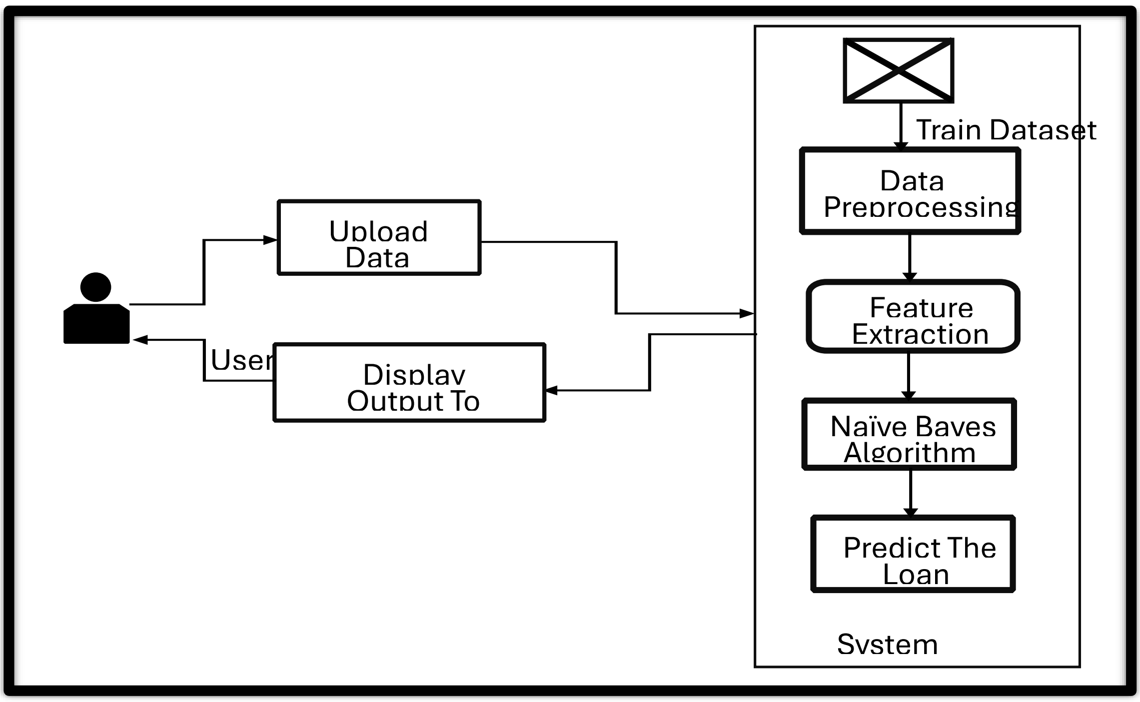

Figure 3.1: Architecture of the Existing System (Sources: Kondeti & Veerabhadra 2022)

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

Constraints of the Existing System

The existing system is therefore constrained with the following;

1. Limited Accuracy and Reliability: Despite its effectiveness, the existing system often struggles with achieving high levels of accuracy and reliability, leading to occasional misclassifications of loan applicants.
2. Dependence on Data Quality: The system heavily relies on the quality and completeness of the data. Missing, inconsistent, or outdated data can significantly reduce the model's predictive performance.
3. Lack of Feature Diversity: The existing system is constrained by the limited range of features used for prediction. It may not account for qualitative factors such as market trends, applicant behavior, or sudden changes in financial circumstances.

Analysis of the Proposed System

The python software will be used to train the neural network model with the Back propagation gradient-based algorithm using the credit dataset. The dataset is sub-divided into training (70%), validation (15%) and testing (15%). In all the cases, the samples have been randomly chosen as to cover the specified percentages. In the context of credit scoring, numerous studies have proven that Artificial Neural Network perform remarkably better than any other statistical approach, such as logistic regression or discriminant analysis. The parameter is customer place of work, customer monthly salary, loan amount, transaction history and credit history are captured for the system to predict if the customer is eligible for the loan. In the credit worthiness prediction system, after training the Artificial Neural Network (ANN) model using data such as monthly salary, loan amount, credit history, and transaction history, a Support Vector Machine (SVM) was utilized to display the final result. The SVM acted as a classifier, leveraging the output of the trained ANN model to categorize individuals into creditworthy or non-creditworthy classes.

During the process, the ANN model learned patterns from the input data through several hidden layers, adjusting its weights and biases based on error feedback. Once training was completed, the model produced predictions for each individual in the dataset. These predictions, which represented complex relationships between the features, were then passed to the SVM for final classification.

The SVM received the output and applied its hyperplane to classify the data points. By transforming the results into a high-dimensional space, the SVM made clear distinctions between individuals, identifying whether they were creditworthy or not. It used a margin to separate the two categories, ensuring that the classification was as accurate as possible. The SVM effectively displayed the prediction results by providing a clear decision boundary, making it easier to interpret the creditworthiness of individuals based on the trained ANN model's output.

The justification of the proposed system is that it is user friendly, uses deeper learning algorithm such as Neural networks and improves on larger dataset

Architecture of the Proposed System

The NN used has three layers: an input layer, a hidden layer, and an output layer. The input layer represents neurons receiving input stimulus. Then the information is transferred to the next level of layer, known as the hidden layer. Information is weighed before being sent to the next level of layers, depending on the size of the connections amongst neurons. Information is sized as per the processing unit or a transfer function represented in Figure 3.2., Seb, (2019), At first, the client logs in and transfers the information. Furthermore, that information is used in preparing datasets. From that point onward, information preparation is performed on those datasets. Further, this pre-processed information goes for feature extraction, where cognizable proof of significant qualities of the information is done and chosen. At that point, we characterize that information using strategies like the k-mean algorithm to make its prediction. Figure 3.2 Architecture of the Proposal System. Unlike the existing system it trained with more machine learning algorithms.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

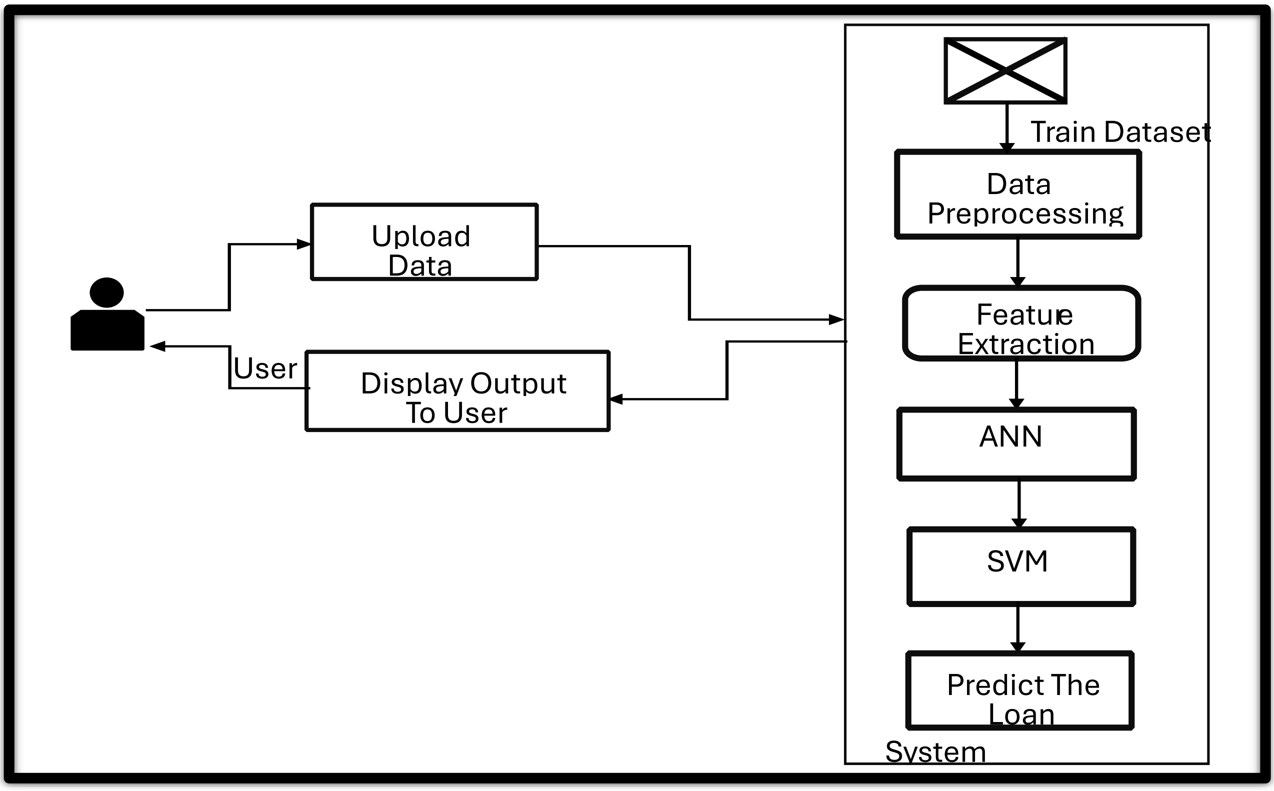

Figure 3.2: Architecture of the Proposed System

The architecture of an ANN typically consists of an input layer, one or more hidden layers, and an output layer. For credit worthiness prediction, the input layer represents the features of the dataset. Each neuron in this layer corresponds to an individual feature. The hidden layers consist of neurons that apply a set of weights and activation functions to transform the inputs into a format that the output layer can use. The output layer, in this case, would have a single –neuron that produces a score representing the borrower's credit worthiness.

Mathematical Model

In this model, we have multiple inputs that feed into a network of interconnected neurons across multiple layers, allowing the model to learn patterns from the data and make predictions.

Let the input data be a vector X=[x1,x2,x3,…,xn], where x1,x2,…,xn, x_1, x_2, …, x_n x1,x2

,…,xn are the financial parameters such as:

x1: Place of work

x2: Loan history

x3: Monthly salary

x4: Loan amount

x5: Transaction history x6: Credit history … xn

Let Wi and bi be the weight matrices and bias terms at layer i

Let f( ) be the activation function commonly a non-linear function like ReLU (Rectified Linear Unit) for hidden layers and sigmoid or softmax for the output layer, depending on the task.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

The model's forward pass can be represented mathematically as (3.1):

$$Z^{(i)} = f(W^{(i-1)} \cdot X^{(i-1)} + b^{(i-1)}) \tag{3.1}$$

where Z(i) is the output of the neurons at layer i, with the final layer's output representing the predicted creditworthiness score.

Training the Model

Training the ANN involves adjusting the weights and biases to minimize the difference between the predicted credit worthiness and the actual known outcomes. This is achieved through a process called backpropagation, which uses the gradient descent algorithm to update the weights and biases. The loss function, often the Mean Squared Error (MSE) for regression tasks, quantifies the prediction error:

ANN falls under the supervised machine learning algorithms category. It can be used for classification, as well as for regression.

Artificial Neural Network Trained Process

The learning (training) process of a neural network is an iterative process in which the calculations are carried out forward and backward through each layer in the network until the loss function is minimized. The entire learning process can be divided into three main parts – see table 3.1. The predictive system is centred on the provision of convolutional neural network which shows number of hidden neurons (hidden unit size), number of training epochs, letter size, maximum number of trials.

Table 3.1 Trained Table

System Parameter	Value
Hidden unit size	30 by 70
Epochs	100
Learning rate	0.01
Soft-max sample temperature	0.1

High Level Input Model (figure 3.3)

The input parameters are customer place of work, customer monthly salary, amount, transaction history and credit history. Any smart device, like a phone, desktop, or laptop, can be used to capture the information into the system. Figure 3.3 is the input design of the proposed system

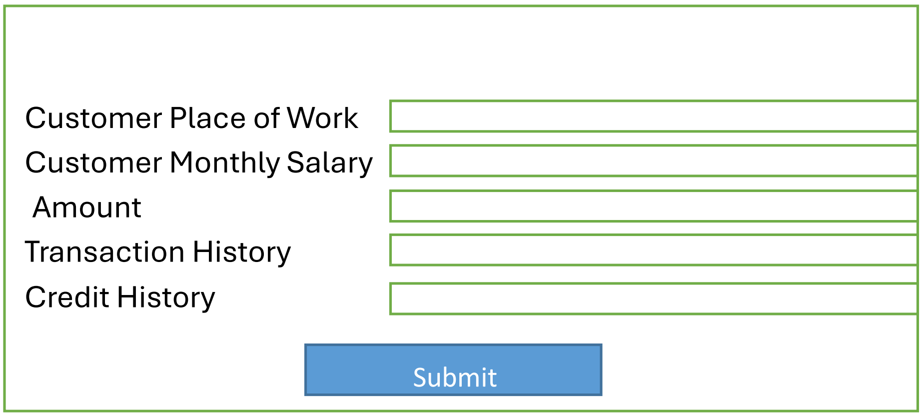

Figure 3.3: Input Design of the System

INTERNATIONAL JOURNAL OF LATEST TECHNOLOGY IN ENGINEERING, MANAGEMENT & APPLIED SCIENCE (IJLTEMAS)
ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

High level process design

The Unified Modeling Language (UML) is a standardized general-purpose demonstration language for object-oriented software engineering. UML makes use of a variety of realistic documentation techniques to create visual representations of article-structured programming frameworks. UML unifies procedures from data demonstrating, business demonstrating, object demonstrating, and segment demonstrating and may be used across the product development lifecycle and across a variety of execution technologies

USE-CASE diagram

The case diagram of the proposed system is shown in figure 3.4. The proposed system allows the user to provide the training data and the threshold constant. Training Neural network requires initialization of input weights and the first input weight will be adjusted until the optimal prediction is achieved. Produce a two-column that summarizes activities and responsibilities of service providers and customers. Identify the various activities and responsibilities associated with each type of service provider and customer. Treat these activities and responsibilities as use cases in a use case diagram.

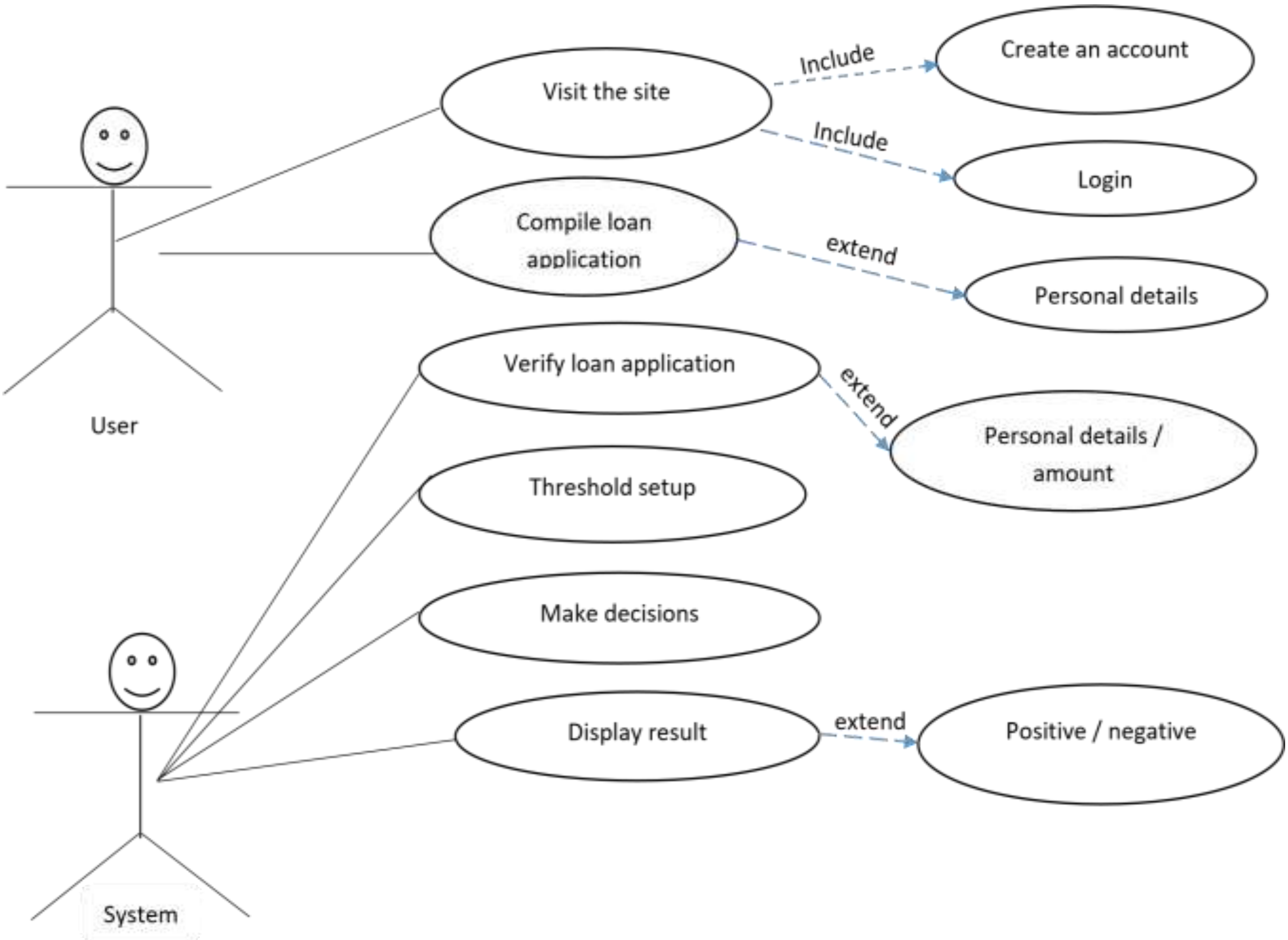

Figure. 3.4 A Use-Case Model of the Proposed System

Sequence Diagrams

Both the loan applicants and the bank employees profit from the strategy's dramatic reduction in sanctioning time. Because of the banking industry's expansion, more people were applying to loans at banks. However, as of right now, no bank can guarantee whether the customer who is selected for a loan application is secure or not, Figure 3.5 Sequence Diagram of the proposed system.

INTERNATIONAL JOURNAL OF LATEST TECHNOLOGY IN ENGINEERING, MANAGEMENT & APPLIED SCIENCE (IJLTEMAS)

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

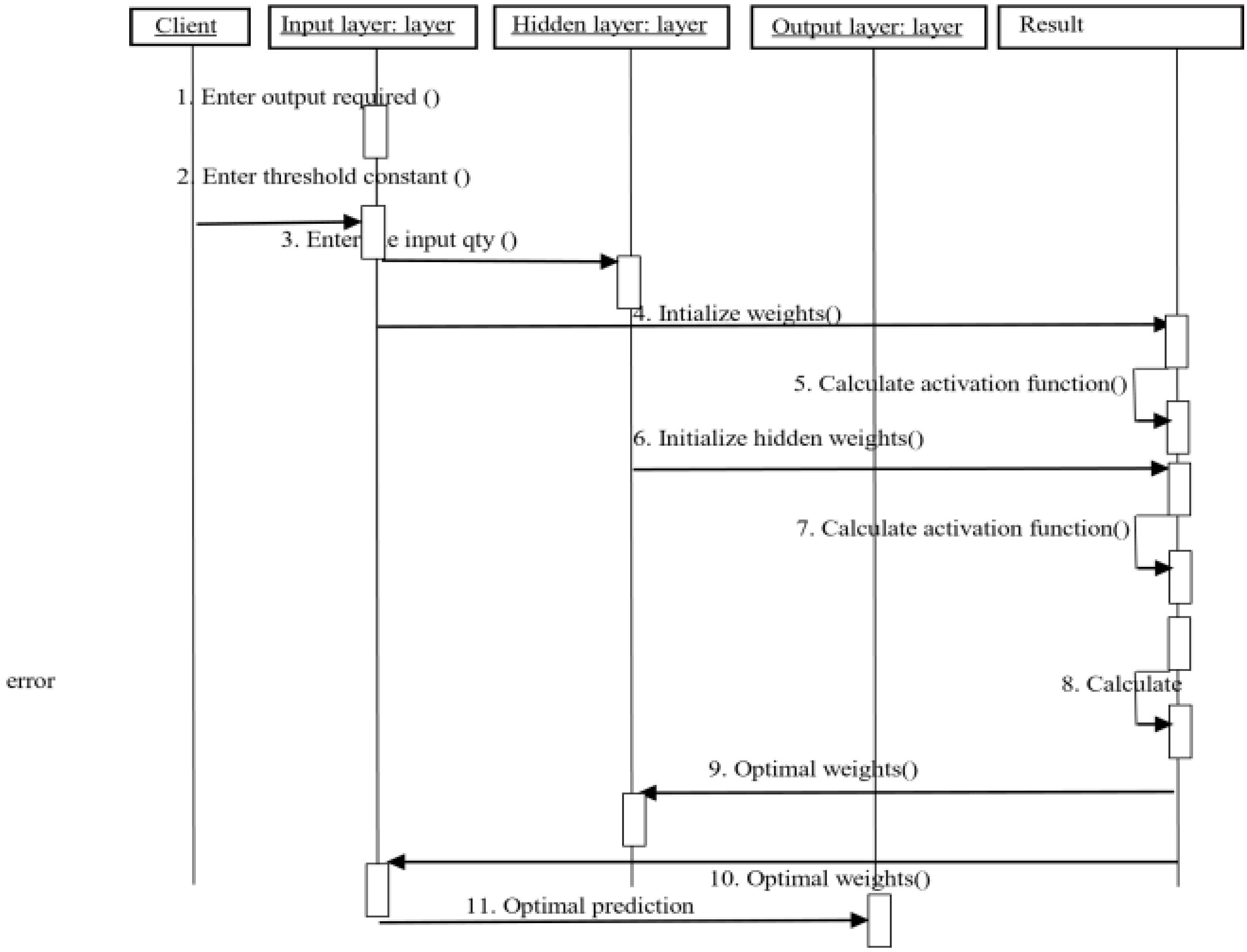

Figure 3.5 Sequence Diagram of the proposed system

High Level output model

The output design figure 3.7, shows the process of predicting the result. It will display the current location.

Congratulation, Mr. Brown! Your loan has been approved.

Figure 3.7: Output form of the proposed system

Database Design of the proposed System

Choice of Development Environment

To accomplish the desired outcome, the implementation of the system was carried out using a combination of tools and technologies which include the following:

1. Python
2. Keras and tensorflow
3. Tweepy
4. Apache Kafka and Zookeeper
5. Flask
6. Plotly

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

System Testing

The software was thoroughly tested using V and V at all levels of the testing. All the input and output pages were thoroughly tested starting from Login screen – figure 4.1.

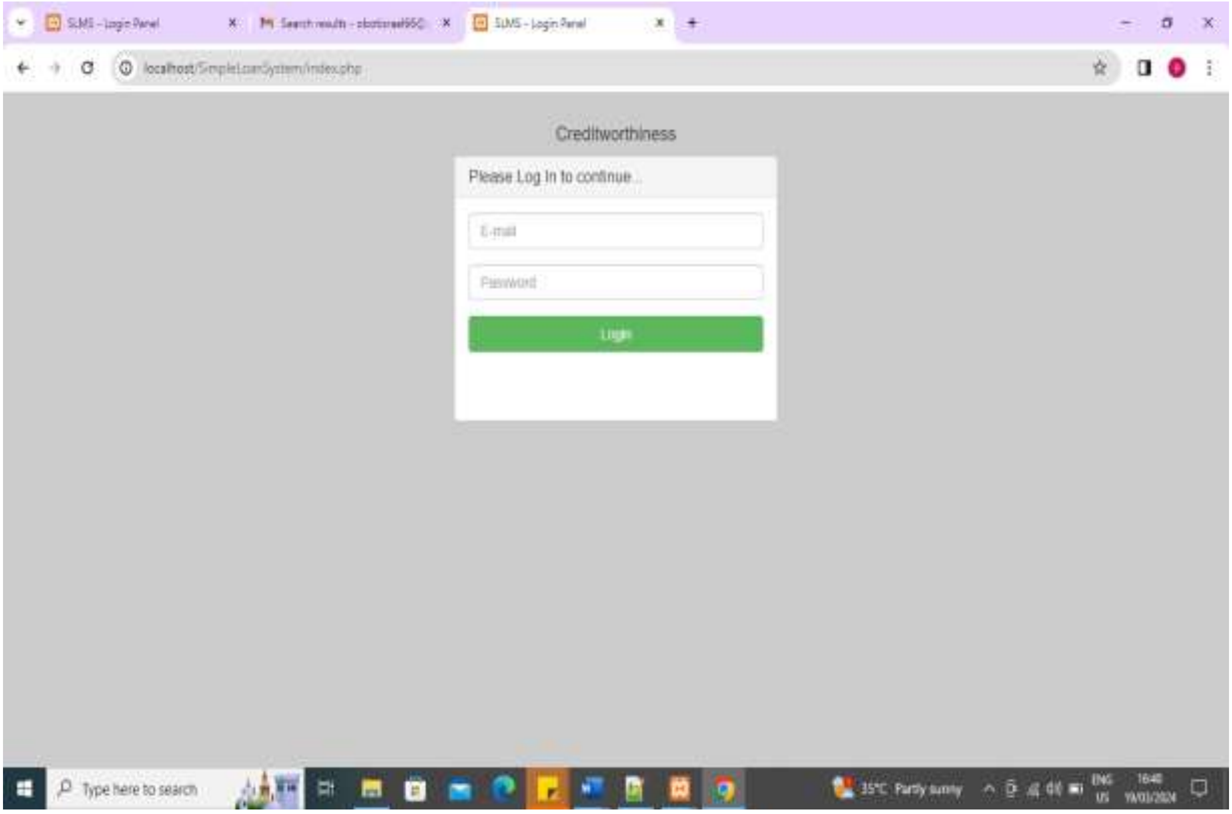

Figure 4.1 Login Page of the System

Admin dashboard

Figure 4.2 Admin Dashboard Page. Admin can add user, register user for loan and view all the record in the database. Admin can also check the list of people that has paid their debt

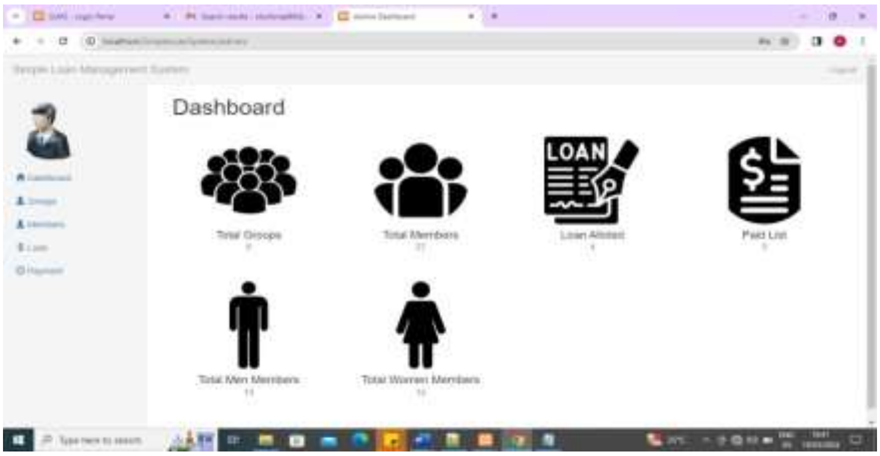

Figure 4.2 Admin Dashboard Page

Loan Page

After user has been register with the system, user can request for loan. The system capture user's information like credit history, amount , place of work , monthly salary, etc., for the trained model ANN to used and evaluate if the people are worthy or not. Figure 4.3 Loan Page

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

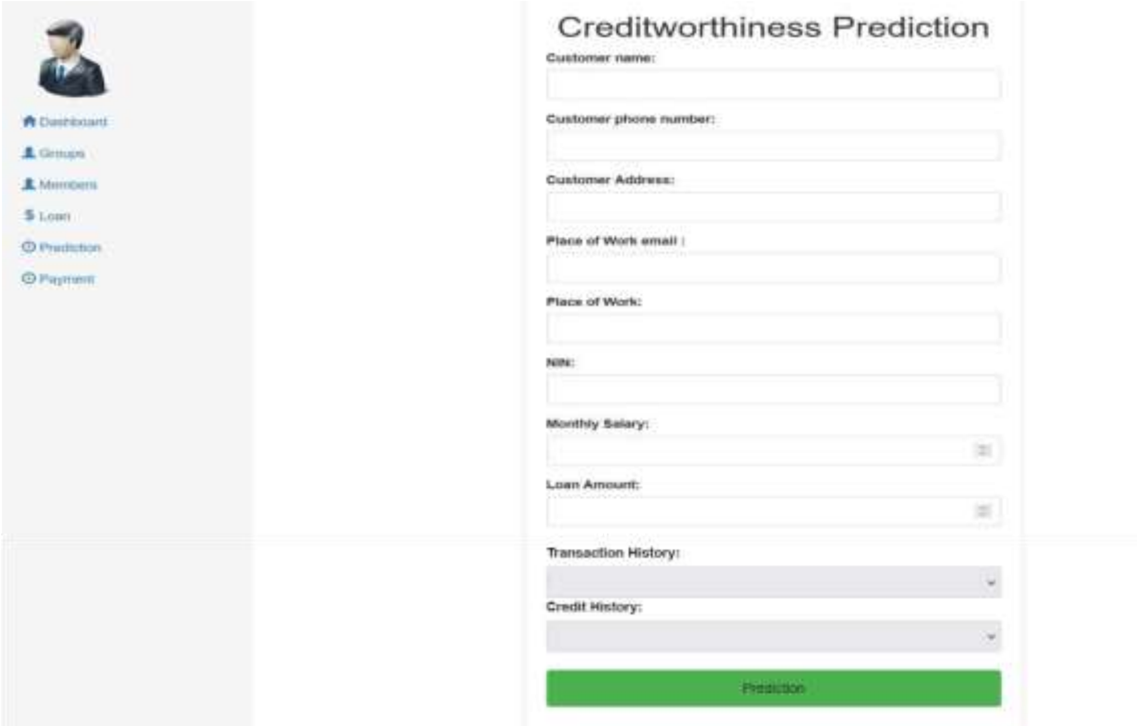

Figure 4.3 Loan Page

RESULT AND DISCUSSION

The performance of the model was evaluated using metrics such as accuracy, precision, recall, and F1-score, which provided insight into the overall efficiency of the creditworthiness prediction system.

Table 4.1: The following table summarizes the metrics and accuracy achieved during the testing phase:

Metric	Value
Accuracy	92.50%
Precision	90.80%
Recall	93.70%
F1-Score	91.20%

Table 4.1 shows the models percentage of the metrics. The model achieved a high accuracy of 92.5%, indicating that the majority of predictions were correct. Precision and recall values, both above 90.8%, demonstrated the model’s reliability in distinguishing creditworthy individuals from non-creditworthy ones. The F1-score of 91.2% reflected a balance between precision and recall, ensuring that the model performed well across all evaluation metrics.

The analysis utilized data points such as Customer name, NIN, Place of Work, Loan History, Monthly Salary, Loan Amount, Transaction History, and Credit History to determine the predicted creditworthiness.

For each customer, the system processed their unique NIN along with the Place of Work to understand the employment context. Loan History and Tansaction History provided insights into previous financial behavior, while Monthly Salary and Loan Amount helped assess the individual's current financial capacity. The Credit History offered a historical perspective on the customer’s credit management.

Using these parameters, the system predicted the creditworthiness of each customer. The predictions ranged across various categories, such as high, average, or low creditworthiness, based on the compiled data – see table 4.2. This predictive analysis allowed for a comprehensive assessment of each individual's ability to handle credit, thereby assisting in making informed lending decisions.

INTERNATIONAL JOURNAL OF LATEST TECHNOLOGY IN ENGINEERING, MANAGEMENT & APPLIED SCIENCE (IJLTEMAS)

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

Table 4.2 Result (Outcome of the predicted model)

Customer Name	Customer Phone Number	Customer Address	Email	Place of Work	NIN	Monthly Salary	Loan Amount	Transaction History	Credit History	Predicted
Ufoma	7E+09	102, Ekewan Road	vernd@yahoo.com	Verbdnic	4.1E+10	₦500,000	₦50,000	Good	Excellent	Approved
Ajonogu	9.1E+09	12, Ugbowo Road	ruffint.@gmail.com	Ruff Int.	2.4E+10	₦300,000	₦500,000	Average	Poor	Not Approved
Ikosa Uku	2.8E+10	204 Festac Town	oluandsons@gmail.com	Olu and Sons	1.2E+10	₦45,000	₦7,500	Good	Good	Approved
Amole	5.3E+10	1 Clitter Avenue	clitterhouse@yahoo.com	Clitter House	4.6E+10	₦600,000	₦100,000	Good	Excellent	Approved
Ikeguru	4.6E+10	109 Idoni Road	akpugari@gmail.com	AkpuGarri Ltd.	4.3E+10	₦250,000	₦240,000	Good	Poor	Not Approved

Figure 4.5 is a sample graph plot representing the predicted creditworthiness of the 10 customers during testing.

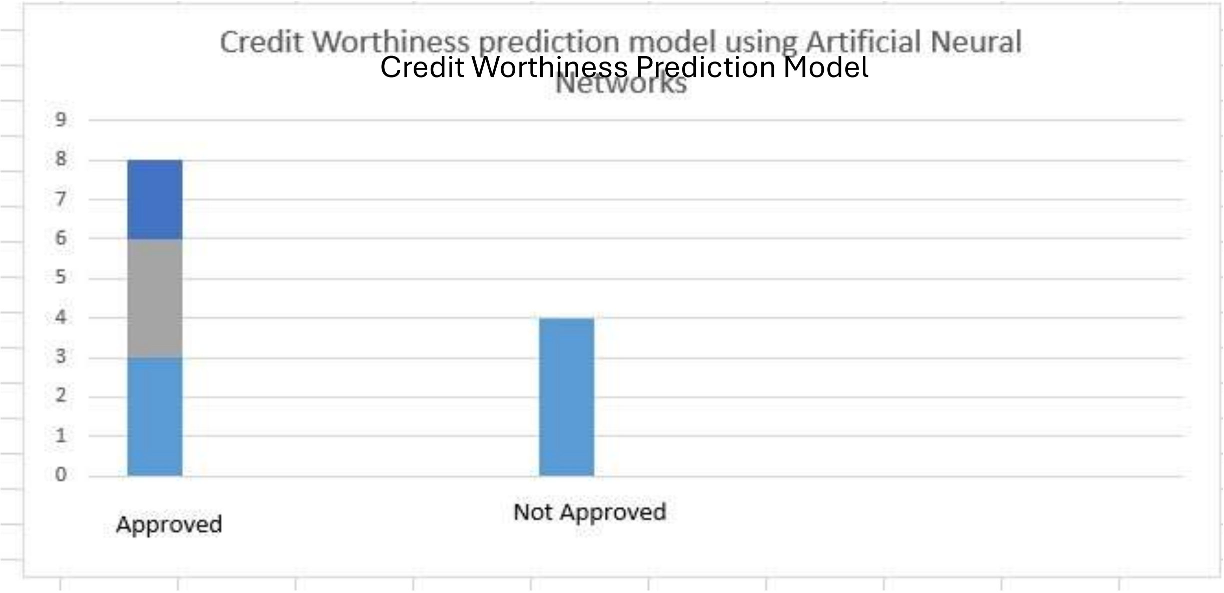

Figure 4.5 graphical result of the system

Table 4.3 Comparison Evaluation.

Metric	Naïve Bayes	ANN	SVM
Accuracy	85.20%	92.50%	89.70%
Loss	35.20%	18.30%	26.40%
F1 Score	83.30%	91.20%	88.60%
Recall	80.20%	93.70%	89.50%
Precision	86.10%	90.80%	88.40%

Explanation of Metrics:

Table 4.3 shows the relative merits of the three models used in terms of Accuracy, Loss, F1 Score Recall and Precision. This is plotted in table 4.6 for further visualization

1. **Accuracy:** Measures the overall percentage of correct predictions out of all predictions.

INTERNATIONAL JOURNAL OF LATEST TECHNOLOGY IN ENGINEERING,
MANAGEMENT & APPLIED SCIENCE (IJLTEMAS)
ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

Higher accuracy indicates a more reliable model.

2. **Loss:** A measure of error in the model's predictions. Lower loss indicates better model

3. **F1 Score:** The harmonic mean of precision and recall, providing a balanced measure of the model's performance. A higher F1 score means the model has a good balance of precision and recall.

4. **Recall:** Measures the model's ability to correctly identify positive instances (e.g., accurately identifying creditworthy individuals). High recall means the model is less likely to miss creditworthy cases.

5. **Precision:** Indicates how many of the predicted positive instances are correct (e.g., when the model predicts a person is creditworthy, how often is it right). High precision means fewer false positives.

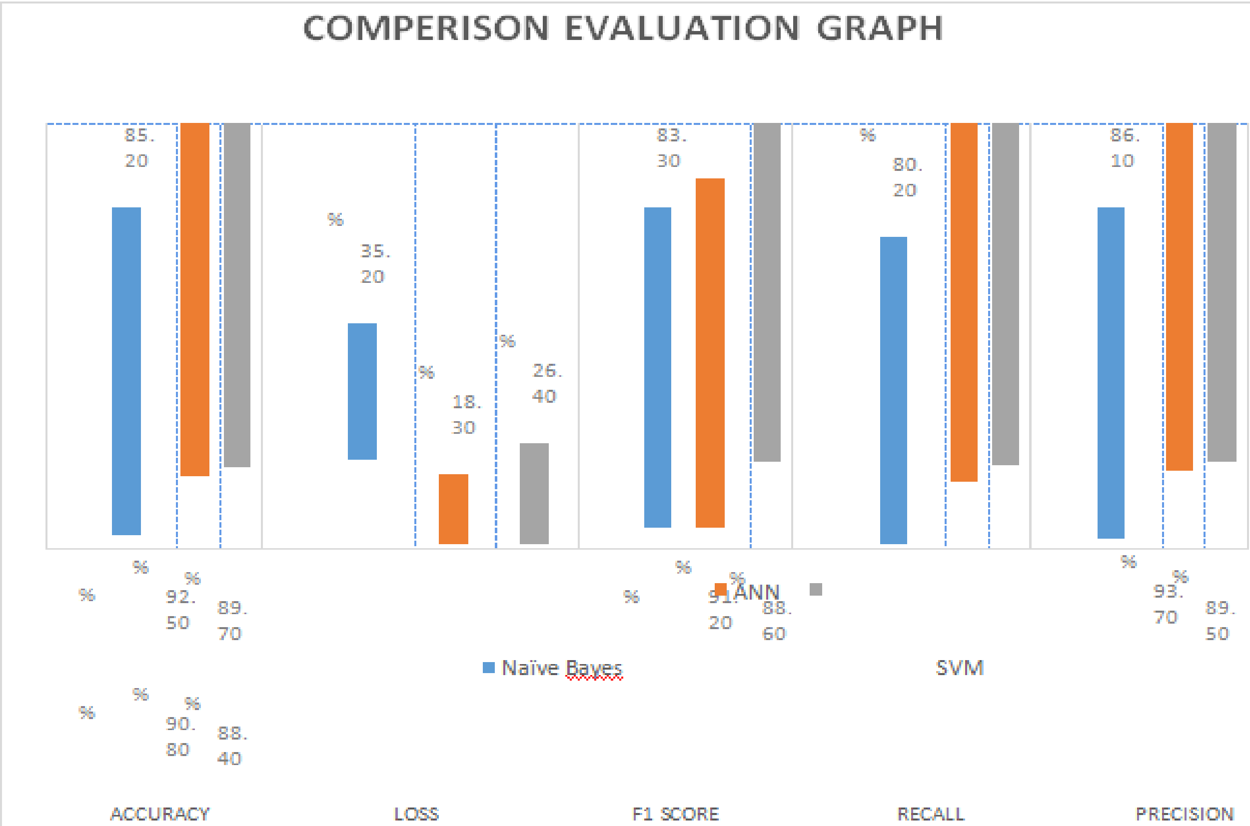

Figure 4.6 graph of the comparison evaluation

Documentation

The researchers tested the implemented software for the following purposes.

1. Interface testing: The interface test focused on three key areas which are the application server, web server, and database server. And they all passed the test cases.
2. Functionality Testing: We carried functionality testing on the implemented system to check if the initial build works as per its system design and analysis. This includes form validation, cookie and session testing, HTML and CSS validation, and database connection check-up
3. Compatibility and performance Testing: In compatibility testing, we check whether or not our web design is compatible with a variety of browsers and devices. This includes browser and OS compatibility testing, along with mobile browsing and printing options testing. Our implemented system looks great in all browsers and compatible with all leading operating systems. And lastly, we carried out performance testing and our implemented system is scalable and capable to withstand multiple users. To install the application, the following steps are required;

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

System Requirements

To be able to access and run the application, users require computer hardware and software with the following minimum requirements and characteristics:

Hardware Requirements

i) Processor: at least Intel(R) Celeron(R) N4020 CPU @ 1.10GHz 1.10 GHz recommended.

ii) Hard disk: at least 1GB of available space required on the system drive, 0.5 GB of available space required on the installation drive.

iii Display super VGA (1024 x 768)

iv. RAM: Minimum requirement of4GB

v CD-ROM Drive

vi USB port enabled

Software Requirements

i Windows 7, 8, 10 and about

ii Remix Solidity IDE

iii Solidity SDK

iv Offline Script Test Kit

v Google Chrome Web browser

SUMMARY

Banks and financial organizations are really facing the challenge of identifying risk factors, which should be considered while advancing the loans/credit to customers. So, Banks have to realize that ML potential can help them to focus their financial decisions, resource utilization and to make smarter decisions. The best model based on ML algorithm algorithms is chosen by data scientist to decide many aspects to generate more useful results. Here in this work, Predictive modelling for detection of Credit Worthiness of Bank Customer is done. The model is based on ANN model of ML. The model is completely built and tested over desktop platform. Building a Predictive model with ML for credit worthiness detection on existent cloud platform and measuring its effectiveness will be the chief focus of this work. The performance of proposed model is shown by comparing its efficacy against three other popular ML algorithms. The learners are compared on basis of various parameters and the proposed model found to be better in terms of prediction rate and other parameters. The findings of work have a lot of implications. In future, we intend to build up an automated risk assessment system over cloud for financial organizations that will incorporate key features to determine credit worthiness of customers.

CONCLUSION

In this project, the researchers have presented an approach to prediction of artificial neural network to prediction of bank customers' credit worthiness. After determining the required variables, the collected data formed input to the model. Results show that out of 50 non-creditworthy cases, 78.0% are predicted as true negative and 22.0% are predicted as false negative. Out of 81 creditworthy cases, 95.1% are true positive and 4.9% are classified as false positive. Overall, 88.5% of the predictions are correct and 11.5% are wrong

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

classifications. Based on the conclusion drawn, the project recommends that banks should adopt artificial neural network technology in determining who gets loan because of its predictive ability.

REFERENCE

1. Ajayi O. Bolaji, (2016), JABU International Journal of Social and Management Sciences Volume 8, Number 2.
2. Ampountolas, A. C., Antonakis, & Sfakianakis, M. E., (2021). Assessing naïve Bayes as a method for screening credit applicants. Journal of Applied Statistics, 537-545.
3. https://doi.org/10.1080/02664760802554263
4. Atiya, D., Djeundje, V. B., Crook, J., Calabrese, R., & Hamid, M., (2023). Enhancing credit scoring with alternative data. Expert Systems with Applications, 163, 113766.
5. Baesens, D., Opitz, R., & Maclin, R., (2022). Popular ensemble methods: An empirical study. Journal of Artificial Intelligence Research, 169-198.
6. Beattie, S., Hand, D. J., & Henley, W. E., (2020). Statistical classification methods in consumer credit scoring: A review. Journal of the Royal Statistical Society: Series A (Statistics in Society), 160(3), 523-541.
7. Bulus, Yusuf D., Okorie George, Adedeji, Kayode E., (2024), Monetary Policy Transmission in Nigeria: assessing the Effectiveness of the Banking Lending Channel, West African Journal of Monetary and Economic Integration, December 2024, 2(23).
8. Burgers, D. H., & Wolpert, D. H., (2022). Stacked generalization. Neural Networks, 5(2), 241259.
9. Charles K. K., (2025), International Journal of Recent Research in Commerce Economics and Management (IJRRCEM) Vol. 12, Issue 1, pp: (1-12), Month: January - March 2025, Available at: www.paperpublications.org
10. Dennis, Robertson, Bertil, Ohlin, (1930), Loanable funds Theory, Economic Online, https://www.economicsonline.co.uk
11. Djeundje, A., & Abhishek., (2021). Enhancement of classification techniques using principal component analysis and class imbalance handling methods in credit card defaulter detection. International Journal of Forensic Engineering, 1. https://doi.org/10.1504/IJFE.2021.117383
12. Edelman, Shimamura, K., Fukushima, S., Koura, A., Shimojo, F., Misawa, M., Kalia, R. K., Nakano, A., Vashishta, P., Matsubara, T., & Tanaka, S. (2021). Guidelines for creating artificial neural network empirical interatomic potential from first-principles molecular dynamics data under specific conditions and its application to α-Ag2Se. The Journal of Chemical Physics, 151(12), 124303.
13. Elmasry, V. B., Djeundje, Crook, J., Calabrese, R., & Hamid, M., (2022). Enhancing credit scoring with alternative data. Expert Systems with Applications, 163, 113766.
14. Hand, N., & Karthiban, R., (2020). A review on machine learning classification techniques for bank loan approval. In International Conference on Computer Communication and Informatics (ICCCI), IEEE, 1-6. https://doi.org/10.1109/ICCCI.2019.8822014
15. Hornik, T., Lee, T., & Chen, I., (2021). A two-stage hybrid credit scoring model using artificial neural networks and multivariate adaptive regression splines. Expert Systems with Applications, 2(21), 1-23.
16. Jentzsch, S., Semerikov, I., Teplytskyi, I., Yechkalo, Y., Markova, O., Soloviev, V., & Kiv, A., (2021). Using spreadsheets as learning tools for computer simulation of neural networks. SHS Web of Conferences, 75(1), 04018.
17. Kanarachos, S., Elreedy, D., & Atiya, A. F.c, (2021). A comprehensive analysis of the synthetic minority oversampling technique (SMOTE) for handling class imbalance. Information Sciences, 32-64. https://doi.org/10.1016/j.ins.2019.07.070
18. Kondeti, Durga Bhavani, & Veerabhadra Rao, A., (2022). Prediction of loan approval using machine learning algorithm. Journal of Engineering Sciences, 13(08). 377-888
19. Laitinen, E. K., Trinkle, B. S., & Baldwin, A. A., (2022). Research opportunities for neural networks: The case for credit. Intelligent Systems in Accounting, Finance and Management, 23(3), 240-254.
20. Li, O., & Ekaterina, V., (2019). Methodology and models for individuals' creditworthiness management using digital footprint data and machine learning methods. Mathematics, 9(15), 1820. https://doi.org/10.3390/math9151820

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

21. Meuller, R., Mukhamadiev, R., Staroverova, N., & Shustrova, M. (2021). Specifics of project management system development for large organizations. In International MultiConference on Industrial Engineering and Modern Technologies (Far East Con), Vladivostok, 1-8. https://doi.org/10.1109/FarEastCon50210.2020.9271201
22. Mills, A., Ampountolas, A., Nyarko, T., Date, P., & Constantinescu, C., (2020). A machine learning approach for micro-credit scoring. Risks, 9(3), 50. https://doi.org/10.3390/risks9030050
23. Moro, B., Blagus, R., & Lusa, L., (2021). SMOTE for high-dimensional class-imbalanced data. BMC Bioinformatics, 14, 106. https://doi.org/10.1186/1471-2105-14-106
24. Olokoyo O. Felicia, Taiwo N. Joseph, Akinjare A. Victoria, (2016), The impact of Banks' activities on Economic Development of Nigeria, International Business Management, Medwell Journals, 10(22), 5341-5351, 2016.
25. Pang, S. V., & Sorokin, A. S. (2022)., Ispolzovanie nejrosetevyh modelej v povedencheskom skoringe. Applied Computer Science, 2(56), 92-109.
26. Rajkumar, S. T., & Vasarhalyi, M. A. (2020). Predicting credit card delinquencies: An application of deep neural networks. In Handbook of Financial Econometrics, Mathematics, Statistics, and Machine Learning (pp. 4349-4381).
27. Rmbrust, F., Shen, F., Zhao, X., Kou, G., Fawaz, E., & Alsaadi, F., (2020). A new deep learning ensemble credit risk evaluation model with an improved synthetic minority oversampling technique. Applied Soft Computing, 98, 106852.
28. Saunders, C., Cornett, V., Vukovic, D., Vyklyuk, Y., Matsiuk, N., & Maiti, M., (2020). Neural network forecasting in prediction Sharpe ratio: Evidence from EU debt market. Physica A: Statistical Mechanics and its Applications, 542, 123331.
29. Schaeck, Klaus, Martin Čihák, and Simon Wolfe., (2009). "Are Competitive Banking Systems More Stable?" Journal of Money, Credit, and Banking 41 (4): 711–34.
30. Shao V. T., Mfaneli N., Lesley M., Bonginkosi T., (2025), Applications and Competitive Advantages of SMEs Performance: a Systemic Review, business 2025, 5(2), 22, https://doi.org/10.3390/business 5020022
31. Seb, (2019), Understanding Basic Neural Network Layers and Architecture, Deep Learning, Machine Learning
32. Sipos, Gusarova, O. M., Komarov, P. I., & Denisov, D. J. E., (2020). Bulletin of the Altai Academy of Economics and Law, 5, 107–113.
33. Suleiman, S., Burodo, M. S., & Yusuf, G., (2018). An application of artificial neural networks in predicting customers' loan status in Nigeria: A case study of First Bank Nigerian PLC. An International Journal of Novel Research in Marketing Management and Economics, 5(3), 20-26.
34. Tamayo, V., & Guijarro, F. (2020). Credit risk management: A multicriteria approach to assess creditworthiness. Mathematics and Computers in Simulation, 200, 2009–2015. https://doi.org/10.1016/j.matcom.2019.07.001
35. Tang, A. T. O., (2022). Types of machine learning algorithms. In New Advances in Machine Learning (pp. 1-12). InTech.
36. Adolphus Toby, (2023), Journal of Global Economics and Business, https://DOI.org/10.98534/JGEB.V2017.152
37. Wu, G., Wang, J., Hao, J., & Jiang, H., (2020). A comparative assessment of ensemble learning for credit scoring. Expert Systems with Applications, 38(11), 223–230. https://doi.org/10.1016/j.eswa.2011.12.036.
38. Wikiversity, (retrieved, 2025), Artificial neural network - Wikiversity
39. Yu-Chang Hu, Jake Ansell, (2007), Measuring Retail Company Performance Using Credit Scoring Techniques, 2007, European Journal of Operational Research, 183(3),1595-1606, DOI: 10.1016/j.ejor.2006.09.101
40. Zhong G., & Banasik J., (2018). Not if but when will borrowers default. Journal of the Operational Research Society, 69(8), 1185–1190. https://doi.org/10.1057/palgrave.jors.2600851