ESTIMATION OF PROBABILITY OF DEFAULT OF A FINANCE HOUSE IN NIGERIA: A COMPARATIVE ANALYSIS BETWEEN LOGISTIC REGRESSION AND NEURAL NETWORK MODEL

By
Ajijola, Lukman Abolaji and Fawehinmi Oluwadamilola Helen
Department of Actuarial Science and Insurance, Faculty of Management Sciences, University of Lagos. Email: lajijola@unilag.edu.ng

Abstract
Effective credit management is essential for economic stability, while poor management can have destabilizing effects. Accurate credit risk assessment, including probability of default (PD), is crucial for financial institutions to prevent substantial losses from non-performing loans. The purpose of the study is to investigate the probability of default in low default portfolio of a finance house in Nigeria. In order to fulfil this purpose, two different models for estimating the probability of default, the logistic regression and the neutral network models, were considered. The data analysis was steered through SPSS version 23 software. Through a comparative analysis of logistic regression and neural network models, we determined that the neural network model is superior in predicting the probability of default. The study findings unveiled that the primary factors influencing the probability of default are the Tenor of loan and Annual nominal interest rate. The study recommends considering adjustments to terms to mitigate risk, adjusting risk strategies and implementing the neutral network model for ongoing prediction of default probabilities.

Keywords: Probability of default, Low default portfolio, Credit risk, Logistic regression, Network neutral.

1. Introduction
Loan lending has had a substantial impact on the daily lives of both enterprises and individuals (Uzair, Hafiz, Asim, & Nowshath, 2019). Obtaining credit has become nearly inevitable due to the growing number of financial constraints on individuals, organisations, and corporations (Perera & Premaratne, 2016). As individuals utilise credit to fulfil various desires, financial institutions, especially banks, have started to make increased revenue for themselves. Prudent management of credit can have a positive impact on the economy and effectively address issues connected to cash flow (Calli & Coşkun, 2021). Individuals worldwide rely on credit transactions for several purposes, including overcoming financial barriers to achieve their aspirations. Furthermore, it is imperative for a bank to fulfil its loan provision responsibility in order to sustain its operations (Osayeme, 2000). However, when used irresponsibly, credit can have detrimental effects and destabilise the economy (Calli, 2019).

The success of financial organisations, particularly banks, credit unions, fintech companies, and other lending institutions, relies on the accurate assessment and efficient control of credit risk in comparison to other inherent risks (Chandran, 2022; Gieseche, 2004). When lending institutions inaccurately determine the cost of credit, they often neglect to consider all associated expenses, which raises the level of risk and makes it more challenging to manage and reduce those risks (Sudi & Maniagi, 2020). These flaws arise due to the inadequate credit rating and viability evaluation of these companies. Defaulting becomes prevalent due to an inadequate assessment of the credit's probable adverse consequences and their effect on the institution. This exposes an institution to potential harm or danger. An inadequate credit risk assessment could lead to substantial financial losses for the lenders (Kiveu, 2015). Banking institutions have experienced a significant increase in non-performing loans, despite the rise in their loan portfolios. Although portfolio diversification has improved, credit risk remains a significant threat to profitable lending. The Probability of Default is the most accurate and often used approach for evaluating credit risk. A defaulter is an individual who is unlikely to repay the loan amount or who will be delinquent on their loan payments by more than 90 days. Hence, the computation of the Probability of Default is a crucial stage in evaluating the creditworthiness of individuals applying for loans (Sudhamathy, 2016). The most accurate estimation of the likelihood of default and the amount of exposure can be achieved through the utilisation of machine learning techniques. The credit risk associated with loans made by commercial banks to individuals and businesses can be described using several quantitative metrics. These metrics include the probability of loan default, acceptable risk, average risk, potential losses in case of default, typical size of those losses, and the maximum losses that can be incurred (Konovalova, Kristovska, and Kudinska, 2016).

Neural Network models and other Machine learning are used to create models that predict loan default. These models are important for generating credit risk evaluation models that can automate or assist in making credit decisions (Zakrzewska, 2007). By utilising machine learning algorithms, it is possible to construct a novel model by employing anonymized historical data. This model can be trained to enhance predictions for various types of risks, such as credit risks and other risks like the likelihood of early repayment leading to losses in interest income or the possibility of money laundering, among others. Financial institutions can utilise a reliable model to predict the probability of a borrower returning a loan before the scheduled date. This model can then be used to develop procedures that include specific preventive steps to be taken before such early repayment happens. (Zoran, 2019). Machine learning techniques have been proven to achieve successful outcomes in credit risk analysis (Bailey, Cao, Kuchler, Stroebel, & Wong, 2018; Freedman & Jin, 2017). The utilisation of machine learning for constructing a credit risk management model is not novel, however it has shown substantial growth in recent years. Machine learning has also gained more interest due to the complexity involved in estimating credit risk. Furthermore, when analysing the application of machine learning in credit risk, it is essential to go beyond solely calculating default probability (PD) and also analyse its applications in projecting exposure amount. In the light of this background, this study aimed to estimate the probability of default of a finance house in Nigeria by using the better model the logistic regression and neural network model.

2. Literature Review

The topic of credit risk has garnered considerable interest as a result of the concerns stemming from the 2008 financial crisis (Matoussi, 2010). Credit risk is a significant problem for commercial banks as it has the potential to affect their ability to maintain their business operations (Al-abedallat & Kattel, 2016). Credit risks refer to the potential for bank borrowers or counterparties to fail to meet their obligations as specified in the agreement (Luy, 2010; Basel Committee on Banking Supervision, 2004). Credit risk management is a preventive measure against credit hazards. This encompasses the process of making strategic decisions before to granting credit, carrying out the necessary steps to complete credit obligations, and overseeing and reporting on these operations (Taiwo, Ucheaga, Achugamonu, Adetiloye, Okoye & Agwu, 2017; Konovalova, Kristovska & Kudinska, 2016).

Nazre, Nasimul, and Md. Umar (2021) analysed a dataset from a bank that included data on consumers who had defaulted on their credit. They employed supervised machine learning algorithms to predict a customer's credit capacity and identify the clients most likely to restore their credit in the short term. In addition, they employed feature scaling techniques to enhance the performance and effectiveness of several machine learning models. The precision of various Machine Learning techniques was subsequently assessed and contrasted with the precision of other classification algorithms. Subsequently, the final mandatory functionalities were incorporated. A study revealed that Random Forest outperforms other classifiers and has a credit recovery prediction accuracy of 89%.

Aniceto, Barboza, and Kimura (2020) conducted a study to assess the effectiveness of machine learning techniques in classifying borrowers using a loan dataset from a Brazilian bank. They developed prediction models for Random Forest, AdaBoost, Support Vector Machine, Bagging, and Decision Trees in order to evaluate their performance compared to a target based on a Logistic Regression model. The analysis of comparisons is conducted utilising the standard categorization performance metrics. The results indicate that Adaboost and Random Forest models exhibit superior performance compared to other models. In addition, SVM models exhibit subpar performance when employing both nonlinear and linear kernels. Their results suggest that it is beneficial to allow banks the opportunity to enhance default forecasting models by exploring machine learning methods.

Luo's (2020) study examined the precision of rating accuracy in credit evaluation using five decision-support approaches. Based on the investigation, the RF algorithm has been determined to be the most effective, with error rates exceeding 5%. The mistake rate for defaulting companies decreased by 22.6%, leading to the conclusion that ANN is the second most effective classifier.

Dawood, Elfakhrany, and Maghraby (2019) conducted a study to examine the application of k-means, improved k-means, fuzzy c-means, and neural networks in analysing bank clients' behaviour for the purpose of evaluating their creditworthiness. The primary objective of this study was to prioritise neural network classification to enhance the speed and accuracy of clustering execution. The Neural Network classifier emerged as the superior choice when the accuracy ratio was employed as the comparison criterion in this experiment.

The study conducted by Supriya, Pavani, Saisushma, Kumari, and Vikas (2019) found that the Decision Tree model achieved the greatest accuracy rate of 81.1% when compared to other machine learning models in the context of loan prediction. The researchers employed decision trees, K Nearest Neighbour, support vector machines, and gradient boosting algorithms to analyse a dataset containing both qualitative and category data, consisting of 12 variables.

Damrongsakmethee and Neagoe (2019) conducted a case study that examined the successful application of Artificial Neural Networks (ANN) in credit risk assessment. According to the authors, the analysis of credit data from Germany and Australia using ANNs resulted in an overall accuracy of approximately 81.2% and 90.85%, compared to 78.67% and 89% obtained using a mixed model. Therefore, the ANNs were determined to be more accurate.

Tudor, Bara, and Opera (2017) conducted a comparative analysis of data mining approaches using financial data from Romanian banks to predict the probability of credit default in the banking sector. The data collection comprises 18,239 cases, out of which 1,489 are marked as NPL. Logistic Regression demonstrated superior performance in predicting defaulters compared to other models, including Naive Bayes and Support Vector Machines.

3. Material and Methods

Material

This study employed a descriptive research approach to monitor and describe the probability of default. The aim was to provide an explanation for the observed condition. For the purpose of this study, a secondary source of data that included the low default data set consisting of credit customers was utilised. The data was sourced from a Finance house in Nigeria. The data set for this study contains five hundred (500) samples of client loans, each of which is represented by 6 variables. The variables include gender, years of work, annual nominal interest rate, tenor of loan, loan type and probability of default. This study comprises one response variable which is probability of default. The explanatory or predictor factors are gender, years of work, annual nominal interest rate, tenor of loan, loan type.

Methods

This section covered the methods of data analysis. This involves both descriptive statistics of data and utilization of both Logistics regression and neural network analysis for modelling the probability of default.

Logistics Regression

The objective of logistic regression is to classify an observation by determining the likelihood that the observation belongs to a specific class. The logistic function can be employed to compute the posterior probability of a client belonging to the default class, based on an input value of k.

$$P(J_i = 1|K_i = k_i) = \frac{e^{\beta_0+\beta^T k_i}}{1+e^{\beta_0+\beta^T k_i}} \quad (1)$$

where the parameters of the linear models are β_0 and β. β_0 signifies the intercept and β denote the coefficients vectors, $\beta = [\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \ldots, \beta_p]^T$. The logistic function from the above equation (3) is gotten from the relationship between the log-odds of $P(J_i = 1|K_i = k_i)$ and a linear transformation of k_i, i.e.

$$\log \frac{P(J_i = 1|K_i = k_i)}{1-P(J_i = 1|K_i = k_i)} = e^{\beta_0+\beta^T k_i} \quad (2)$$

The prediction class is given as:

$$\hat{y}_t = \begin{cases} 1, & if \quad P(J_i = 1|K_i = k_i) \geq h \\ 0, & if \quad P(J_i = 1|K_i = k_i) < h \end{cases} \quad (3)$$

where h is a threshold parameter of the decision boundary. Additionally, the log-likelihood of J_i is maximized in order to determine the parameters β_0 and β. Hence, equation 3 can be represented as:

$$P(k_i;\ \beta_0, \beta) = P(J_i = 1|K_i = k_i;\ \beta_0, \beta) = \frac{1}{1+e^{-(\beta_0+\beta^T k_i)}} \quad (4)$$

Because $P(J_i = 1|K_i = k_i;\ \beta_0, \beta)$ fully characterizes the conditional distribution, the multinomial distribution is suitable as the likelihood function.

Neural Network

An input layer, an output layer, as well as many hidden layers make up a neural network (see Figure 1 below)

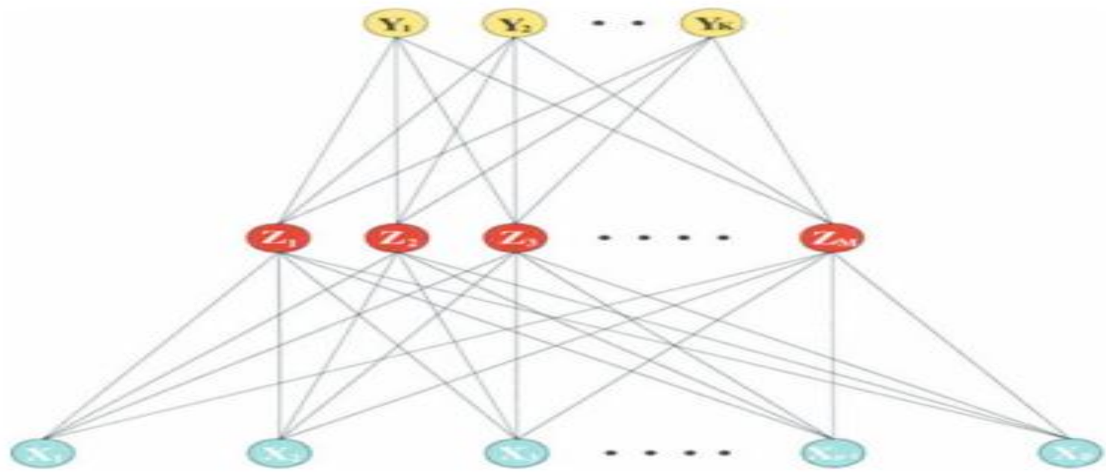

Figure 1: Structure of a single hidden layer, feed-forward neural network (Hastie, Friedman and Tibshirani, 2001).

The input layer consists of p independent variables, denoted as $x_1, x_2, \ldots, x_p$. Each x_i represents the ith observation, where $x_i = [x_{i1}, x_{i2}, \ldots, x_{ip}]$, and can take any value T. In K −class classification problems, K represents the number of target measurements Y_k, where k ranges from 1 to K. Additionally, K is used as a binary variable to represent each subsequent class, with a value of either 0 or 1. The target Y is a function that is formed by combining a variable Z_m, which was created by combining the inputs in a linear manner.

$$C_k = \beta_{0k} + \beta_k^T Z, k = 1, 2, \ldots K \tag{5}$$

where $Z = [Z_1, Z_2, \ldots, Z_M]^T$ and $\beta_k = [\beta_{k1}, \ldots, \beta_{km}]^T$. Additionally, a sigmoid function $\sigma_k(C)$ is adopted to the compute the final transformation of the vector $C = [C_1, C_2, \ldots, C_K]^T$

$$g(x_i) = \sigma_k(C) = \frac{1}{1+e^{-C_k}} \tag{6}$$

Since the full collection of weights is unknown, the goal is to determine their ideal values to ensure that the model better approximates the training set of data.

Model Performance Criteria

The evaluation metrics employed to know the most appropriate model for investigating the probability of default are Accuracy, Precision, Recall and F-measure.

- Accuracy = (True Positive + True Negative)/(True Positive +True Negative + False Positive + False Negative)
- Precision = True Positive /(True Positive + False Positive)
- Recall = True Positive /(True Positive + False Negative)
- F-Measure = (2 × Precision × Recall)/(Precision + Recall)

The performance of the model is greater when the values of Accuracy, Precision, Recall, and F-measure are higher.

4. Results

This section comprehensively presents and analyses the collected data, beginning with a detailed descriptive analysis of the socio-demographic characteristics of credit clients, such as gender, years of working, annual nominal interest rate (%), tenor of loan and loan type. Following this, the chapter delves into probability of default modelling and forecasting, employing both Logistic Regression and Neural Networks.

Demographic Data of the Credit Clients

Table 1: Socio Demographic Analysis of Credit Clients

Variable	Label	Frequency	Percent
Gender			
	Male	353	70.6
	Female	147	29.4
Years of Work			
	Less than 5 years	254	50.8
	5-9 years	124	24.8
	10-14 years	71	14.2
	Above 14 years	51	10.2
Annual Nominal Interest Rate (%)			
	42.00	380	76.0
	46.20	1	.2
	48.00	58	11.6
	54.00	61	12.2
Tenor of Loan			
	3 Months	26	5.2
	4 Months	24	4.8
	5 Months	8	1.6
	6 Months	112	22.4
	7 Months	19	3.8
	8 Months	23	4.6
	9 Months	48	9.6
	10 Months	8	1.6
	11 Months	2	.4
	12 Months	230	46.0
Loan Type			
	New Loan	486	97.2
	Top up Loan	14	2.8

Source: Finance House, 2024

Table 1 presents a range of demographic information about the credit clients involved in this study. The table shows a significant gender disparity among credit clients, with males constituting 70.6% (353 clients) and females making up 29.4% (147 clients). This suggests a male-dominated client base, which could be due to various socio-economic factors or preferences in loan uptake. Most clients have relatively short work experience, with 50.8% (254 clients) having less than 5 years of work experience. This is followed by 24.8% (124 clients) with 5-9 years of experience, 14.2% (71 clients) with 10-14 years, and only 10.2% (51 clients) with over 14 years. This distribution indicates a younger workforce seeking credit, which might imply different risk profiles and financial needs compared to those with longer work histories. A majority of clients (76.0%, or 380 clients) are subject to an annual nominal interest rate of 42.00%. The rest of the clients are distributed among higher rates: 46.20% for one client, 48.00% for 11.6% (58 clients), and 54.00% for 12.2% (61 clients). The concentration at 42.00% suggests that this is the standard rate offered, while the others might be due to special circumstances or different credit profiles. The loan tenor shows a diverse range, with the highest concentration of clients (46.0%, or 230 clients) opting for 12-month loans. This indicates a preference for longer repayment periods. Other notable tenors include 6 months (22.4%, or 112 clients) and 9 months (9.6%, or 48 clients). Shorter tenors (3-8 months) and other durations are less popular, suggesting that clients might prefer spreading their repayments over a longer period for better financial management. A vast majority of the loans are new loans, accounting for 97.2% (486 clients), while top-up loans are minimal at 2.8% (14 clients). This implies that most clients are either first-time borrowers or prefer to take new loans rather than topping up existing ones.

Modelling the Probability of Default Using Logistics

To model the probability of default using logistic regression, which is suitable for a binary response variable, several pre-diagnostic tests are essential to ensure the validity of the model. These include testing for multicollinearity using the Tolerance and Variance Inflation Factor (VIF) method, where a high VIF (above 10) indicates problematic multicollinearity, and testing for outliers using the Mahalanobis distance criterion, which identifies outliers based on their multivariate distance from the mean. Once these diagnostics confirm the assumptions are met, the logistic regression model is specified, with the probability of default modeled as a function of predictor variables. Significant predictors are selected, and model parameters are estimated using maximum likelihood estimation.

Table 2: Multicollinearity Test

Variables	**Collinearity Statistics**	
	Tolerance	**VIF**
Gender	.976	1.024
Years of Work	.975	1.026
Annual Nominal Interest Rate	.978	1.022
Tenor of Loan	.995	1.005
Loan Type	.992	1.008

a. Dependent Variable: Probability of Default

Table 2 presents the results of the multicollinearity test for the predictor variables in the logistic regression model, showing that all variables have high tolerance values (close to 1) and low VIF values (well below 10), indicating minimal multicollinearity. Specifically, Gender has a tolerance of 0.976 and a VIF of 1.024, Years of Work has a tolerance of 0.975 and a VIF of 1.026, Annual Nominal Interest Rate has a tolerance of 0.978 and a VIF of 1.022, Tenor of Loan has a tolerance of 0.995 and a VIF of 1.005, and Loan Type has a tolerance of 0.992 and a VIF of 1.008. These results confirm that the predictor variables are sufficiently independent, ensuring the logistic regression model will produce stable and reliable estimates without being affected by multicollinearity.

Table 3: Test for Outliers

	Min	Max	Mean	Std. Deviation	N	Chi Square Critical Value at 0.001
Mahal. Distance	.681	44.099	4.990	6.017	500	20.46

Table 3 provides the results of the test for outliers using Mahalanobis distance in the logistic regression model. The Mahalanobis distances range from 0.681 to 44.099, with a mean distance of 4.990 and a standard deviation of 6.017. Given a critical chi-square value of 20.46 at $\alpha = 0.001$, any observation with a Mahalanobis distance exceeding this value is considered an outlier. With the maximum distance of 44.099 surpassing the critical value, it indicates the presence of outliers in the dataset. Approximately 14 outliers were identified and subsequently removed from the dataset to ensure that the assumptions of logistic regression, particularly regarding the influence of outliers, are upheld. This step is crucial for maintaining the model's accuracy and reliability in predicting the probability of default effectively.

Table 4 presents the logistics regression analysis for modelling the probability of default. The omnibus tests evaluate the overall significance of the logistic regression model for predicting default probabilities. The Chi-square statistic of 129.667 with 4 degrees of freedom yields a highly significant p-value of .000, indicating that the model as a whole effectively predicts the likelihood of default based on the selected predictor variables. This robust statistical significance assures that the model's variables collectively contribute to explaining variation in default outcomes, supporting its reliability in practical applications. The model summary provides key metrics to assess the logistic regression model's goodness of fit. The -2 Log likelihood of 537.605 serves as a measure of model fit, where a lower value indicates a better fit to the data. The Cox & Snell R Square value of 0.234 suggests that approximately 23.4% of the variance in the probability of default is explained by the independent variables included in the model. Meanwhile, the Nagelkerke R Square, adjusted for the model's complexity, slightly improves to 0.314, indicating that these variables collectively explain about 31.4% of the variance. These metrics offer insights into the model's explanatory power and its adequacy in capturing the relationship between predictors and default probability.

The classification table summarizes the logistic regression model's predictive performance in classifying default and non-default cases. It compares the model's predictions with the actual observed outcomes and calculates the percentage of correct predictions overall. The model achieves an overall percentage correct of 68.9%, indicating moderate accuracy in predicting default probabilities. This table is crucial for assessing the model's predictive ability and understanding its strengths and limitations in practical applications, such as credit risk assessment and management. The logistic regression coefficients, standard errors, Wald statistics, p-values, and odds ratios for each predictor variable included in the model. These coefficients quantify the direction and magnitude of each variable's effect on the log-odds of default. For instance, variables like Tenor of Loan show significant effects on default probability, with coefficients and odds ratios providing insights into how changes in these predictors influence the likelihood of default.

Table 4: Logistics Regression Result for Estimating Probability of Default

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	129.667	4	.000
	Block	129.667	4	.000
	Model	129.667	4	.000

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	537.605[a]	.234	.314

a. Estimation terminated at iteration number 4 because parameter estimates changed by less than .001.

Classification Table[a]

	Observed		Predicted: Probability of Default: No default	Predicted: Probability of Default: Default	Predicted: Percentage Correct
Step 1	Probability of Default	No default	188	83	69.4
		Default	68	147	68.4
	Overall Percentage				68.9

a. The cut value is .500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	Gen	.181	.235	.594	1	.441	1.199
	Years_Work	-.007	.104	.005	1	.943	.993
	Interest	-.047	.026	3.205	1	.073	.955
	Tenor	.377	.038	96.969	1	.000	1.457
	Constant	-1.894	1.229	2.375	1	.123	.151

a. Variable(s) entered on step 1: Gen, Years_Work, Interest, Tenor.

Modelling the Probability of Default Using Neural Network

The neural network model's performance in estimating the probability of default is summarized in Table 5. For the training sample, the cross-entropy error is 182.385 with a 27.0% rate of incorrect predictions. The model's training process was halted after one consecutive step with no decrease in error, indicating convergence. The training time was brief, lasting only 0:00:00.07. For the testing sample, the cross-entropy error is 72.875 with a 24.6% rate of incorrect predictions, showcasing the model's generalizability and accuracy in predicting default probabilities on unseen data. The table also provides the classification results for both training and testing samples, indicating the model's accuracy in predicting default and non-default cases. In the training sample, the model correctly classified 66.3% of non-defaults and 81.5% of defaults, resulting in an overall accuracy of 73.0%.

Table 5: Neural Network Result for Estimating Probability of Default

Model Summary		
Training	Cross Entropy Error	182.385
	Percent Incorrect Predictions	27.0%
	Stopping Rule Used	1 consecutive step(s) with no decrease in error[a]
	Training Time	0:00:00.07
Testing	Cross Entropy Error	72.875
	Percent Incorrect Predictions	24.6%
Dependent Variable: Probability of Default		
a. Error computations are based on the testing sample.		

Classification				
		Predicted		
Sample	Observed	No default	Default	Percent Correct
Training	No default	128	65	66.3%
	Default	28	123	81.5%
	Overall Percent	45.3%	54.7%	73.0%
Testing	No default	57	21	73.1%
	Default	14	50	78.1%
	Overall Percent	50.0%	50.0%	75.4%
Dependent Variable: Probability of Default				

Independent Variable Importance		
	Importance	Normalized Importance
Gender	.106	15.5%
Years of Work	.089	12.9%
Annual Nominal Interest Rate	.119	17.3%
Tenor of Loan	.686	100.0%

In the testing sample, the model's accuracy improved slightly, correctly classifying 73.1% of non-defaults and 78.1% of defaults, with an overall accuracy of 75.4%. These results highlight the neural network model's effectiveness in distinguishing between default and non-default cases, making it a valuable tool for credit risk assessment. The importance of each predictor variable in the neural network model is detailed in Table 5, with Tenor of Loan emerging as the most influential variable, having a normalized importance of 100.0%. Annual Nominal Interest Rate follows with a normalized importance of 17.3%, while

Gender and Years of Work have lower normalized importances of 15.5% and 12.9%, respectively. These values indicate that Tenor of Loan significantly impacts the probability of default, while other variables contribute to a lesser extent. Understanding the importance of these variables can guide risk management strategies and inform the development of targeted interventions to mitigate default risk.

Table 6: Comparison of Model Performance Criteria

	Logistic Regression	Neural Network
Accuracy	68.9%	75.4%
Precision	73.4%	80.3%
Recall	69.4%	73.1%
F-measure	71.3%	76.5%

Table 6 compares the performance of the logistic regression and neural network models in predicting the probability of default. The neural network model achieves a higher accuracy rate of 75.4% compared to the logistic regression model's 68.9%. This indicates that the neural network model is more effective overall in correctly classifying default and non-default cases, showcasing its superior predictive power. Precision measures the proportion of true positive predictions among all positive predictions made by the model. The neural network model has a precision of 80.3%, outperforming the logistic regression model's precision of 73.4%. This higher precision suggests that the neural network model is better at minimizing false positives, thereby making more accurate positive predictions regarding defaults.

Recall, or sensitivity, indicates the proportion of actual positives correctly identified by the model. The logistic regression model has a recall of 69.4%, while the neural network model has a slightly lower recall of 73.1%. Despite this difference, both models demonstrate a strong ability to identify true default cases, with the neural network model having a slight edge. The F-measure, or F1 score, is the harmonic mean of precision and recall, providing a balanced metric that considers both false positives and false negatives. The neural network model achieves an F-measure of 76.5%, compared to the logistic regression model's 71.3%. This higher F-measure reflects the neural network model's overall better performance in balancing precision and recall, making it a more robust tool for predicting default probability.

Overall, the neural network model demonstrates superior performance across all evaluated criteria: accuracy, precision, recall, and F-measure. These results suggest that the neural network model is more effective and reliable for predicting the probability of default, offering improved predictive accuracy and a better balance between precision and recall compared to the logistic regression model.

Forecasting the Probability of Default

Table 7: Prediction of Probability of Default Using Neural Network

S/N	Gender	Years of Work	Annual Nominal Interest Rate	Tenor of Loan	Loan Type	Probability of Default
1	Male	Above 14 years	42	9	New Loan	Default
2	Male	5-9 years	46.2	12	New Loan	Default
3	Male	10-14 years	42	6	New Loan	No default
4	Male	Less than 5 years	42	12	New Loan	Default
5	Male	5-9 years	48	9	New Loan	Default
6	Female	Above 14 years	48	12	New Loan	Default
7	Male	Above 14 years	42	12	New Loan	Default
8	Female	10-14 years	48	4	New Loan	No default
9	Male	5-9 years	48	12	New Loan	Default
10	Male	Above 14 years	48	12	New Loan	Default
11	Male	10-14 years	48	12	New Loan	Default
12	Female	Less than 5 years	48	12	New Loan	Default
13	Male	10-14 years	48	8	New Loan	No default
14	Male	Less than 5 years	48	6	New Loan	No default
15	Male	10-14 years	42	6	New Loan	No default
16	Female	5-9 years	42	9	New Loan	Default
17	Male	5-9 years	42	12	New Loan	Default
18	Female	5-9 years	42	3	New Loan	No default
19	Female	Less than 5 years	42	12	New Loan	Default
20	Female	10-14 years	42	12	New Loan	Default
21	Male	Above 14 years	42	12	New Loan	Default
22	Male	Less than 5 years	48	6	New Loan	No default
23	Male	Less than 5 years	48	12	New Loan	Default
24	Male	Above 14 years	48	3	New Loan	No default
25	Male	Above 14 years	42	9	New Loan	Default

Table 7 presents the predicted probabilities of default for 25 credit clients using a neural network model. These predictions are based on various socio-demographic and loan-specific variables, including gender, years of work, annual nominal interest rate, tenor of loan, and loan type. The table provides a detailed look into how these variables interplay to influence the likelihood of default. The sample includes both male and female clients, with predictions of default and no default distributed across both genders. This indicates that the neural network model does not exhibit a strong gender bias in its predictions. Males are slightly more prevalent in the sample, and the default outcomes for males and females are mixed. This suggests that gender, in isolation, is not a definitive predictor of loan default within the scope of this model, aligning with the model's focus on more financially deterministic factors. Clients' years of work experience range from less than 5 years to above 14 years. The model predicts defaults across all experience levels, though it appears that those with fewer years of work experience have a higher frequency of default predictions. For example, clients with less than 5 years of work experience and 5-9 years of experience often show up in the default predictions, suggesting that less work experience may correlate with higher default risk. However, defaults are also predicted for clients with extensive work experience, indicating that other factors also play a critical role.

Interest rates among the clients range from 42% to 48%. A notable observation is that higher interest rates, particularly at 48%, are frequently associated with default predictions. This pattern underscores the financial burden that higher borrowing costs can place on clients, increasing their likelihood of default. The model’s sensitivity to interest rates reflects the significant impact that borrowing costs have on a client's financial stability and ability to meet repayment obligations. Loan tenors in the table vary from 3 months to 12 months, with longer tenors (9 to 12 months) often corresponding to default predictions. This suggests that the duration of the loan significantly influences the likelihood of default, as longer loan tenors may lead to extended financial commitments that clients might struggle to maintain over time. For instance, several defaults are predicted for clients with a 12-month loan tenor, indicating that such extended durations increase the risk of financial strain and subsequent default.

All loans analyzed are new loans. This consistency eliminates variability that might arise from different loan types, allowing a focused analysis on how other variables influence default predictions. By standardizing the loan type, the model can more accurately assess the impact of demographic and financial variables on default risk. The neural network model predicts a combination of default and no default outcomes across the sample. Specific combinations of characteristics, such as higher interest rates and longer tenors, are more likely to result in default predictions. For instance, a male client with an above 14 years work experience, an interest rate of 48%, and a 12-month loan tenor is predicted to default. This aligns with financial risk principles, where higher costs and extended obligations increase the likelihood of default.

The results in the table demonstrates the neural network model's capability to predict default probabilities based on a multifaceted analysis of socio-demographic and loan-related factors. The model highlights the significant impact of higher interest rates and longer loan tenors on default risk, providing valuable insights for risk management in credit lending. By identifying high-risk profiles, lenders can implement targeted strategies to mitigate default risks, such as adjusting interest rates, modifying loan terms, or offering additional support to clients with riskier profiles. This detailed analysis underscores the importance of a comprehensive approach in credit risk assessment, integrating multiple factors to enhance predictive accuracy and financial stability.

In summary, the neural network model outperformed the logistic regression model in predicting the probability of default. It achieved higher performance metrics across the board: accuracy (75.4% vs. 68.9%), precision (80.3% vs. 73.4%), recall (73.1% vs. 69.4%), and F-measure (76.5% vs. 71.3%). The most significant factor influencing the probability of default, as indicated by the neural network model, is the Tenor of Loan, with a normalized importance of 100.0%. Longer loan tenors substantially increase default risk. The neural network model demonstrated effective prediction capabilities, correctly identifying a significant proportion of default and non-default cases in the test data. For example, the model predicted defaults for clients with higher interest rates and longer loan tenors, which aligns with established financial risk principles.

5. Conclusion

This study investigated the most appropriate model for predicting the probability of default, identified the key factors influencing default risk, and forecasted future probabilities of default using the identified model. Through a comparative analysis of logistic regression and neural network models, we determined that the neural network model is superior in predicting the probability of default. It achieves higher accuracy, precision, recall, and F-measure, making it a more reliable tool for this purpose. The analysis revealed that the primary factors influencing the probability of default, as suggested by the neural network model, are the Tenor of Loan and the Annual Nominal Interest Rate. Longer loan durations and higher interest rates significantly increase the likelihood of default, reflecting the financial strain imposed on borrowers over extended periods and by higher borrowing costs. Demographic factors such as Gender and Years of Work also play a role but are less influential compared to loan-specific variables.

Using the neural network model, we can forecast future probabilities of default by applying it to new data and monitoring trends in relevant variables. By identifying high-risk profiles, lenders can tailor their risk management strategies, such as adjusting loan terms or providing additional support to vulnerable borrower segments. This proactive approach helps mitigate default risks and enhances the overall stability of the lending institution. In conclusion, the neural network model's superior performance and ability to capture complex relationships between variables make it the most appropriate tool for modeling the probability of default. The insights gained from this model can guide lenders in making informed decisions, ultimately leading to more effective credit risk management and improved financial outcomes. We recommends considering adjustments to terms to mitigate risk, adjusting risk strategies and implementing the neutral network model for ongoing prediction of default probabilities.

References

Al-abedallat, A.Z. (2016). Factors affecting credit risk: an empirical study of the Jordanian commercial banks. *European Scientific Journal December*, 12(34), 307-321.

Aniceto, M.C., Barboza, F. & Kimura, H. (2020). Machine learning predictivity applied to consumer creditworthiness. *Future Business Journal*, 6(1), 1-14.

Bailey, M., Cao, R., Kuchler, T., Stroebel, J. and Wong, A. (2018). Social connectedness: Measurement, determinants, and effects. *Journal of Economic Perspectives*, 32 (3), 259–280.

Basel Committee on Banking Supervision (2004) Bank failures in mature economies. *Working 13, Basel Committee on Banking Supervision, Basel.*

Calli, B. A. (2019). A psychometric and financial factors based framework suggestion for an integrated credit risk assessment information system (Sakarya University). Sakarya University. https://acikerisim.sakarya.edu.tr/handle/20.500.12619/68672

Calli, B.A. & Coşkun, E. (2021) A longitudinal systematic review of credit risk assessment and credit default predictors. *SAGE Open*, pp. 1–19.

Chandran, B. (2022). Credit Risk of Vehicle Loans: A Machine Learning based Prediction. Retrieved from Credit Risk of Vehicle Loans : A Machine Learning based Prediction | by Bharath Chandran | The Startup | Medium on 29th June, 2024.

Damrongsakmethee T. & Neagoe, V. (2019) Decision tree enhanced with AdaBoost versus multilayer perceptron for credit scoring modeling. In: Silhavy R, Silhavy P, Prokopova Z (eds) Computational statistics and mathematical modeling methods in intelligent systems. CoMeSySo 2019. Advances in intelligent systems and computing, 1047. Springer, Cham, 216–226.

Dawood, E.A.E., Elfakhrany, E. & Maghraby, F.A. (2019). Improve profiling bank customer's behavior using machine learning. *IEEE Access*, 7, 109320- 109327.

Freedman, S. & Jin, G.Z. (2017) The information value of online social networks: lessons from peer-to-peer lending. *International Journal of Industrial Organization*, 51, 185–222.

Gieseche, K. (2004). *Credit risk modelling and valuation: an introduction*. Credit Risk: Models and Management, Cornell University, London.

Kiveu, C.W (2015) Appraisal of credit applicant using logistic and linear discriminant models with principal component analysis. *Unpublished dissertation submitted school of mathematics at the University of Nairobi.*

Konovalova, N., Kristovska, I. and Kudinska, M. (2016) Credit risk management in commercial banks. Polish Journal of Management Studies, 13(2), 90-100.

Luo, C. (2020). A comprehensive decision support approach for credit scoring. *Ind Manag Data Syst*, 120 (2), 280–290.

Luy, D. D. (2010) Evaluation of credit risk management policies and practices in a Vietnamese joint-stock commercial bank's transaction office. *Business Economics and Tourism*, pp. 5-16

Matoussi, H (2010). Credit–risk evaluation of a Tunisian commercial bank: logistic regression vs neural network modeling. *Accounting and Management Information Systems*, 9 (1), 92 –119.

Nazre, I.T., Nasimul, H. & Md.Umar, F. (2021). *A concern of predicting credit recovery on supervised machine learning approaches*. 12th ICCCNT 2021, July 6-8, 2021 - IIT – Kharagpur.

Osayameh, R.O. (1991) Lending and credit administration model for commercial banks. *Nigerian Financial Review*, 4 (2), 17-34.

Perera, H.A.P.L. & Premaratne, S.C. (2016) An Artificial Neural Network Approach for the Predictive Accuracy of Payments of Leasing Customers in Sri Lanka. *Full Paper Proceeding BESSH-2016*, 285 (2), 1-11

Sudhamathy G. (2016). Credit risk analysis and prediction modelling of bank loans using R. *International Journal of Engineering and Technology* (IJET), 8(5),1954-1966.

Sudi, M.P. & Maniagi M. (2020). Effect of credit risk management practices on financial performance of micro-finance institutions in Nairobi. *International Journal of Recent Research in Social Sciences and Humanities*, *7*(3), 22-39.

Supriya, P., Pavani, M., Saisushma, N., Kumari, V.N. & Vikas, K. (2019) Loan prediction by using machine learning models. *International Journal of Engineering and Techniques*, 5(2), 114-148.

Taiwo, J. N., Ucheaga, E.G., Achugamonu, B.U., Adetiloye, K., Okoye, O. & Agwu, E. (2017) Credit risk management: implications on bank performance and lending growth. *Saudi Journal of Business and Management Studies*, 2(5), 584-590.

Tudor, L.A., Bara, A. & Opera, V.S. (2017). Comparative analysis of data mining methods for predicting credit default probabilities in a retail bank portfolio. *Latest Trends in Information Technology*, 117-122.

Uzair, A., Hafiz I.T.Z., Asim, S. & Nowshath, K.B. (2019) An empirical study on loan default prediction models. *Computational and Theoretical Nanoscience*, 16, 3483–3488.

Zakrzewska, D. (2007). On integrating unsupervised and supervised classification for credit risk evaluation. *Information Technology and Control*, 36(1), 99-102.

Zoran, E. (2019). Predicting default loans using machine learning (OptiML). 27th Telecommunications forum TELFOR 2019, Serbia, Belgrade, November 26-27, 2019. 1-4.