doi.org/10.54473/IJTRET.2025.9606
International Journal of Trendy Research in Engineering and Technology
Volume 9 Issue 6 December 2025
ISSN No. 2582-0958

BANK CUSTOMER CHURN PREDICTION USING MULTILAYER PERCEPTRON

O.C. Nor*, P. O. Odion
Department of Computer Science, Nigerian Defence Academy, Kaduna
*Correspondence E-mail: nororseercollins@gmail.com

Received 28 October 2025 Received in revised form 02 November 2025 Accepted 05 November 2025
Available Online 08 November 2025

ABSTRACT

Customer churn is a significant challenge for banks, as acquiring new customers is far more costly than retaining existing ones. Retaining existing customers is essential for banks to ensure profitability and growth. This study focuses on analyzing customer churn within the banking sector by developing an effective churn prediction model using Multilayer Perceptron. The model aims to provide banks with insights to reduce churn, improve customer retention, and optimize retention strategies. The dataset used includes 51,784 customer records (of a leading commercial bank in Nigeria). Data preprocessing involved using Python libraries for label encoding and feature scaling. Performance evaluation was carried out using a confusion matrix and the Area Under the Receiver Operating Characteristic Curve. The results showed exceptional performance, with the model achieving 100% accuracy, precision, F-score, and recall. Both the training and test data ROC curves reached a value of 1.0, confirming optimal model performance. The study concludes that MLP are highly effective for building a reliable customer churn prediction model with impressive accuracy. This model can assist banks in identifying at-risk customers. By leveraging such predictive insights, banks can improve decision-making, reduce churn, ultimately boost customer retention and optimizing financial returns.

Keywords: Bank, Customer churn, Customer retention, Multilayer Perceptron, Prediction

I.INTRODUCTION

According to Elena [1], Customer churn is a tendency of customers to cancel their subscriptions to a service they have been using and, hence Customer churn occurs when a customer discontinues using that service. The percentage of customers who leave within a specified time frame is known as the customer churn rate. The customer growth rate, which tracks new clients, is the reverse to this. In addition to regular churn, which is commonly occurring in all businesses, there are other indicators that show that something has gone wrong in the firm and has to be corrected.

The factors for Customer churn include inadequate or poor-quality customer support, unfavorable customer experiences, transitioning to a rival offering superior terms or pricing tactics, client preferences altered, long-term clients feel unsatisfied, the service failed to fulfill customers' anticipations, financial concerns, safeguarding against fraud in customer payments. In instances of customer churn due to inadequate service, the firm's Reputation can be severely harmed due to unfavorable feedback from dissatisfied former customers on social platforms or evaluation sites.

The volume of data has been increasing quickly due to technological advancements, making it the 21st century's oil. However, oil is only valuable until it is processed into fuel. Data mining is the process of uncovering relevant and significant information that is concealed in data using a variety of novel approaches and strategies. Support Vector Machine (SVM), Linear Regression, Genetic Algorithm, Decision Tree, and Neural Network were among the strategies and tactics [2]. Customer churn, which is the loss of customers due to them switching to competitors, is a major issue for businesses. An organization can gain significant knowledge into retaining and growing its client base by anticipating consumer churning behavior. Over the past few decades, a variety of customer churn prediction models have been created. Modern machine learning classifiers like logistic regression, random forest, and linear regression are used in most sophisticated models [3]. The ability of the business to anticipate likely churners and react to them promptly is crucial for retaining existing clients. Organizations should identify possible churn indicators, meet customer demands, and rebuild and re-establish customer loyalty in order to reduce the expenses associated with acquiring new clients [4].

Almost everyone in a developed economy needs the services of financial institutions on a regular basis, if not more frequently. Financial organizations need to acquire new clients while retaining existing ones since customer retention is closely correlated with profit margins. According to Business Week, organizations can see a 140% gain in earnings by reducing the customer defection rate by 5%, while Harvard Business Review estimates a 25% to 85% boost in profitability. Selective or customized marketing techniques are currently taking the place of traditional mass marketing strategies thanks to customer relationship management (CRM) technologies [5]. These targeted marketing strategies entail determining a portion of current clients who are most likely to discontinue utilizing the business's goods or services (churn). The past few years have seen a rise in the study of consumer marketing and management research on churn prediction, as it is anticipated that existing customers' churn will lead to business losses and a corresponding decrease in profit.

In the business and marketing sector, it is commonly acknowledged that the cost of recruiting new clients is roughly six to seven times higher than the cost of keeping existing ones. The importance of client retention tactics as a way to maximize resource allocation and raise overall profitability is highlighted by this industry insight. Reducing customer attrition has a significant impact on financial institutions' capacity to increase both their core competitiveness and profits. Therefore, to take timely action to retain consumers and prevent additional clients from departing, financial institutions need to rapidly improve their ability to predict customer turnover [6]. Developing a more precise and effective customer churn prediction model is essential for efficient customer churn control. Statistical and data mining techniques have been utilized to construct churn prediction models. The data mining techniques can be used to discover interesting patterns or relationships in the data and predict or classify the behavior by fitting a model based on available data. In other words, it is an interdisciplinary area with a general objective of predicting outcomes and employing sophisticated data processing algorithms to discover mainly hidden patterns, associations, anomalies, and/or structure from extensive data stored in data warehouses or other information repositories [7].

Artificial Neural Network also called Neural Network (NN) is a complex network that comprises a large set of simple nodes known as neural cells. Artificial Neural Networks (ANN) are algorithms based on brain function and are used to model complicated patterns and forecast issues. The Artificial Neural Network (ANN) is a deep learning method that arose from the concept of the human brain Biological Neural Networks. Artificial Neural Network was proposed based on advanced biology research concerning human brain tissue and neural system and can be used to simulate neural activities of information processing in the human brain [8]. Artificial Neural Network have been used in finance for portfolio management, credit rating and predicting bankruptcy, forecasting exchange rates, predicting stock values, inflation and cash forecasting.

II.AIM AND OBJECTIVES OF THE WORK

The primary aim of this study is to predict accurate customer churn in the banking sector using multilayer perceptron. The specific objectives are:

a) To prepare relevant dataset with categorization and prediction model-related features.
b) To develop a prediction model for customer turnover using Multilayer Perceptron Artificial Neural Network.
c) To evaluate the accuracy, precision, recall, f-score and AUROC of the proposed customer churn prediction model using performance measures.

III.REVIEWS OF RELATED LITERATURES

In the global context, various works have been done in the area of customer churn prediction for numerous industries. Due to its importance in numerous industries, churn prediction has been one of the key topics of interest for data mining researchers. In order to predict churning, customers as well as the causes of churn, numerous approaches had been made. Because those details can help businesses in lowering customer churn.

Customer Churn analytics and management services help organizations identify key churn drivers to develop churn prevention strategies to improve customer retention, loyalty and Life time value [9]. According to [9], identifying these key (churn) drivers helps organizations design churn strategies to enhance customer loyalty by developing a customer focused churn management, retention and revival strategy.

Market Equations -A Research & Analytics Outsourcing Company [9] analyzed churn on a large automobile company in United States. From the They have come up with a neural network-based approach for churn prediction. The expectation of the analysis was to identify the customer segments with diminished level of satisfaction with the company’s product. They have come up with three main steps in detecting churn.

Churn Trend Analysis: Aim was to understand trends in customer data, churn Profiling: Was to identify segments which are more prone to defect, churn Scoring and Segmentation: Assign each segment a churn score and with the use of a predictive model, find the most prone defectors.

With the analysis of data, Market equations tried to identify the factors for churn and reasons for a customer to leave or stay with a particular service provider and to find any demographic patterns or trends in customers.

Sandeepkumar & Mundada [10], have considered multilayered neural network which is also known as Deep Feed Forward Neural Network (DFNN) to perform predictive analytics on customer attrition in the banking sector. The dataset was collected from the UCI machine learning archive which has total 10,000 customer data with 14 dimensions of features. The model use optimized one hot encoding and Tukey outliers' algorithms to perform data cleaning and preprocessing. The model was compared with the machine learning algorithms such as Logistic regression, Decision tree and Gaussian Naïve Bayes algorithm. The results show that Enhanced Deep Feed Forward Neural Network (DFNN) model performs best in accuracy when compared with other machine learning algorithms.

Tariq et al, [11], Lalwani et al [12] and Garimella et al [13] have used the same dataset to predict the churn of telecom customers. Tariq et al, [11] used CNN, and Garimella et al [13] used Deep CNN. In contrast, Lalwani et al [12] used several machine learning techniques such as AdaBoost, Support Vector Machine (SVM), Logistic Regression (LR), Naïve Bayes (NB), Random Forest (RF), Decision Trees (DT), and XGBoost. These researchers used Confusion Matrix outputs accuracy, Area Under the Curve (AUC), and loss margin. Garimella et al [13] used the maximum dice coefficient and the Jaccard coefficient as performance indicators along with its accuracy. Tariq et al, [11] and Garimella et al [13] did not compare their models with any other model; they only presented the performance results. On the other hand, Lalwani et al [12], compared all the machine learning techniques and found AdaBoost had the best AUC and accuracy.

Cenggoro et al [14] did not compare the suggested model with any other model and instead used a Vector Embedding Model for loss estimation for a telecom dataset of 3333 consumers. They only provided the suggested model's accuracy and F1-score, which shows that the model does a good job of differentiating between churning and non-churning clients. Various machine learning models, including logistic regression (LR), decision tree (DT), k-nearest neighbor (KNN), random forest (RF), were used in this study to estimate the likelihood that a client will leave. Performance measures including memory, accuracy, and others are compared [15].

Rosa [16] put forth a brand-new approach for measuring and forecasting client attrition in the banking sector that made use of artificial neural networks. From January 2017 to December 2017, data on 1588 customers were retrieved using SAS Base from the bank's data warehouse. The research ignored other machine learning techniques like Decision Trees, Logistic Regression, or Support Vector Machines in favor of constructing neural networks alone.

Elyusufi and AitKbir[17] have analyzed several specific machine learning models that had been put forth in the literature to address this issue and contrasted them with some recently developed models that are based on ensemble learning techniques. As a result, they developed predictive churn strategies that analyzed customer history data, determined who was active after a specific period, and then developed models that pinpointed the points at which a client might stop using a certain firm service. When completing the training step with traditional models like multi-layer perception neural networks, ensemble learning methods were also employed to locate pertinent features to reduce their quantity. The proposed methodologies could obtain accuracy levels of up to 89% when other research studies using the same dataset only manage to reach 86%. According to this study, the SVM model has an accuracy rating of 86.18%, which is rather good. A high accuracy for a particular model would mean that the model can predict the choice that a client can make (leave the bank/remain with the bank).

Muneer et al [18] have created a strategy for predicting client attrition using the three intelligence models random forest, adaboost, and support vector machine (SVM). The method yields the best results when the Synthetic Minority Oversampling Technique (SMOTE) is employed to address the unbalanced dataset and the combination of under sampling and oversampling. On SMOTED data, the method produced excellent results, with a 91.90% F1 score and an overall accuracy of 88.7% when using RF. Additionally, the results of the experiment showed that RF delivered successful results for all feature-selected datasets. SMOTE was used in this investigation to lessen the discrepancy in sample sizes. This increases the accuracy of the forecasted results as well.

Surviving in the fiercely competitive telecommunication industry while retaining loyal customers has necessitated the prediction of potential churn customers using predictive modeling techniques. Efficient predictive models play a vital role in identifying loyal customers.

Allocating resources for the retention of these customers can stem the outflow of dissatisfied consumers considering switching companies. Their study introduced an artificial neural network (ANN) approach to predict customers who are likely to switch to other operators. The model utilized various attributes, including demographic data, billing information, and usage patterns from telecom companies' datasets. Unlike other prediction methods, the results achieved using the ANN-based approach displayed an accuracy of 79% in predicting telecom churn in Pakistan. This accuracy underscores the ANN's effectiveness in pinpointing churn factors, enabling proactive measures to eliminate the root causes of churn [19].

Churn prediction based on Artificial Neural Networks (ANNs) have been successful, however, they are affected by the noise or outliers present in such datasets. The effect of such noise, and number of training samples on churn prediction was investigated. Two filters were applied to the data, the Genetic Algorithm (GA) and K-means filter. The filtered data were used to train an ANN model and tested with a 30% unfiltered data. The performance shows that the training performance improved when noise was filtered while the testing performance was affected by the unbalanced data caused by filtering [20].

Research by Nguyen et al [21] has investigated the impact of customer segmentation on customer churn prediction accuracy in the banking sector using various machine learning models. Their study employs k- means clustering for segmentation and models like k-nearest neighbors, logistic regression, decision tree, random forest, and support vector machine. Results indicate strong performance with the random forest model achieving 97% accuracy, and after segmentation, mean accuracy is notable. Logistic regression exhibits the lowest accuracy (87.27%), while random forest excels (97.25%). The study highlights that customer segmentation's impact varies based on the dataset and selected models, providing practical applications across industries.

Customer Churn (CC) is a major issue and important concern for large organizations and businesses alike. Telecom industries are attempting to improve methods to predict possible customer churn due to the immediate impact on revenue, particularly in the telecom sector [22]. Rajendran et al [22] have discussed the various ML algorithms used to construct the churn model that helps telecom operators to predict customers who are likely to churn. They compared their experimental results to predict the best model among various techniques. As a result, the use of the Random Forest combined with SMOTE-ENN outperforms best result than others in terms of F1-score. According to their analysis, the maximum prediction was 95 percent based on F1-score.

Amuda and Adeyemo [23] have created a predictive model based on artificial neural networks to foresee customer churn in financial organizations. Their model sought to do away with manual feature engineering, in contrast to more conventional classifiers like Logistic Regression and Decision Trees. They used Python to build a multi-layer perceptron model using data from a Nigerian financial organization that had fifty thousand customers. To prevent overfitting, they used Dropout and L2 regularization. Python's performance was equivalent when compared to Neuro Solution Infinity program. With ROC curve graphs at 0.89 and 0.85, respectively, the attained accuracy rates were 97.53% and 97.4% [23].

In a study by Adebiyi et al [24] on customer churn in the Nigerian mobile telecommunications sector, churn and retention rates were modeled and predicted using Markov chains. According to the findings, MTN had the highest retention rate (86.11%), followed by Etisalat (67.5%), GLO (70.51%), and Airtel (67%). An optimized weighted soft voting ensemble model was created for the suggested churn prediction system to identify potential churn consumers in the high-churn-rate telecom industry [25]. To improve input data quality and deal with imbalanced data challenges, the framework incorporates procedures like exploration data analysis, data preprocessing, feature engineering, and data sampling. In a real-world database examination, the ensemble weights were optimized using Powell's optimization algorithm, yielding an accuracy score of 84% and an F1 score of 83.42% [25]. Comparative analysis demonstrated superior prediction accuracy compared to existing customer churn prediction systems, encompassing both machine learning and deep learning models.

V.METHODOLOGY

To anticipate customer turnover and do away with the manual feature engineering procedure during the data pretreatment stage, this research created a predictive model employing the Multilayer Perceptron architecture; Predict potential churners and non-churners using the model. The study also aims to enhance customer retention strategies within the banking industry. By accurately predicting potential churners through the ANN model, the study aims to provide actionable insights that empower banks to implement proactive retention strategies, thereby reducing customer attrition.

4.1Method of Data Collection

Data from one of the commercial banks in Nigeria was used. Secondary method of data collection was used to collect the data from the bank's database. The dataset used contains fifty-one thousand seven hundred and eighty-four (51,784) bank customers' data with several attributes. Some of these attributes are shown in table 1.

Table 1: Attributes of the dataset collected

Features	**Description**
Gender	The gender of the customer either Male or Female
Age	The age of the customer
Marital_status	This stand for the marital status of the customer group in the following ways: Single, Married, Divorced
Branch-code	The branch code of the financial institution
Region	The region where the financial institution is situated
Branch	The branch of the financial institution
Closing_Balance	The amount remaining in the customer's account
Opening_Balance	The amount in the customer's account at the start of the account
Credit	The money deposited into a customer's account or an increase in their available Balance
Debit	The money withdrawn or spent by a customer from their account
Account_type	Whether the customer's account type is Saving, current etc
Customer_Status	Whether the customer is active or not

4.2Proposed System Architecture

The Multilayer Perceptron (MLP) Proposed Model is shown in Figure 1. Figure 1 shows the steps for the MLP model, which are:

1. Getting client data, including transaction history, out of a bank database.
2. Data pre-processing to remove noise and transform numerical data from categorical data.
3. Feature scaling and data normalization
4. Data separation (validation, test, and training sets)
5. To create output, start at the input layer by sending train data patterns through the network.
6. To reduce the error rate, compute the network output using the cost function.
7. Update the model by determining the derivative for each weight in the network.
8. To get the expected class label, compute the network output and apply the threshold function.
9. A model evaluation.

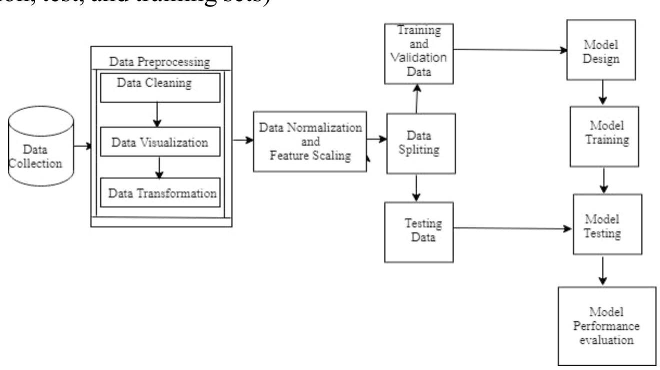


Fig.1.Proposed Model of Multilayer Perceptron

The proposed system layout is shown in Figure 1. A dataset listing the bank's clients' active and inactive statuses will be created using the customer data that has been entered into the proposed architecture. The proposed model was validated using the dataset of bank customers. There are several stages to the proposed architecture. Data visualization and cleansing are accomplished using preprocessing procedures. Data division: 25 percent of the entire dataset was utilized for testing and validating the model's performance, while 75 percent was used for training the model. The output for customer prediction is obtained by combining the train model with the test dataset after it was developed through classification using supervised machine learning, such as a multilayer perceptron. Finally, the customer prediction model determines whether the customer will churn.

4.3Data Preparation

In order for the raw data to be fed into the modeling algorithm and produce the model, it was converted into a final dataset during the data preparation step. Data cleaning, outlier removal, missing value removal, data transformation, and data exploration utilizing Python programming language libraries were among the many tasks completed during this period to save computation time and standardize the data.

Handling Missing Values

The dataset used contains fifty-one thousand, seven hundred and eighty-four (51,784) bank customers, missing values were like 5%, and these missing values were removed amounting to the data of two thousand, seven hundred and twenty-six (2726) bank customers.

Normalizing Data

Bringing values into line with a single scale is the aim of normalization. Every numerical column's skew and kurtosis were measured; if the values were outside the range of +/-2, the variable was considered to have skewed data. To normalize these attributes, the sklearn MinMaxScaler() method was used.

Feature Selection

The dataset is subjected to feature selection due to its vast number of dependent variables. Feature selection is used to identify relevant features for the model's creation. The Correlation Matrix with Heatmap technique was used to pick features for this investigation. A correlation was discovered between the independent and dependent variables as well as between the dependent variables themselves. The degree to which one variable is dependent upon another was quantified using correlation. If the correlation is greater than the 0.5 criterion, the variables will not be used because this might affect the accuracy of the model [26]. The degree of relationship with interpretation is displayed in the table below.

Table 2: Correlation Table [26]

Size of Correlation	Interpretation
.90 to 1.00 (-.90 to -1.00)	Very high positive (negative) correlation
.70 to .90 (-.70 to -.90)	High positive (negative) correlation
.50 to .70 (-.50 to -.70)	Moderate positive (negative) correlation
.30 to .50 (-.30 to -.50)	Low positive (negative) correlation
.00 to .30 (.00 to -.30)	negligible correlation

Feature selection was applied to the dataset. The feature selection was used to find relevant features for the model construction. Feature selection improves the accuracy of the model. It trains the model faster and reduced the complexity of the model.

Encoding

Both continuous and categorical variables are present in the dataset. Only numerical input is accepted by a small number of machine learning algorithms, such as Multilayer Perceptron (MLP) of artificial Neural Networks. For this reason, label encoding was used to transform the category data into numerical data input using label encoding.

Multilayer Perceptron

Finally, the Multilayer Perceptron was evaluated. Multilayer Perceptron is a nonlinear predictive model which is learned through training and the structure is like a biological neural network. Neuron acts as the basic building blocks of the network. The output depends on the activation function of the neuron. Here in this research relu activation function was used based on previous research. The relu (Rectified Linear Unit) activation function was computationally less expensive. It takes input and then each input is multiplied by a weight. Then all the weighted inputs were summed up together with a bias. Finally, the sum was passed through an activation function. The most common activation function used for Neural Network was 'Sigmoid' function. This function is useful for binary classification as it outputs in the range of 0 and 1 [27].

V. RESULTS AND DISCUSSION

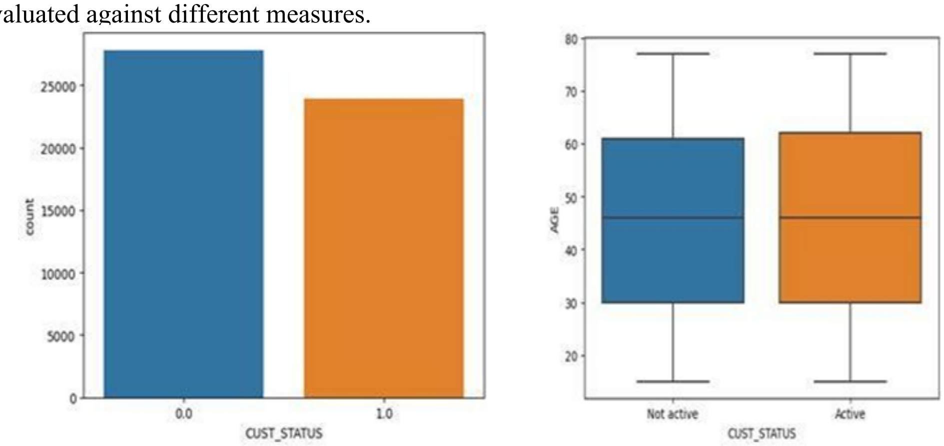
In this section, the results of the experiment for the churn prediction of customers is presented and discussed. This section also assessed the performance of the machine learning methodology of Multilayer Perceptron evaluated against different measures.

Exploratory Data Analysis

Exploratory data analysis (EDA) is used by data scientists to analyze and investigate data sets and summarize their main characteristics, often employing data visualization methods. It helps determine best to manipulate data sources to get the answers you need, making it easier for data scientists to discover patterns, spot anomalies, test a hypothesis, or check assumptions, this is done after the data preprocessing stage. Also, report and insight is gotten from the data after exploratory data analysis (EDA) stage [28].

The Customers Distribution

(a) Customer Status and age distribution of non-churner (Active) and churner (Not Active) customers in our dataset is quite balance which is shown in figure 2 (a), figure 2 (b) shows the distribution of non-churner (Active) and churner (Not Active) customers in our dataset, in the figure, the number of customer that are likely to leave (churn) the bank and the once that are likely to stay (Non-Churn) is the same.



(a) (b)

Figure-2 (a) Non-churner (Active) and churner (Not Active) Customer's Distribution (b) Non-churner (Active) and Churner (Not Active) Customer's Distribution Based on Age.

(b) Gender and Marital Status

In the dataset, both male and female customers have equal chances of churning as shown in figure 3(a), but in terms of loyalty, male customers are more loyal to the bank based on the dataset as indicated in figure 3(a). Figure 3(b), shows the total number of both male and female customers, in the figure, the percentage of male customers is 51.16% which is about twenty six thousand four hundred and eighty (26480) customers and the percentage of female customers is 48.84% which is about twenty five thousand two hundred and seventy nine (25279) customers, that shows that male customers are more than female customers in the bank.

International Journal of Trendy Research in Engineering and Technology
Volume 9 Issue 6 December 2025
ISSN No. 2582-0958

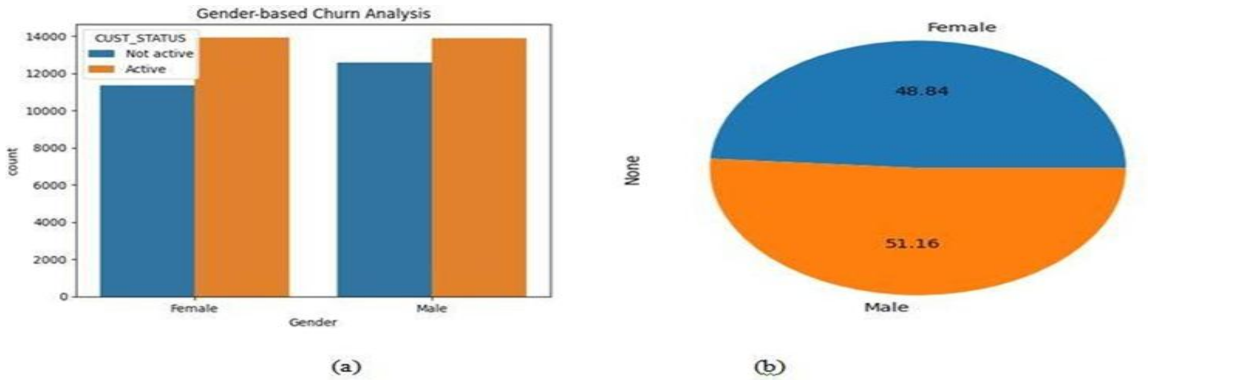


Figure 3.(a) Non-churner (Active) and Churner (Not Active) Customer's Distribution based on Gender. (b) Gender of the Customer's Distribution.

Correlation Analysis

Correlation analysis, also known as bivariate, is primarily concerned with finding out whether a relationship exists between variables and then determining the magnitude and action of that relationship [29].

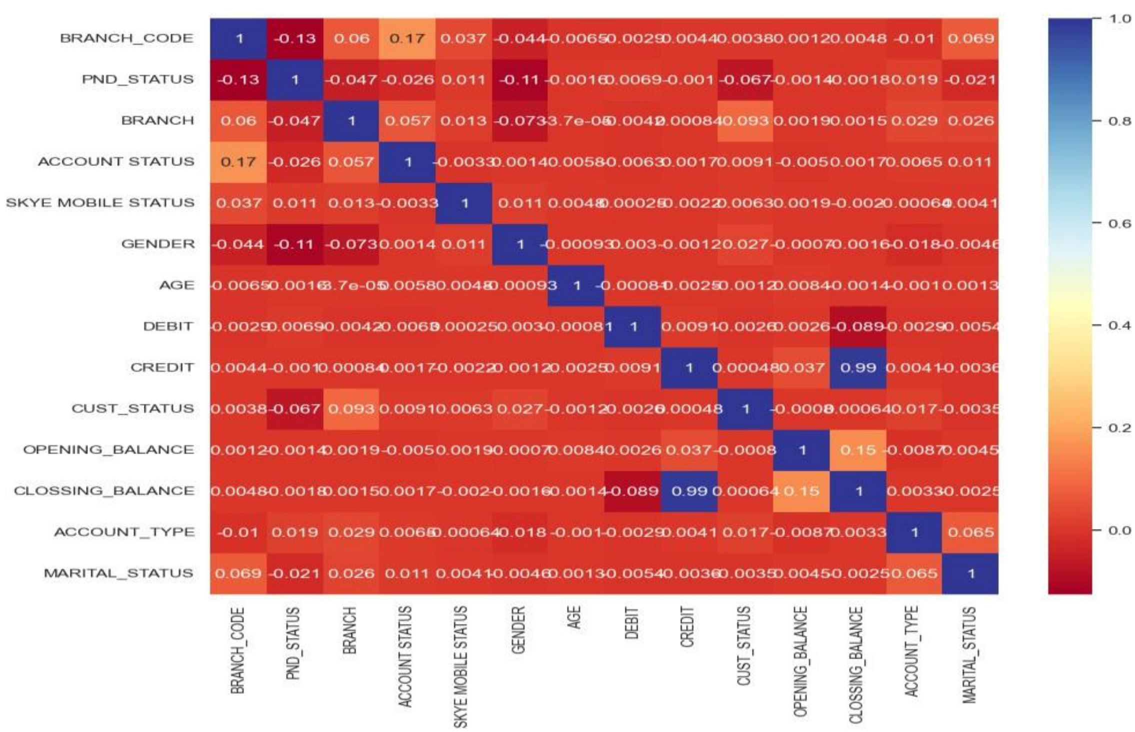


Figure 4: Correlation heatmap of the variables

Correlation Analysis of the data was done to analyze the correlation between the independent and dependent variable and to identify multicollinearity between the independent features. The 'Spearman' correlation method was used to identify correlation as the dissertation consists of both continuous and categorical variables. The correlation heatmap was generated as shown in the below figure and also the correlation matrix listed in Figure 4.

Implementation

The implementation was carried out in Python and the result is presented in Figure 5, Figure 6 and Figure 7. It can be seen that training loss was decreasing while training accuracy was increasing and the validation loss was decreasing while validation accuracy was increasing. These show that the model was neither overfitting nor underfitting.

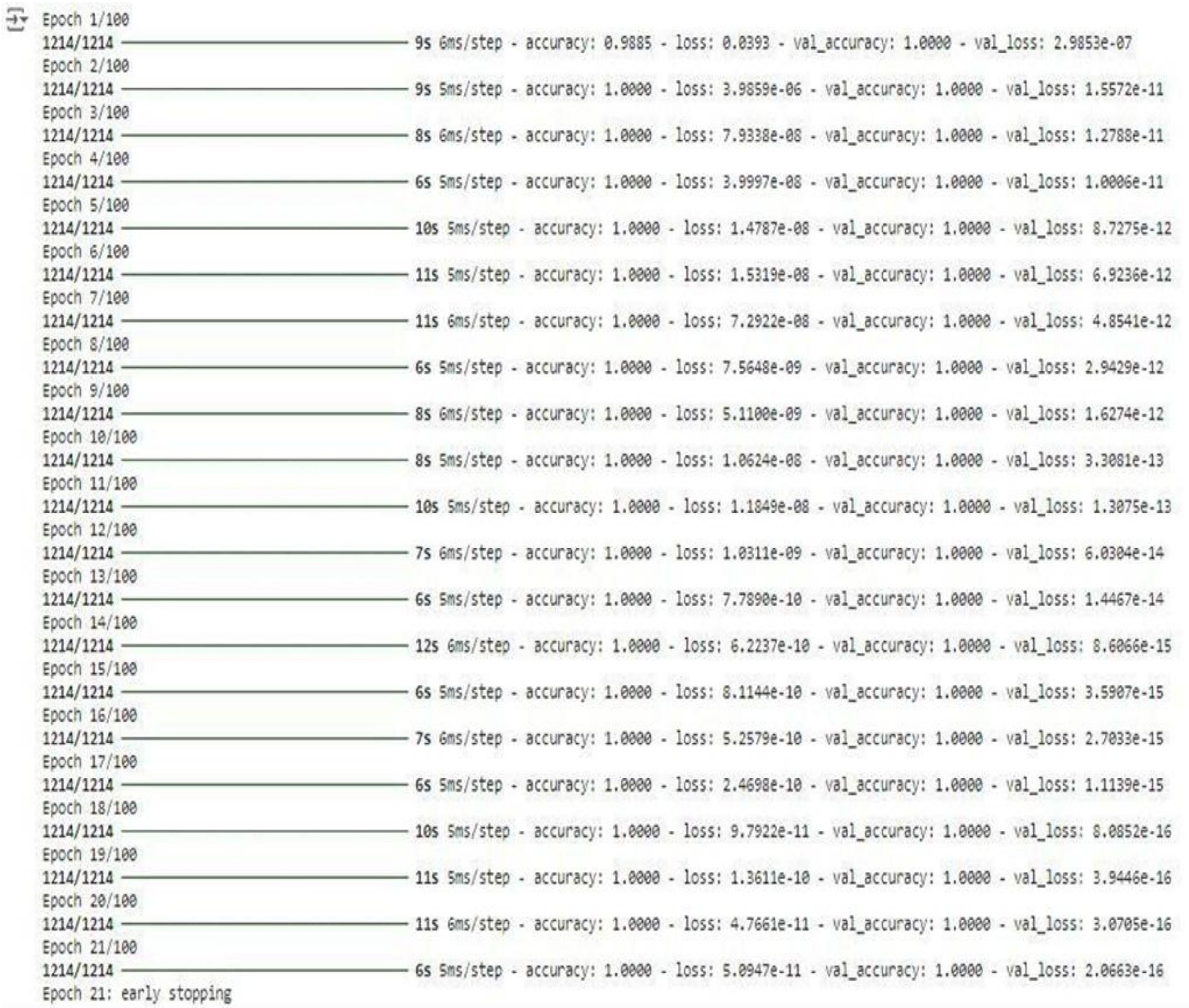


Figure 5: Learning Process

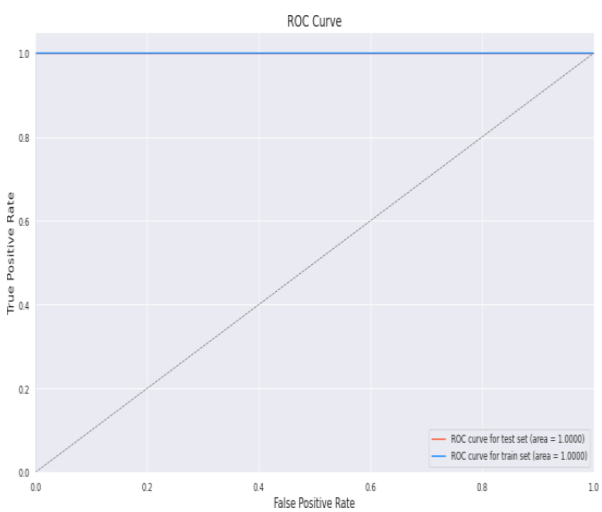


Figure 6: Training Accuracy and Validation Accuracy Graph

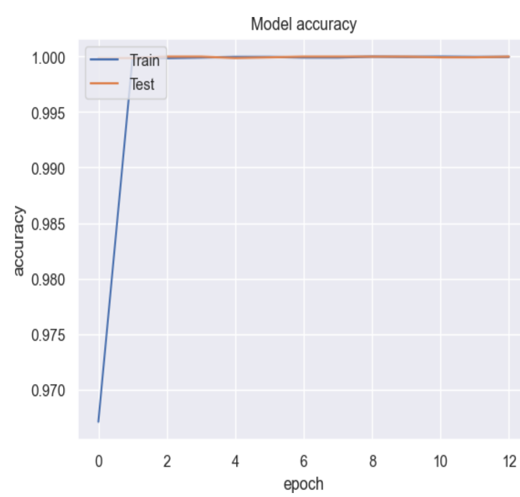


Figure 7: Training Loss and Validation Loss Graph

Model Performance Evaluation

Based on a set of test data for which the true values were known, the confusion matrix was used to describe the performance of the Multilayer Perceptron classifier. Using it, you can visualize the model's performance. The confusion matrix table summarized the model's predictions. By class Churn (Not Active) and Non Churn (Active), the numbers of correctly classified and incorrectly classified (misclassified) items were summarized with count values. Performance evaluation is about providing insight not only into the errors made by a model, but also into the types of errors.

The other metric AUC is a measurement of the entire area under the Receiver Operating Characteristic (ROC) curve, and it is one of the metrics used to evaluate model performance, together with the ROC curve [30]. The AUC value ranges from 0 to 1, with a value near 1 indicating a more accurate model. The distributions of TN and TP do not intersect when the area under the ROC curve is large, indicating that the classes have been successfully separated [31].

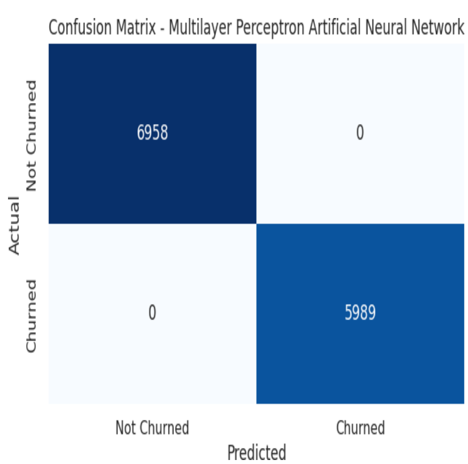


Figure 8: Confusion Matrix for the Model

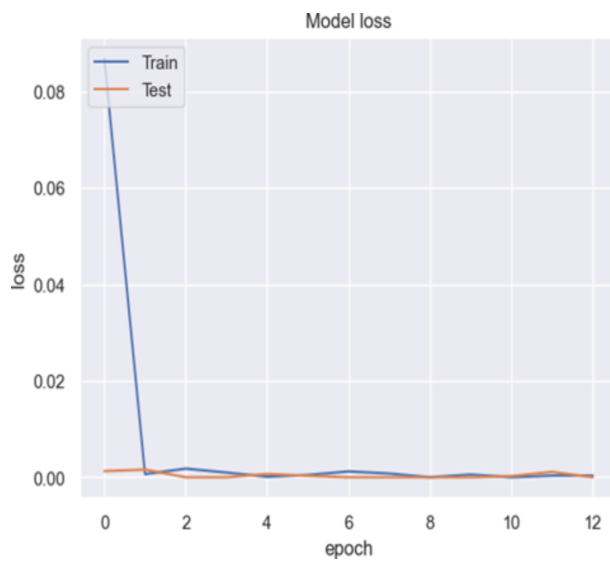


Figure 9: Area under the Receiver Operating Characteristic Curve (AUROC)

Figure 8 shows the confusion matrix of the Multilayer Perceptron Artificial Neural Networks model which clearly depicts the correct and incorrect counts of both churn (Not Active) and non-churn (Active). Here the correctly predicted churners and non-churners count 5,989 and 6,985 respectively. No Non-Churners who are wrongly predicted as churners and also no churners who are wrongly predicted as non-churners.

As depicted in Figure 9, both the train and test data ROC curve values reached 1.0, indicating optimal model performance. To mitigate over fitting, under fitting, and network complexity, dropout was implemented in the network, streamlining its functionality to a minimal level.

Comparison between the proposed model with some selected state-of-the-art models

The accuracy, precision, recall, and F1-Score metrics and Receiver Operating Characteristic (ROC) were used as a basis for the comparison between the proposed approach (model) with selected state-of-the-art models. Table 3 depicts the performances obtained.

Table 3: Models comparison

Architecture	Accuracy	Precision	Recall	F1-Score	ROC
ANNs **(Amuda & Adeyemo, 2019)**	97.53%	97.7%	99.8%	98.8%	89.0%
ANNs **(Rosa, 2019)**	98.78%	79.02%	51.28%	99.78%	Not Provided
Multi-layer Perceptron artificial-Neural Network (**our Model**)	100%	100%	100%	100%	100%

VI.CONCLUSION

The objective of this study was to predict customers churn using data obtained from a Nigerian financial institution. Three sets of data, a training set, a test set, and a validation set were created by extracting the data from the bank database. The dataset was split up as follows: 75% were employed for training, 25% was used for testing and validating the model. EDA was used, and the text (categoricaal) data was converted into numerical data after 5 percent of the missing values were removed during data pre-processing. Additionally, feature scaling was used to lengthen the algorithm's computation time.

REFERENCES

[1].Elena, K. (2022). Data Science in Marketing: Customer Churn Rate Prediction. Retrieved from https://www.datacamp.com/blog/data-science-in-marketing-customer-churn-rate-prediction

[2].Taiwo & Adeyemo. (2012). Data Mining Technique for Predicting Telecommunications Industry Customer Churn Using both Descriptive and Predictive Algorithms Customer Churn Using both Descriptive and Predictive Algorithms. Computing, Information Systems & Development Informatics Journal 3(2)27-34

[3].Fujo S. W., Subramanian S., & Khder, M. A. (2022). Customer Churn Prediction in Telecommunication Industry Using Deep Learning. Inf.Sci.Lett.11(1)185–198 http://dx.doi.org/10.18576/isl/110120

[4].Mitkees, I. M., Ibrahim, A., & Elseddawy, B. (2017). Customer Churn Prediction Model using Data Mining techniques, European journal of operational research, 27, 262–268.

[5].García, D., Nebot, À. & Vellido, A. (2017). Intelligent data nalysis approaches to churn as a Businessproblem: a survey. Knowledge Information System 51, 719–774. https://doi.org/10.1007/s10115-016- 0995-z

[6].Gallo, A. (2014). The Value of Keeping the Right Customers. Retrieved on 11/12/2023 from https://hbr.org/2014/10/the-value-of-keeping-the-right-customers

[7].Prasdika, P., & Sugiantoro, B. (2018). A Review Paper on Big Data and Data Mining Concepts and Techniques., International Journal on Informatics for Development,7(1),36–38. https://doi.org/10.14421/ijid.2018.07107

[8].Markopoulos, A.P., Manolakos, D.E. & Vaxevanidis, N.M. Artificial neural network models for the prediction of surface roughness in electrical discharge machining. *J Intell Manuf* **19**, 283–292 (2008). https://doi.org/10.1007/s10845-008-0081-9

[9].Market Equations. (n.d.). Customer Churn Analytics and Management. Retrieved from http://www.marketequations.com/casestudy/outsourcing-segmentation-analysis.html.Accessed:10-10-2023.

[10].Sandeepkumar, H, & Mundada, M. R. (2019). Enhanced Deep Feed Forward Neural Network Model for the Customer Attrition Analysis in Banking Sector. Int..Jl. Intelligent Systems and Applications, 11(7), 10–19. https://doi.org/10.5815/ijisa.2019.07.02

[11].Tariq, M. U., Babar, M., Poulin, M. & Khattak, A. S. (2021), Distributed model for customer churn prediction using convolution neural network, Journal of Modeling in Management, 17(3), 853-863. https://doi.org/10.1108/JM2-01-2021-0032

[12].Lalwani, P., Mishra, M. K., Chadha, J. S. & Sethi, P. (2022). Customer churn prediction system: A machine learning approach, International Journal of Computer Applications, 104(2), 271–294

[13].Garimella, B., Prasad, G. V. S. N. R. V. & Prasad, M. H. M. K. (2021). Churn prediction using optimized deep learning classifier on huge telecom data, Journal of Ambient Intelligence and Humanized Computing 14, 2007–2028. https://doi.org/10.1007/s12652-021-03413-4

[14].Cenggoro, T. W., Wirastari, R. A., Rudianto, E., Mohadi, M. I., Ratj, D., & Pardamean, B. (2021).,Deep Learning as a Vector Embedding Model for Customer Churn, Procedia Computer Science. 179 (7)624–631 http://dx.doi.org/10.1016/j.procs.2021.01.048

[15].Kaur, I., & Kaur, J. (2020). Customer churn analysis and prediction in banking industry using machine learning. Proceedings of the Sixth International Conference on Parallel, Distributed and Grid Computing, Wak-naghat, India, 434–437. https://doi.org/10.1109/PDGC50313.2020.9315761

[16].Rosa, N. (2019). Gauging and foreseeing customer churn in the banking industry: a neural network approach. Thesis. Faculty of Informatics, Department of Computer Science, St. Mary's University, Addis Ababa, Ethiopia

[17].Elyusufi, Y., & AitKbir, M. (2022). Churn prediction analysis by combining machine learning algorithms and best features exploration. International Journal of Advanced Computer Science and Applications,13(7),615–622. https://doi.org/10.14569/IJACSA.2022.0130773

[18].Muneer, A., Ali, R. F., Alghamdi, A., Taib, S. M., Almaghthawi, A., & Abdullah, E. A. (2022). Predicting customers churning in banking industry: A machine learning approach. Indonesian Journal of Electrical Engineering and Computer Science, 26(1), 539–549. https://doi.org/10.11591/ijeecs.v26.i1.pp539-549

[19].Khan, Y., Shahryar. S., Abid. N., Sheeraz. A, Nadeem. S & Sabir.H, (2019). Customers Churn Prediction using Artificial Neural Networks (ANN) in Telecom Industry, International Journal of Advanced Computer Science and Applications,10(9), 234-242

[20].R. Yahaya, O. A. Abisoye and S. A. Bashir, "An Enhanced Bank Customers Churn Prediction Model Using A Hybrid Genetic Algorithm And K-Means Filter And Artificial Neural Network," *2020 IEEE 2nd International Conference on Cyberspac (CYBER NIGERIA)*, Abuja, Nigeria, 2021, pp. 52-58, doi: 10.1109/CYBERNIGERIA51635.2021.9428805. http://repository.futminna.edu.ng:8080/jspui/handle/123456789/6891

[21].Nguyen, V., Ngoc Le & Hoang, T., (2023). Customer Churn Prediction in the Banking Sector Using Machine Learning-Based Classification Models. Interdisciplinary Journal of Information. 18, 087-105 https://doi.org/10.28945/5086

[22].Rajendran, S., Devarajan, R & Elangovan, G., (2023). Customer Churn Prediction Using Machine Learning Approaches. Conference: 2023 International Conference on Artificial Intelligence and Knowledge Discovery in Concurrent Engineering (ICECONF). http://dx.doi.org/10.1109/ICECONF57129.2023.10083813

[23].Amuda. K. A. & Adeyemo. A. B (2019). Customers Churn Prediction in Financial Institution Using Artificial Neural Network. International Multi-Conference on ICT Applications At:Obafemi Awolowo University, Ile Ife 12(13), 54-60. https://arxiv.org/ftp/arxiv/papers/1912/1912.11346.pdf

[24].Adebiyi, S. O, Oyatoye, E, O & Mojekwu, J, N. (2015). Predicting Customer Churn and Retention Rates in Nigeria's Mobile Telecommunication Industry Using Markov Chain Modeling. Acta Univ. Sapientiae, Economics and Business, 3 (2015) 67–80. https://doi.org/10.1515/auseb-2015-0004

[25].Khoh, W.H., Pang, Y.H., Ooi, S.Y., Wang, L. &, Poh, Q.W.(2023). Predictive Churn Modeling for Sustainable Business in the Telecommunication Industry: Optimized WeightedEnsemble Machine Learning. Sustainability. 15(11), 8631. https://doi.org/10.3390/su15118631

[26].Mukaka, M. (2012). Statistics Corner: A guide to appropriate use of Correlation coefficient in medical research. Malawi Medical Journal: The Journal Of Medical Association Of Malawi, 24(3), 69-71.

[27].Zorich, A. (2018). Predicting Customer Churn in Banking Industry Using Neural Networks. Interdisciplinary Description of Complex Systems, 14. 116-124. https://doi.org/10.7906/indecs.14.2.1

[28].Xenonstack (2018). Data Preparation Process, Preprocessing and Data Wrangling. Retrieved on 10-Oct-2023 from https://xenonstack.medium.com/data-preparation-process-preprocessing-And-data-wrangling-6c4068f2fcd1.

[29].Jabeur, S. B., Gharib, C., Mefteh-Wali, S., & Arfi, W.B. (2021). CatBoost model and artificial intelligence techniques for corporate failure prediction. Technological Forecasting and Social Change, 166(8), 120658. https://doi.org/10.1016/j.techfore.2021.120658

[30].Mai, F., Tian, S., Lee, C., & Ma, L. (2019). Deep learning models for bankruptcy prediction using textual disclosures. European journal of operational research, 274(2), 743-758.