**CUSTOMER CHURN ANALYSIS USING MFNN**

**Rishikesh.C, Vishnu.N, Vishva. K, Dr. G. Uma Maheshwari, Dr. A. Meenakshi**

Department of Computer Science and Engineering, Kamaraj College of Engineering and Technology, Anna University, Virudhunagar, Tamil Nadu, India

E-mail: [21ucs080@kamarajengg.edu.in](mailto:21ucs080@kamarajengg.edu.in)

**1. ABSTRACT:**

In the banking sector, customer attrition happens when customers break off their relationship with the bank after using its services and products for a while. Considering this, maintaining customers is crucial in the fiercely competitive banking industry of today. Furthermore, a strong clientele helps draw in new business by instilling trust and encouraging recommendations from existing customers. Because of these issues, banks should make lowering client attrition a top priority. This study looks at bank data to predict which users are most likely to stop utilizing the bank's services and start paying for them by predicting their patterns through their state and behavior. This study examines the data and predicts the customer churn pattern by cleaning the uncorrelated variables in the dataset, converting the categorical variables into binary numeric variables, and finally splitting the 10,000-customer data into eighty percent of training dataset and twenty percent of testing dataset using a complex state of art machine learning algorithm namely **Multilayer Feedforward Neural Network (MFNN)** and presents a predictive analysis based on numerous evaluation measures with eighty six percent accuracy and fourteen percent error rate. The bank will identify trends through the analysis of this data and work to keep customers about to leave. Monitoring and managing customer disclosure and retention in a competitive banking sector every financial year can provide financial institutions with valuable insight. This knowledge may help these institutions to develop targeted strategies for differentiation from their competitors, which can increase retention of customers in future.

**2. INTRODUCTION:**

Customer churning, also known as customer attrition, is the process where customers exit their commitment to an organization. In the context of banking, this happens when customers withdraw their accounts and the services of a bank. Analyzing and managing customer retention is competitive for Financial Institutions to preserve their reputation and keep their finances consistent. Customer attrition can have a substantial monetary effect on financial institutions, which might result in revenue loss across a range of banking services. Thus, developing and upholding long-term relationships with clients is vital to a bank's development. Banks can identify clients who are expected to exit and put retention measures in place by receiving information about attrition patterns. This strategy increases the average lifetime value of clients and boosts bank profitability. Additionally, a bank's reputation and brand perception suffer from customer attrition. High attrition rates might be a sign of deeper problems, such ineffective customer service, antiquated systems, or a dearth of features and goods that rival competitors. For banks to overcome these obstacles and improve their entire customer experience, it is imperative that they comprehend and manage client attrition. Banks can gain important insights into the demands and preferences of their customers by tracking and managing customer attrition in the banking sector. This information can assist banks in creating focused strategies to set themselves apart from rivals and improve client retention.

**3. LITERATURE REVIEW:**

Joana Dias . , et al [1] “Machine learning for customer churn prediction in Retail banking”. The paper has dealt with the issue of predicting the period the customer will churn. This work uses six different methods to identify the churners six months in advance. STOCHASTIC BOOSTING is used for getting better results.

Kristof Coussement., et al [3] This work implements customer churn prediction models that is based on Convolutional Neural Network which incorporates textual information. It investigated the value added by incorporating the textual data to the customer churn prediction. (CCP) models. It validates how textual data can be integrated into a predictive model. It helps customers in making an informed decision in the investing process of text mining.

Manas Rahman., et al [5] “Machine Learning Based Customer Churn Prediction in Banking” works on the most critical issue of customer churn in the bank and the loss of engagement in bank activities. The customer behavior is analyzed by KNN, DECISION TREE, RANDOM FOREST classifiers to find the likelihood of churn.

Lee, Yong Jae., et al [7]” Machine learning for enterprises: Applications, algorithm selection, and challenges” deals with the tradeoff between the accuracy and interpretability of machine learning algorithms to select the right algorithm for right task outline three cases of machine-learning development in financial services.

Salma Karray., et al [14] “Applying hybrid machine learning algorithms to assess customer risk-adjusted revenue in the financial industry”. This work focuses on forecasting the customers’ RISK-ADJUSTED REVENUE (RAR) it is a measure of return-on-investment. It is one of the most critical issues in financial decision-making. Traditional prediction methods can’t provide higher accurate prediction by using both unsupervised and supervised learning machine learning algorithms. The usage of both labeled and unlabeled has increased the accuracy of the prediction model.

Aishwarya Saxena., et al [15] Analyzing customer churn in banking using data mining framework. enables bank administrators to analyze the activities of the client. The data collected from the analysis can be used later to form tactics to make plans for appropriate actions. The prediction models include logistic regression, support vector machines, random forest, and decision tree.

Aurélie Lemmens., et al [16] Managing Churn to Maximize Profit has dealt with the problem of customer defection by defining a function that works on loss function that is based on profit that predicts the fiscal impact of retention intervention. It coordinates the aim of increasing the campaign profit with an estimation algorithm.

Anik.et al [17] investigated customer churn in banking using a machine learning algorithm and visualization application for data science and management. This paper analyzes dataset from a bank and gives a prediction about which users will most likely stop using the bank’s services and change their relationship with the bank to paying customers. It is focused on core machine learning approaches, including SUPPORT VECTOR MACHINE (SVM), LOGISTIC REGRESSION, RANDOM FOREST, AND EXTREME GRADIENT BOOSTING (XGBOOST).

In this comprehensive analysis, we embark on an intricate exploration of a dataset brimming with details about bank customers, intricately intertwined with indicators of their departure from the banking institution. The mission is to meticulously dissect each column, decipher its significance, and subsequently delve into key categorical variables such as exit status, gender, credit card ownership, active membership, and geographical location. Through this exhaustive examination, we aim to gain a profound understanding of the dataset, unveiling invaluable insights to inform churn prediction models and devise effective customer retention strategies.

**4. PROBLEM STATEMENT:**

In the competitive world of banking, maintaining the customer is crucial. Banks need to tackle all the challenges to retain their customers in a rapidly changing market. To make the bank profitable, it's vital to understand why customers leave (churn) and find ways to keep the customer. By using previous data to predict exited customers and find insights to keep customers integrated with the bank. This research initially analyses ten thousand customers data by various features and techniques. The goal is to analyze how the proposed model will predict the exited customers optimally.

**5. OBJECTIVES:**

* To preprocessed the given dataset by cleaning and transforming respective attributes
* To identify patterns and indicators that lead to customer churn.
* To develop a MFNN predictive model that can anticipate percentage of customers are at risk of leaving the bank optimally.

**6. METHODOLOGY:**

**6.1. EXPLORATORY DATA ANALYSIS:**

The description of the dataset is displayed in **Table.1.** For further exploration is provided below to infer the information about the customers. **Geography** offers the distribution of exited customers across different countries is as follows: In France, there are 4204 exited customers, in Germany, there are 814 exited customers, and in Spain, there are 413 exited customers. Similarly, for customers who did not exit, in France, there are 1695, in Germany, there are 810, and in Spain, there are 2064. **Gender** categorized as the average count of customers who exited is 1139 for females and 898 for males. Similarly, for customers who did not exist, the count is 3404 for females and 4559 for males.

**Age** represents the average age of customers who exited is approximately 44 years, while the average age of customers who did not exit is approximately 37 years. **The Tenure (average) of** exited customers is approximately 5 years, while the average tenure of non-exited customers is approximately 4 years. **Balance**, the average account balance of customers who exited is approximately **91,107** units, while the average balance of customers who did not exit is approximately **72,765** units.

**The NumberOfProducts** represents the count of exited customers based on the number of products is as follows: 1 product - **1409**, 2 products - **348**, etc. Similarly, for customers who did not exist. The **HasCreditCard** is a binary indicator the count of exited customers based on whether they had a credit card is **1424**, and not having a credit card is **613**. Similarly, the count of non-exited customers having a credit card is **5631**, and not having a credit card is **2332**. The **IsActiveMember** specifies the count of exited customers based on their active membership status is as follows: **735** are Active Members, and **1302** are Not Active. Similarly, for customers who did not exit, **4416** are Active Members, and **3547** are Not Active.

The **Estimated Salary** represents the average estimated salary of customers who exited, with income greater than 1**00,000,** is **1044**, and for those who did not exit 3966. On the other hand, the average salary of customers with income less than **100,000** who exited is **993**, and for those who did not exit, it is **3997**.

|  |  |
| --- | --- |
| **Features** | **Description** |
| **RowNumber** | A value is consecutively assigned to each row to identify the overall entry. |
| **CustomerId** | It is used to identify customers clearly by assigning unique values, thereby avoiding confusion. |
| **CreditScore** | A numerical representation of customer creditworthiness based on their credit history. It indicates that higher scores imply lower risk for lenders. |
| **Geography** | It specifies the geographical representation of the customer's location, indicating where they reside. |
| **Gender** | It provides gender details for further churn analysis categorized by male or female. |
| **Age** | shows the age of the customer, which it affects their banking preferences and behaviors |
| **Tenure** | It shows how long the customer has been with the bank, which shows customer loyalty and retention |
| **Balance** | It indicates money the customer retains in their bank account, showing whether they're financially stable or active. |
| **NumOfProducts** | It tells number of banking products the customer has, infers about what they need and how involved they are with banking. |
| **HasCreditCard** | This tells if the customer has a credit card from the bank, which assess credit risk. |
| **IsActiveMember** | A binary value indicating whether the customer is an active member of the bank or not |
| **EstimatedSalary** | This provides an estimate of the customer's annual salary, helping to cluster them by income and plan finances accordingly. |
| **Exited** | This indicates that the customer has ended their relationship with the bank or not. |

**Table.1. Description of Attributes in the Dataset**

|  |  |
| --- | --- |
| **STEP 1** | Analyze the count of exited (7963) and not exited (2037) in the dataset. |
| **STEP 2** | Data preprocessing was done by cleaning the duplicate data and dropping the unrelated data (i.e., Row Number, Customer ID, Surname). |
| **STEP 3** | Transform the categorical data into numerical data or indicator variables Geography and gender data into 0’s and 1’s). |
| **STEP 4** | Split the data into 80% for training and 20% for testing and preprocessing the data using SCIKIT-LEARN. |
| **STEP 5** | Build a sequential Neural network model with four layers, RELU, and SIGMOIDAL activation functions by fitting theprocessed array of the training dataset. |
| **STEP 6** | Predict the metrics with 86% accuracy and 14% error. |
| **STEP 7** | Give the new input array to predict whether the customer will exit or not. |

**Table.2. ALGORITHM**

**6.2. DATA PRE-PROCESSING:**

Data Pre-Processing is an iterative process, and it often involves collaboration between data scientists, analysts, and domain experts. Automated tools and scripts can be used to streamline some aspects of data cleaning, but human expertise is crucial for making context-specific decisions. The goal is to prepare a clean, accurate, and reliable dataset for further analysis or modeling.

In this paper, there are 10,000 customer cases with zero null cells and the dataset needs to eliminate the uncorrelated fields for fitting in the model i.e., Surname, Customer ID, and Row number.

**6.3. VARIABLE CONVERSION:**

In pandas, the **get\_dummies()** function is used to convert categorical variables into numerical variables through a process called **one-hot encoding**.

**One-hot encoding** is a technique where categorical variables are converted into binary (0 or 1) vectors to represent distinct categories. Each category becomes a binary column, and for each observation, the corresponding column is marked with a 1 if the category is present and 0 otherwise. This conversion process allows you to use categorical information in models that expect numerical input, and it is a common preprocessing step in machine learning.

In this dataset the fields namely Geographical location and gender stated in categorical data, which need to be transformed into numeric data without manipulating the data. When you apply the **pd.get\_dummies** function to this Data Frame, specifying the categorical column (e.g., **'gender'**), it transforms each category into a new binary (0 or 1) column. For each unique category, a new binary column is created. If there are three categories (Germany, Spain, France), the data will transform with two new rows depicting the binary representation of the data. In the new columns, '1' is placed in the row corresponding to the original category, and '0's are placed in the rows for the other categories. This creates a binary representation of the categorical data. The original categorical column is then often removed from the Data Frame, as the information is now captured in the binary columns. The resulting Data Frame contains the original numerical columns along with the new binary columns representing the categories. This transformed dataset is suitable for machine learning algorithms that require numerical input.

**6.4. TOOLS AND FRAMEWORKS:**

**6.4.1 TENSORFLOW:**

TensorFlow is an open-source framework based on python and java for deploying complex Machine Learning algorithms in real-time application. This online cloud-based platform is developed by Google.

**6.4.2. SCIKIT LEARN:**

Scikit Learn which is also known as sklearn, is an python based open-source library which is used for data analysis and machine learning modules. This library is developed by David Cournapeau in 2007

**6.4.3. KERAS:**

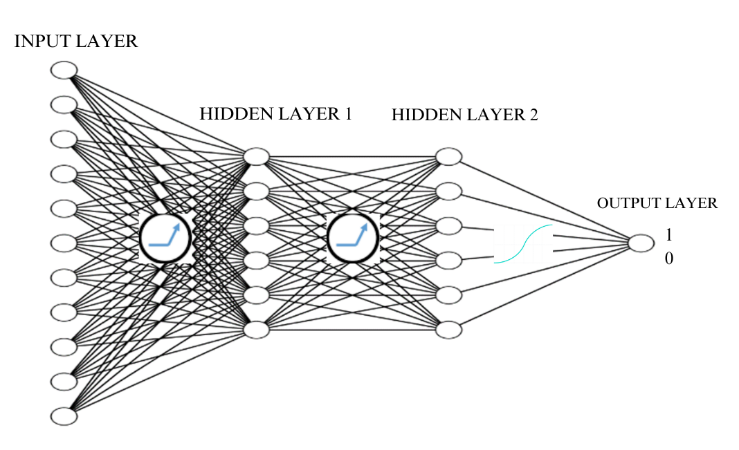
Keras is an open – source backend library for Artificial Neural Network in python interface developed by Francois Chollet - Google. It is an Interface for libraries such as TensorFlow, Theano, CNTK.

**6.4.4. MATPLOTLIB:**

It is one of the python based visualization libraries, which creates statics, dynamic and interactive visualization such as graphs, charts 2D animations, plots like line, bar, scatter, histogram, etc.

**6.5. MODEL BUILDING:**

A **Multilayer Feedforward Neural Network**, also known as **MFNN** or a Multilayer Perceptron (MLP), is an artificial neural network. In this method information flows in forward direction only, it starts from the input layer and parses through hidden layers (if any) gives the output in the output layer of the MFNN. The input layer consists of nodes (or neurons) that represent the features of the input data. In the input layer, every node corresponds to the specific feature of the input. Between the input and output layers, there can be one or more hidden layers. The model of MFNN is represented in **Fig.1.**

****

**Fig.1. – MFFNN MODEL**

Each hidden layer contains nodes that transform the input data using weighted sum connecting the features, weights and biases. During the training process the Neural Network learns the most complex patterns, Relationships of every parameter in the data. Each connection has an associated weight assigned to it. In Neural Network weight with a certain value is assigned to each node. Biases are additional parameters added to each node, allowing the network to account for offsets. Each node in the hidden layers and the output layer typically applies an activation function to its weighted sum of inputs. The activation functions such as Hyperbolic tangent (tanh), Rectified linear unit (ReLU), Sigmoid, Rectified linear unit (ReLU), SOFTMAX function are few examples. The output layer produces the final prediction or classification. The algorithm to execute the optimal model for this analysis in **Table.2.**

**6.5.1. NEURONS AND LAYERS:**

The neural Network consists of three types of layers namely Input layer, Hidden layers, and Output layer. Each node in a every layer represents a unique Neuron of the Neural Network. In this architecture there are eleven input neurons, twelve neurons in hidden layers (i.e., six neurons in each of two layers) and finally one neuron in the output layer. Layers such as sequencial and Densed layers are been used from Keras library.

**6.5.2. NEURON WEIGHTED CALCULATION:**

The weighted calculations are derived by the summation of the total set of input values with its corresponding weights and biases in the association of two neurons of different nodes which indicates mathematically in (1).

|  |  |  |
| --- | --- | --- |
|  | **Y = + b** | (1) |

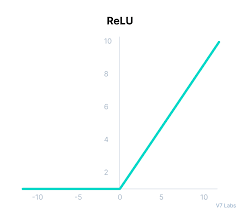
**6.5.3. ACTIVATION FUNCTION:**

The activation function is used categories the output binomially or multinomially from the weighted sum calculations of combination of neurons to predict the instances. In this research, there are two activation functions are been used namely ReLU and sigmoid.

**6.5.3.1. ReLU ACTIVATION FUNCTION:**

Rectified Linear Unit, or ReLU. In deep learning, this is the activation unit that is used the most. R(x) is equal to max(0, x) in (2). As a result, R(x) = 0 if x < 0 and x ≥ 0 if x ≥ 0. When considering gradient descent, it also improves convergence more than sigmoid or tanh activation functions. The ReLU function displays as a graphical representation in **Fig.2.**

|  |  |  |
| --- | --- | --- |
|  | **f(x) = max(0,x)** | (2) |

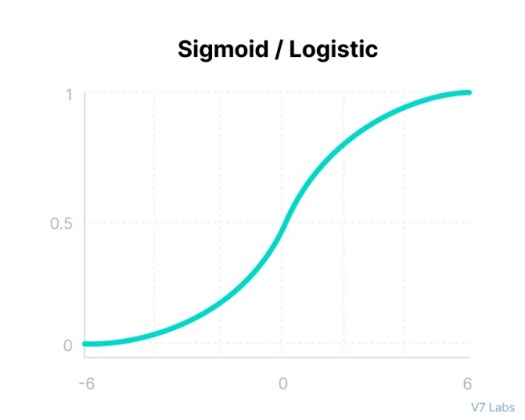


**Fig.2. Graphical representation of ReLU model**

**6.5.3.2. SIGMOID ACTIVATION FUNCTION:**

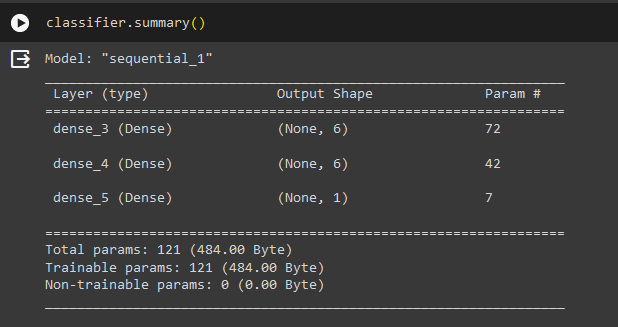
The sigmoid function is a mathematical function that modifies values in the interval between 0 and 1. Also referred to as the Sigmoidal Curve, it is an S-shaped curve that is mostly employed with non-linear activation functions in (2). The Sigmoid function displays as a graphical representation in **Fig.3.**

|  |  |  |
| --- | --- | --- |
|  | **f(x) = 1/(1+e^-x)** | (3) |

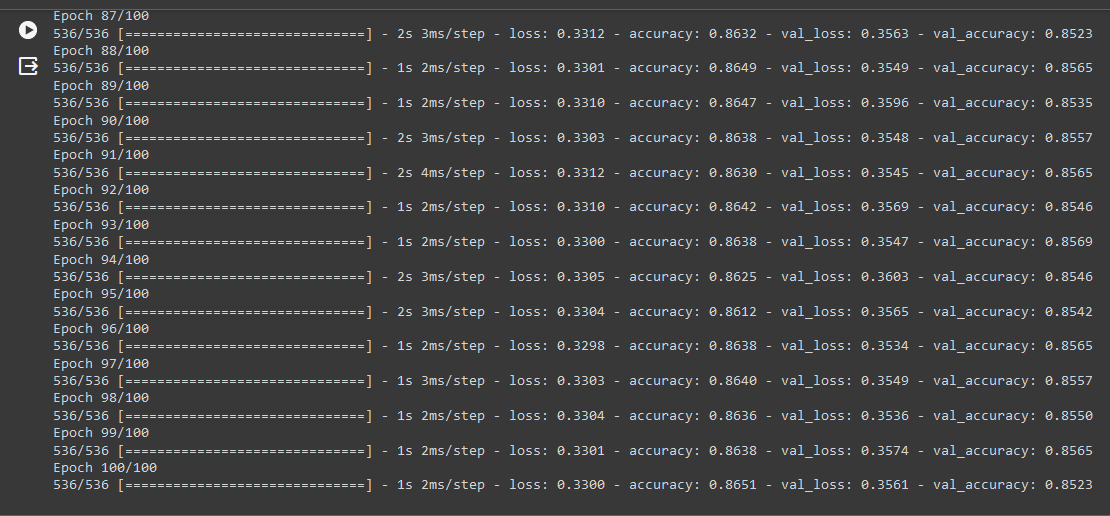


**Fig.3. Graphical representation of sigmoid curve model**

The model is built as a **Four-Layered Structure** compromised of eleven nodes in the input layer representing the transformed array of the eleven attributes mapped with six nodes of hidden layer 1 with ReLU activation function, weighted ratio of **72 (11\*6+6)**, and these six nodes of hidden layer 1 mapped with another six nodes of hidden layer 2 with ReLU activation function, weighted ratio of **46 (6\*6+6)**. Finally, the six nodes of hidden layer 2 are mapped with 1 node of the output layer with the sigmoid activation function, the weighted sum of **7 (6\*1+1)**. In conclusion, Finally the parameters associated in the connecting paths of each neuron of every layers is 121 in **Fig.4**.



**Fig.4. Representation of parameters with their respective layers in TensorFlow**



**Fig.5. Display of actual and expected accuracy and loss respectively over 100 epochs**

**6.6. PERFORMANCE EVALUATION:**

**6.6.1 CONFUSION MATRIX:**

It is one of the methods for evaluating the performance classification model in machine learning algorithm It usually represent in the form of a square matrix where each row represents predicted class, and each column represents actual class. It consists of various performance metrics such as ACCURACY, PRECISION, RECALL, SPECIFICITY, F1-SCORE, and so on. These metrics help in understanding how well the model is performing. The following are components of a confusion matrix:

|  |  |
| --- | --- |
| **TP** | True Positives are the cases were the number of instances that were predicted by the model as positive, and it also belongs to positive class. |
| **FP** | False Positives represents the number of instances that were predicted by the model as positive, but it belongs to the negative class. |
| **TN** | True Negatives are the cases where the number of instances that were correctly predicted by the model as negative, and it belongs to negative class |
| **FN** | False Negatives represents the number of instances that were wrongly classified by the model as negative, but it belongs to the positive class. |

**6.6.2 RATIO OF ACCURACY:**

It is the proportion of the True-Positive and True-Negative of the metrics to the sum of all four metrics. It is used to calculate the correctly predicted values from both parameters of the dataset

**ACCURACY=**

**6.6.3. RATIO OF ERROR:**

It is the proportion of the False-Positive and False-Negative of the metrics to the sum of all four metrics. It is used to calculate the wrongly predicted values from both parameters of the dataset

**ERROR=**

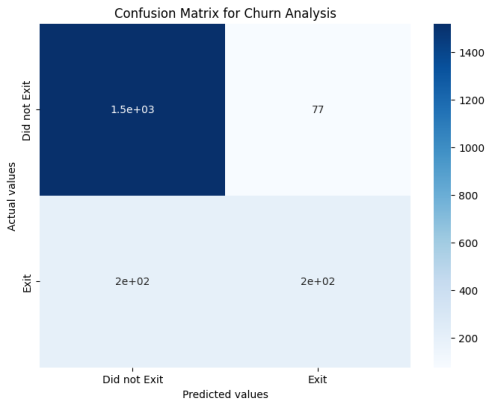
**7. RESULTS:**

This research is examined in the Google Collaboratory, an open-source software for executing complex Machine Learning algorithms with the specification of 8 GB RAM and local storage, i3 10th Gen. Processor. The execution of the MFNN model over 100 epochs represents the respective actual and expected values of accuracy and error in **Fig.5.** The **Credit Score** emerges as a critical metric in analysis, and it indicates the credit score of each customer, ranging from 300 to 850, with higher scores indicating better creditworthiness. The average credit score of customers who exited is about 646, while that of customers who did not exit is about 651.

This model has predicted the customers who would exit and not exit by the corresponding variables. The churn analysis using a multilayer feed-forward neural network preprocessed the dataset and converted the categorical data into binary numeric data in pandas. The MFNN Model predicted the confusion matrix which is displayed in **fig.6.** for churn analysis and it attained a prediction model with an accuracy of 86% and error rate of 14% in **fig.7.** and **fig.8.** respectively. Finally the proposed model is compared with the best model predicted with the same dataset in the work of **Pahul Preet Singh ., Fahim Islam Anik ., et al 2024 [17] in Table 3.** The insights gained from the analysis will help the bank take proactive measures, such as personalized retention strategies, to mitigate churn and enhance overall customer satisfaction.

**7.1. COFUSION MATRIX OF MFNN MODEL:**

This confusion matrix interprets the True-Positive = 1518, False-Positive = 77, False-Negative = 203 and True-Negative = 202. The output of the Confusion Matrix displayed



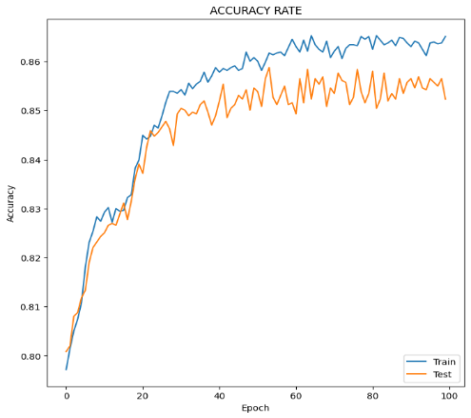
**Fig.6 – CONFUSION METRIXOF MFNN**

**7.2. ACCURACY RATE:**

It is proposed as the ratio of 1721 to 2000. The Accuracy Rate has been predicted as 86.05%.

**ACCURACY= =**

**ACCURACY= 86.05%**



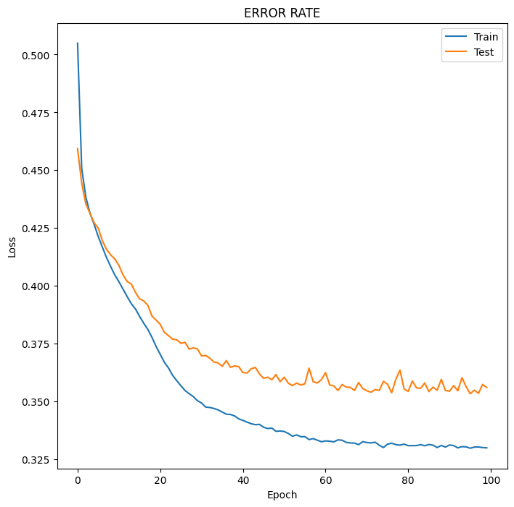
**Fig.7. – ACCURACY RATE of MFNN**

**7.3. ERROR RATE:**

It is proposed as the ratio of 279 to 2000. The Error Rate has been predicted as 13.95%

**ERROR= =**

**ERROR = 13.95%**



**Fig.8. – Error rate of the MFNN**

**Table 3: Comparison of Proposed MFNN Model vs Previous Algorithms**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **PREVIOUS WORK** | | **PROPOSED WORK** |
| **MODEL** | RANDOM\_FOREST | XG BOOST | MFNN |
| **ACUURACY RATE** | 0.8440 | 0.8520 | 0.8605 |
| **ERROR RATE** | 0.1660 | 0.1480 | 0.1395 |

The results of the respective algorithms (i.e., random forest, XG Boost algorithms) of the previous work kin this dataset by **Pahul Preet Singh ., Fahim Islam Anik ., et al 2024 [17].**

**8. CONCLUSION/ FUTURE WORK:**

This research concludes by predicting the Customers behavior exiting the bank. Among them with their financial behavioral pattern with 86% accuracy from the **Multi Layered** **Feedforward Neural Network Model (MFNN)**, there is a scope for improving the accuracy in terms of back propagation of the neural network model. The attributes provided are more tailored to the profile of a consumer than their recency—that is, the measurements that track behavior right before churning. In conclusion, studying each detail about customers and their behavior helps us understand why some of them leave the bank. By knowing their demographics, how they manage their finances, and how active they are, banks can develop strategies to retain them. Based on that, we predict the churn rate, enabling banks to tackle customer attrition effectively and build lasting relationships while ensuring steady business growth

These measures can be derived in the future by researchers, which will enable them to better identify churning tendencies by tracking a change in customer churning behavior right before they leave. In addition, they can go further proceed by incorporating complex algorithms, especially **Feedback Recurrent Neural Network**s, leading to more impactful predictions

**9. REFERENCE:**

[1] Dias, J., Godinho, P., Torres, P., 2020. Machine learning for customer churn prediction in

retail banking. In: International Conference on Computational Science and its

Applications. Springer, Berlin, pp. 576–589

[2] Al-Mashraie, M., Chung, S.H., Jeon, H.W., 2020. Customer switching behavior analysis in

the telecommunication industry via push-pull-mooring framework: a machine

learning approach. Comput. Ind. Eng. 144 (Jun.), 106476

[3] De Caigny, A., Coussement, K., De Bock, K.W., et al., 2020. Incorporating textual

information in customer churn prediction models based on a convolutional neural

network. Int. J. Forecast. 36 (4), 1563–1578.

[4] Lemmens, A., Gupta, S., 2020. Managing churn to maximize profits. Market. Sci. 39 (5),

956–973.

[5] Rahman, M., Kumar, V., 2020. Machine learning-based customer churn prediction in

banking. In: 2020 4th International Conference on Electronics, Communication and

Aerospace Technology. IEEE, pp. 1196–1201.

[6] Schaeffer, S.E., Sanchez, S.V.R., 2020. Forecasting client retention—a machine-learning

approach. J. Retailing Consum. Serv. 52 (Jan.), 101918.

[7] Lee, I., Shin, Y.J., 2020. Machine learning for enterprises: applications, algorithm

selection, and challenges. Bus. Horiz. 63 (2), 157–170.

[8] Vo, N.N., Liu, S., Li, X., et al., 2021. Leveraging unstructured call log data for customer

churn prediction. Knowl. Base Syst. 212 (Jan.), 106586.

[9] Guliyev, H., Tatoglu, F.Y., 2021. Customer churn analysis in banking sector: evidence

from explainable machine learning models. J. Appl. Mic. Econ. 1 (2), 85–99.

[10] Baghla, S., Gupta, G., 2022. Performance evaluation of various classification techniques

for customer churn prediction in E-commerce. Microprocess. Microsyst. 94 (Oct.),

104680.

[11] Geiler, L., Affeldt, S., Nadif, M., 2022. An effective strategy for churn prediction and

customer profiling. Data Knowl. Eng. 142 (Nov.), 102100.

[12] Van Der Donckt, J., Van der Donckt, J., Deprost, E., et al., 2022. Plotly-resampler:

effective visual analytics for large time series. In: 2022 IEEE Visualization and Visual

Analytics (VIS). IEEE, pp. 21–25.

[13] De Lima Lemos, R.A., Silva, T.C., et al., 2022. propension to customer churn in a financial

institution: a machine learning approach. Neural Comput. Appl. 34 (14),

11751–11768.

[14] Machado, M.R., Karray, S.,et al 2022. Applying hybrid machine learning algorithms to assess customer risk-adjusted revenue in the financial industry. Electron. Commer. Res.

Appl. 56 (Nov.), 101202.

[15] Aishwarya Saxena., Anushi singh .,et al 2023. Analyzing customer churn in banking: A data mining framework.

[16] Aurélie Lemmens.,et al. 2023. Managing Churn to Maximize Profit.

[17] Pahul Preet Singh ., Fahim Islam Anik ., et al 2024. Investigating customer churn in banking: A machine learning approach and visualization app for data science and management.