International Journal of Innovative Research in Engineering and Management (IJIREM)
ISSN(Online): 2350-0557, Volume-12, Issue-1, February 2025
https://doi.org/10.55524/ijirem.2025.12.1.5
Article ID IJIRE-1374, Pages 34-40
www.ijirem.org

Credit Card Churn Prediction: An Analytical and Model-Driven Study

Viswadhanush B R

MBA Scholar, School of Management, SASTRA Deemed to be University, Thanjavur, Tamil Nadu, India

Received: 4 January 2025 Revised: 19 January 2025 Accepted: 4 February 2025

ABSTRACT- This study compares the performance of three machine learning models namely Logistic Regression, Naive Bayes, and K-Nearest Neighbors (KNN) in forecasting customer churn within financial institutions. Predicting customer churn is important for banks and the financial sector because once firms develop targeted interventions to improve customer satisfaction and loyalty, they can be kept for the long run. A set of performance metrics like accuracy, precision, recall, F1 score, and area under the curve were used in the analysis for the better comparison of the prediction capabilities of models.

Logistic Regression resulted as the one which performed best. It gained the highest possible accuracy of 89.6% and its recall rate was very good at 96.8%. Also, its AUC score shows high discriminative power at 0.91. Naive Bayes, gaining less accuracy by producing 86.7% yet showed good precision at 92.0%, and having good recall rates, 92.2% and thus forms another competitive selection. Its AUC score of 0.83 establishes its efficiency to differentiate churners from non-churners. KNN's accuracy was good with 89.2% along with excellent recall rate 97.6%. Moreover, AUC score at 0.86 enhances the reliability as well as performance in the role of a good prediction model for churners.

Promising results may indicate more research on such advanced techniques. More models that would improve the performance may be neural network models in this regard such as Artificial Neural Network, Feed Forward Neural Network, Multi-layer Perceptron in the sub-field of neural networks and the models applied to Deep Learning that may involve models such as Convolutional Neural Network, Recurrent Neural Networks, Long Short-Term Memory. These techniques are expected to further enhance predictive accuracy by capturing complex patterns in large datasets.

This would greatly enhance customer retention for banks and financial institutions in their implementation of machine learning models. The accurate prediction of churn will allow organizations to engage at-risk customers proactively, providing customized interventions to enhance satisfaction and loyalty. The results of this study further highlight the value of machine learning in changing customer relationship management and leading to long-term customer retention and organizational success.

KEYWORDS- Predictive Analytics, Comparative Perspective, Credit Card, Customer Churn, Customer Retention, Customer Relationship.

I. INTRODUCTION

Customer churn is the method by which the customers terminate an association with their service provider. The issue had gained great deal of significance as a concern among the trading firms, including banking and the financial sectors also. It happens to be significantly more expensive keeping a new customer where as it might be easier acquiring a new customer hence the area on churn prediction got much importance and focus. With increased competition, financial institutions need to use advanced tools to identify at-risk customers who are likely to leave and design targeted strategies to retain them. Machine learning can be an effective solution for the same, enabling businesses to analyze large datasets and find complex patterns in customer behavior that traditional methods fail to capture. This research takes into account the prediction of customer churn in the banking sector through the application of three of the widely used machine learning models, namely Logistic Regression, Naive Bayes, and K-Nearest Neighbors. This prediction relies on behavioral, transactional, and demographic variables.

Traditional rule-based or statistical approaches often tend to fail in capturing the non-linear relations influencing customer churn. Machine learning, on the other hand, presents a strong alternative by discovering subtle patterns in data. Such models enhance predictive accuracy, thus providing actionable insights into customer dissatisfaction, allowing for tailored retention strategies that promote loyalty. This study evaluates three of the most widely used machine learning algorithms that have been known to be simple yet effective in carrying out predictive tasks. Logistic Regression is a linear model. Logistic Regression fits perfectly with problems in binary classification such as churn prediction and valued for interpretability. It's extremely relevant to business use cases. Naive Bayes is a probabilistic classifier that relies on Bayes' theorem, which has been found to be computationally efficient and has a good performance whenever the independence of features holds. K-Nearest Neighbors is a non-parametric model that classifies data points according to the majority class among their nearest Neighbors. It is quite effective for complex datasets. The dataset for this analysis consists of 10,127 credit card customers with 21 attributes divided into demographic, account-specific, activity-based, and transactional metrics. The demographic attributes are Age, Gender, Education Level, Marital Status, Income Range, and Dependents.

Other account-specific metrics, such as Card Type, Tenure Months, and Total Relationship Count, reveal insights related to customer engagement. Activity-based metrics such as Inactive Months, Contact Count, Credit Limit, and Credit Utilization Ratio give clues about usage patterns. Finally, transactional variables such as Total Transaction Amount, Count, and Quarterly Changes in Transactions are comprehensive of customer behavior. The target variable, Churn Status, classifies customers as either retained or attrited, forming the basis of the analysis. Model evaluation is based on commonly used metrics: accuracy, precision, recall, F1-score, and Area Under the Curve. These metrics assess the ability of the models to classify churned customers correctly, minimize false positives, and maintain a balance between precision and recall. Logistic Regression emerged as the best-balanced model, with the highest scores for all metrics, and an AUC of 92%, proving that it classifies churned from retained customers better.

Naive Bayes gave high precision and recall values. Its ability in probabilistic classification was very good, but AUC was moderate at 84%, proving it is mediocre at separating classes. K-Nearest Neighbors showed good recall, with high accuracy in identifying churned customers, but its slightly lower precision than Logistic Regression, along with an AUC of 87%, indicated the trade-offs in performance. The findings point toward the trade-off of each of the models adopted, and with business priorities around minimizing false positives or maximizing at-risk customer detections, the appropriate algorithm would come into play. Besides, it unfolds that the overall prediction of churning depends very much on elements like Credit Utilization Ratio, Total Transaction Amount, and customers' demographics. This study emphasizes the change that machine learning can affect to prevent a higher churn amongst customers that significantly impact profits and competitiveness in the banking sector. Not only do high churn rates cause difficulties in direct revenue terms, but they also taint organizational reputation. Machine learning can be employed in financial institutions to bring about proactive strategies, including personalized offers and enhanced support structures that will enhance customer relationships with long-run stability.

However, there are still challenges, such as data quality issues and the interpretability of models. Future experiments should be based on the application of more complex neural networks, like Artificial Neural Networks, Feedforward Neural Networks, and Multilayer Perceptrons, in detecting even more complex patterns in customers' information. Techniques like these are flexible and have higher predictive accuracy. Deep learning approaches such as Convolutional Neural Networks, Recurrent Neural Networks, and Long Short-Term Memory models should be further applied for sequential data analysis given that customer behavior changes over time. The main objectives of this study are to evaluate the predictive capabilities of Logistic Regression, Naive Bayes, and K-Nearest Neighbors for customer churn prediction, assess model performance in terms of metrics such as accuracy, precision, recall, F1-score, and ROC-AUC, and offer strategic recommendations for reducing churn and enhancing customer loyalty. In that context, the strength and limitation of such models identify this research and thus contributes toward further knowledge regarding churn prediction, building further on research work based on more sophisticated methodologies.

II. LITERATURE REVIEW

Predictive modeling has become a very important tool in dealing with credit card customer churn, and many studies have been done that contribute to the field. Itoo et al. [1] in their study "Comparison and analysis of logistic regression, Naïve Bayes and KNN machine learning algorithms for credit card fraud detection" compared these base models and found Logistic Regression to outperform others in accuracy, sensitivity, specificity, precision, F-measure, and AUC. The study emphasized the importance of data preprocessing in improving model performance. [1]

Miao and Wang [2] in their work "Customer Churn Prediction on Credit Card Services using Random Forest Method" showed that the most important predictors of churn were transaction amounts and revolving balances. They pointed out the potential of predictive modeling in providing actionable insights for customer retention.

Arram et al. [3] have introduced a novel dataset in "Credit card score prediction using machine learning models: A new dataset." The authors showed the superior performance of MLP Neural Networks by applying several algorithms on it. The paper emphasized early default prediction as one of the major areas for financial institutions.

Amuda and Adeyemo [4] in their work "Customers Churn Prediction in Financial Institution Using Artificial Neural Network" demonstrated artificial neural networks to be so efficient in churn prediction, with the study from them showing the rewards of neural networks in automatically performing feature engineering.

Kaur and Kaur [5] proposed "Machine Learning Approach for Credit Card Fraud Detection (KNN & Naïve Bayes) with emphasis of Feature Selection, Preprocessing toward Enhanced Fraud Performance" where, based on experimentations, demonstrated that the Hybrid Naïve Bayes along with KNN enhances the precision of detection as well as recall performance.

Tiwari et al. [6] in "Credit Card Fraud Detection using Machine Learning: A Study" provided a detailed review of several machine learning techniques, discussed the strengths and weaknesses of methods applied in fraudulent activities.

Mena et al. [7] discussed, in their paper "Churn Prediction with Sequential Data and Deep Neural Networks. A Comparative Analysis," the ability of deep learning techniques, such as LSTM models, to enhance the predictive capability.

Ram Kumar et al. [8] proposed, in "Automation of Credit Card Customer Churn Analysis using Hybrid Machine Learning Models," a hybrid model that integrates Logistic Regression, KNN, and Decision Trees for improving accuracy substantially. It has been put to use in the study to establish the potential of hybrid models for automation in churn analysis that brings customer retention strategies.

Despite advancements in predictive modeling for credit card customer churn, notable gaps remain. There is a lack of comprehensive comparative analyses of foundational models like Logistic Regression, Naïve Bayes, and KNN specifically for churn prediction, even though these models are interpretable and easier to implement in practical scenarios. Most works focus on some few performance metrics such as accuracy and AUC. However, in the case of dealing with imbalanced datasets, common in churn prediction, critical metrics such as precision, recall, and F1 score are ignored. Hybrid approaches which involve simpler algorithms for balancing performance with interpretability have not been investigated much. Very few works address the credit card customer segment. Addressing these gaps can improve the effectiveness and applicability of churn prediction models while providing actionable insights for financial institutions.

III. RESEARCH METHODOLOGY

This study employs a quantitative research method to analyze the case of customer churn in the banking sector. The dataset is of structured secondary data and contains 10,127 credit card customers. The source of the data set is the publicly available repository and has 21 variables for customer demographics, account details, and transactional behavior. Key characteristics are Age, Gender, Marital Status, Income Range, Card Type, Total Transaction Amount, Credit Utilization Ratio, and Churn Status, which categorizes the customers as churned or attrited. Such a wide set of variables creates a robust foundation for understanding why customers churn.

The sampling design would ensure that churned and retained customers are proportionately represented. This will reduce the bias in model training and testing. Although the dataset is secondary, its detailed and structured nature makes it ideal for this analysis. The data was prepared by removing irrelevant columns such as Customer ID and redundant features to focus on attributes directly related to churn prediction. The categorical variables Gender, Education Level, and Card Type were converted to numerical using Label Encoding. Numerical features are standardized using StandardScaler. Preprocessing ensures it improves the machine learning algorithms in distance-based computation-dependent algorithms.

Statistical methods were used to realize the study goals. EDA gave an insight into the distribution of variables and the relationship between them. Three machine learning models were used: Logistic Regression, Naive Bayes, and K-Nearest Neighbors (KNN). These models were picked based on classification performance and the degree to which data was handled by each model in different ways. The dataset had been divided between training (80%) and testing (20%) subsets with the aim that the models have been trained from the majority while being tested in a separate and unseen subset of data.

The performances of the models are assessed using key metrics, such as accuracy, precision, recall, F1 score, and Area Under the Curve (AUC). These have been selected because they give an overall view of the strengths and weaknesses of the models. This way, the whole analysis is robust and reliable, bringing out actionable advice to improve customer retention in the banking industry.

IV. ANALYSIS AND RESULTS

The analysis was conducted using Python 3 in the Google Colab Notebook.

A. Importing Necessary Libraries

Efficiency in analysis and model building requires importing many critical libraries. This import library will allow the use of extensive data handling capabilities with NumPy and Pandas for convenient preprocessing and smooth analysis. The following libraries provide various tools to create detailed and informative graphical representations of the data: Matplotlib, Seaborn, and Plotly. SciPy assists with the statistical analysis and offers various functions that might be used to probe and validate any patterns in the dataset. Finally, implementation of machine learning models happens to be done through Scikit Learn, providing an algorithm set and evaluation metrics. All three together make data analysis and visualization, and prediction modeling smooth processes.

B. *Loading and Viewing the Dataset*

The dataset is loaded using the Pandas function pd.read_csv() to ensure efficient handling of structured data. Initial exploration is conducted with data.head(), which gives a glimpse of the first few rows to understand the structure of the dataset. To confirm completeness, data.tail() is used to view the last few rows to ensure that no unexpected truncation occurred during loading. In addition, data.info() gives a summary of the dataset: number of entries, column names, data types, and null values. All these steps confirm that the dataset is ready for preprocessing and analysis.

C. Data Preprocessing

It consists of cleaning and preprocessing of data. Relevant columns to use in models would be added here, dropping unwanted columns, which are 'Customer ID', 'Total Revolving Balance', and 'Total Amount Change (Q4 to Q1)'. Since 'Gender' and 'Marital Status' contain categorical data, LabelEncoder encodes it as numerical value suitable for processing on machine learning models. data.isnull().sum() counts for missing values for a complete dataset. Features (X) and the target variable (y) are separated, with 'Churn Status' as the target. The dataset is then standardized using StandardScaler after being split into training and testing sets for optimal model training.

D. Exploratory Data Analysis (EDA)

The churn distribution presented in Figure 1 shows approximately 16.1% of the customers have churned, with the remaining 83.9% staying loyal. This is a pretty good customer retention rate; however, 16.1% is still an area in which improvement may be made regarding retaining customers. This could be due to dissatisfaction or unmet needs, implying that the churn can be minimized if these factors are addressed through improved services or targeted retention strategies. The understanding of the reasons behind this churn is important for improving long-term customer loyalty and business sustainability.

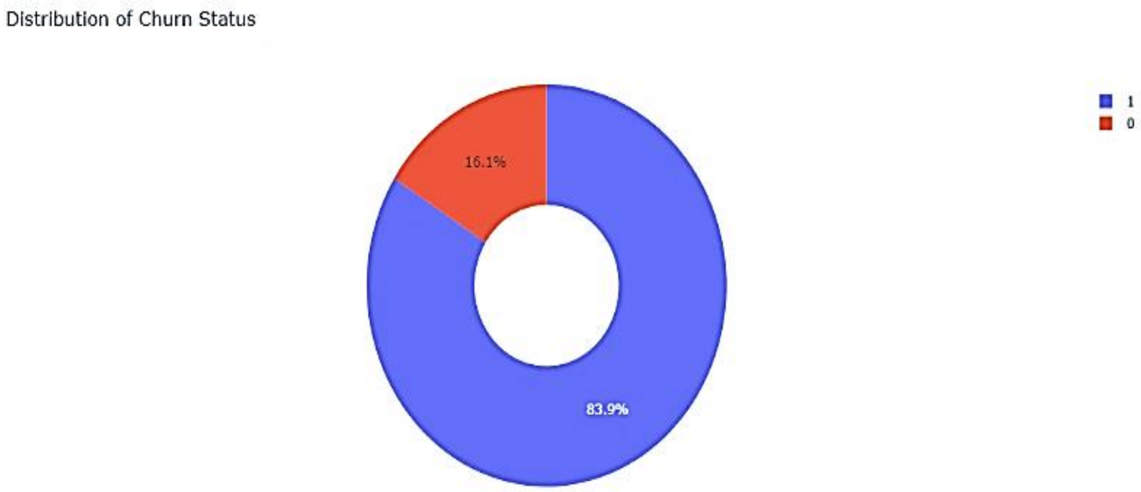

Figure 1: Distribution of Churn Status (Source: Computed data)

Figure 2 Spearman correlation heatmap showing the strength and direction of relationships between variables. As displayed in the heat map, the 'Churn Status' was moderately negatively correlated to 'Total Transaction Count' with a value of -0.38, where fewer transactions were accompanied by higher levels of churns. 'Transaction Count Change (Q4 to Q1)' too was negatively correlated with 'Churn Status', with a score of -0.31 indicating that reduced frequencies of transactions tend to predict risk of churn. In addition, very high positive correlations between 'Tenure Months' and 'Customer Age' (0.77) reveal demographic patterns about churn.

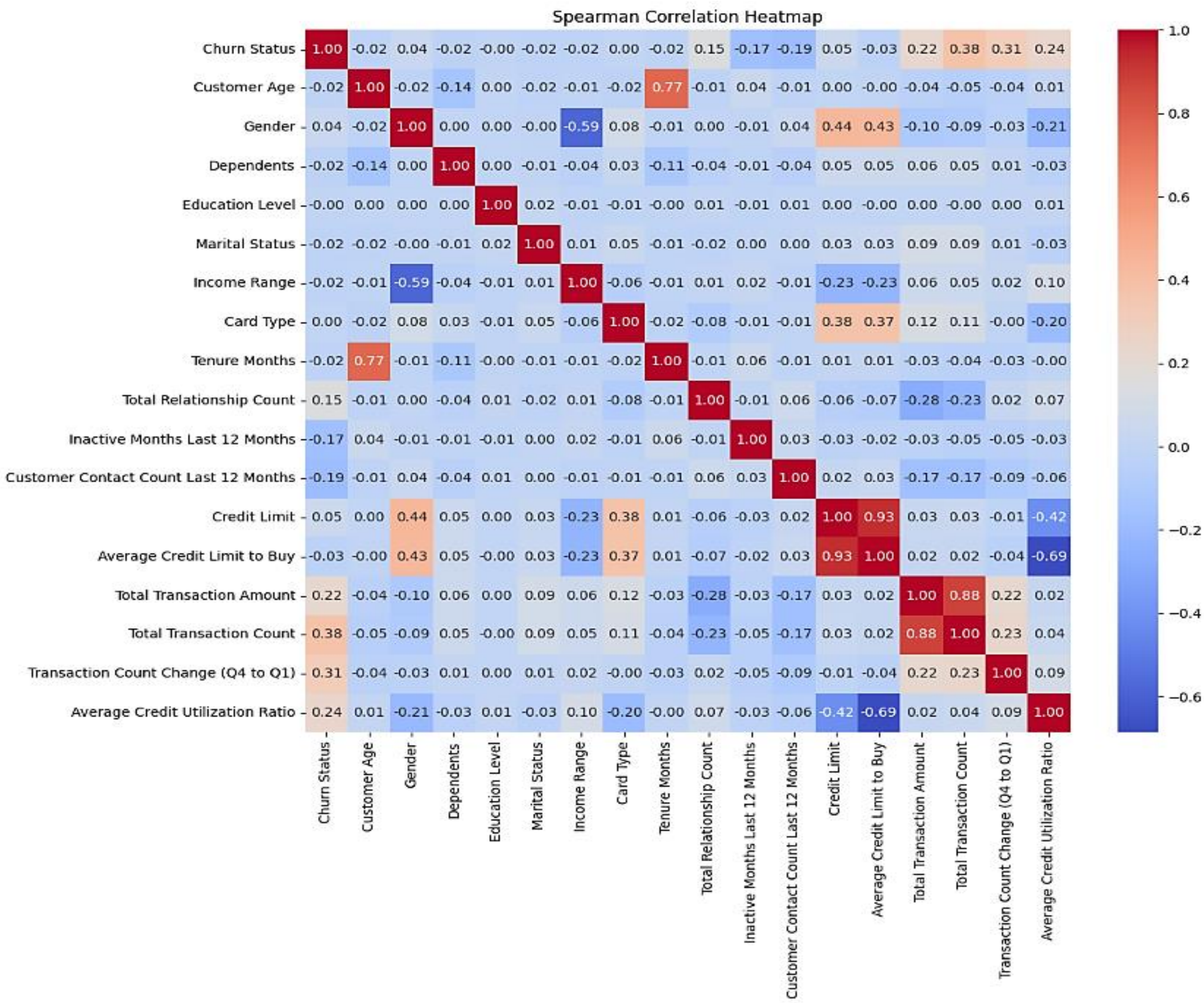

Figure 2: Spearman Correlation Heatmap (Source: Computed data)

Figure 3 presents the age distribution of customers, with a focus on churn patterns. The data reveals that the majority of customers are in the 40–55 age range, and churn is observed consistently across all age groups. Notably, younger customers, particularly those aged 30–40, show a slightly higher churn rate, which suggests that this demographic may be more likely to leave. This means that retention strategies targeting the younger generation should be developed since they are likely to churn because of different needs or preferences from the older generation.

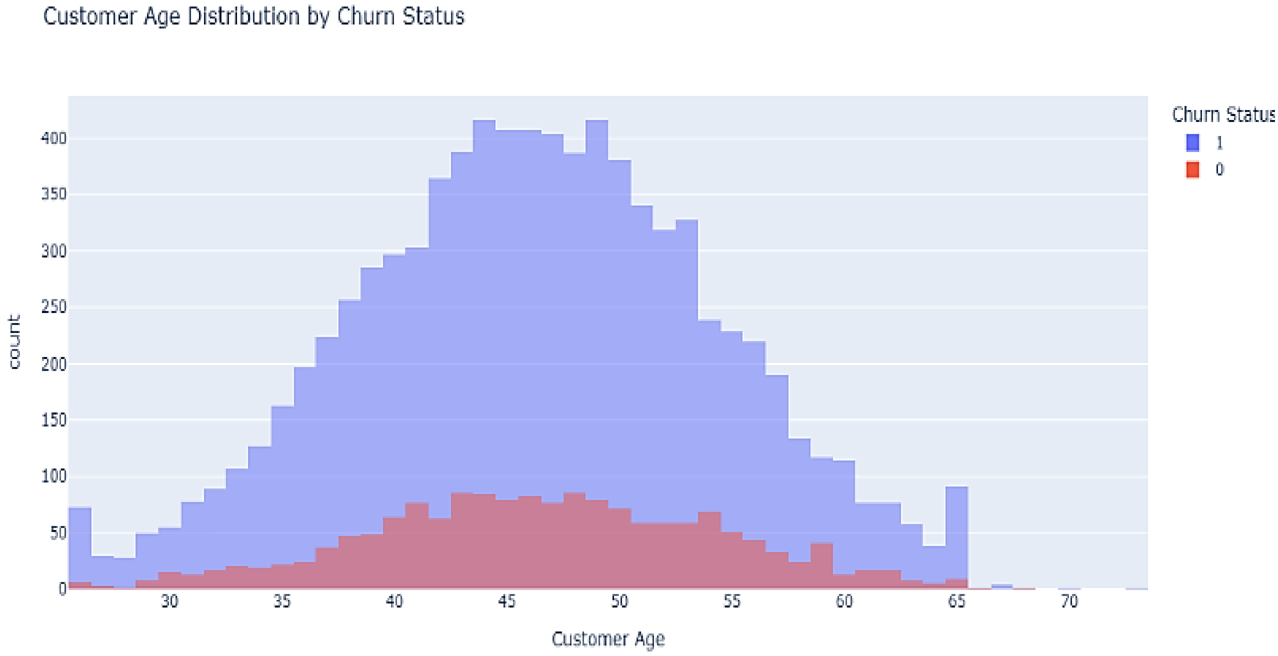

Figure 3: Customer Age Distribution by Churn Status (Source: Computed data)

E. Predictive Modelling

- Logistic Regression- The Logistic Regression model has a great performance with an accuracy of 89.6%. This means that the model is very good at making correct predictions. With a precision of 91.3%, it minimizes false positives, ensuring that predictions of churn are largely reliable. It has a high recall of 96.8%, which means it gets most of the actual churn cases right, very important for a churn prediction task. The F1 score of 0.94 is good, as it balances precision and recall. In addition, the AUC score is 0.91, indicating that the model is good at distinguishing between churn and non-churn.
- Naive Bayes- The Naive Bayes model has a good accuracy of 86.7%, which implies that the model is reliable in terms of predictive power. The model is also good at minimizing false positives since its precision is 92%, and this will be important in avoiding incorrect predictions of churn. The recall of 92.2% shows that the model is effective at identifying most churn cases while maintaining a balance between precision and recall. The F1 score of 0.92 further demonstrates this balance. Nonetheless, with an AUC score of 0.83, that's a tiny bit lower than logistic regression, one would say there's a little margin to improve its differentiation between these two classes.
- K-Nearest Neighbors(KNN)- The K-Nearest Neighbors (KNN) model has an accuracy of 89.2%, which is very good because it gives a good predictive power. The model has also achieved a precision of 90.4%. This means that most of the churn cases predicted are right ones with less false positives, and the model excels in its recall value that is very high at 97.6%, and thus makes KNN very sensitive to churning. The F1 score of 0.94 further shows the good balance of precision and recall of the model. The AUC score of 0.86 indicates that KNN is a strong classifier, though slightly behind Logistic Regression in terms of overall performance.

F. Model Comparison (Accuracy, Precision, Recall, F1 Score)

Model comparison shows that models have different strengths. Logistic Regression is the strongest in terms of accuracy at 89.6% and recall at 96.8%. Naive Bayes has impressive precision at 92% and recall at 92.2%, avoiding false positives, and correctly predicting positive cases. KNN performs well in recall at 97.6% with closely competitive accuracy at 89.2% and precision at 90.4%. Balances between precision and recall across different F1 scores that are given within the models-Logistic regression and KNN lead at 0.94. As summarized in Figure 4, metrics highlight what all models are actually capable of-to predict churn with customers.

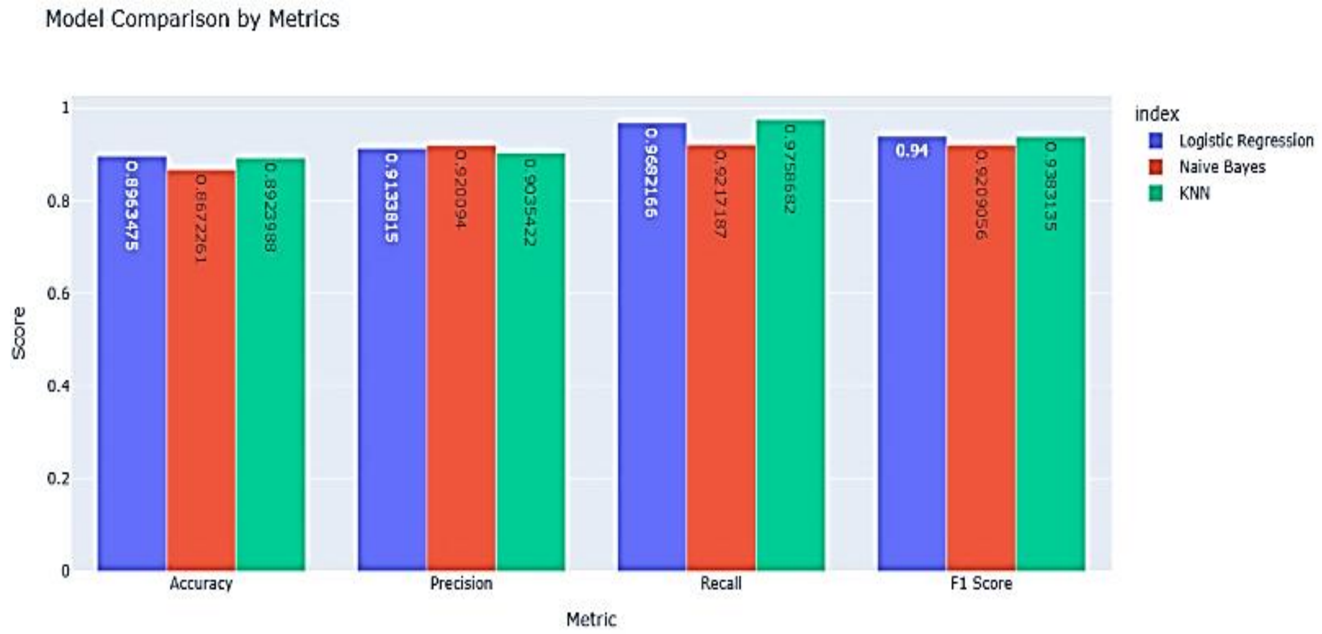

Figure 4: Model Comparison by Metrics (Source: Computed data)

G. Model Comparison (AUC)

The ROC curves of the three models give a good view of how discriminative the models are. Logistic Regression scores the highest AUC score at 0.92, proving that it is the best-in-class discrimination. KNN comes next with an AUC score of 0.87, which proves to be pretty strong, while Naive Bayes has an AUC of 0.84, indicating scope for improvement. The ROC curve analysis, illustrated in Figure 5, validates Logistic Regression as the best model for predicting customer churn. These insights assist in selecting the most effective model for practical application.

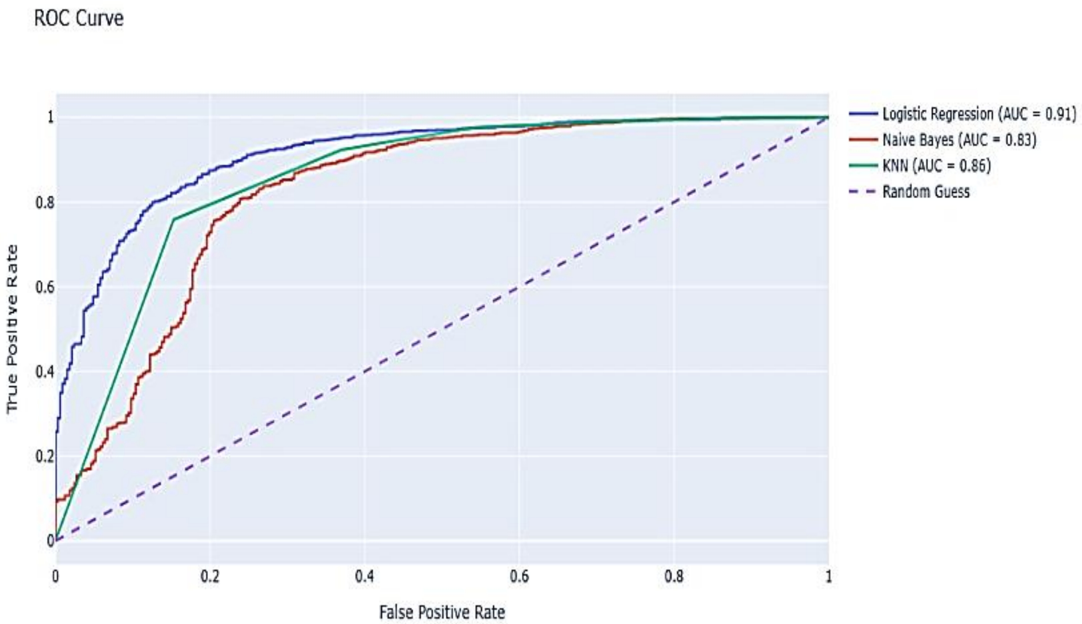

Figure 5: ROC Curve(Source: Computed data)

V. DISCUSSION

The analysis of customer churn in the banking sector with machine learning models reveals several critical findings. Logistic Regression was the most balanced model, which presented high accuracy, precision, recall, and F1 score; this indicated that it could well differentiate between churned and retained customers. Its AUC of 92% reflects the capability of this model to make accurate predictions across various thresholds. The Naive Bayes model has performed satisfactorily, mainly in precision and recall, proving its efficiency in probabilistic predictions. However, the AUC of 84% it achieved shows only a little capability in separating classes between churned and non-churned ones. The KNN model performed well, with outstanding recall, which indicates its strength in correctly identifying churned customers. However, the slightly lower precision than Logistic Regression suggests that it may yield more false

positives. With an AUC of 87%, KNN performs competitively but fails to beat Logistic Regression on balance overall. The results highlight the trade-offs of various models, where each has strengths and weaknesses in different business priorities, whether it is to minimize false positives or maximize churn identification.
The research shows that customer demographic factors, account attributes such as Credit Utilization Ratio, and transactional behaviors like Total Transaction Amount have a high impact on churn prediction. This indicates the importance of using extensive customer data to get accurate modeling. Even though the models performed well with good accuracy and evaluation scores, it is not advisable to limit to only three classification models. Adding more advanced techniques and expanding the scope of the dataset may enhance the predictive capabilities.

VI. CONCLUSION AND FUTURE SCOPE

This study proves the capabilities of machine learning models for predicting the accurate scenarios of customer churn in the banking sector. Leverage the key factors such as demographics of customers, account-specific metrics, and transactional behavior by this study to evaluate the performance of Logistic Regression, Naive Bayes, and K-Nearest Neighbors models. Logistic Regression performed the best of all the models in terms of precision and recall, with Naive Bayes also performing very competitively especially in specificity and sensitivity. The K-Nearest Neighbors highly did well in terms of recall but at slight trade-offs in terms of precision.
The findings shows that machine learning has transformed the role into preventing customer churn, thereby being a significant concern for banks and financial institutions. By achieving a better level of accuracy while identifying at-risk customers, banks can intervene by providing personalized offers or enhancing the support system to make it more satisfying, and in turn increase loyalty. Data-driven decisions by financial institutions using machine learning help them achieve long-term success when it comes to customer retention strategies.
However, predictive modeling on churn is very much in the nascent stages, and there remains a lot to be explored. More research in future could be to implement advanced types of neural networks, such as Artificial Neural Network, Feedforward Neural Network, Multilayer Perceptron which are known to be flexible as well as more sensitive to complicated patterns in the data. These involve techniques, including Convolutional Neural Networks, Recurrent Neural Networks, and Long Short-Term Memory models, that promise to better reveal the characteristics of sequential and temporal data, mainly of changing customer behavior over time.
The application of ensemble learning techniques, such as Random Forest, Gradient Boosting Machines, and XGBoost, can improve prediction performance by tapping into the collective strengths of the algorithms. In combination with the use of large and diverse data sets, it is possible that these advanced methods will result in more accurate and reliable predictions.
In conclusion, the study presents how machine learning models can predict churn and build proactive retention strategies by having significant potential. Further research can contribute to making solutions even more effective and accurate by developing further these techniques, integrating various sources of data, and thereby assisting financial institutions to achieve sustainable customer retention and long-term competitiveness.

CONFLICTS OF INTEREST

The author declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

ACKNOWNLEDGMENT

I am deeply grateful to Dr. Nalini R, Assistant Professor at School of Management, SASTRA Deemed to be University, Thanjavur, Tamil Nadu, India, for her constructive feedback, which has significantly enhanced the quality of this research.

REFERENCES

[1] F. Itoo, Meenakshi, and S. Singh, "Comparison and analysis of logistic regression, naïve Bayes and KNN machine learning algorithms for credit card fraud detection," *Int. J. Inf. Technol.*, vol. 13, pp. 1503–1511, 2021. Available from: https://doi.org/10.1007/s41870-020-00430-y

[2] X. Miao and H. Wang, "Customer churn prediction on credit card services using random forest method," *Adv. Econ., Bus. Manag. Res.*, vol. 211, 2022. Available from: https://doi.org/10.2991/aebmr.k.220307.104

[3] Arram, M. Ayob, M. A. A. Albadr, A. Sulaiman, and D. Albashish, "Credit card score prediction using machine learning models: A new dataset," *arXiv Preprint*, 2023. Available from: https://doi.org/10.48550/arXiv.2310.02956

[4] K. A. Amuda and A. B. Adeyemo, "Customers churn prediction in financial institution using artificial neural network," *arXiv Preprint*, 2019. Available from: https://doi.org/10.48550/arXiv.1912.11346

[5] P. Tiwari, S. Mehta, N. Sakhuja, J. Kumar, and A. K. Singh, "Credit card fraud detection using machine learning: A study," *arXiv Preprint*, 2021. Available from: https://doi.org/10.48550/arXiv.2108.10005

[6] G. Mena, A. De Caigny, K. Coussement, K. W. De Bock, and S. Lessmann, "Churn prediction with sequential data and deep neural networks: A comparative analysis," *arXiv Preprint*, 2019. Available from: https://doi.org/10.48550/arXiv.1909.11114

[7] M. Kaur and S. Kaur, "Machine learning approach for credit card fraud detection (KNN & naïve Bayes)," in *Proc. Int. Conf. Innov. Comput. Commun. (ICICC) 2020*, 2020. Available from: http://dx.doi.org/10.2139/ssrn.3564040.

[8] R. P. Ram Kumar, B. Sahithi, K. Neeharika, M. Shivaleela, D. Singh, and K. R. K. Reddy, "Automation of credit card customer churn analysis using hybrid machine learning models," *E3S Web Conf.*, vol. 430, 2023. Available from: https://doi.org/10.1051/e3sconf/202343001034.