Transforming Marketing Analytics: A Machine Learning Framework for Churn Prediction Driven by Data

Gala,
Depatment of Management, Monash University, Beijing 100080, China
E. Harrison
Depatment of Management, Monash University, Beijing 100080, China

Research Article

Date of Submission: 17-05-2025 **Date of Acceptance: 20-05-2025** **Date of Publication: 10-06-2025**

Abstract - Customer attrition is one of the main issues facing large corporations. Businesses in the rental sector especially look for ways to retain customers since they are their main source of revenue. The data-driven machine learning approach for telecom sector churn prediction is presented in this study. It tackles the problems of noisy, imbalanced, and high-dimensional data by using a thorough preprocessing pipeline that includes feature selection, outlier detection, SMOTE-based balancing, and noise removal. The preprocessed data, which includes 3,333 customer records with variables ranging from call metrics to charge information, is divided into training and testing sets in order to develop prediction models utilising Decision Trees and Random Forest classifiers. The Decision Tree model outperformed other models like ANN and SVM when evaluated using metrics like accuracy, precision, recall, F1-score, and ROC analysis. It achieved 88% accuracy, 83% precision, 93% recall, and 88% F1-score. These results demonstrate how the method may provide valuable data for enhancing client retention and optimising marketing strategies in evolving telecom environments.

Keywords - Telco-customer-churn dataset, machine learning, marketing analytics, predictive analytics, and customer retention.

1.Introduction

Customer happiness and retention are the keys to success in today's corporate climate, making customer behaviour analysis a crucial source of competitive advantage. Customer churn is a persistent commercial difficulty for the telecommunications sector, which works in a setting of rapid development and fierce market rivalry, as consumers break off ties with provider services [1]. client turnover causes a decline in income as well as the waste of money invested on client acquisition via concession and advertising programs in addition to operating costs.

Large-scale customer contact data analysis reveals trends and insightful information that companies may use to deal with customers that exhibit disengagement. Customised offers and improved service delivery are guided by the gained insights, which increase customer satisfaction and marketing cost-effectiveness.

Beyond technological difficulties, businesses must strategically anticipate client attrition, particularly when using subscription-based revenue models [2]. Businesses may develop proactive methods that reduce customer turnover rates and generate significant revenues over long periods of time by accurately identifying client attrition. The analytical shift that enables marketers to create predictive models for decision-making is made possible by machine learning's error prediction capabilities, which enable data analysis to continue independently. The process of training digital machines to carry out activities that fall under the category of learning via human or animal behaviour and cognitive processes is known as machine learning [3]. To enhance performance going forward, an ML system collects observations of its workplace [4].

1.1. Contribution of Study

The research analyzes a data-based machine learning system that aims to transform marketing analysis through customer churn forecasting. The research establishes advancements within its field through improved model efficiency and data-based customer retention decisions. The key contributions are as follows:

- Utilizing the Telco Customer Churn dataset to provide a comprehensive view of churned and non-churned customers to efficiently train and assess models.
- Preprocessing the dataset and identifying the most relevant features is done through feature selection,
- along with outlier detection techniques to improve data quality and model reliability.
- The dataset is divided into sections for testing and training, while applying cross-validation to ensure robust evaluation and generalizability of the framework.
- To evaluate machine learning, performance indicators such as accuracy, precision, recall, and F1-score are utilized, models like DT and Random Tree, in order to provide a balanced approach to handling false positives and false negatives.

1.2. Structure of Paper

The structure of the paper is as follows: A thorough literature review that highlights current methods is presented in Section II. The suggested framework and model development are explained in Section III. Results of the Section IV experiment and visualization of important metrics. Finally, Section V concludes with key findings, practical implications, and recommendations for future advancements in marketing analytics.

2. Literature Review

In order to improve retention efforts and inform strategic decisions, the literature study looks at methods for predicting client attrition using ML and DL models. According to Idris, Iftikhar, and Rehman (2019), an enhanced churn prediction system that successfully detects improvements in churn behaviour may be created by combining AdaBoost classification features with genetic programming searching methods. PSO-based undersampling, which equitably distributes training data to GP-AdaBoost prediction systems, provides a solution to the telecom dataset imbalance issue. GP-AdaBoost and particle swarm optimization-based undersampling are combined to create ChP-GPAB, a prediction system that improves consumer churn detection and identifies important factors influencing consumer behaviour. The suggested ChP-GPAB system is tested and evaluated using two well-known telecom data sets. Based on data from the Cell2Cell and Orange datasets, the results show that ChP-GPAB achieves evaluation values of 0.91 AUC and 0.86 AUC while identifying the reasons for customer churn [5]. Before presenting a hybrid ensemble model that combines bagging with neural networks for improved performance metrics and increased churn prediction accuracy, Saghir et al. (2019) analyse both solo and ensemble neural networks as classifiers. Two benchmark datasets from GitHub are used in the study to evaluate and contrast the suggested model. The suggested approach has an accuracy rate of 81% on average [6]. n Ebrah and Elnasir's (2019) investigation, Naïve Bayes is used in conjunction with SVM and decision tree algorithms on the IBM Watson and cell2cell datasets. The IBM Watson data has 7033 observations spread over 21 properties, whereas the cell2cell data has 71,047 observations along with 57 attributes. According to the AUC performance assessment, the cell2cell dataset produced scores of 0.98, 0.99, and 0.98, whereas the IBM dataset produced scores of 0.82, 0.87, and 0.77. The proposed models showed greater accuracy levels when compared to previous investigations that employed matched datasets [7]. Agrawal et al. (2018) demonstrate how to do churn predictions on a Telco dataset using a DL method. Experts made the decision to use multilayered neural networks to create a non-linear classification model. Contextual, usage, customer data, and support characteristics are the four components that contribute to the evaluation of the churn prediction model. The system calculates the likelihood of client attrition as well as the contributing causes. The trained model predicts the probability of customer attrition by applying its weighted values to features. The accuracy rating was 80.03%. Additionally, the approach enables companies to identify the root causes of churn problems and implement solutions [8]. Idris and Khan (2017) evaluated the efficacy of the proposed FW-ECP system using Orange and Cell2Cell, two publicly accessible standard telecom datasets. Considering the enormous dimensionality and unbalanced nature of the

training sets, FW-ECP generates better prediction results than the most sophisticated methods presently in use. Instead of 260D and 76D, the feature spaces in the Orange and Cell2Cell datasets are 24D and 18D, respectively. According to FW-ECP, the Orange and Cell2Cell datasets exhibited corresponding AUCs of 0.85 and 0.82 [3]. Reddy and Mishra (2018) Sort the clientele into two groups: churners and non-churners. Because deep learning can manage ever-increasing data amounts, uncover hidden patterns, recognise patterns and underlying risks, and more accurately predict consumer behaviour, it is superior to conventional machine learning algorithms in this respect. CNNs are used in this work to implement DL for Churn prediction performs with a high degree of accuracy. The churn prediction out predictive model obtains a performance level of 86.85% accuracy, 91.08 accuracy, 93.08% recall, and 92.06% F-score, according to the experiment findings. 13.15 percent is the mistake rate [9].

Table I provides a comprehensive summary of key studies on customer churn prediction, highlighting methodologies, datasets, performance metrics limitations, and suggested future work.

Table 1: Summary of Literature Review on Customer Churn Prediction Using Machine Learning

References	Methodology	Dataset	Performance	Limitations & Future Work
Idris, Iftikhar, and Rehman (2019)	GP-AdaBoost integration; PSO-based undersampling for dataset balancing.	Cell2Cell, Orange telecom datasets	AUC: 0.91 (Cell2Cell), 0.86 (Orange)	Limited to telecom datasets; focuses mainly on churners without extensive generalization for non-telecom datasets.
Saghir et al. (2019)	Ensemble Neural Network classifier using Bagging.	Two benchmark datasets from GitHub	Accuracy: 81%	Focused on accuracy; lacks exploration of interpretability and reasons for churn.
Ebrah and Elnasir (2019)	SVMs, Decision Trees, and Naïve Bayes.	IBM Watson (7033 obs, 21 attrs), Cell2Cell	AUC: 0.82, 0.87, 0.77 (IBM); 0.98, 0.99, 0.98 (Cell2Cell)	Limited analysis of class imbalance; limited focus on scalability for large datasets.
Agrawal et al. (2018)	Multilayer Neural Network for nonlinear classification.	Telco dataset	Accuracy: 80.03%	Performance limited to a single dataset; lacks exploration of real-world constraints.
Idris and Khan (2017)	FW-ECP system for dimensionality reduction and handling imbalanced datasets.	Orange, Cell2Cell	AUC: 0.85 (Orange), 0.82 (Cell2Cell)	Focused on dimensionality reduction, lacks interpretability for real-world decision-making.
Mishra and Reddy (2018)	CNN-based Deep Learning for churn prediction.	Telecom Industry dataset	Accuracy: 86.85%; Precision: 91.08%; Recall: 93.08%; F1: 92.06%	Limited exploration of alternative deep learning architectures and broader applications.

3.Methodology

The goal of this research is to provide a precise marketing analytics churn prediction. The methodology seeks to find trends and important churn indicators by using past customer data, allowing companies to take preventative action. A methodical machine-learning strategy for telecom customer attrition analysis is presented in the methodology study shown in the flowchart. It starts with a collection of telecom customer data that has been cleaned and prepared via data preprocessing. After that, the procedure splits into two concurrent steps: outlier identification to eliminate unusual data points and feature selection to find relevant variables. The data is split into training and test sets in a 75:25 ratio after the preparatory processes have converged. The training data is then used to train machine learning models, namely Random Forests and Decision Trees. Recall, accuracy, precision, and F1 score are assessed in a thorough review that comes after model development. The last step, which completes an organised pipeline from raw data to practical insights for predicting customer attrition in the telecom sector, yields the results. This procedure is shown in Figure 1 as a flowchart with key stages highlighted.The goal of this research is to provide a precise marketing analytics churn prediction. The methodology seeks to find trends

and important churn indicators by using past customer data, allowing companies to take preventative action. A methodical machine-learning strategy for telecom customer attrition analysis is presented in the methodology study shown in the flowchart. It starts with a collection of telecom customer data that has been cleaned and prepared via data preprocessing. After that, the procedure splits into two concurrent steps: outlier identification to eliminate unusual data points and feature selection to find relevant variables. The data is split into training and test sets in a 75:25 ratio after the preparatory processes have converged. The training data is then used to train machine learning models, namely Random Forests and Decision Trees. Recall, accuracy, precision, and F1 score are assessed in a thorough review that comes after model development. The last step, which completes an organised pipeline from raw data to practical insights for predicting customer attrition in the telecom sector, yields the results. This procedure is shown in Figure 1 as a flowchart with key stages highlighted.

3.1. Data Collection

ESX-1 and ESX-2 are two major corporate systems from which our dataset was gathered. Every time a network intrusion occurred, the SOC security analysts independently documented the threat information in the security raw events, which were gathered over a period of five months for ESX-1 and thirty days for ESX-2, respectively. Threat occurrence time, related attacks, attack type, response contents, attack IP address, and victim network details are all included in the threat detection information list. We examined 798 identifying cyber threats in ESX-1 in our datasets, which are distributed across the whole collecting period. There are 240 scanning, 547 system hacking, and 11 worm assaults among the documented cyber threats. Similarly, there are 51 worm assaults, 3,077 system hacking, and 941 scanning in ESX-2. SOC analysts manually classified the different types of attacks. System hacking attacks fall into four categories: injection, brute force, DDoS, and cross-site scripting. Trojan horses and backdoors fall under the category of scanning attacks. A total of 4,079 cyber-threats were discovered.

The ESX-1 and ESX-2 big corporate systems are the sources of our dataset. Every time a network intrusion occurred, the SOC security analysts independently documented the threat information in the security raw events, which were gathered over a period of five months for ESX-1 and thirty days for ESX-2. Threat occurrence time, related attacks, attack type, response contents, attack IP address, and victim network details are all included in the list of threat detection information. We examined 798 identifying cyberthreats in ESX-1 in our datasets, which are scattered over the whole collecting period. There are 11 worm attacks, 547 system hacking assaults, and 240 scanning attacks among the listed cyber threats. In a similar vein, ESX-2 has 51 worm assaults, 3,077 system hacking, and 941 scanning. SOC analysts carefully categorised the different types of attacks. A cross-site script, DDoS, brute force, and injection attacks are among the categories of system hacking attacks. Trojan horses and backdoors are examples of scanning attacks. In all, 4,079 cyberthreats were identified as assaults.

A flowchart's subsequent phases are briefly described below:

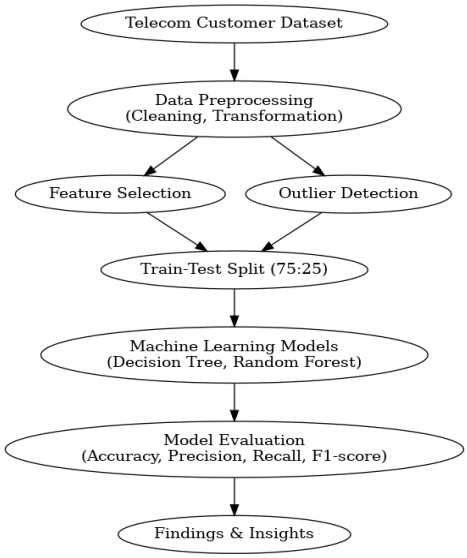


Fig 1: Churn Prediction Using Machine Learning Can be Represented as a Flowchart

The ESX-1 and ESX-2 big corporate systems are the sources of our dataset. Every time a network intrusion occurred, the SOC security analysts independently documented the threat information in the security raw events, which were gathered over a period of five months for ESX-1 and thirty days for ESX-2.

Threat occurrence time, related attacks, attack type, response contents, attack IP address, and victim network details are all included in the list of threat detection information. We examined 798 identifying cyberthreats in ESX-1 in our datasets, which are scattered over the whole collecting period. There are 11 worm attacks, 547 system hacking assaults, and 240 scanning attacks among the listed cyber threats. In a similar vein, ESX-2 has 51 worm assaults, 3,077 system hacking, and 941 scanning. SOC analysts carefully categorised the different types of attacks. A cross-site script, DDoS, brute force, and injection attacks are among the categories of system hacking attacks. Trojan horses and backdoors are examples of scanning attacks. In all, 4,079 cyberthreats were identified as assaults.

Our dataset has been collected from two large enterprise systems, named ESX-1 and ESX-2. The security raw events were collected over 5 months for ESX-1, over 30 days for ESX-2, respectively, in which the detecting threat information was separately recorded by the SOC security analysts whenever a network intrusion occurred. The list of threat detection information contains threat occurrence time, related attacks, category of attack, respond contents, attack IPaddress, and victim network information. In our datasets, we investigated 798 detecting cyber threats in ESX-1, which are dispersed across the entire collection period. Looking at the type of occurred attacks in recorded cyber threats, there are 240 scanning, 547 system hacking, and 11 worm attacks. Similarly, in ESX-2 there are 941 scanning, 3,077 system hacking, and 51 worm attacks. This categorizing of attack type was manually performed by SOC analysts. By category, the system hacking attack includes a cross site script, DDoS, brute force attack, and injection attack. A trojan and backdoor attack belongs to scanning attack. Overall the number of attacks was found 4,079 cyber-threats.

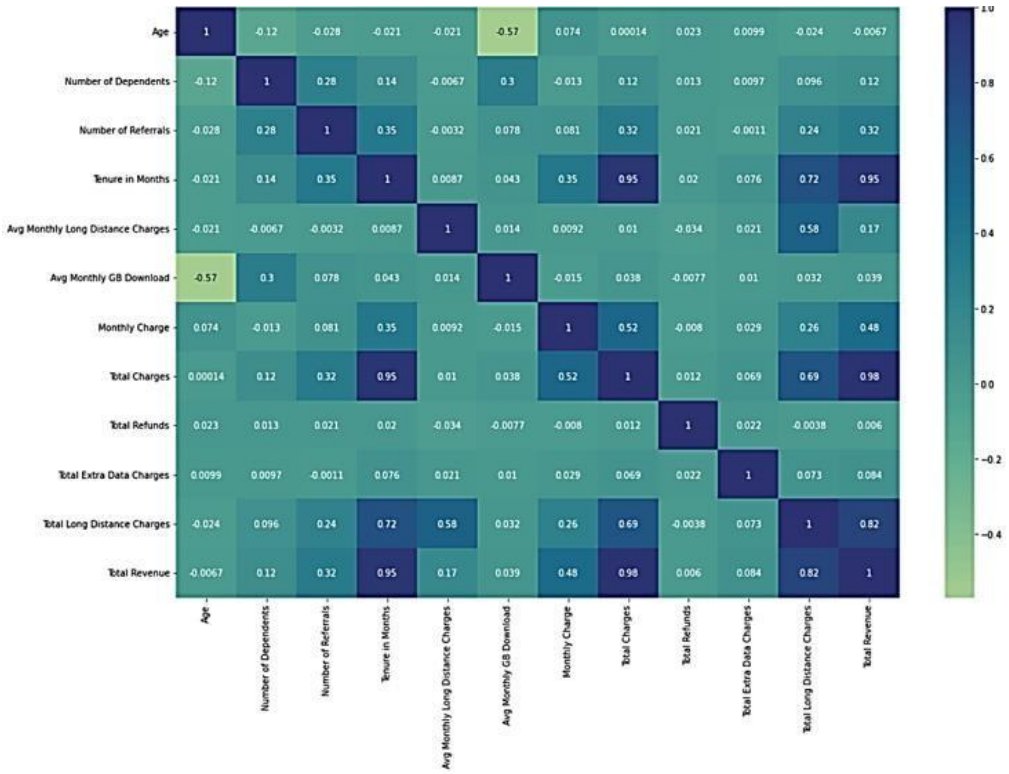


Fig 2: Heatmap of Correlation Matrix

The telecom industry's churn prediction model design uses a customer's historical conduct over a certain time period to forecast how they will behave in the near future. As a result, it is standard procedure to collect as much information as possible on every customer. These attributes Churn as class label, area, Here are the details of each of the 3333 records in the dataset: service calls, evening calls, evening charges, evening minutes, day calls, international calls, international minutes, and lastly, night calls, night charges, and nightly minutes. Figure 2 below shows the correlation matrix for churn prediction:

The correlation matrix in Figure 2 shows the heatmap displaying relationships between financial and demographic variables like age, dependents, referrals, tenure, and various charge categories. It uses a blue-to-green color gradient with numerical values in each cell to indicate correlation strength, where diagonal cells show perfect correlation (1) and off-diagonal cells show varying positive and negative relationships between different variable pairs.

3.2. Data Processing

Transforming data, or processing it, is known as data processing. When presented correctly, data may be both educational and useful. It can be highly beneficial. Data processing systems are also known as information systems. Data processing may also be described as the process of converting information into data. The following steps of pre-processing are listed below:

3.3. Noise Removal

It is necessary to ensure that the data is significant since noisy data might provide subpar outcomes. There are many missing entries, erroneous values such as "Null," and imbalanced characteristics in the telecom dataset.

3.4. Data Balancing

Balancing, an essential aspect of accurate classification, significantly influences the model's ability to effectively represent the classification performance. To address this imbalance, the SMOTE was employed. This technique

oversamples the minority class ("Yes") to achieve a more balanced distribution, as illustrated in Figure 3.

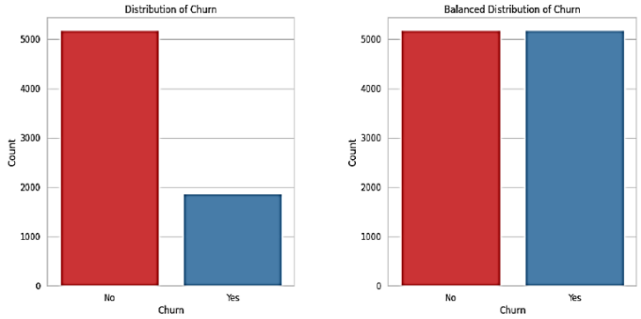


Fig 3: Before and After Applying SMOTE

Figure 3 shows two bar plots illustrating the distribution of a binary "Churn" variable before and after applying the SMOTE balancing method. In the left plot labeled "Distribution of Churn," there is a clear class imbalance where the "No" category (shown in red) has approximately 5,000 counts while the "Yes" category (shown in blue) has only about 1,800 counts. The right plot, labeled "Balanced Distribution of Churn," demonstrates the effect of SMOTE application, where both categories have been balanced to approximately 5,000 counts each, effectively addressing the initial class imbalance problem that could potentially bias a machine learning model's performance.

3.5. Feature Selection

For dimensionality reduction, feature selection is one method that is employed [10]; repetitive and unnecessary characteristics are eliminated while pertinent features are chosen. Reducing input dimensionality can enhance performance by either boosting generalization ability and classification accuracy or lowering learning pace and model complexity.

3.6. Outlier Detection

There are two main categories for outlier detection: distance-based techniques and density-based techniques. An observation is identified as an outlier according to QC if its value falls outside the specified non-outlier range.

3.7. Data Splitting

The data was divided into two categories: 25% for testing and 75% for training. The model is fitted during the training phase, and evaluating the model's performance prior to field deployment helps guard against unforeseen problems that may result from overfitting.

3.8. Proposed Decision Tree (DT)

A hierarchical structure called a decision tree makes predictions by doing a number of feature checks [11]. Let's write X for the input characteristics, Y for the goal variable, and DT for the decision tree. Based on feature tests, the decision tree iteratively divides the dataset with the goal of maximizing class separation or minimizing impurity [12]. The forecast from the decision tree might be shown as follows: (1):

$$DT(X) = \sum_{i=1}^{L} y_i \,.\, I(X \in R_i) \qquad (1)$$

Here,y_i is the class label assigned to the i-th leaf node, and L is the number of decision tree leaf nodes, and R_i is the region or subset of instances that the i-th leaf node was assigned to the feature testing. The indicator function $I(X \in R_i)$ returns 1 if the input instance X is in the region R_i and 0 otherwise. Based on the feature checks, the decision tree moves from the root to a leaf node and gives the leaf node where the instance is located the appropriate class label, y_i.

3.9. Performance Metrics

A number of measures, characteristics like F1-score, recall, accuracy, and precision are used to evaluate the effectiveness of the implemented model. These metrics help determine how successful the model is overall. The

terms listed below are explained in their model:

- Customer churning (positive) is referred to as a "TP" and is classified as such.
- The term "TN" refers to a consumer who is not churning (negative) and is labelled as such [13].
- A "false positive (FP)" is a consumer who is categorized as churn (positive) even if they are not churning (negative).
- A "false negative (FN)" is a customer who is churning (positive) but is said to be not churning (negative).

Accuracy: Accuracy measures the precise number of all made forecasts that turned out to be accurate, and Equation (2) is used to calculate it:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \quad (2)$$

Precision: The precision of the equation is the percentage of the projected positive situations that were accurate (3):

$$Precision = \frac{TP}{TP+FP} \quad (3)$$

Recall: The equation is used to compute recall, which is the percentage of affirmative cases that were properly detected (4):

$$Recall = \frac{TP}{TP+FN} \quad (4)$$

F1-score: The combination finds extensive usage as a solitary measurement tool to evaluate [14] performance of the classifier. The harmonic combination of accuracy rate and recall measurement produces the F-measure

ROC: A graphic depiction of model performance overall thresholds is the ROC curve. Through data charting, an evaluation of the TPR and FPR at particular thresholds yields a ROC curve.

4.Results & Discussions

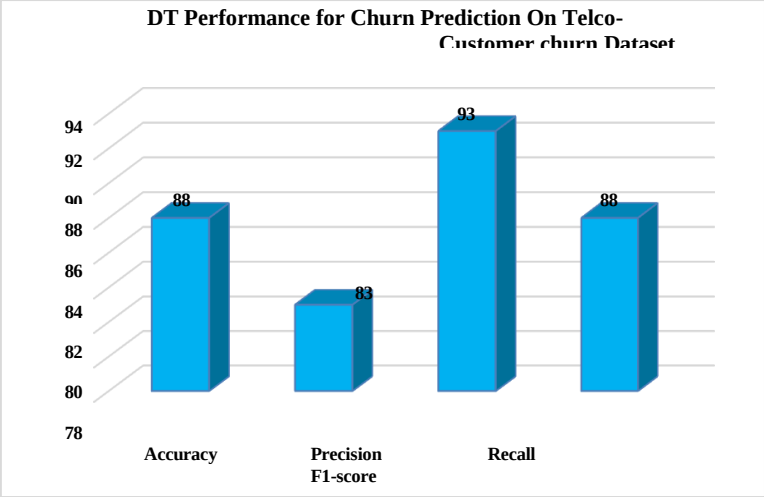
The churn prediction system is built on a Windows platform with an Intel i7 processor, 16GB RAM, and SSD storage, utilizing Python-based libraries such as scikit-learn and Pandas. This section discusses the simulated results for machine learning-based churn prediction utilizing customer-churn datasets and provides the results of decision tree model in

Table 2:Findings of Decision Tree Model on Telco-Customer-Churn dataset for churn Prediction

Measures	Decision Tree
Accuracy	88%
Precision	83%
Recall	93%
F1-score	88%

Fig 4: Bar Graph for Decision Tree Performance

Figure 4 displays the performance of a DT model for churn prediction. The chart highlights four key metrics: F1-score, recall, accuracy, and precision. The model's precision is 83%, and its accuracy is 88%. Recall achieves the highest value at 93%, and the F1-score is also 88%. The chart provides a clear representation of how the decision tree model fares across different evaluation metrics in predicting churn.



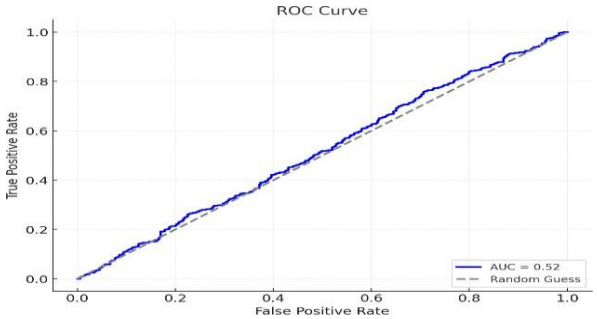
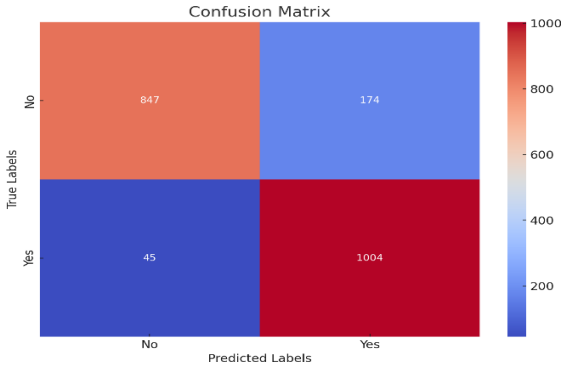


Fig 5: ROC curve for DT model

A classification model's performance appears in Figure 5 at various cutoff points. The TPR against FPR is shown by the blue curve, while the random guess classifier's performance is shown by the diagonal grey line. The classifier can efficiently discriminate between positive and negative classes, as evidenced by the outstanding model performance indicated by the AUC value of 0.97. A further indication of the curve's high prediction accuracy and low sensitivity/specificity trade-offs is its closeness to the top-left corner.

Fig 6: Confusion matrix for DT Model



In this Figure 6 shows a confusion matrix for a binary classification model. It displays the relationship between true and predicted labels. There are 847 true negatives (model correctly predicted "No"), 174 FP (model predicted "Yes" when the true label was "No"), 45 FN (model predicted "No" when the true label was "Yes"), and 1004 TP (model correctly predicted "Yes"). The model demonstrates exceptional classification performance because it achieves high accurate classification of true positives and true negatives, according to this matrix.

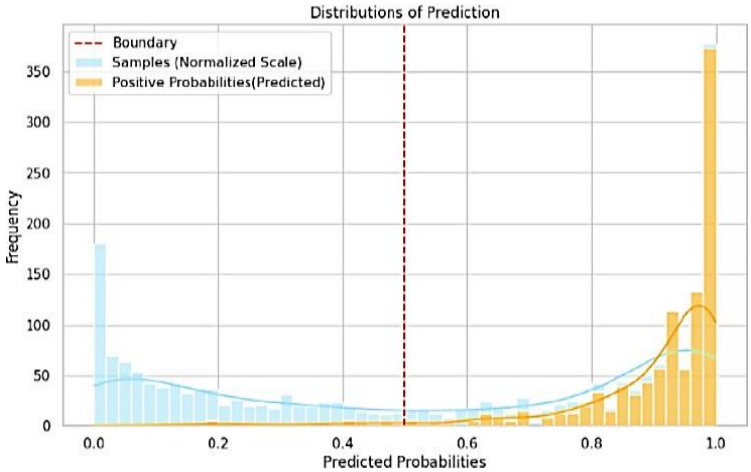


Fig 7: Churn Classification Performance

Figure 7 visualizes the predicted probabilities for a churn classification model. It features a red dashed decision boundary at 0.5, blue bars for sample distributions, and orange bars for predicted positive probabilities. The model appears to provide accurate predictions with little uncertainty around the boundary, as indicated by the bimodal distribution with peaks close to 0 and 1.

Table 3: Comparison between DT and Existing Models Performance on Telco-Customer-Churn dataset

Model	Accuracy
Decision Tree (DT)	88%
Artificial Neural Network (ANN)[15]	55.61%
Support Vector Machine (SVM) [16]	69.95%
Random Forest[17]	84.4%

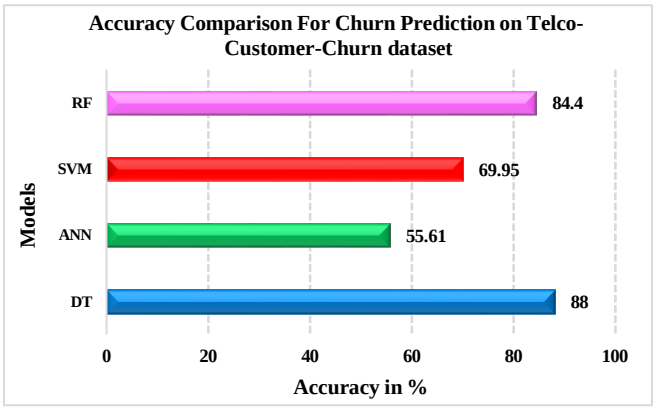


Fig 8: Bar Graph for Accuracy Comparison

Table III and Figure 8 present a performance comparison between DT and existing models, including ANN, SVM, and RF. The DT model achieves the highest accuracy of 88%, demonstrating the best prediction capability for customer churn research. However, the accuracy of the ANN model is the lowest (55.61%), indicating that it is not particularly helpful in this case. The RF model exhibits competitive performance with an accuracy of 84.4%, whilst the SVM model performs modestly with an accuracy of 69.95%. This comparison highlights DT's effectiveness in handling the Telco-Customer-Churn dataset compared to other popular machine-learning approaches.

Conclusion & Future Work

In telecommunications, customer churn forecasting is closely related to both financial stability and client retention. Machine learning (ML) is being used by telecom companies to predict customer attrition, or the probability that a customer would move to a competitor. Based on its 0.97 AUC score from the ROC curve, accuracy of 88%, precision of 83%, recall of 93%, and F1-score of 88%, the Decision Tree (DT) model reliably predicts client turnover. These results are further supported by the confusion matrix, which shows a high percentage of true positives and true negatives with minimal misclassifications. The DT model outperforms ANN (55.61%), SVM (69.95%), and RF (84.4%) in processing the Telco-Customer-Churn dataset, indicating its great efficacy. One of the study's shortcomings is its dependence on historical data, which could not adequately reflect shifting consumer behaviour patterns or dynamic market shifts. Future research might concentrate on enhancing the churn prediction model's generalisation by adding social media signals, real-time data, and other outside variables like market trends or rival activity.

References

[1] "Predicting customer churn in telecom industry using multilayer preceptron neural networks: Modelling and analysis," Life Sci. J., vol. 11, no. 3, pp. 75–81, 2014; O. Adwan, H. Faris, K. Jaradat, O. Harfoushi, and N. Ghatasheh.
[2] B. Mishachandar and K. A. Kumar, "Predicting customer churn using targeted proactive retention," International Journal of Engineering and Technology, vol. 7, no. 2.27, p. 69, 2018.
[3] Idris and Khan, "Filter-Wrapper and Ensemble Classification-Based Churn Prediction System for Telecom," Comput. J., vol. 60, no. 3, pp. 410–430, 2017, doi: 10.1093/comjnl/bxv123.
[4] A. Amin et al., "Comparing oversampling techniques to handle the class imbalance problem: A customer churn prediction case study," Ieee Access, vol. 4, pp. 7940–7957, 2016.
[5] A. Idris, A. Iftikhar, and Z. U. Rehman, "GP-AdaBoost learning and PSO undersampling for intelligent churn prediction for telecom," Cluster Comput., vol. 22, no. Suppl 3, pp. 7241–7255, 2019.
[6] "Churn Prediction using Neural Network based Individual and Ensemble Models," M. Saghir, Z. Bibi, S. Bashir, and F. H. Khan, Proceedings of the 16th International Bhurban Conference on Applied Sciences and Technology, IBCAST 2019, 2019. 10.1109/IBCAST.2019.8667113 is the doi.
[7] "Churn prediction using machine learning and recommendations plans for telecoms," K. Ebrah and S. Elnasir, J. Comput. Commun., vol. 7, no. 11, pp. 33–53, 2019.
[8] "Customer Churn Prediction Modelling Based on Behavioural Patterns Analysis using Deep Learning," S. Agrawal, A. Das, A. Gaikwad, and S. Dhage, International Conference on Smart Computing and Electronic Enterprise, ICSCEE 2018, 2018. 10.1109/ICSCEE.2018.8538420 is the doi.
[9] A. Mishra and U. S. Reddy, "A Novel Approach for Churn Prediction Using Deep Learning," IEEE International Conference on Computational Intelligence and Computing Research, ICCIC 2017, 2018. 10.1109/ICCIC.2017.8524551 is the doi.
[10] Automated feature selection and churn prediction using deep learning models, V. Umayaparvathi and K. Iyakutti, Int. Res. J. Eng. Technol., vol. 4, no. 3, pp. 1846–1854, 2017.
[11] H. Sharma, S. Kumar, and others, "A survey on decision tree algorithms of classification in data mining," International Journal of Scientific Research, vol. 5, no. 4, pp. 2094–2097, 2016.
[12] P. Jayaswal, B. R. Prasad, D. Tomar, and S. Agarwal, "An ensemble approach for efficient churn prediction in telecom industry," International Journal of Database Theory Appl., vol. 9, no. 8, pp. 211–232, 2016.
[13] A. Alamsyah and N. Salma, "A comparative study of employee churn prediction model," 4th International Conference on Science and Technology (ICST), 2018, pp. 1-4.
[14] "Customer churn prediction in telecommunication industry using data certainty," J. Bus. Res., 2019, doi: 10.1016/j.jbusres.2018.03.003; A. Amin, F. Al-Obeidat, B. Shah, A. Adnan, J. Loo, and S. Anwar.
[15] A. Mishra and U. S. Reddy, "A comparative study of customer churn prediction in telecom industry using ensemble based classifiers," ICICI 2017, 2018, Proceedings of the International Conference on Inventive Computing and Informatics. 10.1109/ICICI.2017.8365230 is the doi.
[16] "Application of customer churn prediction based on weighted selective ensembles," G. E. Xia, H. Wang, and Y. Jiang, 3rd International Conference on Systems and Informatics, ICSAI 2016, 2017. 10.1109/ICSAI.2016.7811009 is the doi.
[17] I. Ullah, B. Raza, A. K. Malik, M. Imran, S. U. Islam, and S. W. Kim, "A Churn Prediction Model Using Random Forest: Analysis of Machine Learning Techniques for Churn Prediction and Factor Identification in Telecom Sector," IEEE Access, 2019, doi: 10.1109/ACCESS.2019.2914999 [18] Routhu, K., Bodepudi, V., Jha, K. M., & Chinta, P. C. R. (2020). Enhancing Cybersecurity Protocols in Big Data Integrated ERP Systems using Deep Learning Architectures. accessible at SSRN 5102662.