IJMRT: Volume (6), Issue 5, 2024
International Journal of Multidisciplinary Research Transactions
(A Peer Reviewed Journal)
www.ijmrt.in

ISSN (Print):2663-2381
ISSN (Online):2663-4007

Mobile Communication Services: Machine Learning-Based Customer Churn Prediction

K Mohanappriya[1*], C Vennila[2]

[1]Assistant Professor, Department of Computer Science and Engineering, Saranathan College of Engineering, Trichy, India.
[2]Professor, Department of Electronics and Communication Engineering, Saranathan College of Engineering, Trichy, India.

[*]Corresponding author

DoI: https://doi.org/10.5281/zenodo.12590630

Abstract

Customer churn is a significant challenge for large telecom firms, directly impacting their revenue. This paper presents a churn prediction model tailored for the telecom sector, utilizing novel approaches to feature engineering and selection through machine learning techniques applied on large datasets. Evaluation of the model's performance using the Area Under Curve (AUC) metric yielded a promising result of 93.3%. Specifically, the proposed model leverages algorithms such as Gradient Boosting, Extreme Gradient Boosting (XG Boost), and K- means clustering to achieve an impressive accuracy score of 95.74%. These findings underscore the efficacy of machine learning algorithms in accurately anticipating customer churn and provide valuable insights into its underlying causes. This project focuses on using Machine learning algorithms to predict customer churn in mobile communication services. It involves collecting and preprocessing a comprehensive dataset, performing feature engineering to enhance predictive power, and optimizing the model through training. Evaluation metrics such as accuracy, precision, recall, and F1 score are used to assess model performance. Feature importance analysis helps identify key factors influencing churn.

Keywords: Churn prediction, Machine learning, XG Boost algorithm.

1. Introduction

1.1. Machine Learning

Machine learning is a branch of artificial intelligence (AI) that focuses on developing algorithms and techniques that enable computers to learn from data and make predictions or decisions without being explicitly programmed. The core idea behind machine learning is to enable computers to learn from experience (i.e., data) and improve their performance over time [1]. There are several types of machine learning approaches:

Supervised Learning: In supervised learning, the algorithm learns from labeled data, where each example in the dataset is associated with a corresponding label or target variable. The goal is to learn a mapping from input features to output labels, allowing the algorithm to make predictions on new, unseen data. Examples of supervised learning algorithms include linear regression, logistic regression, decision trees, random forests, support vector machines (SVM), and neural networks.

Unsupervised Learning: Unsupervised learning involves learning from unlabeled data, where the algorithm must identify patterns, structures, or relationships within the data without explicit guidance. Clustering algorithms, such as k-means clustering and hierarchical clustering, and dimensionality reduction techniques, such as principal component analysis (PCA) and t-distributed stochastic neighbor embedding (t-SNE), are common unsupervised learning methods.

Semi-Supervised Learning: Semi-supervised learning combines elements of supervised and unsupervised learning by using a small amount of labeled data in conjunction with a larger amount of unlabeled data. This approach can be useful when labeled data is scarce or expensive to obtain.

Reinforcement Learning: Reinforcement learning involves training an agent to interact with an environment and learn optimal actions through trial and error. The agent receives feedback in the form of rewards or penalties based on its actions, allowing it to learn a policy that maximizes cumulative rewards over time. Applications of reinforcement learning include game playing, robotics, and autonomous vehicle control.

Deep Learning: Deep learning is a subfield of machine learning that focuses on training deep neural networks with multiple layers of interconnected nodes (neurons). Deep learning has achieved remarkable success in various domains, including computer vision, natural language processing, speech recognition, and recommendation systems.

By leveraging the power of machine learning, organizations can extract valuable insights from data, automate decision-making processes, and develop intelligent systems that improve efficiency, productivity, and innovation across various industries.

1.2. Deep Learning System

A deep learning system refers to a type of artificial neural network (ANN) architecture that consists of multiple layers of interconnected nodes (neurons), enabling the system to learn complex patterns and representations directly from raw data. Deep learning has emerged as a powerful approach for solving a wide range of tasks in fields such as computer vision, natural language processing, speech recognition, and reinforcement learning [2]. Here are some key components and characteristics of a deep learning system:

Neural Network Architecture: Deep learning systems typically consist of multiple layers of neurons, including input, hidden, and output layers. Each layer performs transformations on the input data, with deeper layers learning increasingly abstract and hierarchical representations of the data. Common architectures include convolutional neural networks (CNNs) for image processing, recurrent neural networks (RNNs) for sequential data, and deep feedforward networks for general-purpose tasks.

Feature Learning: Deep learning systems automatically learn relevant features and representations directly from the data, eliminating the need for manual feature engineering. This ability to extract hierarchical features enables deep learning models to achieve superior performance on tasks such as image classification, object detection, and speech recognition.

Activation Functions: Activation functions introduce non- linearity into the neural network, enabling the model to learn complex mappings between inputs and outputs. Common activation functions include sigmoid, tanh, ReLU (Rectified Linear Unit), and variants like Leaky ReLU and ELU (Exponential Linear Unit).

Regularization Techniques: To prevent overfitting and improve generalization, various regularization techniques are applied to deep learning models. These include dropout, L1 and L2 regularization, batch normalization, and data augmentation.

Hyperparameter Tuning: Deep learning models often have many hyperparameters that need to be tuned to achieve optimal performance. These include learning rate, batch size, network architecture, number of layers, layer sizes, and regularization parameters. Hyperparameter tuning techniques such as grid search, random search, and Bayesian optimization are used to find the best combination of hyperparameters.

Model Evaluation: Deep learning models are evaluated on performance metrics specific to the task at hand, such as accuracy, precision, recall, F1-score, mean squared error (MSE), or others. Evaluation is typically performed on a held-out validation set or through cross-validation to assess the model's generalization performance.

1.3. Exploratory Data Analysis

Exploratory Data Analysis (EDA) is an essential step in the data analysis process that involves examining and understanding the characteristics, patterns, and relationships present in a dataset. EDA helps analysts gain insights into the data, identify trends, anomalies, and potential relationships between variables, and inform subsequent steps in the analysis pipeline [3]. Here are some key aspects of exploratory data analysis are shown in fig 1.

Data Inspection: Begin by inspecting the dataset to understand its structure, size, and format. Check for missing values, outliers, or inconsistencies in the data. Understand the meaning and type of each variable (e.g., numerical, categorical, ordinal) and their distributions.

Summary Statistics: Compute summary statistics such as mean, median, mode, standard deviation, minimum, maximum, and percentiles for numerical variables. For categorical variables, calculate frequency counts and proportions for each category. These statistics provide an overview of the central tendency, spread, and variabilityof the data.

Data Visualization: Visualize the data using various graphical techniques to uncover patterns and relationships. Common visualization methods include histograms, box plots, scatter plots, bar charts, pie charts, and heatmaps. Visualization helps identify trends, clusters, outliers, and distributions in the data that may not be apparent from summary statistics alone.

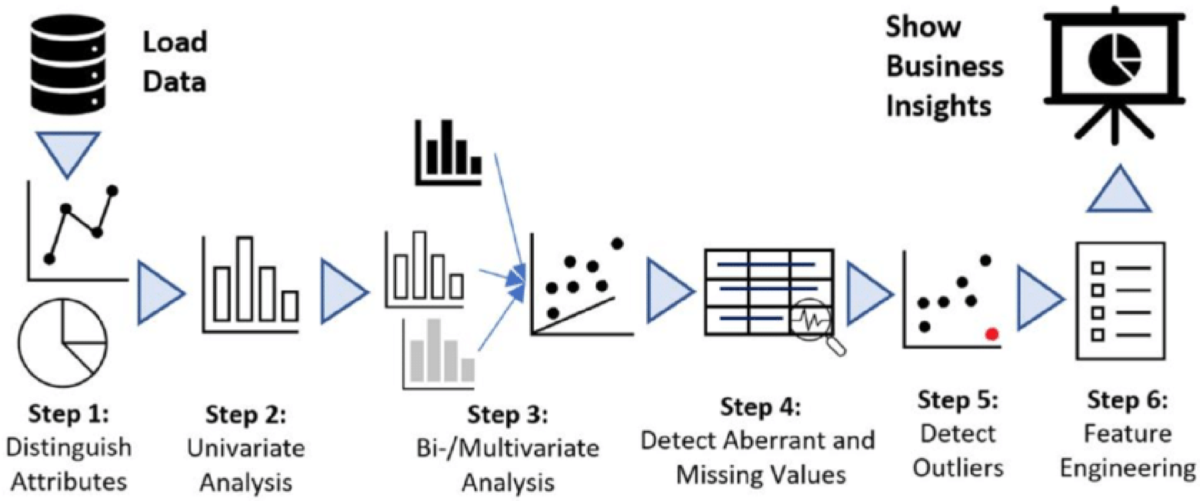


Figure.1. Exploratory Data Analysis

Correlation Analysis: Explore correlations between pairs of variables to understand their relationships and dependencies. Calculate correlation coefficients such as Pearson correlation for numerical variables and use correlation matrices or heatmaps to visualize correlations among multiple variables. Correlation analysis helps identify potential predictors or confounding variables for further investigation.

Data Cleaning and Preprocessing: Address missing values, outliers, and inconsistencies in the data through techniques such as imputation, outlier detection, and data transformation. Standardize or normalize numerical variables to ensure they have similar scales. Handle categorical variables by encoding them appropriately for modeling.

Hypothesis Testing: Formulate hypotheses about the relationships between variables based on exploratory findings and domain knowledge. Conduct statistical tests such as t-tests, chisquare tests, ANOVA, or non- parametric tests to evaluate the significance of observed differences or associations in the data.

Interactive Exploration: Use interactive tools and dashboards to explore the data dynamically, allowing for iterative exploration and drill-down analysis. Interactive visualization libraries such as Plotly, Bokeh, or Tableau enable users to explore and interact with data visually.

Document Findings: Document key insights, observations, and decisions made during the exploratory analysis process. Summarize important findings, trends, and patterns discovered in the data. This documentation serves as a reference for subsequent analysis steps and helps communicate findings to stakeholders.

Overall, exploratory data analysis plays a crucial role in understanding the characteristics and properties of a dataset, guiding subsequent analysis steps, and informing decision-making processes in data-driven projects.

1.4. Customer Churn

Customer churn prediction is a vital component in the arsenal of customer relationship management strategies for businesses across various industries, particularly in sectors like telecommunications, subscription-based services, and e- commerce. It refers to the process of identifying customers who are likely to discontinue their engagement with a company's products or services. By leveraging historical data on customer behavior, interactions, and demographics, along with advanced analytics techniques such as machine learning, businesses can develop predictive models to anticipate and mitigate churn effectively. The ability to forecast customer churn empowers organizations to proactively intervene with targeted retention efforts, personalized offers, and improved customer experiences, thereby maximizing customer lifetime value and sustaining business growth. In an increasingly competitive marketplace where acquiring new customers can be costly, accurate churn prediction enables businesses to prioritize resources, optimize marketing strategies, and foster stronger customer relationships, ultimately driving long-term profitability and sustainability.

1.5. Motivation and Contribution

The main objective on determining the effectiveness of the factors, i.e. lower and upper distance between the samples, are considered by the proposed model for the CCP. Further, we demonstrate a novel solution pertaining to the telecommunication sector showing the hidden factors considered for predicting the customer churn. Finally, we investigate the effects of both types of samples: those samples that are low distance and the upper distance (in terms of relevance) to the majority samples in given publicly available dataset. As a result of the study, we found that lower distance test set (LDT) samples have obtained best performance as compare to upper distance test set (UDT) samples in term of increased in the accuracy, f-measures, precision and recall when the uncertain sample size increases. Because the classification performance on upper distance samples remain almost the same when the size of samples increased in the test set. The CCP is binary classification problem where all the customers are divided into two possible behaviors:

- Churn, and
- Non-Churn.

Further, the churn behavior can be classified into the following sub-categories (a) voluntary customer churn, in which a customer decides to leave the service or even company, and (b) involuntary customer churn, in which the company or service provider decides to terminate a contract with customer. This study only addresses the voluntary customer churns due to difficulty in predicting the later type of customer churn, and also because it is easier to filter out the voluntary customer churn by simple queries. On the other hand, the literature revealed that existing studies have been published but still there is no agreement on choosing the best approach to handle CCP problem.

2. Problem Definition

2.1. Churn Prediction

Customer churn is a major problem and one of the most important concerns for large companies. Due to the direct effect on the revenues of the companies, especially in the telecom field, companies are seeking to develop means to predict potential customer to churn. Therefore, finding factors that increase customer churn is important to take necessary actions to reduce this churn. The main contribution of our work is to develop a churn prediction model which assists telecom operators to predict customers who are most likely subject to churn. The model developed in this work uses machine learning techniques on big data platform and builds a new way of features' engineering and selection. In order to measure the performance of the model, the Area Under Curve (AUC) standard measure is adopted, and the AUC value obtained is 93.3%. Another main contribution is to use customer social network in the prediction model by extracting Social Network Analysis (SNA) feature. Customer churn refers to the phenomenon of customers ceasing to do business with a company or brand. In the context of the telecommunication industry, churn occurs when customers end their subscription with a particular telecom service provider and switch to another provider or terminate the service altogether. Churn can be a significant problem for companies as it can result in lost revenue, decreased market share, and increased costs associated with acquiring new customers to replace those who have left. Customer churn is a crucial problem in the telecommunication sector, and it has been the focus of research in recent years. According to Coussement. Many dynamic and competitive businesses within a marketplace view their consumers as one of their most valuable assets. At an incredibly fast rate, telecom companies produce a large amount of data in the modern world. Numerous telecom service providers are vying for more customers by competing in the market. The competition in the telecommunication sector is intense, and telecom companies need to retain their customers to remain profitable.

2.2. Current Framework

In the existing system a normal customer churn analysis can be done with emotion prediction. When it comes to a business environment, customer attrition simply refers to customers switching services. Subscriber churn or customer churn is similar to attrition, when a customer switches from one service provider to another anonymously. In machine learning terms, churn prediction is a supervised (i.e. labeled) problem: Given a predetermined forecast horizon, one goal is to predict the number of subscribers that will churn over that time frame. Churn Prediction identifies churners in advance, before they leave the network. Therefore, the CRM department is able to prevent subscribers from churning in the future by implementing retention policies that attract and retain likely churners. Thus, the company would not suffer a potential loss. A mobile subscriber's past calls, along with his or her personal details and business information, are inputs into this problem. A list of churners is also provided for the training phase. When a model has been trained to the highest level of accuracy, it must be able to predict the churners from the real dataset which does not include any churn labels. The knowledge discovery process categorizes this problem as predictive modeling or data-mining.

The Demerits in existing system

- Emotion based analysis cannot be done in this system.
- The identification of the emotions are kept unknown.
- Customer churn analysis is completely low.

2.2.1. Customer Churn Prediction Using Data mining for an Iranian Payment Application

Customer Relationship Management (CRM) and data-driven marketing have become of paramount importance in this age of evolved markets and fierce competition among businesses. One of the most important branches of CRM is retaining existing customers. Since customer acquisition is about 5 to 6 times more costly than retaining customers, achieving an accurate model for customer churn prediction is essential to devise marketing retention strategies. Therefore, in this study, ensemble models are proposed to predict customer churn. Since customer churn is a rare occurrence in an organization and causes an imbalanced distribution in the target variable, ensemble learning algorithms, one of the most efficient and widely used methods, have been used to deal with this problem[7]. With regard to the case study, the dataset was generated on demographic and 13-month transactions of users of an Iranian payment application. In this study, the best model to predict customer churn is the bagging version of Decision Tree, reaching the highest accuracy, f-measure and AUC.

2.2.2. LSTM Model to Predict Customer churn in Banking sector with smooth Data preprocessing

In any organization, the technique used to acquire new users is by Customer Relationship Management systems. In order to achieve more profitability with increase in customer retention is by maintaining a healthy association with them. Customer Churn is also known as Customer Loyalty or Retention. The inspiration behind churn forecast is to categorize and discover clients into churner & non- churner. A churned client implies there is a greater chance of the client is around to take off from the organization. A novel software can be utilized to discover the clients who will donate increased benefits for the organization. Moreover churn forecast can maintain a strategic distance from the misfortune of income by holding the existing clients [3]. A few procedures are accessible for churn prediction with ensemble and hybrid models. This paper points to anticipate client Churn in banking sector with LSTM model and the data is preprocessed using SMOTE technique to overcome imbalanced information. The work is an extension to predicting customer loyalty in banking sector using Mixed Ensemble and Hybrid model. This paper proposes an accurate way to predict customer churn using LSTM model and the data is preprocessed using SMOTE technique. In this way the framework is more valuable for organizations to discover the clients with more chances to become churn. The results of the evaluation indicated that this is to be the case, the proposed systems for churn prediction performs with an accuracy of 88% and which is much better than the system without SMOTE technique.

2.2.3. Enhancing Customer Churn Prediction in Digital Banking using Ensemble Modeling

There are many purchases in the digital banking platform occurred daily. The electronic banking contains various transactions with different purchase behavior. Customer attrition from a digital store to another has become a challenge to the business owners. So, Businesses should measure their customer churn rate at regular intervals, as it is an important metric. Digital Banks started in building intelligent models to increase the customers satisfaction. Customer churn prediction (CCP) is an important strategy involved in the customer relationship management (CRM) strategies to forecast the probability of their attrition. This paper presents a classifier-based model to predict customer churn behavior based on the customer profile data [4]. In this research, several supervised classification algorithms have been applied and compared like, KNN (k-Nearest Logistic Regression, AdaBoost, Gradient Boosting and Random Forest. A unified voting is next applied for the set of classifiers in order to produce a high prediction accuracy. The model performance is enhanced by applying hyperparameters optimization and tuning. The most important features are defined and ranked using the Random Forest model. In the presented experimental case study with a dataset of size 10K, the model achieved an accuracy value 87%.

2.2.4. Predicting Customer churn Prediction in Telecom Sector using various Machine Learning Techniques

Customer churn analysis and prediction in telecom sector is an issue now a days because it's very important for telecommunication industries to analyze behaviors of various customer to predict which customers are about to leave the subscription from telecom company. So data mining techniques and algorithm plays an important role for companies in today's commercial conditions because gaining a new customer's cost is more than retaining the existing ones [2]. In this paper we can focuses on various machine learning techniques for predicting customer churn through which we can build the classification models such as Logistic Regression, SVM, Random Forest and Gradient boosted tree and also compare the performance of these models.

2.2.5. Warning Model of Customer Churn based on Emotions

Emotion is an important catalyst which affects a customer in the process of buying service, it shows satisfaction of customers for goods and services, and changes the loyalty of customers to the enterprise in the future. Customer complaints often reflect the quality of service objectively and customer emotion. Therefore, the paper use customer complaints of information service enterprise for data analysis, and try to establish the warning model of customer churn based on emotions. Churn prediction is a crucial aspect of customer relationship management, particularly in industries where customer retention is essential for sustainable growth, such as telecommunications, subscription services, and finance. Churn refers to the phenomenon where customers cease their engagement with a company's products or services. Predicting churn involves leveraging data analytics and machine learning techniques to identify patterns and indicators that precede customer defection. By analyzing historical customer data including usage behavior, transaction history, demographic information, and interactions with the company, businesses can build predictive models to forecast which customers are at risk of churning [5]. These models allow organizations to proactively intervene with targeted retention strategies, such as personalized offers, loyalty programs, or proactive customer service interventions, to prevent customer attrition. Ultimately, churn prediction enables businesses to optimize their customer retention efforts, minimize revenue loss, and maximize the lifetime value of their customer base.

2.3. Proposed System

This study proposes to develop a customer churn prediction model using machine learning techniques to predict customer churn in the telecommunication sector. In this research, some of the addressed questions will be; analysis of the most important feature for customer churn, which type of customers are leaving more, and which machine learning model is the best one for result analysis and prediction. We explored classification techniques, compared their accuracy, as well as other metrics, precision, recall, f1-score, True/False Positive Rates. Data Preprocessing checks for missing values, correlated variables, and outliers; AUC and LGBM for hypothesis generation; data scaling to improvise data accuracy; train and test dataset generation; training models for cross-validation and plotting data accuracy results from test data. This study only addresses the voluntary customer churns due to difficulty in predicting the later type of customer churn, and also because it is easier to filter out the voluntary customer churn by simple queries. On the other hand, the literature revealed that existing studies have been published but still there is no agreement on choosing the best approach to handle CCP problem. To the best of our knowledge, there is no state-of-the-art study to focus on the uncertain samples for building CCP model. The Merits of Proposed System

- Identifies the multiple model of customer churn
- Accuracy of the prediction is high
- Analysis over multiple churn reasons are verified

3. Modules Description

The Data Flow Diagram describes the work in detail to predict the Customer Churn in Fig 2

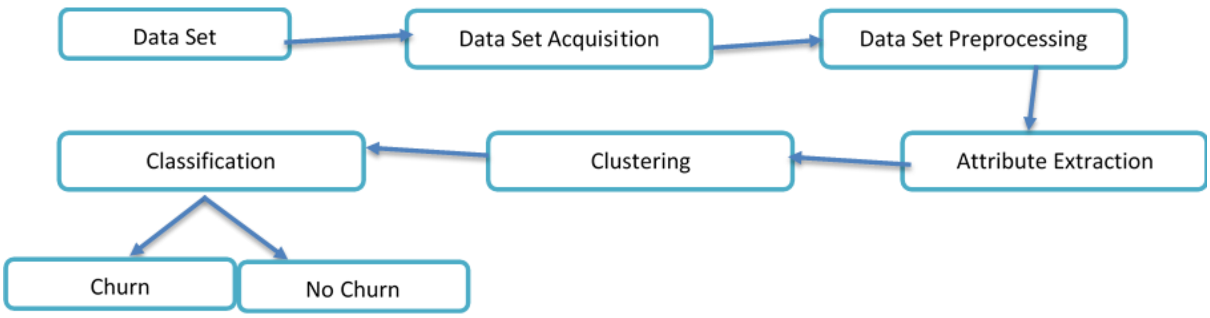


Figure.2. Data Flow Diagram

3.1. Dataset Access

The data used in this research is collected from multiple systems and databases. Each source generates the data in a different type of files as structured, semi-structured (XML-JSON) or unstructured (CSV-Text). Dealing with these kinds of data types is very hard without big data platform since we can work on all the previous data types without making any modification or transformation. By using the big data platform, we no longer have any problem with the size of these data or the format in which the data are represented.

3.2. Preprocessing

The generated dataset was unbalanced since it is a special case of the classification problem where the distribution of a class is not usually homogeneous with other classes. Te dominant class is called the basic class, and the other is called the secondary class. Te data set is unbalanced if one of its categories is 10% or less compared to the other one. Although machine learning algorithms are usually designed to improve accuracy by reducing error, not all of them take into account the class balance, and that may give bad results In general, classes are considered to be balanced in order to be given the same importance in training.

3.3. Missing Values

There is a representation of each service and product for each customer. Missing values may occur because not all customers have the same subscription. Some of them may have a number of services and others may have something different. In addition, there are some columns related to system configurations and these columns have only null value for all customers. Missing and incorrect values are prevalent in the data set. The entire dataset was analyzed and only the most useful features were listed. By listing features, the listing will be more accurate and contain only useful features. For a knowledge-based approach to data selection, feature selection is a crucial step. Out of the dataset here, we chose the features necessary for improving performance and helpful for decision-making, while the rest of the features have less importance.

3.4. Feature Extraction

Explorative Data Analysis (EDA) provides a clear and better understanding of data patterns and potential hypothesis. The distribution of feature is an essential for trend analysis of dataset. The unwanted feature attributes are developed with unwanted analysis of the data can be predicted for a knowledge-based approach to data selection, feature selection is a crucial step. Out of the dataset here, we chose the features necessary for improving performance and helpful for decision-making, while the rest of the features have less importance.

3.5. Clustering

The k-means clustering is a method of vector quantization, originally from signal processing, that aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean (cluster centers or cluster centroid), serving as a prototype of the cluster. This results in a partitioning of the data space into Voronoi cells. k-means clustering minimizes within-cluster variances (squared Euclidean distances), but not regular Euclidean distances, which would be the more difficult Weber problem: the mean optimizes squared errors, whereas only the geometric median minimizes Euclidean distances. For instance, better Euclidean solutions can be found using k- medians and k-medoids. The problem is computationally difficult (NP-hard); however, efficient heuristic algorithms converge quickly to a local optimum. These are usually similar to the expectation-maximization algorithm for mixtures of Gaussian distributions via an iterative refinement approach employed by both k- means and Gaussian mixture modeling. They both use cluster centers to model the data; however, k-means clustering tends to find clusters of comparable spatial extent, while the Gaussian mixture model allows clusters to have different shapes.

3.6. Classification

Many customers have been with the telecom company just for a month, while many others have been with the company for about 72 months. Different contracts may apply to different customers. Due to this, depending on the contract they are in, it might be easier or harder for customers to stay or leave the telecom company. Several extremely important parameters for predictive churn analysis were included in the dataset, and the data is extremely large. 7043 instances of 21 attributes are contained in the dataset. Features include details about demographic information like gender, age, and dependents, services they have signed up for, contract information, payment methods, paperless billing, monthly charges, and a variable in which we anticipate which customers have left within the past month. Input data is in CSV format and visualized using various visual elements such as graphs, helping to identify trends, outliers, and patterns in the data. The analysis starts with data cleaning followed by graphical analysis, machine learning model, estimation and result analysis. To find the uncertain samples in the given dataset, first we calculated the distance of each instance from every other instance of the dataset using Manhattan distance formula. The Manhattan distance is the sum of absolute differences between points.

4. Implementation Work

4.1. Initial Implementation Work

Data Collection: Gather relevant data sources, including customer demographics, transaction history, usage patterns, customer interactions, and churn labels shown in fig 3 (indicating whether a customer has churned or not).

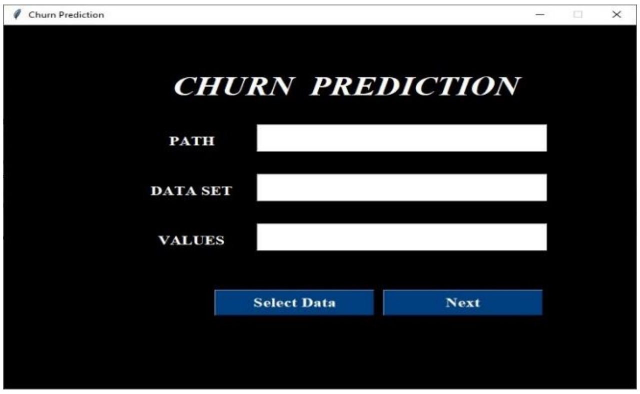


Figure. 3. Data Collection

Data Preprocessing: Clean the data by handling missing values, removing duplicates, and encoding categorical variables. Perform feature engineering to extract meaningful features that capture customer behavior and characteristics shown in fig 4.

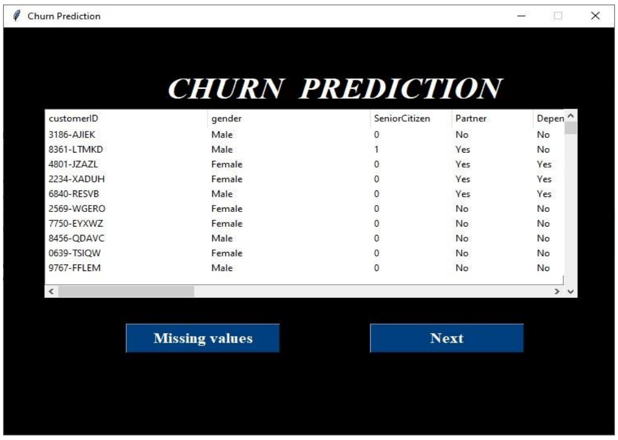


Figure.4. Data Preprocessing

Exploratory Data Analysis (EDA): Explore the data to understand patterns, correlations, and insights that may influence customer churn. Visualize key metrics and relationships between variables to gain insights into customer behavior.

Feature Selection: Select the most relevant features for predicting churn using techniques such as correlation analysis, feature importance ranking, or dimensionality reduction methods.

Model Selection: Choose appropriate machine learning algorithms for building the churn prediction model. Commonly used algorithms include logistic regression, decision trees, random forests, gradient boosting and neural networks.

Model Training: Split the dataset into training and validation sets. Train the chosen machine learning models on the training data using techniques such as cross- validation to optimize hyperparameters and prevent overfitting.

Model Evaluation: Evaluate the performance of the trained models using appropriate evaluation metrics such as accuracy, precision, recall, F1-score, and ROC-AUC. Compare the performance of different models to select the best-performing one.

Model Deployment: Deploy the chosen model into the production environment, integrating it with the company's systems for real-time or batch prediction of customer churn. Implement monitoring and feedback mechanisms to track model performance and retrain the model periodically with new data.

Integration with Business Processes: Integrate the churn prediction system with existing business processes and customer relationship management (CRM) systems. Develop workflows and processes for acting on churn predictions, such as targeted marketing campaigns, retention offers, or customer outreach efforts.

Continuous Improvement: Continuously monitor and update the churn prediction system based on new data, evolving customer behavior, and changes in business conditions. Iterate on the model to improve accuracy and effectiveness over time.

4.2. Gradient Boost Algorithm

Implementing the Gradient Boosting algorithm involves several key steps:

Data Preparation: Prepare your dataset, ensuring it is clean and properly formatted. Split the dataset into training and testing sets.

Initialize Model: Choose a base model (often a decision tree, but can be any other model). Initialize the model with a constant value, usually the mean of the target variable for regression or the log-odds of the target variable for classification.

Calculate Residuals: Calculate the residuals (the difference between the predicted values and the actual values) for the initial model.

Build Trees (Weak Learners): Fit a decision tree (or other weak learner) to the residuals. This tree will learn to predict the residuals. Train the tree on the features and the residuals. The number of trees (or weak learners) is a hyperparameter that you can tune.

Update Model: Update the model by adding the predictions of the newly trained tree to the previous model's predictions. This step involves adding a fraction (learning rate) of the predictions made by the newly trained tree to the existing model's predictions.

Update Residuals: Calculate new residuals by subtracting the updated predictions from the actual values. These new residuals represent the errors that the current model couldn't capture.

Repeat Steps 4 to 6: Repeat steps 4 to 6 for a specified number of iterations (or until a certain condition is met). Each iteration builds a new tree to predict the residuals of the previous model and updates the overall model.

Final Prediction: After all iterations are completed, the final prediction is obtained by summing the predictions of all the trees in the ensemble.

Evaluation: Evaluate the performance of the gradient boosting model on the testing set using appropriate evaluation metrics such as Mean Squared Error (MSE) for regression or accuracy, precision, recall, and F1-score for classification.

Hyper parameter Tuning: Tune hyperparameters such as the number of trees, learning rate, tree depth, and others using techniques like grid search or random search to optimize the model's performance.

Validation: Validate the model's performance using techniques such as cross-validation to ensure its generalizability to unseen data.

5. Result and Discussion

Our paper proves that using computer programs like K- Means clustering and XG Boost classification is really good at guessing when customers might stop using a mobile service. By looking at past customer data, we could predict who might leave soon with high accuracy. This helps phone companies find customers who might leave and try to keep them happy with special offers or better service. If they can keep more customers happy, they can make more money and stay successful. Overall, our project shows that using these computer programs can really help phone companies keep their customers and make more money shown from fig 5 to fig 5.11

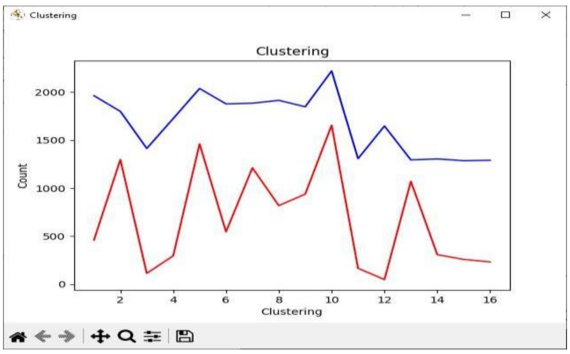


Figure.5. Clustering Values

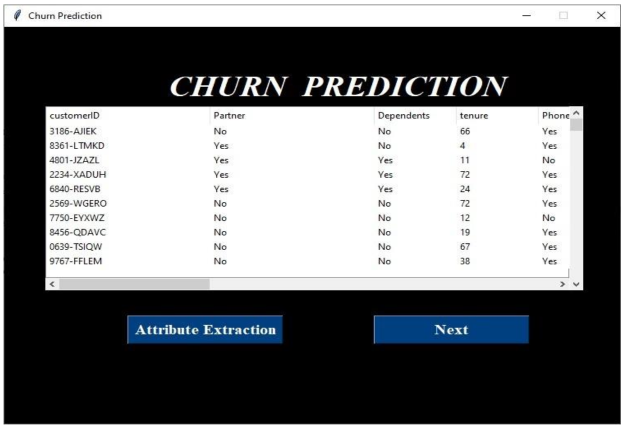


Figure.6. Attribute Selection

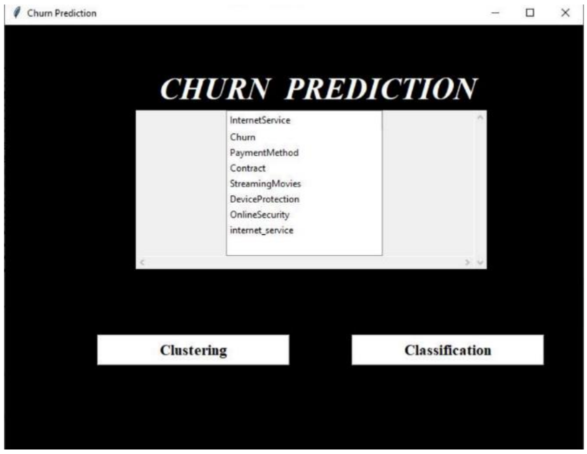


Figure.7. Clustering System

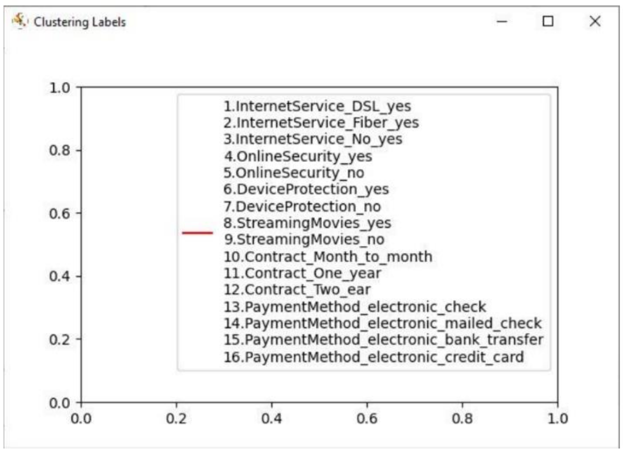


Figure.8. Classification Instances

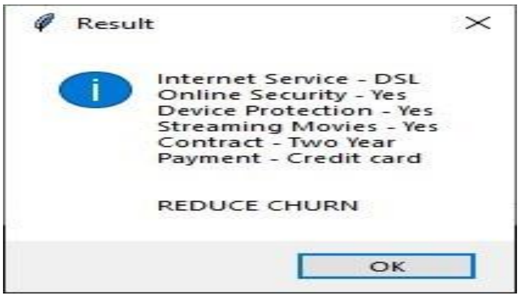


Figure.9. Result

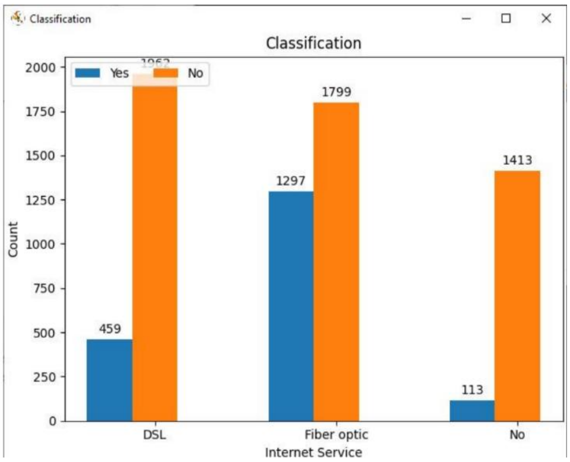


Figure.10. Internet Service

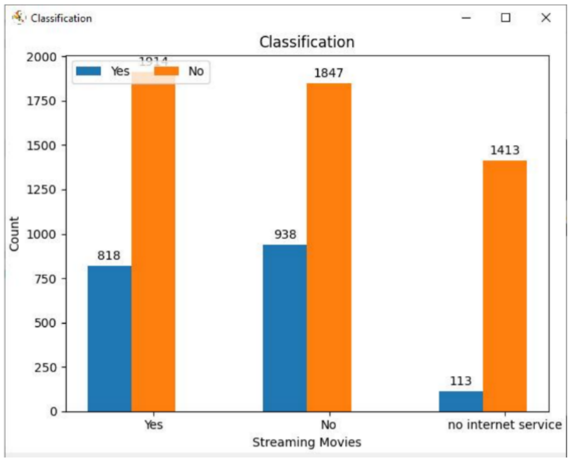


Figure.11. Streaming Movie

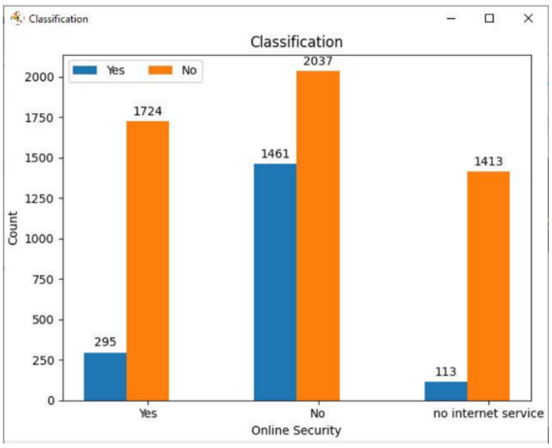


Figure.12. Online Security

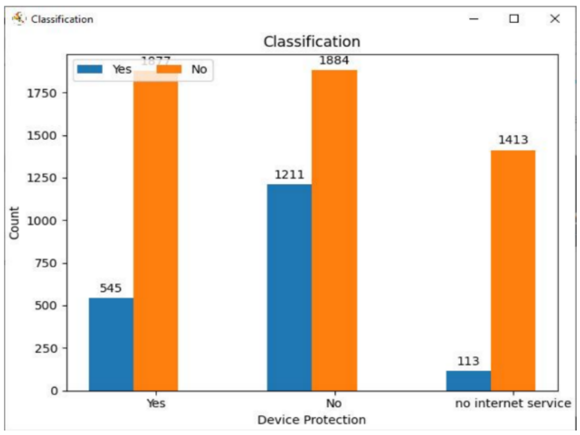
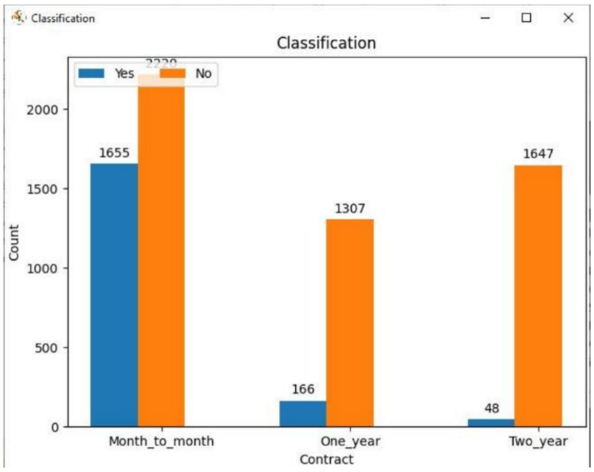


Figure.13. Device Protection



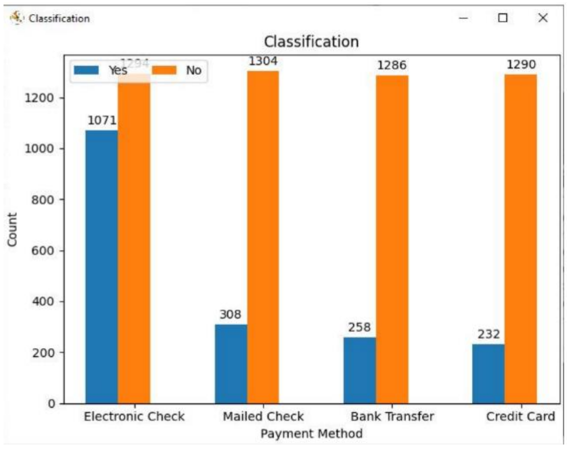


Figure.14. Contract

Figure.15. Payment Method

6. Conclusion

In conclusion, our project focused on developing a machine learning-based churn prediction system for the mobile communication services sector. By leveraging advanced techniques such as K-Means clustering and XG Boost classification, we aimed to forecast the likelihood of customer churn based on historical data encompassing usage patterns, demographics, and service interactions. Through our analysis, we demonstrated the efficacy of machine learning algorithms in predicting customer churn, providing valuable insights for mobile communication service providers to proactively identify at-risk customers and implement targeted retention initiatives. By integrating predictive analytics into business strategies, telecom companies can foster customer loyalty, enhance revenue sustainability, and optimize operational efficiency in a competitive market environment. Looking ahead, there are numerous opportunities for future enhancements, including the integration of external data sources, dynamic model updating, and personalized retention strategies. These improvements will further enhance the accuracy and effectiveness of churn prediction systems, enabling telecom companies to stay ahead of customer churn and drive long-term business success. In summary, our project underscores the importance of data-driven approaches in mitigating customer churn and highlights the potential for machine learning to revolutionize customer retention strategies in the mobile communication services sector. By leveraging predictive analytics and innovative technologies, telecom companies can build stronger relationships with customers, reduce churn rates, and thrive in a rapidly evolving market landscape.

REFERENCES

[1]. Trupti S. Gaikwad; Snehal A. Jadhav; Ruta R. Vaidya; Snehal H. Kulkarni. "Machine learning amalgamation of Mathematics, Statistics and Electronics". International Research Journal on Advanced Science Hub, 2, 7, 2020, 100-108. doi: 10.47392/irjash.2020.72

[2]. Abhishek and Ratnesh ,"Predicting Customer Churn Prediction in Telecom Sector Using Various Machine Learning Techniques", In the proceedings of 2017 International Conference on Advanced Computation and Telecommunication, Bhopal, India, 2017.

[3]. Abinash and Srinivasulu U ,"Machine Learning techniques applied to prepaid subscribers: case study on the telecom industry of Morocco", In the proceedings of 2017 International Conference on Inventive Computing and Informatics , Coimbatore, India, pp. 721-725, 2017.

[4]. Alae and El Hassane , "A Comparative Study of Customer Churn Prediction in Telecom Industry Using Ensemble Based Classifiers", In the proceedings of Intelligent Systems and Computer Vision , Fez, Morocco, 2017.

[5]. Salini Suresh; Suneetha V; Niharika Sinha; Sabyasachi Prusty; Sriranga H.A. "Machine Learning: An Intuitive Approach In Healthcare". International Research Journal on Advanced Science Hub, 2, 7, 2020, 67-74. doi: 10.47392/irjash.2020.67

[6]. Anuj and Prabin ,"A Neural Network based Approach for Predicting Customer Churn in Cellular Network Services", International Journal of Computer Applications, Volume 27, No.11, pp. 26-31, 2011.

[7]. Balasubramanian M, and Selvarani M ,"Churn Prediction in Mobile Telecom System using Data Mining Techniques ", International Journal of Scientific and Research Publications, Volume 4, Issue 4, pp. 1-5, 2014.

[8]. Irfan Ullah , Basit Raza, Ahmad Kamran Malik , Muhamad Imran , Saif Ul Islam and Sung Won Kim., "A Churn Prediction Model Using Random Forest: Analysis of Machine Learning Techniques for Churn Prediction and Factor Identification in Telecom Sector", In the proceedings of IEEE Access, vol. 07, no. 2169-3536, pp. 60134 - 60149, 2019.

[9]. Varshaa a; Vinitha V; Usha Nandhini D; Yogeshwaran R; Soundharya B M. "Artificial intelligence and its applications- A Review". International Research Journal on Advanced Science Hub, 1, 2, 2020, 1-4. doi: 10.47392/irjash.2019.11

[10]. Kavitha V, Hemanth G , Mohan S.V and Harish M , "Churn Prediction of Customer in Telecom Industry using Machine Learning Algorithms", International Journal of Engineering Research & Technology(2278-0181), Vol. 9, Issue 05, pp. 181-184, 2020.

