

Intelligent Customer Retention: Using Machine Learning for Enhanced Prediction of Telecom Customer Churn

Milestone 1: Define Problem / Problem Understanding

Activity 1: Specify the business problem

Business problem related to customer churn in the telecom industry is to identify the main reasons why customers are leaving and develop effective strategies to prevent them from churning. Companies need to analyze customer data and feedback to identify the common reasons why customers are leaving, such as pricing, service quality, or package offerings, and then develop targeted strategies to address these issues and retain customers. Machine learning algorithms can assist in identifying patterns and trends in customer data that can be used to develop effective churn prevention strategies

Activity 2: Business requirements

Accurate Prediction Model: The main business requirement for a telecom company is to develop an accurate churn prediction model. The model should be able to identify the customers who are most likely to churn and provide insights into the reasons for their churn. This information can be used by the company to develop targeted retention strategies

Activity 3: Literature Survey

Understanding the factors behind customer churn is very important in any business. Different market domains have some similar factors like the price offered to the customer, the benefits provided, the number of years for which customer is entangled with the organization, etc. Earlier papers have used Data analysis to understand customer behaviors using the regression models. Some research papers have used clustering to segment the groups of the customer of similar behaviors. One of those research papers is "A Review on Customer Churn Prediction in Telecommunication Using Data Mining Techniques (S. Babu¹, Dr. N. R. Ananthanarayanan²)" where the authors have used the Data mining techniques to predict customer churn. Research on such behaviors helps companies to take the necessary steps to retain customers. In this paper, we are going to address specifically the telecommunications industry and categorize the factors which contribute to it.

M.A.H. Farquad proposed a hybrid approach to overcome the drawbacks of general SVM model which generates a black box model (ie., it does not reveal the knowledge gained during training in human understandable form). The hybrid approach contains three phases: In the first phase, SVM-RFE (**SVM-recursive feature elimination**) is employed to reduce the feature set. In the second phase, dataset with reduced features is then used to obtain SVM model and support vectors are extracted. In the final phase, rules are then generated using Naïve Bays Tree (**NB Tree which is combination of Decision tree with naïve Bayesian Classifier**). The dataset used here is bunk credit card customer dataset (Business Intelligence Cup 2004) which is highly unbalanced with 93.24% loyal and 6.76% churned

customers. The experimental showed that the model does not scalable to large datasets.

Irfan Ullah et al. Identified churn factors that are essential in determining the root causes of churn. By knowing the significant churn factors from customers' data, Customer Relationship Management (**CRM**) can improve productivity. Recommend relevant promotions to the group of likely churn customers based on similar behavior patterns, and excessively improve marketing.

Kavitha V et al. Used a Decision Tree, Random Forest, and XG Boost to predict the customers who are likely to cancel the subscription which can offer them better services and reduce the churn rate. By preprocessing and feature selection the data set for training and testing. For the above mentioned algorithm, it is necessary to do some feature engineering to have more efficient and accurate results.

Krishna Sai and Sasikala implemented an **EDA** using Visualization, statistical tests for feature selection and Data mining methods for predicting the likely churners by utilizing a Logistic Regression Model Here dataset has been analyses by using the data visualization techniques before entering into the modeling process.

Web Chin-Ping Wei and I-Tang Chiu

*proposed the churn prediction technique for customer retention analysis. The author used the decision tree approach C4.5 on customer call details. **Yi-Fan wang, Ding-An chlang and Mei-Hua Hsu** .discussed a Recommender system for customer churn by proposing a decision tree algorithm. Data used for the analysis has covered over 60,000 transactions and of more than 4000 members, over a period of three months. **Jadhav and Pawar** ,designed a decision support system using data mining Technique. The churn behavior of customers is Predicted in advance using this technique. The Authors have used Back propagation algorithm on a customer billing data. **Tomas Philip Rúnarsson, Ólafur Magnússon, Birgis Hrafnkelsson** , Constructed a churn prediction model that can Output the probabilities that customers will churn in the near future. In this paper the training data is used to build classifiers by using machine Learning methods. **N.Kamalraj and A.Malathi** Focused their research on the better*

Understanding of churn prediction using data mining techniques. Telecommunication industry can use this approach to customer retention activities within the context of their Customer Relationship Management efforts. The author uses the **DM technique** on customer details.

Activity 4: Social or Business Impact.

Social impact: By accurately predicting which customers are likely to churn, telecom companies can proactively address their concerns and offer targeted retention strategies. This can lead to increased customer satisfaction and loyalty, which can have a positive impact on the overall reputation of the company.

Business impact: By reducing customer churn, companies can increase their revenue and profitability. Retaining existing customers is often more cost-effective than acquiring new ones, so focusing on customer retention can lead to significant cost savings.