

Customer Churn Prediction

Thorat Anushka Prashant, Patil Saish Subash, Sayyed Sufiyaan Imtiyaz,
Sonavane Dhruv Suhas

anushkathorat6@gmail.com (Diploma last year KKW),

saishp83@gmail.com (Diploma last year KKW),

sayyedsufiyaan875@gmail.com (Diploma last year KKW),

dhruvsonavane45@gmail.com (Diploma last year KKW)

Abstract - Customer churn is a significant concern in the telecom industry as retaining existing customers is more cost-effective than acquiring new ones. Predicting customer churn is crucial for telecom companies to improve retention strategies. This study focuses on developing a churn prediction model using machine learning techniques to help telecom operators identify customers likely to churn. The model utilizes a large dataset containing customer information such as demographics, services, and payment methods. The model was built and tested in a Spark environment. Three machine learning algorithms were implemented: Decision Tree, Random Forest, and XGBoost. The goal is to assist telecom operators in predicting and managing customer churn, ultimately improving customer retention and business performance.

1. INTRODUCTION

The telecommunications industry faces increasing competition, making customer retention a key focus, particularly as companies enter the mature phase of their lifecycle. Customer churn, either accidental (due to economic factors) or intentional (due to better offers from competitors), is a significant challenge. Intentional churn, in particular, is the most concerning for companies. Businesses often use churn analysis to focus retention efforts on the most at-risk customers. This study aims to investigate churn among fixed-telephony subscribers using machine learning. By analyzing customer data, including attributes like gender, tenure, and service usage, three algorithms—Random Forest, Decision Tree, and XGBoost—were implemented to predict churn. The results of this study highlight critical factors influencing churn, offering telecom companies an automated solution for predicting and managing customer attrition.

A. Advantages

- ☐ Improved Customer Retention
- ☐ Increased Revenue
- ☐ Optimized Marketing Efforts
- ☐ Better Customer Insights
- ☐ Cost Efficiency

B. Problem statement

Customer churn leads to lost revenue and higher acquisition costs. Predicting churn helps businesses identify at-risk customers and take proactive measures to improve retention, but developing accurate models can be challenging and resource-intensive.

C. Aim and Objective

Aim

To develop a predictive model that accurately identifies customers at risk of churning, enabling businesses to take proactive steps to retain them and reduce revenue loss.

Objectives

1. Analyze customer behavior and identify key factors contributing to churn.
2. Develop a reliable churn prediction model using machine learning techniques.
3. Implement retention strategies based on predictive insights to improve customer loyalty.
4. Evaluate the model's performance and refine it for better accuracy and effectiveness.

2. MODULE DESCRIPTION

☐ **Data Preprocessing:** The module handles data cleaning, transformation, and feature engineering to prepare the dataset for analysis. It ensures that customer data, such as demographics, usage patterns, and transaction history, is properly structured.

□ **Churn Prediction Model:** Using machine learning algorithms (e.g., decision trees, logistic regression, random forests), this component builds a model to predict the likelihood of customer churn based on historical data.

□ **Insights Generation:** The module generates actionable insights by analyzing the factors that contribute to churn, such as customer behavior patterns, service dissatisfaction, or pricing issues.

- Use the model to generate churn probabilities and suggest actions for each customer (e.g., offer discounts, personalized outreach).

5. Customer Segmentation:

- Segment customers based on churn risk, engagement levels, and other relevant characteristics.
- Tailor retention strategies for different segments, ensuring targeted and effective interventions.

6.

3. PROPOSED SYSTEM

The **Customer Churn Prediction System** is designed to help businesses accurately identify customers at risk of leaving and implement proactive measures to retain them. The proposed system will combine data analytics, machine learning, and customer segmentation techniques to predict churn and optimize retention strategies.

Key Features of the Proposed System:

1. Data Collection and Integration:

- Collect customer data from multiple sources such as transaction history, customer service interactions, demographics, and product usage.
- Integrate data into a centralized database for easy access and processing.

2. Data Preprocessing and Feature Engineering:

- Clean and preprocess raw data by handling missing values, outliers, and inconsistencies.
- Engineer relevant features such as customer behavior patterns, engagement levels, and past interactions to improve model accuracy.

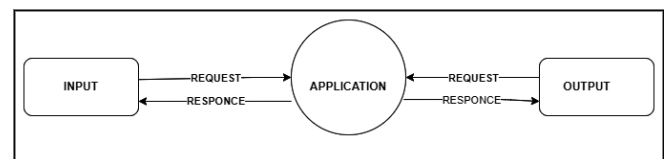
3. Churn Prediction Model:

- Utilize machine learning algorithms (e.g., Random Forest, Logistic Regression, Support Vector Machine) to develop a model that predicts the likelihood of customer churn.
- Train the model using historical data, ensuring it can generalize to future customer behaviors.

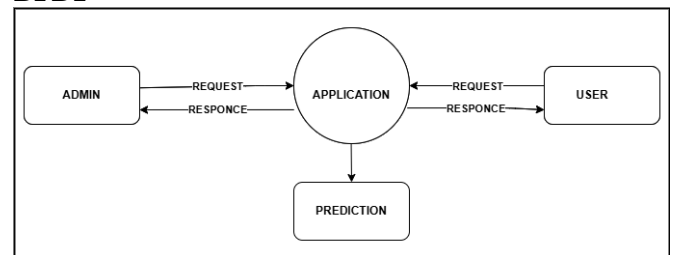
4. Real-time Prediction:

- Implement real-time prediction capabilities, allowing businesses to flag at-risk customers as they interact with the system.

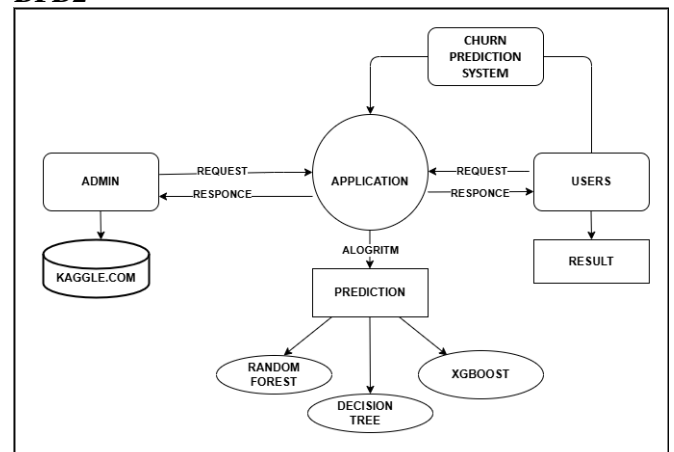
DFD Diagrams DFD0



DFD1



DFD2



7. CONCLUSIONS

The Customer Churn Prediction System provides businesses with a valuable tool to proactively identify customers at risk of leaving. By leveraging machine learning algorithms and data analytics, businesses can better understand customer behavior, anticipate churn, and implement targeted retention strategies. This not only helps in reducing customer attrition but also improves customer satisfaction and loyalty. With continuous monitoring and refinement, the system can adapt to changing customer dynamics, ensuring that businesses stay ahead of churn trends and maintain a strong customer base. Ultimately, the system helps businesses optimize resource allocation, increase profitability, and enhance long-term customer relationships, creating a significant competitive advantage.

6. REFERENCES

- 1) Md Sayedur Rahman, Md Sahrial Alam and Md Ikbal Hosen, "To Predict Customer Churn By Using Different Algorithms", The International Conference on Decision Aid Sciences and Applications, 2022.
- 2) P. Lalwani, M. K. Mishra, J. S. Chadha and P. Sethi, "Customer churn prediction system: a machine learning approach", *Computing*, vol. 104, no. 2, pp. 271-294, 2022.
- 3) J. K. Sana, M. Z. Abedin, M. S. Rahman and M. S. Rahman, "Data transformation based optimized customer churn prediction model for the telecommunication industry", 2022.
- 4) Prasanth Senthana, RMKT Rathnayaka, Banujan Kuhaneswaran and BTGS Kumara, "Development of Churn Prediction Model using XGBoost", IEEE International IOT Electronics and Mechatronics Conference (IEMTRONICS), 2021.
- 5) VLN RajaGopal Kesiraju and P. Deeplakshmi, "Dynamic Churn Prediction using Machine Learning Algorithms - Predict your customer through customer behaviour", International Conference on Computer Communication and Informatics (ICCCI), 2021.