Customer Churn Analysis for Telecom Industry

1. Introduction

In today's highly competitive telecom industry, retaining existing customers is more important than acquiring new ones. Customer churn refers to customers discontinuing their service and switching to competitors. High churn rates lead to revenue loss and increased operational costs. This project focuses on analyzing customer behavior and predicting churn using machine learning techniques to help telecom companies improve customer retention strategies.

2. Abstract

This project aims to predict customer churn in the telecom industry using machine learning. A structured dataset containing customer usage patterns such as call duration, recharge frequency, complaints, and network quality was analyzed. A Random Forest classification model was developed to identify customers likely to churn. The results help in segmenting customers and providing actionable recommendations to reduce churn and improve customer loyalty.

3. Tools Used

- Programming Language: Python
- Libraries: Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn
- Model: Random Forest Classifier
- IDE: Visual Studio Code
- Dataset Format: CSV

4. Steps Involved in Building the Project

Step 1: Data Collection

A telecom customer dataset containing usage and service-related attributes was used.

Step 2: Data Preprocessing

- Checked for missing values
- Filled missing numerical values using mean
- Selected relevant features
- Separated input variables and target variable (churn)

Step 3: Data Splitting

The dataset was divided into training and testing sets using an 80:20 ratio.

Step 4: Model Building

A Random Forest Classifier was trained on the dataset to predict whether a customer would churn or not.

Step 5: Model Evaluation

- Accuracy score
- Classification report (Precision, Recall, F1-score)
- Confusion matrix visualization

Step 6: Result Visualization

- Confusion matrix heatmap
- Feature importance bar chart to identify key churn factors

5. <u>Conclusion</u>

The Customer Churn Analysis project successfully predicts customers who are likely to leave the telecom service. Factors such as complaint count, recharge frequency, and network quality significantly influence churn behavior. The developed model can assist telecom companies in identifying at-risk customers early and taking proactive retention measures. This project demonstrates the practical application of data analytics and machine learning in solving real-world business problems.