Customer Churn Analysis Project

Abstract

This project focuses on analyzing and predicting customer churn using machine learning techniques. By identifying key patterns and reasons behind customer attrition, organizations can take proactive steps to enhance customer retention. The project integrates data cleaning, predictive modeling, and visualization through Power BI to provide a comprehensive analytical solution.

Introduction

Customer churn represents the percentage of customers who stop using a company's product or service over a given period. In a competitive business environment, reducing churn is crucial to maintaining profitability. This project demonstrates a complete end-to-end data pipeline from preprocessing to predictive analytics and visualization designed to understand and forecast customer churn behavior.

Tools Used

- Languages & Libraries: Python
  - pandas
  - numpy
  - sklearn
  - seaborn
  - matplotlib
  - pymysql
  - graphviz
- Database: MySQL
- Visualization: Power BI
- Machine Learning: Random Forest Classifier

Steps Involved in Building the Project

1. **Data Cleaning:**

   Removed missing values, renamed columns, and eliminated duplicates using Python and pandas.

2. **SQL Integration:**

   Loaded cleaned data into a MySQL database for structured storage and querying.

3. **Data Preprocessing:**

   Encoded categorical features and split the dataset into training and testing sets.

4. **Model Building:**

   Trained a Random Forest model to predict customer churn with high accuracy.

5. **Model Evaluation:**

   Evaluated model performance using accuracy score, confusion matrix, and feature importance.

6. **Visualization:**

   Built an interactive Power BI dashboard to visualize churn insights and customer behavior patterns.

Conclusion

This project successfully demonstrates how data-driven methods can predict and explain customer churn. The Random Forest model achieved strong predictive accuracy, while Power BI visualizations provided actionable business insights. By understanding the main drivers of churn such as contract type, internet service, and support quality companies can design targeted retention strategies and improve customer satisfaction.