Customer Churn Prediction

Amulya H S (U03CQ23S0071)

Final Year BCA, DBIMSCA Kumbalgodu

Abstract:

The "Customer Churn Prediction" project focuses on predicting whether a customer is likely to discontinue using a company's products or services. It is a Machine Learning-based system designed to analyze customer data and identify potential churn patterns. The main goal is to help businesses proactively retain customers by understanding the factors leading to churn. Using data such as customer demographics, transaction history, and engagement level, the system predicts churn probability, enabling timely interventions. The project helps in improving customer satisfaction, business revenue, and retention strategies through data-driven insights.

Introduction:

In competitive markets, customer retention is as important as customer acquisition. Many organizations lose a significant percentage of their customers annually, impacting revenue and brand value. The "Customer Churn Prediction" project aims to provide a data-driven solution to identify customers who are likely to leave. The project uses Machine Learning algorithms such as Logistic Regression, Random Forest, and XGBoost to build predictive models. Python, Pandas, NumPy, and Scikit-learn are used for data preprocessing and model training. The prediction system can be integrated into business dashboards to provide real-time churn insights. This project plays a vital role in marketing, telecom, banking, and subscription-based industries, where customer churn has a major financial impact.

Objectives:

• To develop a Machine Learning model capable of predicting customer churn with high accuracy.

• To analyze historical customer data and identify churn-related factors.

• To visualize customer behavior and churn trends using data visualization tools.

• To help organizations take proactive actions for customer retention.

• To deploy the model in a user-friendly web interface for real-time predictions.

Modules:

• Data Collection Module: Collect and preprocess data from customer records.

• Data Analysis Module: Perform statistical analysis and visualization of churn factors.

• Model Training Module: Build and train ML models using classification algorithms.

• Prediction Module: Predict customer churn probability based on new data.

• Dashboard Module: Display insights and churn trends visually using charts and graphs.

Technologies Used:

Frontend: HTML, CSS, JavaScript, Flask / Streamlit

Backend: Python (Machine Learning with Scikit-learn, Pandas, NumPy)

Database: MySQL / CSV dataset

Visualization Tools: Matplotlib, Seaborn, Power BI (optional)

System Requirements:

Hardware Requirements:

Component	Minimum Requirement
Processor	Intel i3 or equivalent
RAM	4 GB (8 GB recommended)
Hard Disk	250 GB HDD / 128 GB SSD
Display	1024x768 or higher

Software Requirements:

Category	Description
OS	Windows 10/11, Linux (Ubuntu)
IDE	VS Code / Jupyter Notebook / PyCharm
Libraries	NumPy, Pandas, Scikit-learn, Matplotlib, Seaborn
Version Control	Git & GitHub

Future Enhancements:

• Integration with real-time customer data APIs.

• Implementation of Deep Learning models for higher accuracy.

• Development of automated retention strategies based on churn risk.

• Integration with CRM dashboards for business insights.

Conclusion:

The Customer Churn Prediction project provides a powerful analytical tool for businesses to understand and predict customer behavior. By leveraging Machine Learning, it helps organizations minimize customer loss, optimize marketing strategies, and enhance long-term customer relationships. The system's predictive capabilities offer significant value in decision-making and operational planning.

Date: _____

HOD