

Customer Churn Prediction

Amulya H S (U03CQ23S0071)

Final Year BCA, DBIMSCA Kumbalgodu

Abstract

The Customer Churn Prediction project aims to identify customers who are likely to discontinue a company's services using Machine Learning techniques. The system analyzes customer data such as demographics, usage patterns, and engagement history to predict churn probability. The project helps organizations proactively retain customers by identifying risk factors and enabling timely interventions. The system also provides churn risk scores, visualization dashboards, and supports prediction for new customer data, thereby improving business decision-making and customer retention strategies.

Introduction

In competitive business environments, retaining customers is as important as acquiring new ones. Customer churn leads to revenue loss and impacts long-term business sustainability. This project proposes a data-driven approach to predict customer churn using Machine Learning algorithms such as Logistic Regression, Random Forest, and XGBoost. The system processes historical customer data, trains predictive models, and generates churn risk predictions. Additionally, the project includes interactive dashboards for analyzing churn patterns and allows prediction for new customer records through data upload. The solution can be integrated into business intelligence systems to support proactive retention strategies in industries such as telecom, banking, and subscription-based services.

Objectives

- Develop a Machine Learning model to predict customer churn accurately.
- Analyze customer behavior and identify churn-related patterns.
- Provide churn risk scores instead of only churn predictions.
- Visualize churn trends and customer segmentation using dashboards.
- Enable real-time churn prediction using uploaded customer data.
- Support businesses in planning effective retention strategies.

Modules

- Data Collection Module: Collects and preprocesses customer datasets.
- Data Analysis Module: Performs statistical analysis and visualization of churn factors.
- Model Training Module: Builds classification models using ML algorithms.
- Prediction Module: Predicts churn probability for new customer data.
- Dashboard Module: Displays churn trends and risk distributions through charts and graphs.
- Data Upload Module: Allows users to upload new datasets for churn prediction.

Technologies Used

- Frontend: HTML, CSS, JavaScript, Flask / Streamlit
- Backend: Python (Pandas, NumPy, Scikit-learn)
- Database: MySQL / CSV datasets
- Visualization Tools: Matplotlib, Seaborn, Power BI (optional)
- Version Control: Git & GitHub

System Requirements

Hardware Requirements

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

Software Requirements

Category	Description
OS	Windows 10/11 or Linux
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 for live churn monitoring.
- Implementation of advanced ensemble and deep learning models for improved accuracy.
- Automated retention strategy recommendations based on churn risk.
- Integration with CRM systems for business analytics and reporting.
- Cloud deployment for large-scale enterprise usage.

Conclusion

The Customer Churn Prediction system provides businesses with an effective analytical tool to predict customer behavior and reduce customer loss. By leveraging Machine Learning and visualization techniques, organizations can optimize retention strategies and improve long-term customer relationships. The system's predictive and analytical capabilities offer valuable support for business decision-making and operational planning.

Date: _____

HOD Signature: _____