

Customer Churn Prediction in the Telecommunications Sector

This project focuses on understanding and predicting customer churn using transaction-level and account-related data from USSD service users. Churn—when customers stop using a service—is one of the biggest challenges in telecommunications, as retaining a customer is often more cost-effective than acquiring a new one.

Objective

The primary objective of this project was to develop a machine learning model that can accurately identify customers who are at a high risk of churning. With early detection, the business can intervene through targeted offers, improved service, or proactive engagement strategies.

Methodology

1. Data Exploration: Performed a detailed exploratory data analysis (EDA) to understand customer demographics, transaction patterns, and behavioral factors such as failed transactions and complaints logged.
2. Data Preprocessing: Encoded categorical variables, scaled numerical features, and handled class imbalance.
3. Model Development: Trained and compared three models—Logistic Regression, Random Forest, and XGBoost—on the prepared dataset.
4. Evaluation: Models were assessed using Accuracy, Precision, Recall, F1-score, and ROC-AUC to balance predictive power and interpretability.
5. Model deployment using Hugging Face and streamlit.SDK in Docker

Results & Impact

- RandomForest outperformed other models, achieving the highest ROC-AUC and Recall, which is critical since false negatives (failing to identify actual churners) are costly for the business.
- Feature importance analysis showed that failed transactions, number of complaints, and tenure were strong churn indicators.
- The insights from this model can directly inform retention campaigns, saving potential revenue loss and improving customer satisfaction.

Link

https://huggingface.co/spaces/ebytt12/Customer_churn

Relevance to Data Science MSE

This project demonstrates the application of data science techniques, including EDA, supervised learning, model evaluation, and interpretability, in a real-world business setting.

It aligns with the objectives of the MSE program by bridging theory and practice, and by showcasing how statistical methods and machine learning can drive business decisions.