

Title: Customer Churn Prediction Using Machine Learning

Author: Francisca Nwakaego Kelechi

Abstract

This report presents a machine learning approach to predicting customer churn using Python, SQL, and Power BI. The project achieved an accuracy of 85%, providing actionable insights for reducing churn and improving customer retention strategies.

Introduction

Customer churn is a critical challenge for businesses, directly impacting revenue and growth. This project aimed to build a predictive model to identify customers at risk of leaving, enabling proactive retention measures.

Methodology

- **Data Collection:** 10,000+ rows of customer transaction and engagement data.
- **Data Cleaning:** SQL and Pandas used to remove duplicates, handle missing values, and normalize variables.
- **Feature Engineering:** Created variables such as frequency of purchases, average spend, and complaint history.
- **Modeling:** Applied Logistic Regression and Random Forest using Scikit-Learn.
- **Visualization:** Power BI dashboards to display churn trends and retention KPIs.

Results

- Achieved **85% accuracy** in predicting churn.

- Identified top 3 risk factors: low engagement frequency, high complaint rate, and reduced monthly spend.
- Dashboards reduced reporting time by 40% and improved visibility across 3 departments.

Conclusion

This project demonstrates the value of combining machine learning with business intelligence tools to address real-world challenges. The methodology can be adapted to other sectors, including NGOs, to analyze migration trends, program participation, or beneficiary retention.

Keywords

Data Science, Machine Learning, Customer Churn, Predictive Analytics, Power BI, SQL