Customer Churn Prediction Project

Machine Learning Approach for Customer Retention

October 2025

By CINDY AKINYI

Project Overview

- Customer churn is a key issue for telecom companies.
- Goal: **Predict which customers are likely to leave** using machine learning.

Benefits:

- Improves retention strategies.
- Reduces revenue loss.
- Enhances customer satisfaction.

Data and Preparation

- Dataset includes customer demographics, service usage, and account info.
- Target variable: **Churn (Yes/No)**

Steps taken:

- Cleaned missing values and encoded categorical data.
- Split into training (80%) and testing (20%).
- Applied feature scaling where needed.

Machine Learning Models

- **Decision Tree Classifier**
  - Tuned using GridSearchCV.
  - Best parameters: max_depth=5, min_samples_split=5.
  - Accuracy: 97%.
- **Logistic Regression**
  - Used as baseline model.
  - Accuracy: 86%.

Key Insights

- Top churn indicators:
  - Total day minutes / charges – heavy users churn more.
  - Customer service calls – multiple calls = dissatisfaction.
  - International plan – higher churn tendency.
- Decision Tree captured complex customer behavior patterns.

Limitations & Recommendations

- **Limitations:**
  - Class imbalance (few churners).
  - Limited feature engineering.
- **Recommendations:**
  - Target customers with high usage or frequent support calls.
  - Improve service quality and pricing for international plans.
  - Use SMOTE or cross-validation for future model enhancement.

Conclusion

- ML models can **accurately predict customer churn**.
- **Decision Tree** achieved the best performance.
- Insights can guide **data-driven retention strategies**.
- Future work: enhance dataset and interpretability.