Understanding Customer Churn : Comprehensive Analysis

By: Purity Kibaki

Outline

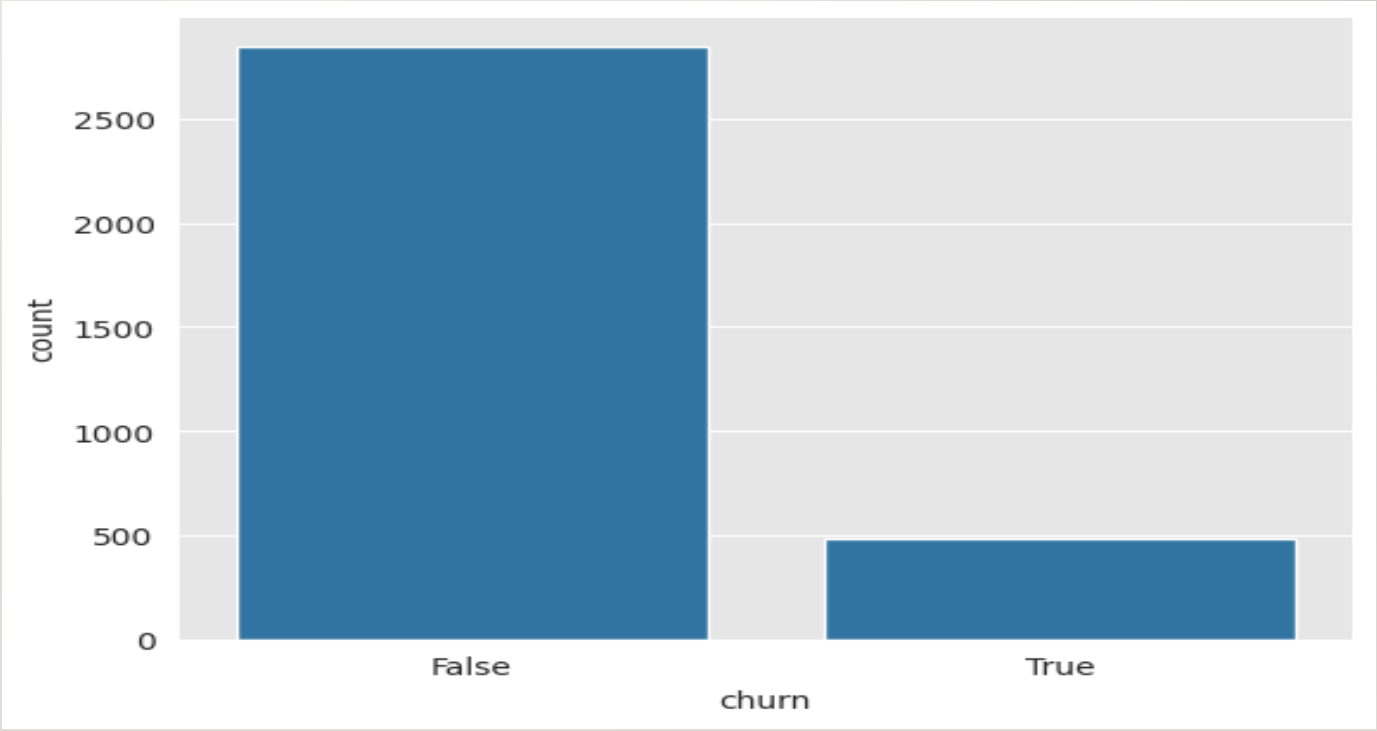
- BACKGROUND
- BUSINESS PROBLEM
- DATA
- METHODOLOGY
- FINDINGS
- CONCLUSIONS

BACKGROUND

- Customer churn impacts every organization. In the case of this telecommunications company, the root causes are initially identified.
- Machine learning techniques are then employed to determine which models yield the most precise forecasts for future predictions.

Business Problem

- Maximize customer retention
- Minimize customer turnover
- Identify causes
- Develop models that will help predict customer turnover



Data

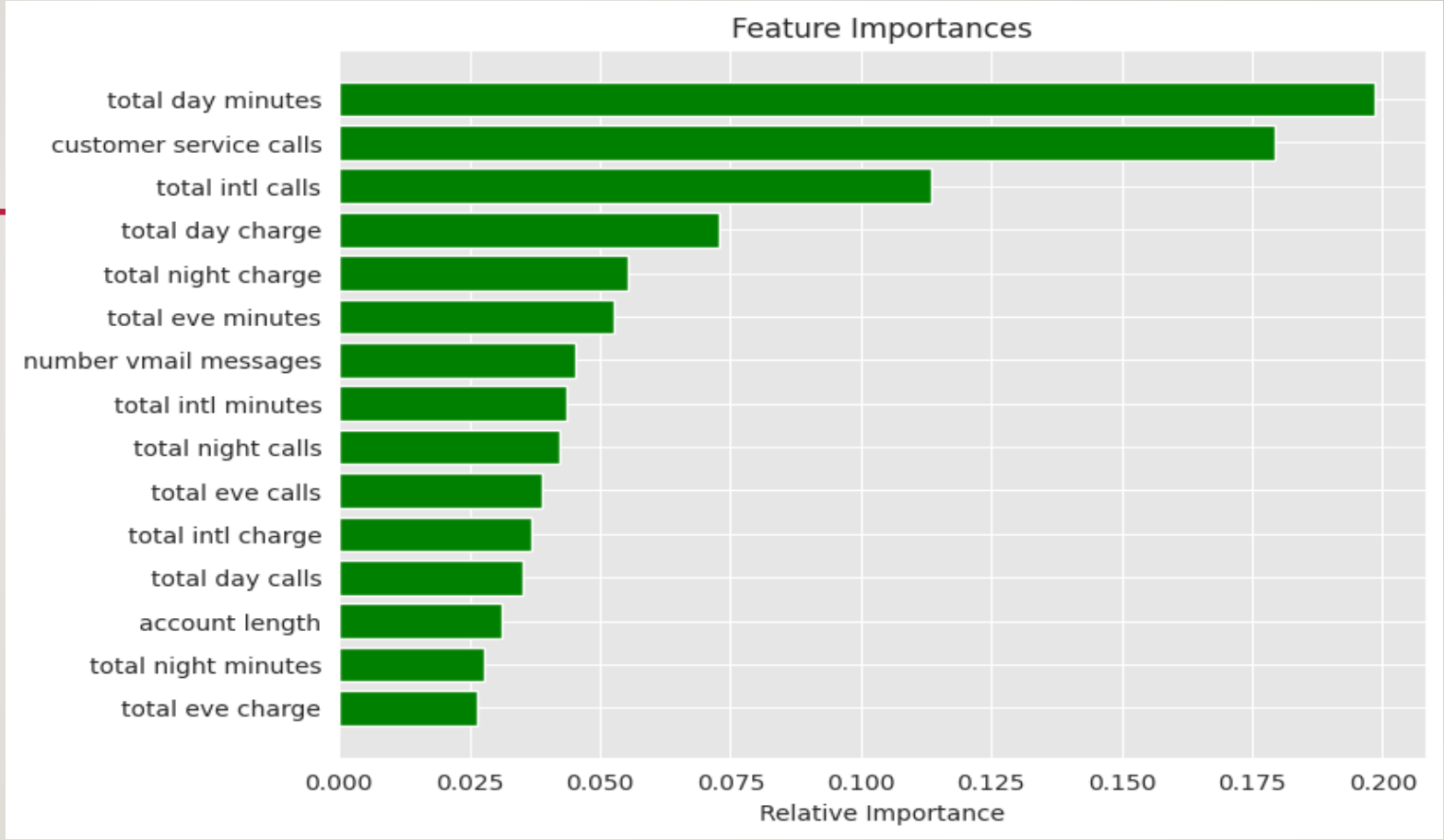
- Using Telecom data set available on Kaggle with 3,333 customers.
- Contains data including
  - daytime data
  - night time data
  - International data
  - Customer service calls
  - International plan
  - Churn/turnover

Methodology

- Identify variables with statistical significance
- Address class imbalance
- Remove non-necessary variables
- Split data into Train and Test set
- Employ Feature selection
- Implement Machine Learning models

FEATURE IMPORTANCE RESULTS

Results



Customer service calls and International Plans top two features for turnover.

Summary of Findings

- Customers with an International Plan are 7 times more likely to leave compared to those without.
- Each additional customer service call increases the likelihood of churn by a factor of 1.55.
- Customers making international calls are less likely to churn, with an odds ratio of 0.92.
- Having a Voicemail Plan decreases the likelihood of churn significantly, with an odds ratio of 0.16.
- F1 score, deciding factor. Random Forests best model.

MODELS FINDINGS

Results

Model Performances:
Logistic Regression: Accuracy score of 0.70.
Decision Trees, and Random Forests: Accuracy score of 0.87.

F1 Scores:
Decision Trees: F1 score of 0.42.
Random Forests: F1 score of 0.55.

Best Model:

Random Forests emerged as the best model based on accuracy and F1 score.

THANK YOU