Telecom Customer Churn Prediction & Analysis

By Shem Omondi

Business Problem

Overview

- Customer churn is costly and retaining is cheaper than acquiring.
- SyriaTel needs to identify at-risk customers before they leave.
- **Goal**: Predict churn risk and guide targeted retention campaigns by building models.

Data understanding

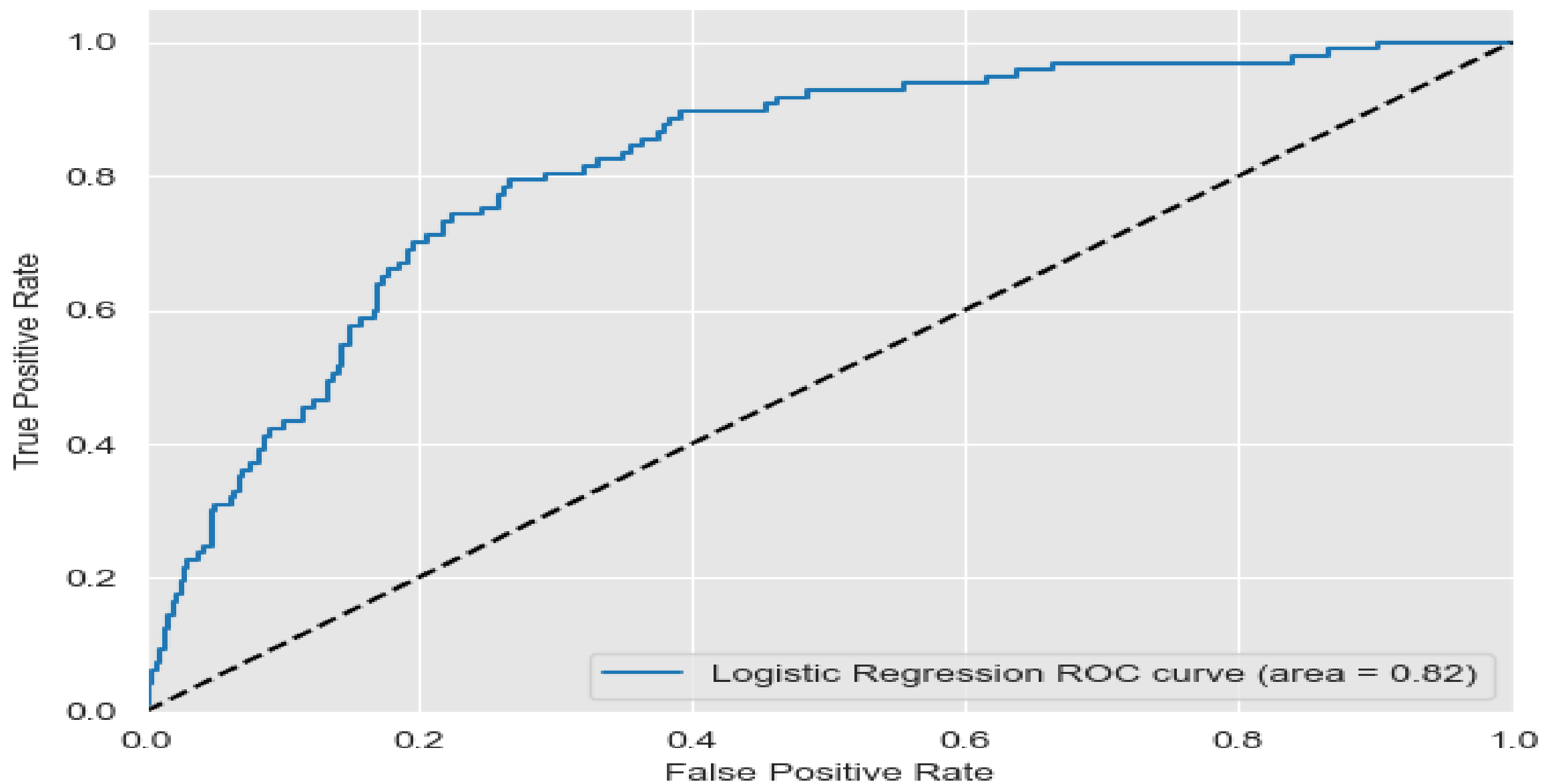Historical customer records of about 3,300 rows.

Some features included:

- Customer demographics (state, area code).
- Service plans (international plan, voicemail plan).
- Usage patterns (day, evening, night, international minutes & calls).
- Customer service interactions (number of service calls).
- Target variable: Churn (whether the customer left or stayed).
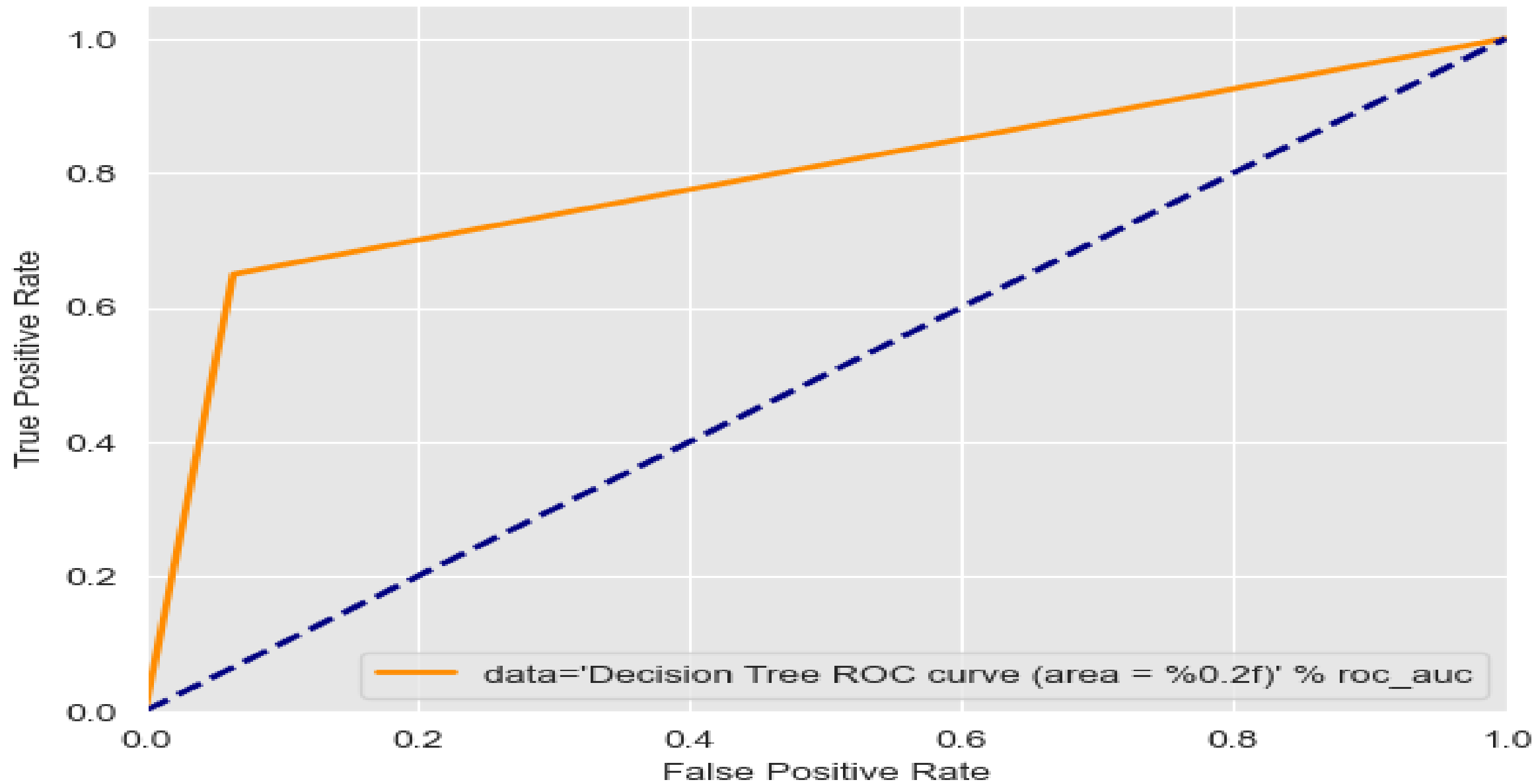
Modeling

We built 3 models to try and capture the relationships In the dataset:

- Logistic regression model which acts as a baseline for comparison.
- Untuned decision tree classifier: decision trees capture nonlinear relationships better.
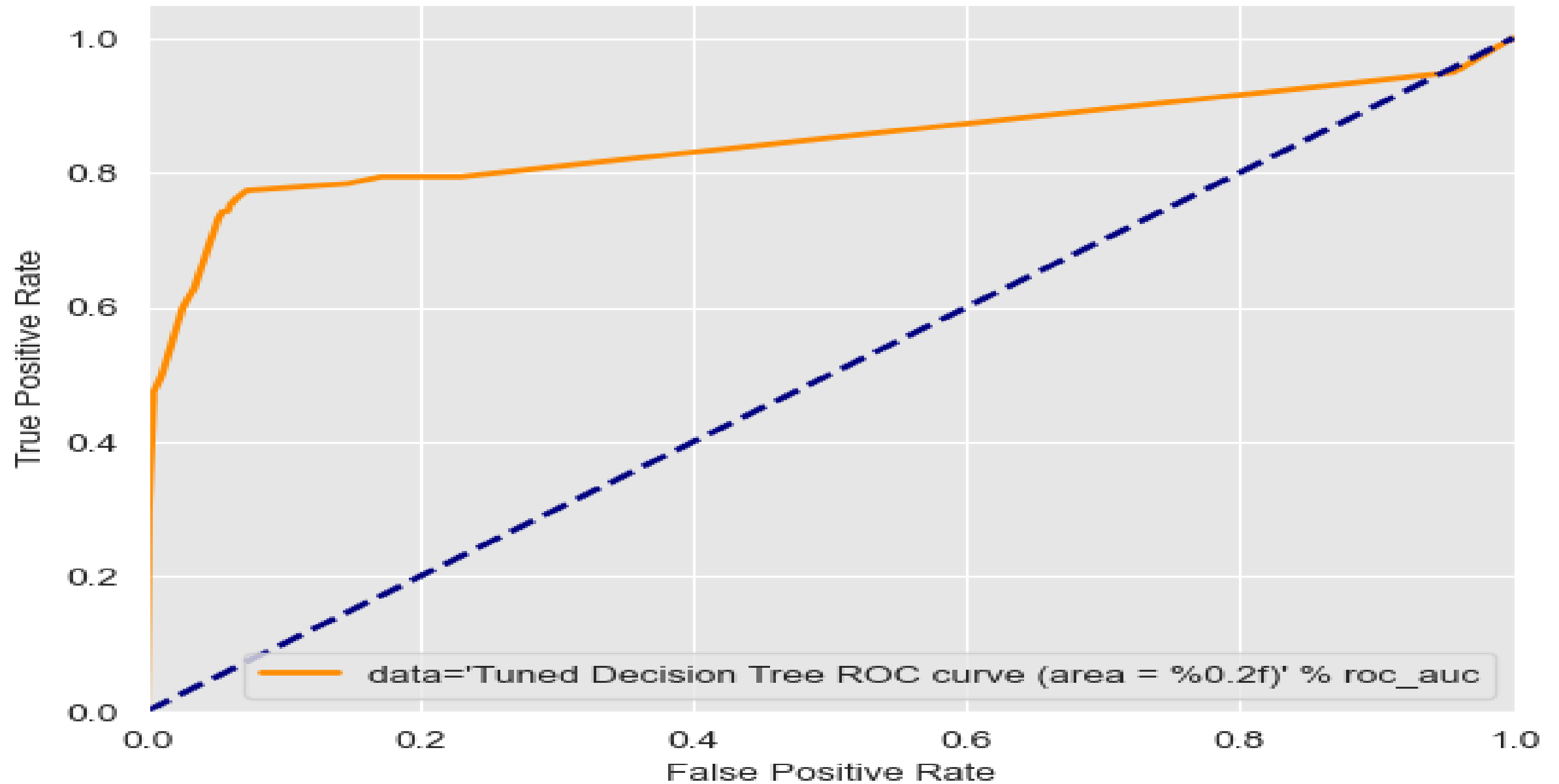- Tuned decision tree classifier: To reduce overfitting.

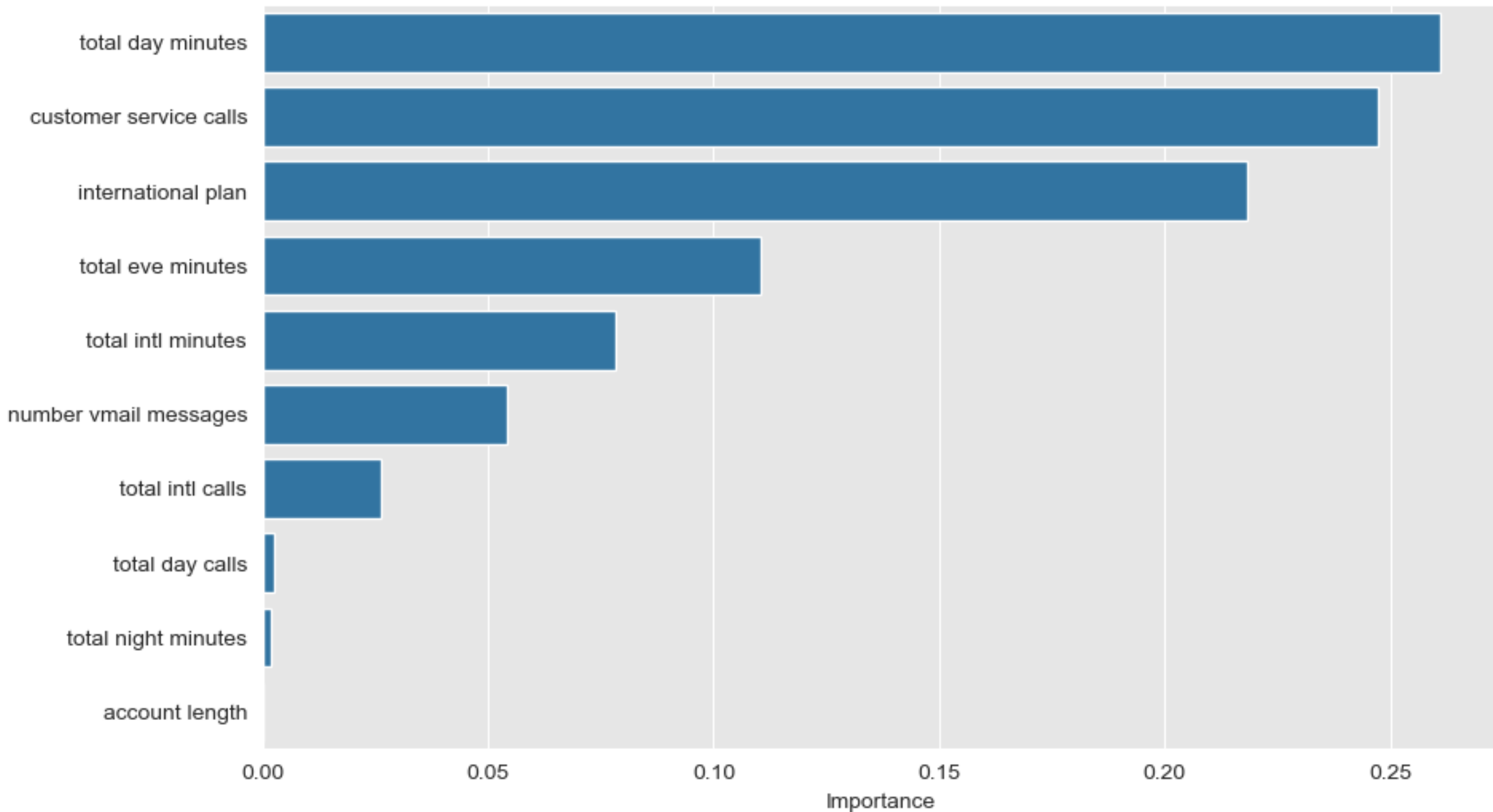ROC Curve

Receiver Operating Characteristic

data='Decision Tree ROC curve (area = %0.2f)' % roc_auc

Receiver Operating Characteristic – Tuned Decision Tree

data='Tuned Decision Tree ROC curve (area = %0.2f)' % roc_auc

Top 10 Feature Importances from Tuned Decision Tree

Model Evaluation

Results:

- Logistic Regression: caught around 7 of 10 churners, but flagged many who did not churn as churners.
- Decision Tree (untuned): moderate performance.
- Tuned Decision Tree: best balance since it catches 3 of 4 churners, around 7 of 10 flagged are true churners.

Recommendations

- Offer loyalty bundles to international plan users.
- Improve customer support for repeated customer service callers.
- Proactively recommend plans for heavy day users.

Next Steps

- Deploy the tuned decision model for operation.
- Retrain the model with recent data to adjust to current markets.

THANK YOU!