Telecom Customer Retention Analysis

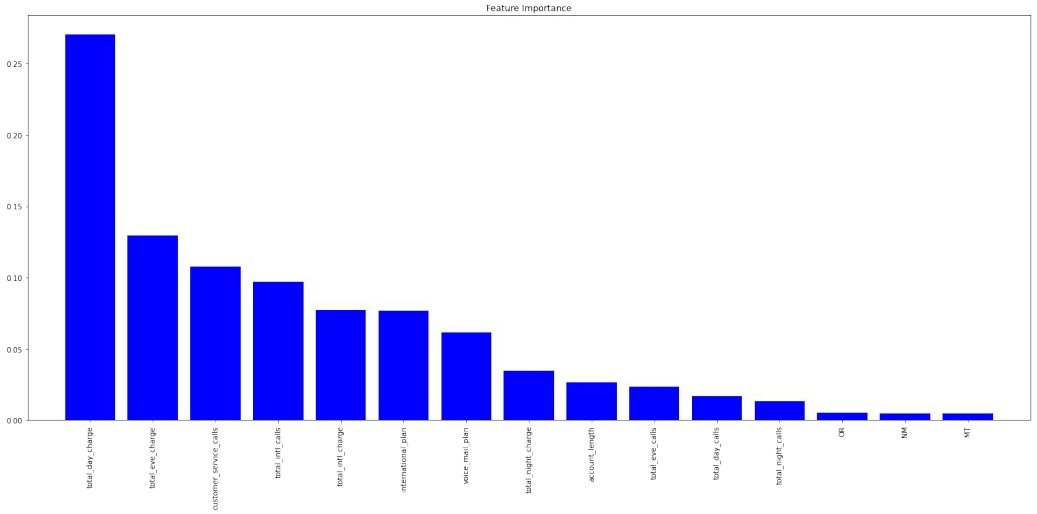
Overview

- Telecom customer subscription
  - SyriaTel
  - Binary Data Classification
    - Classification Decision Tree
    - Logistic Regression Model

- Goals
  - Better predict likelihood of customers canceling subscriptions
    - Churn
    - Accuracy
      - True Positive
      - True Negative

Data Overview

- 3333 Customers
  - Customer profile containing 20 features
  - X Variables
    - Phone charges
    - Customer Service Calls
    - International Calls
  - Y Variable
    - Churn



Classification Modeling

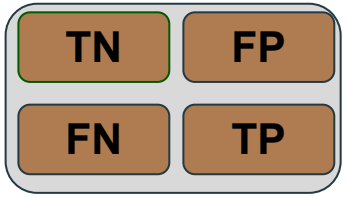
- Decision Tree Classifier
- Logistic Regression Model
- Baselines and Improved Models

- Iterations
  - Feature Scaling
  - Oversampling
  - Decision Tree Pruning
  - Hyperparameter Tuning

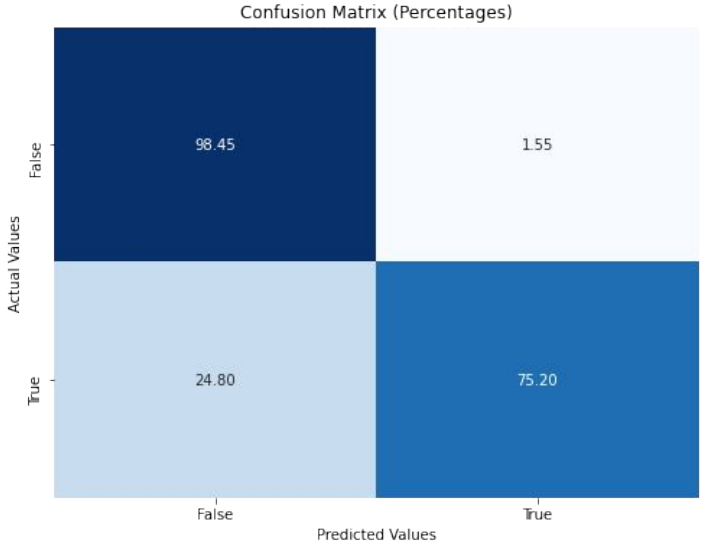
Evaluation

- Evaluate Models
  - Overall Accuracy
  - True Positives
  - True Negatives

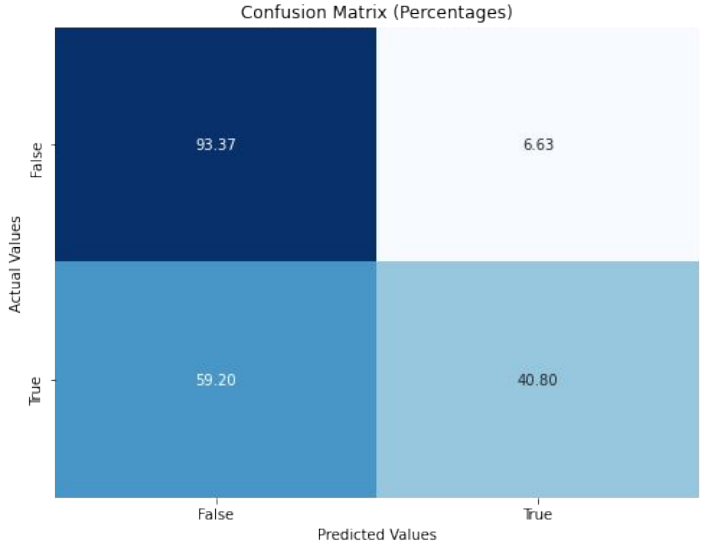
Confusion Matrix Structure



Decision Tree



Logistic Regression



Further Analysis

Improvements

	Decision Tree Classifier	Logistic Regression
Cross Validation	90	85
AUC	88	67

Decision Tree Classifier	Baseline	Improved
Accuracy	91	95
True Positives	71	90

*Feature Scaling and Parameter Tuning

Logistic Regression	Baseline	Improved
Accuracy	84	85
True Positives	44	52

*Class Weighting and Parameter Tuning

Next Steps

- Characteristics of customers who churn:
  - Customers with high bills
  - Customers who make more customer service calls
  - Customers who have international plans

- Use more sophisticated modeling methods and parameter tuning like clustering or GridSearchCV.

Additional Questions

Thank you!

By: Will Byrd

Byrdwcrawford@gmail.com

The Flatiron School