 ceflynn / Phase-3-Syria_Telcom

Public

phase 3 project for Flatiron School Data Science

☆ 0 stars

🔗 0 forks

☆ Star

▼

👁 Unwatch

▼

<> Code

🕒 Issues

🔗 Pull requests

🎬 Actions

📁 Projects


📖 Wiki

🛡 Security

📈

🔑 main ▼

⋮

 ceflynn adding pdf and updating images ...

2 minutes ago ⌚ 8

View code

☰ README.md

✎

SyriaTel Communications Customer Retention

Overview

Flatiron School Data Science Phase 3 project. Focusing on classification.

Business Problem

Syria Telcom is looking to lower its churn rate. Being able to retain customers is essential in the competitive telecommunications industry.

Data Used

data from the following sources can be found in the 'data' folder

data/telcom.csv

https://github.com/ceflynn/Phase-3-Syria_Telcom

1/5

Metric - Recall

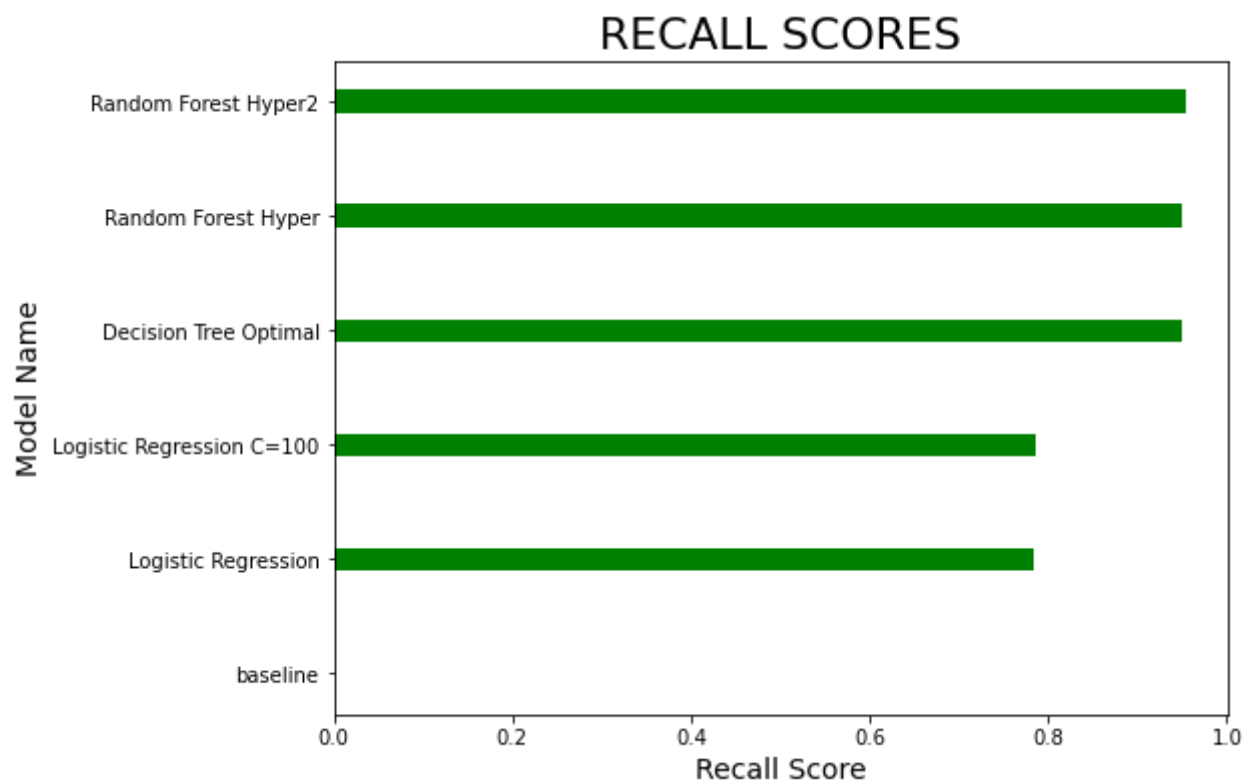
We are primarily concerned with Recall in this situation. Getting as many of the churned customers identified is very important.

False Negatives are the primary concern with Recall. Getting this as low as possible was the goal. Due to the nature of the problem it will be very challenging to get all the churns identified without have a ton of False Positives.

While False Positives are not our primary concern, too many false positives will add additional costs that may negate the savings made by identifying the churned customers.

Iterative Modeling process

Recall Comparison



Final Model

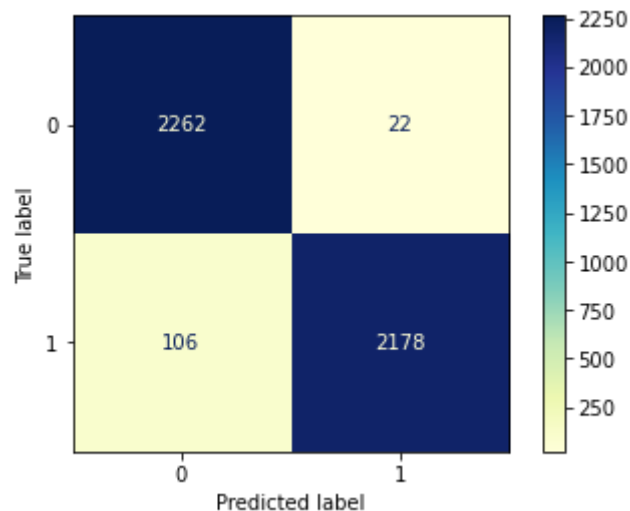
The final model was a Random Forest Classifier with hyper-tuned parameters.

```
# Random Forest pt 2
rfc2_optimal=RandomForestClassifier(criterion='gini',
                                     max_depth=10,
                                     min_samples_leaf=1,
                                     max_features=5,
                                     n_estimators=100,
                                     random_state=42)
```

Confusion Matrix

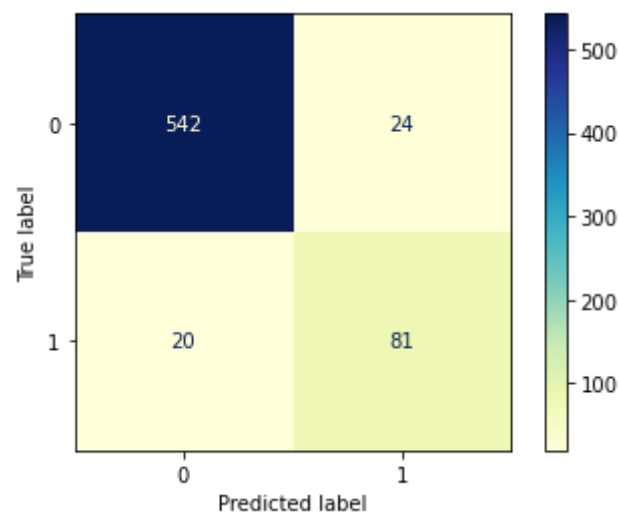
Our final confusion matrix can be seen below for both our training and testing data

Training Matrix



Recall ~95%

Testing Matrix

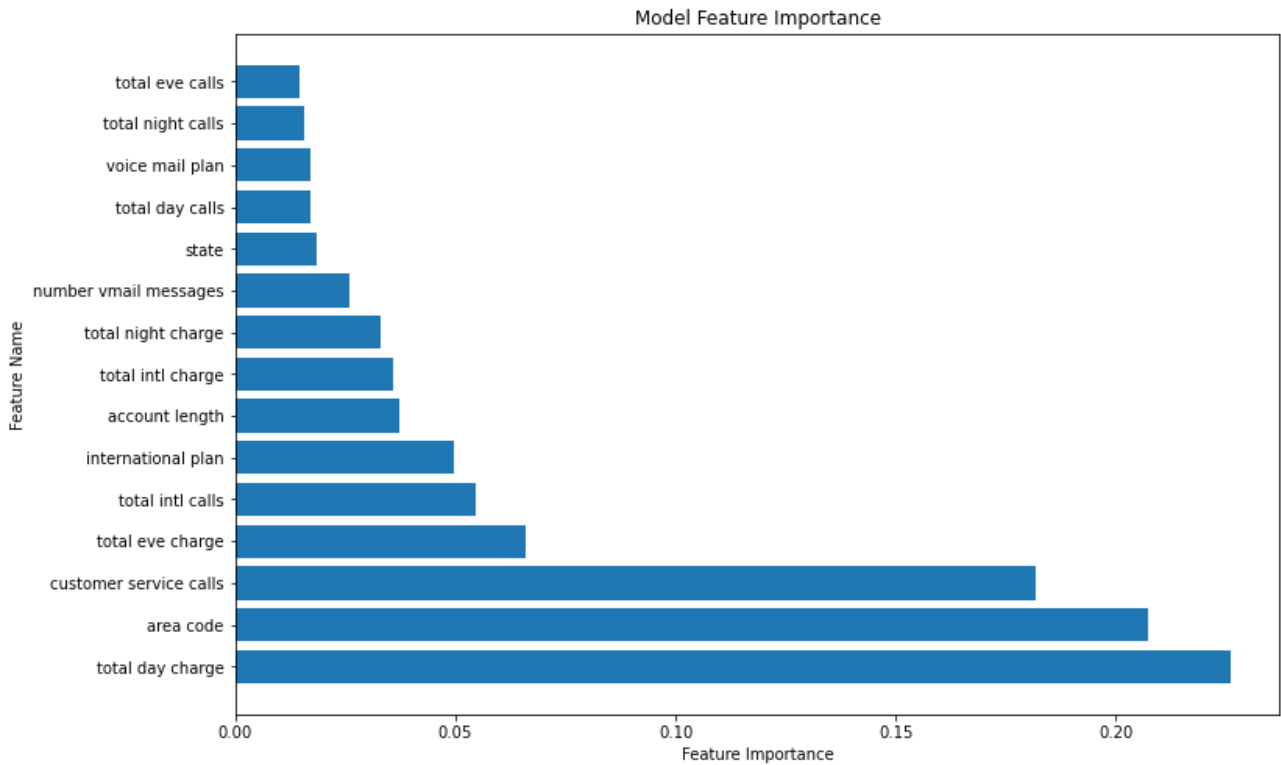


Recall ~80%

Feature Importance

Our top 3 features:

- Total Day Charge
- Area Code
- Customer Service Calls



Deliverables

- [Non-Technical Presentation](#)
- [GitHub Repository](#)
- [Jupyter Notebook](#)

Releases

No releases published
[Create a new release](#)

Packages

No packages published
[Publish your first package](#)

Languages

● Jupyter Notebook 100.0%