



Predicting Churn in Streaming Platforms

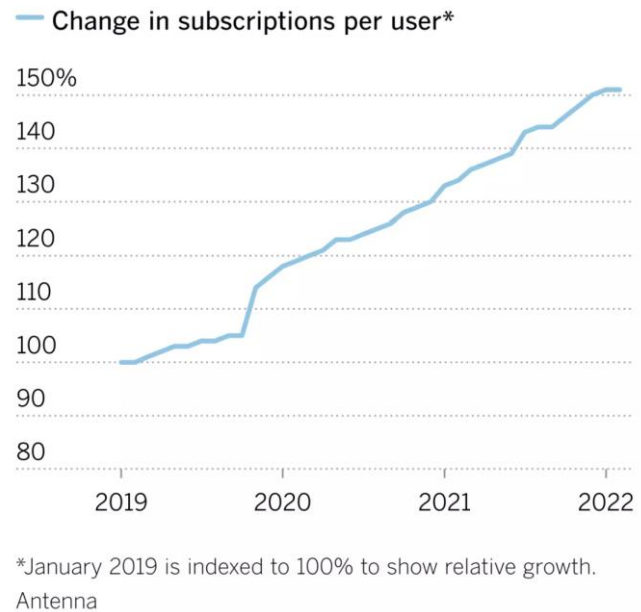
Faraan Khan

The State of the Industry

Although streaming services are experiencing growth, churn remains an important issue that companies seek to resolve*

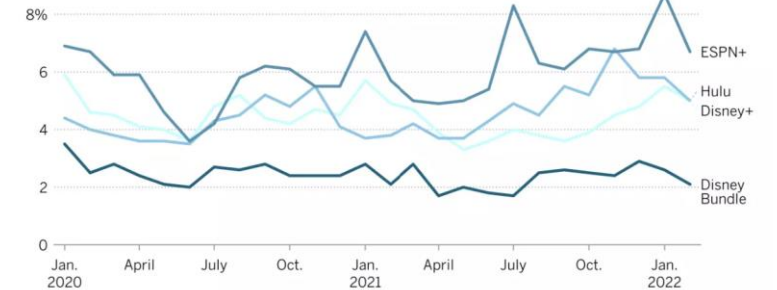
Streaming growth

The average number of video streaming service subscriptions per user has increased 51% in the U.S. since the beginning of 2019.



Disney streaming churn

U.S. churn rates for Disney's standalone streaming services, versus the Disney Bundle.



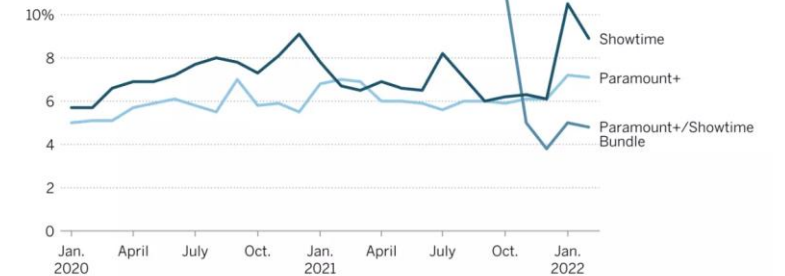
Monthly churn rates exclude involuntary churn (i.e. credit card declines).
Antenna

Ryan Faughnder LOS ANGELES TIMES

Churn rates are lower for the Disney Bundle, compared to individual services.

Paramount and Showtime streaming churn

U.S. churn rates for Paramount+ and Showtime, versus a bundled option.



Monthly churn rates excluded involuntary churn (i.e. credit card declines).
Antenna

Ryan Faughnder LOS ANGELES TIMES

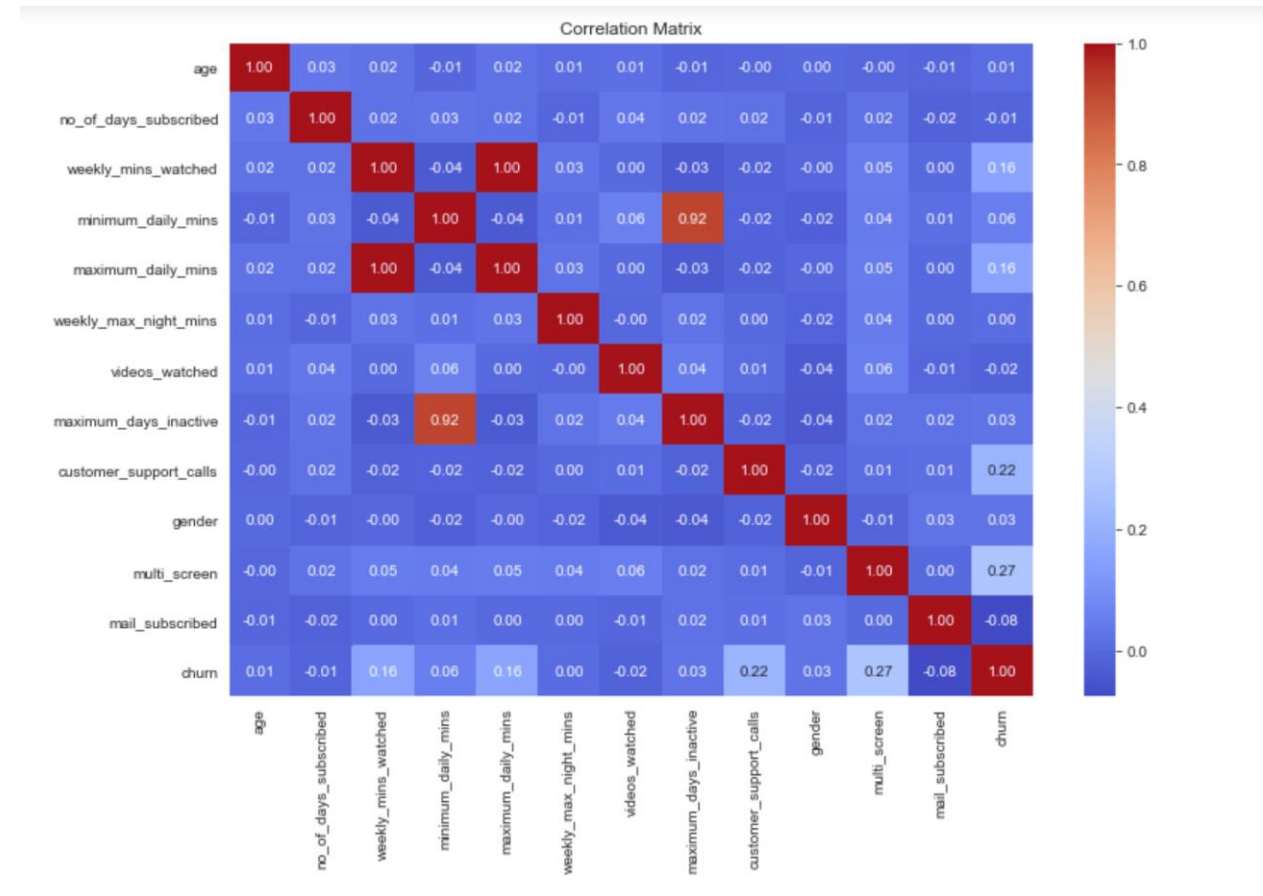
Churn rates were lower for the recently introduced Paramount+ bundle with Showtime.

*<https://www.latimes.com/entertainment-arts/business/newsletter/2022-04-05/here-are-four-charts-that-explain-future-of-streaming-the-wide-shot>

Objective and Data

Build a classification model that helps predict whether a user will churn or not

- Dataset: Kaggle dataset “Churn Modeling for OTT Platforms”
- After SMOTE and Feature Engineering, **2388 rows, 10 columns**
- Target Variable: **Churn**
- Independent Variables:
 - Age
 - Number of Days Subscribed
 - Weekly Minutes Watched
 - Weekly Max Night Minutes
 - Videos Watched
 - Maximum Days Inactive
 - Customer Support Calls
 - Gender
 - Multi-Screen
 - Mail Subscribed
- Optimize Model for **Recall** – minimize false negatives
- Baseline counts: **87% no churn, 13% churn**



Baseline Logistic Regression

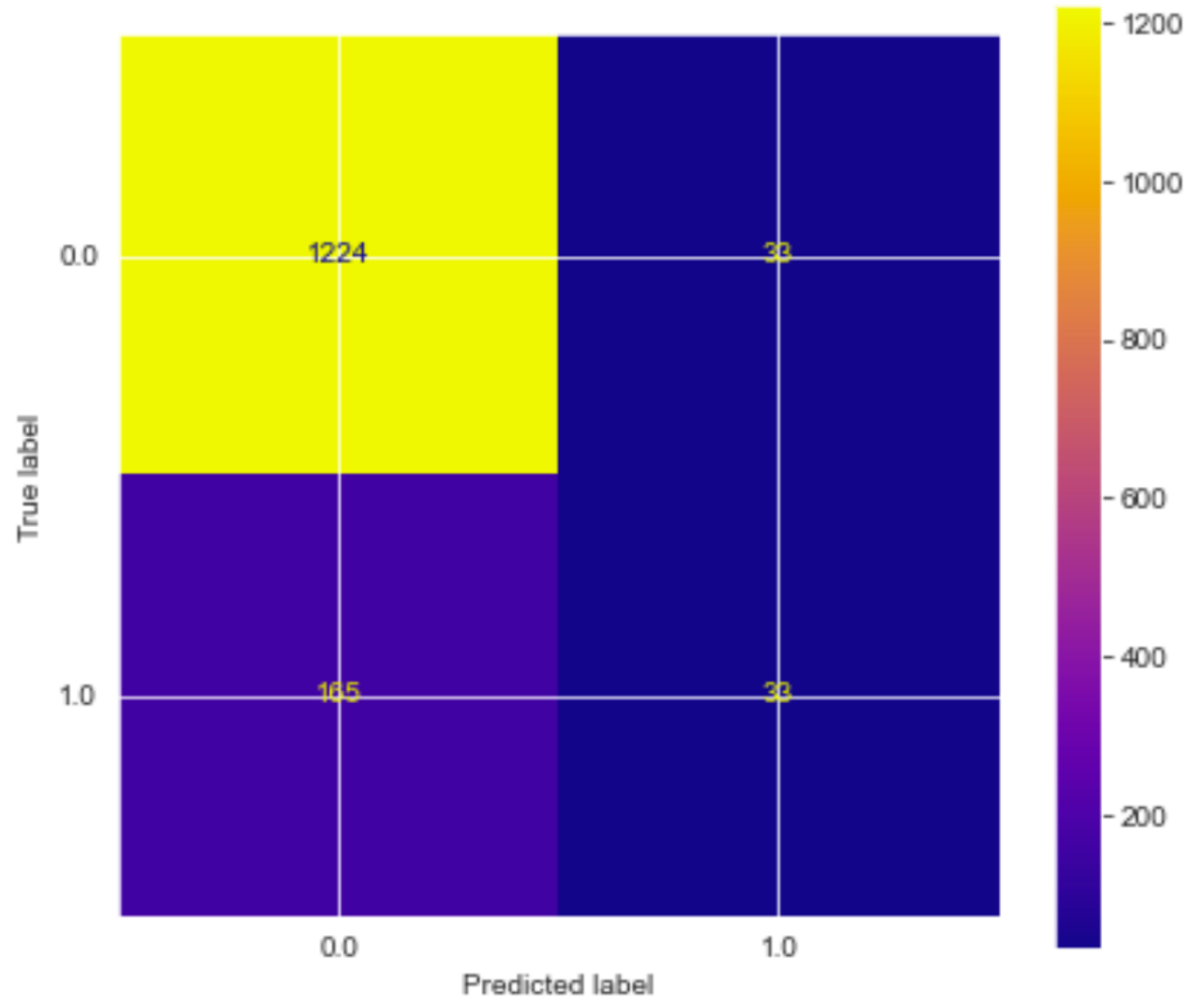
Accuracy: .86

Precision: .39

Recall: .18

F1- Score: .24

Model performs poorly for accuracy; almost worse than the baseline when picked at random



Model Performance post SMOTE

Sampling Strategy Ratio of .9 selected:

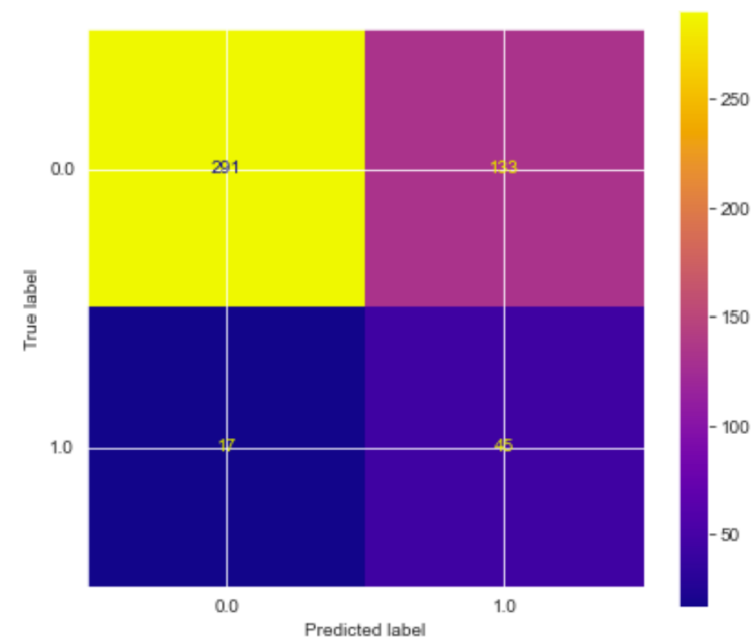
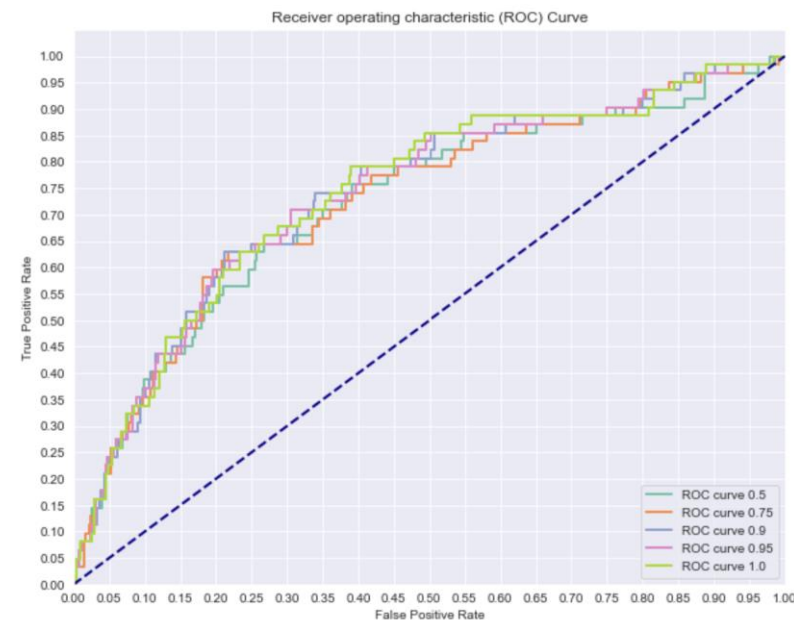
Accuracy: 0.69

Precision: 0.25

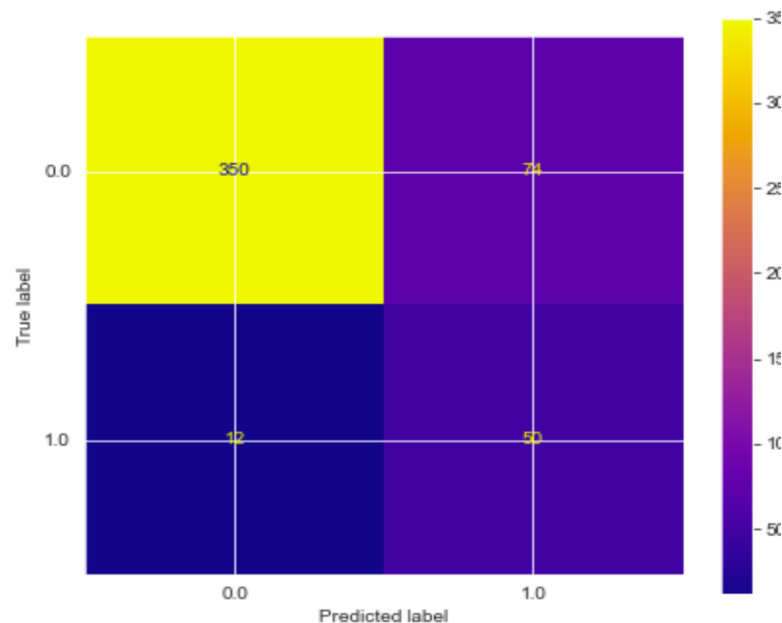
Recall: 0.72

f1 Score: 0.38

Recall and f1 scores dramatically improve

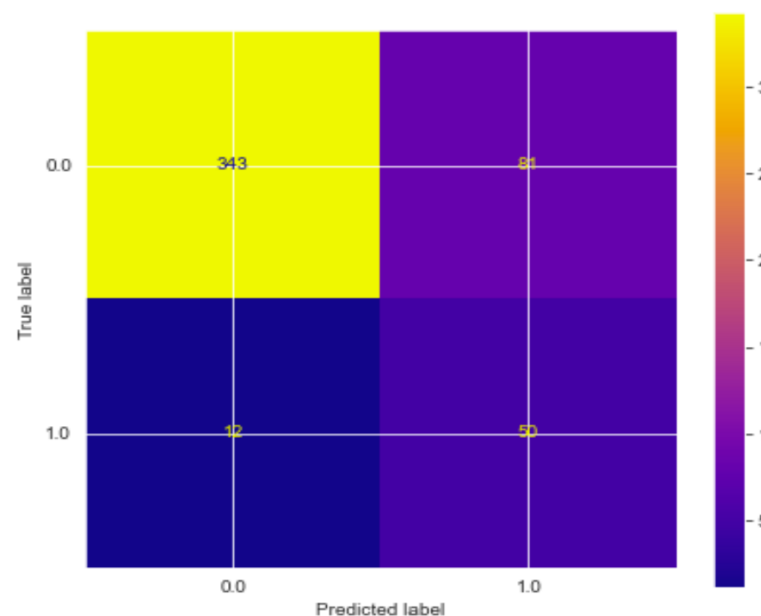


Model Performance After Polynomial Transformation



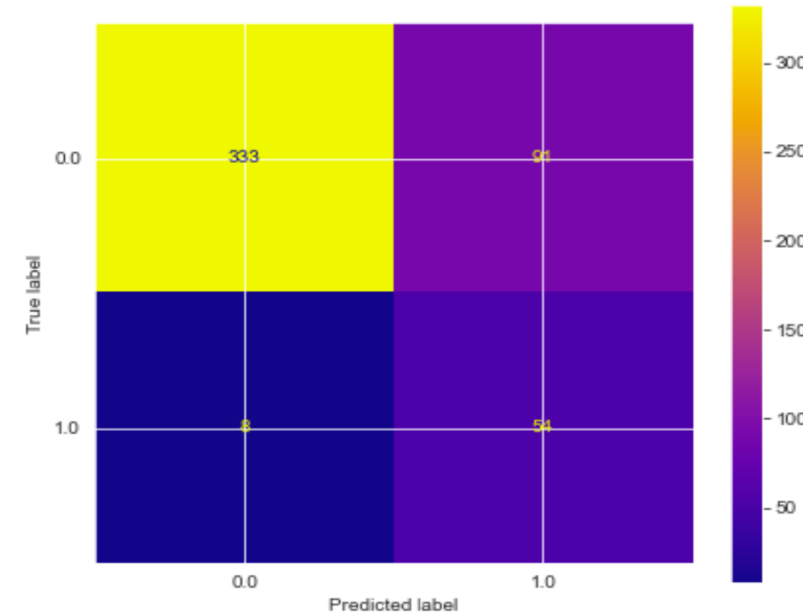
Polynomial Degree 4:

Accuracy: 0.82305
Precision: 0.40323
Recall: 0.80645
F1 Score: 0.53763



Polynomial Degree 3:

Accuracy: 0.80864
Precision: 0.38168
Recall: 0.80645
F1 Score: 0.51813



Polynomial Degree 2:

Accuracy: 0.79630
Precision: 0.37241
Recall: 0.87097
F1 Score: 0.52174

Decision Tree Baseline

Very clearly overfitted:

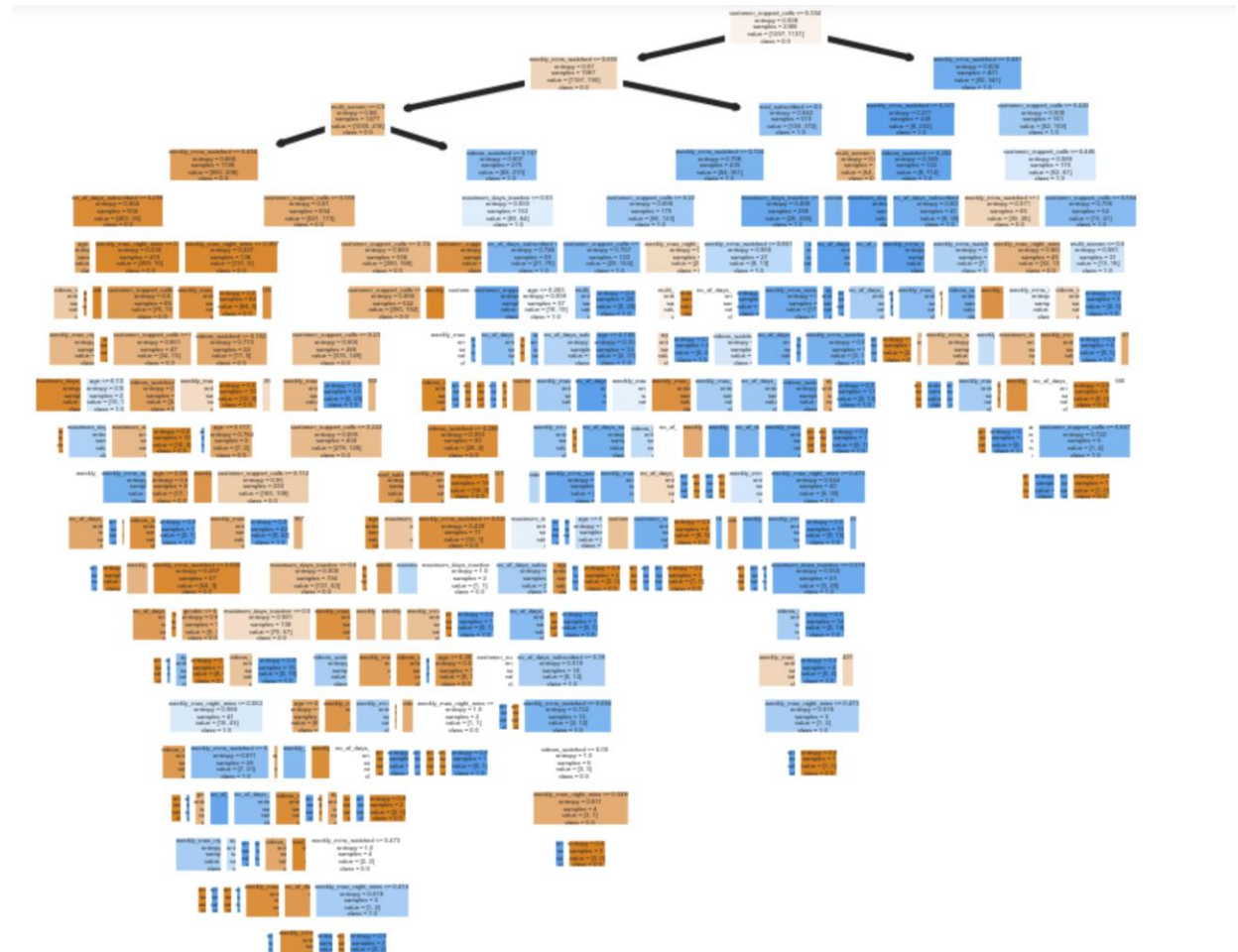
Accuracy: 0.5

Precision: 0.18

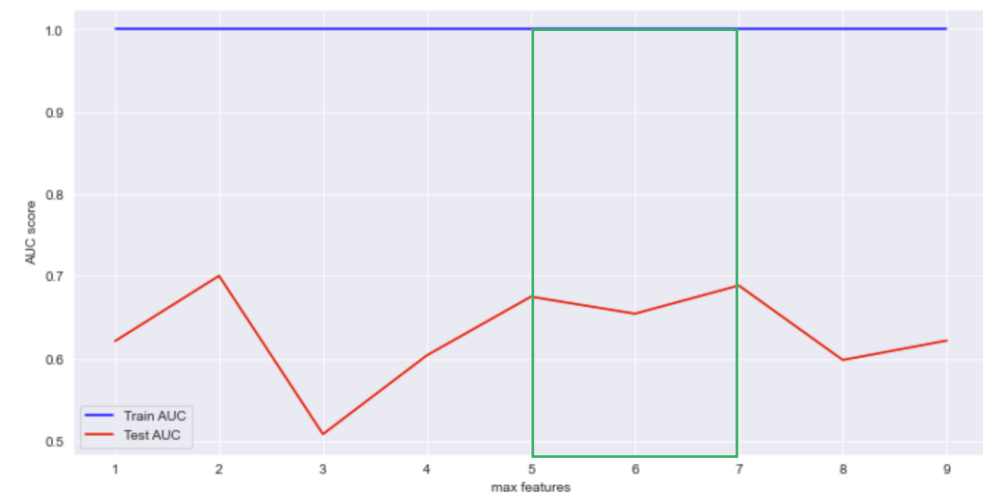
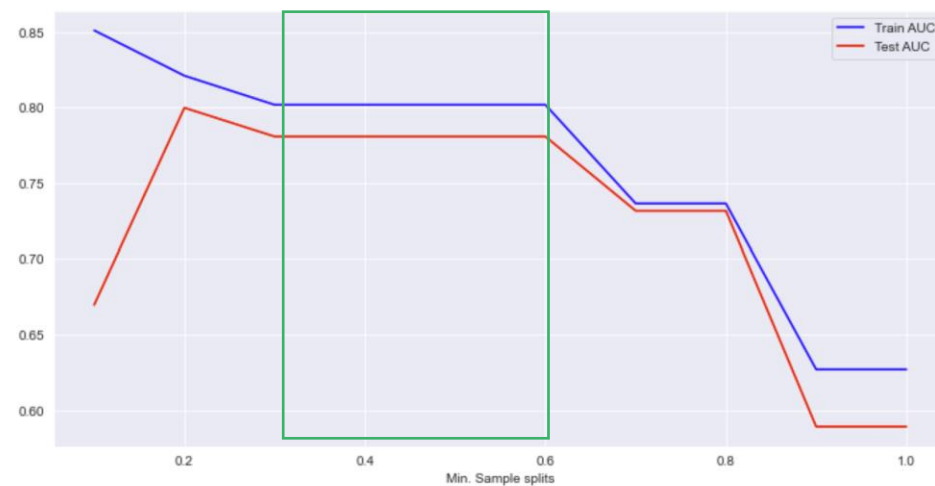
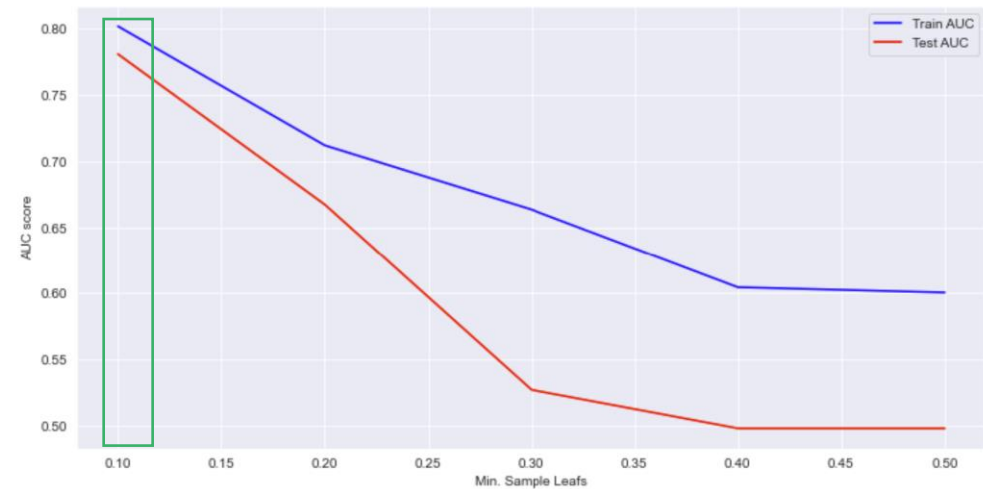
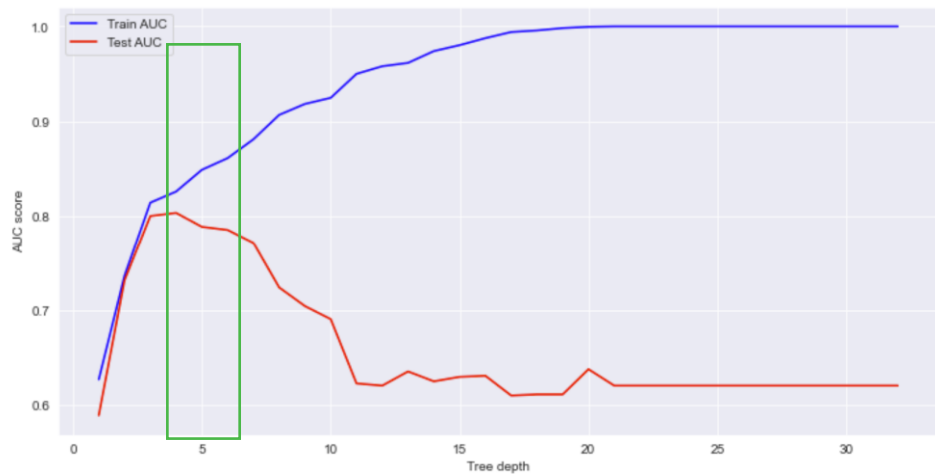
Recall: 0.82

F1 Score: 0.30

ROC AUC: 0.64



Finetuning Hyperparameters



Optimized Decision Tree

Performance Metrics for pruned and hyper tuned tree:

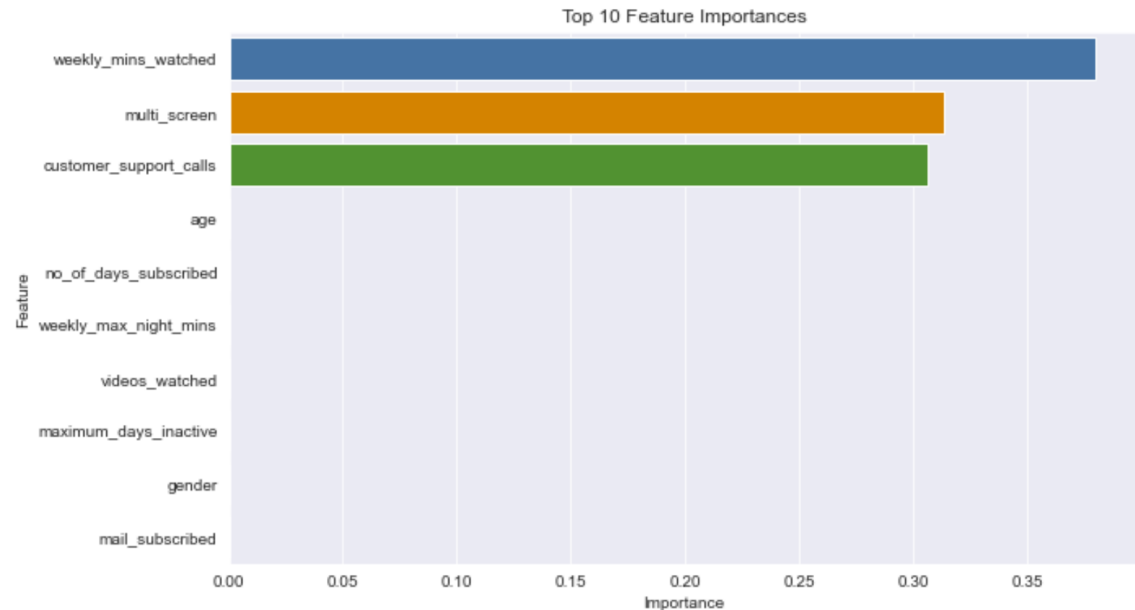
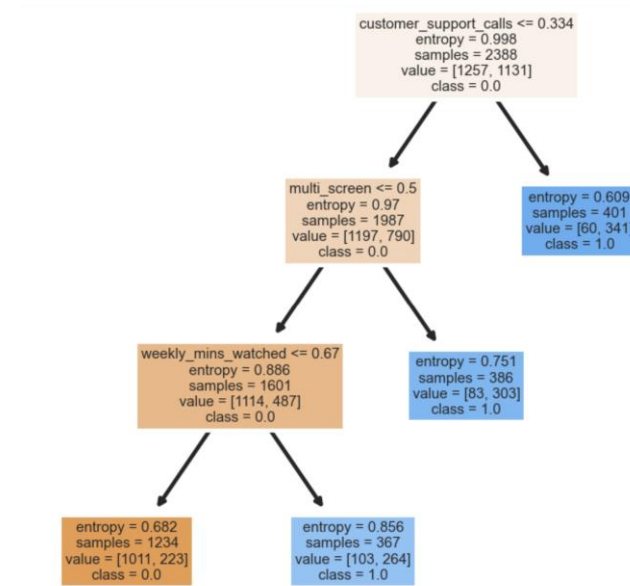
Accuracy: 0.71

Precision: 0.29

Recall: 0.89

F1 Score: 0.44

ROC AUC: 0.79



Recommendations and Next Steps

- Of the two models (optimized logistic regression and decision tree), the decision tree yielded the higher recall score
- The most important features in predicting churn: **weekly minutes watched, multi-screen, customer support calls**
 - Incentivize customers to download the app on multiple devices (exclusive bonus content only accessible using the app)
 - In the long run, approach other companies for partnerships that bundle content and cost
 - Begin A/B testing features like customer support bots to reduce the number of support calls made
 - Identify high risk customers as those approaching the critical support call volume when probability of churn increases and assign a special support member to resolve issues

Questions?