Machine Learning Classification Project

July 11, 2019

Chelsea Power

Project objective:
Create a supervised machine learning classification model to determine if a song will be skipped or not skipped.

Overview

Based on a Spotify list of tracks and user log information from 2018 that included over 160,000 records, I combined these two CSV files into a single dataset to predict if a track will be skipped or not skipped.

Before I trained and tested the data, I reduced the feature set (number of columns) from 51 to 27 by eliminating unnecessary categorical and highly correlated features.

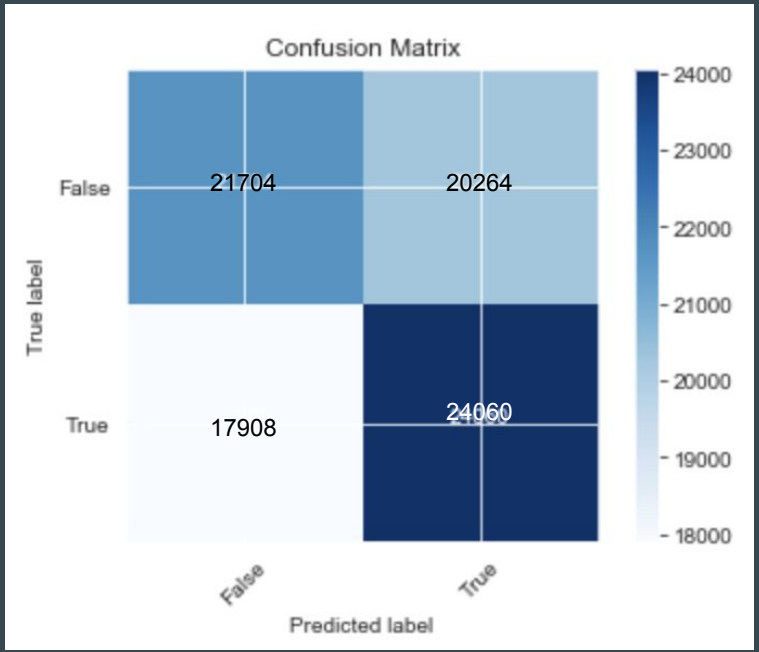
The next few slides will show the classification algorithms used for this project..

The final model was able to perform with **69% accuracy** for whether a track will be skipped or not skipped based on the predictor variable not_skipped.

Confusion Matrix

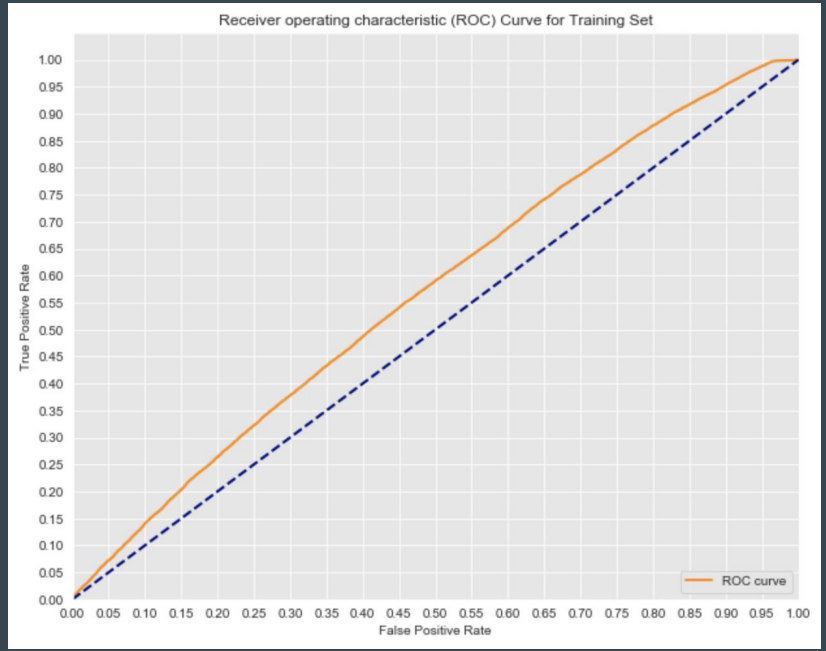
The binary chart to the right shows the accuracy of the classification model based on the test data by comparing true and predicted classes.

1. True Positives: 24060
2. True Negatives: 21704
3. False Positives: 20264
4. False Negatives: 17908



ROC Graph

The orange line in the graph shows how good the model is accurately predicting. The Area under the Curve (AUC) measure is performing at 56%.

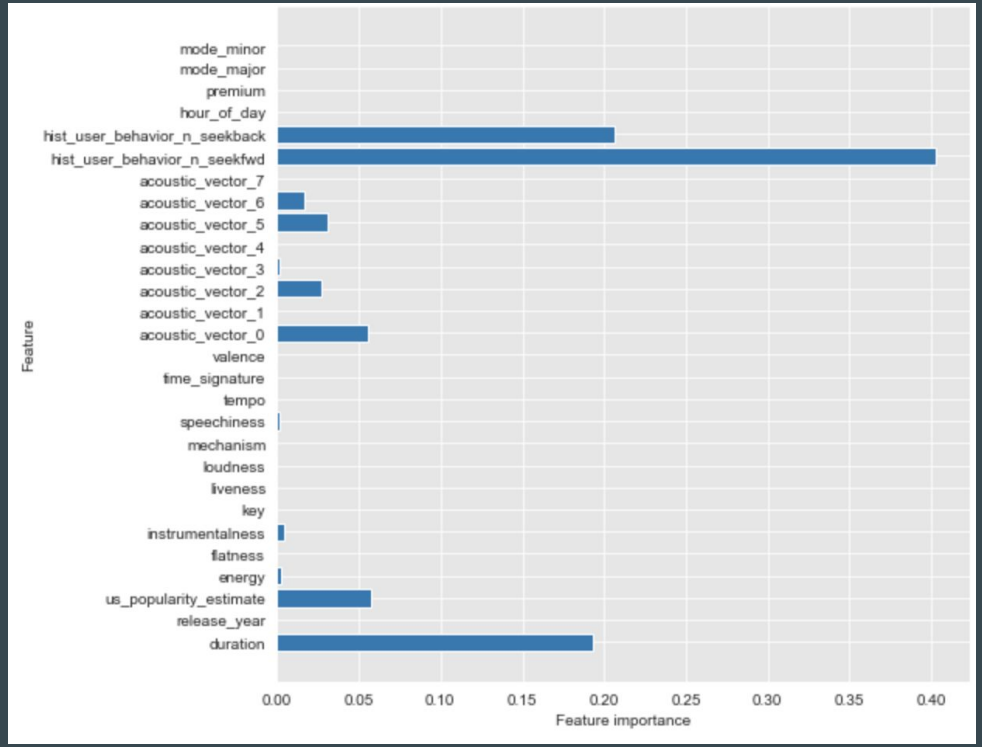


AUC: 0.5681489260366702

Feature Importance

The top five features are:

1. hist_user_behavior_n_seekfwd
2. hist_user_behavior_n_seekback
3. duration
4. us_popularity_estimate
5. acoustic_vector_0



Model Performance

	Precision	Recall	f1-score
False	**0.71**	**0.54**	**0.61**
True	**0.37**	**0.56**	**0.45**
Micro Average	0.54	0.54	0.54
Macro Average	0.54	0.55	0.53
Weighted Average	0.60	0.54	0.56

- Testing accuracy for Decision Tree Classifier: **54.38%**

Precision is the ability of the classifier to not label a true negative observation as a positive.

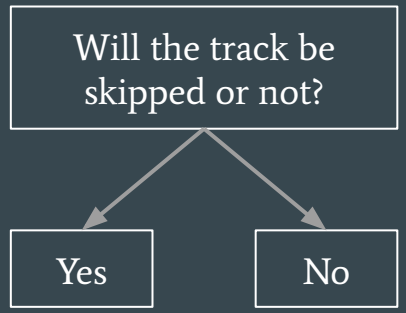
Recall is the ability of the classifier to find positives.

F1-score is the harmonic mean of precision and recall.

Parameter Tuning

Decision Tree

- Testing accuracy: **54.38%**



XG Boost

- Training accuracy: 57.11%
- Validation accuracy: 53.94%

Grid Search

- Mean Training Score: **69.3%**
- Mean Testing Score: 66.87%

Next Steps

After running Grid Search where the accuracy increased to **69.3%**, the model still has potential to be further improved by running **Principal Component Analysis (PCA)**.

Another consideration would be to re-run this project using a different predictor variable to see if the accuracy can be increased further.