Predicting a Song's Success Using Machine Learning

—

Patricio Contreras
19 February, 2021

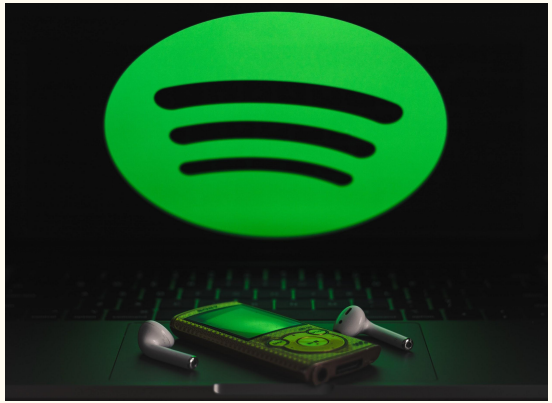
Outline

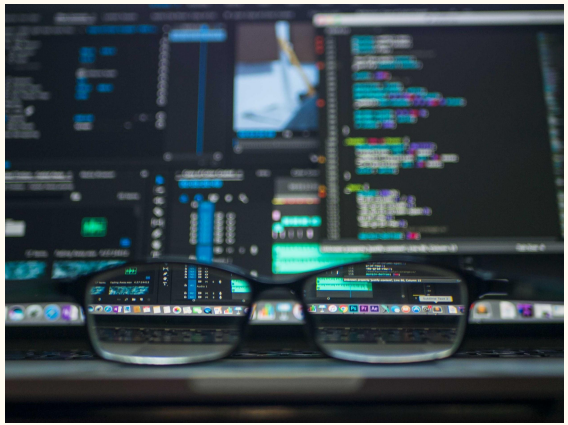
- Business Problem
- Data and Methods
- Results
- Conclusions and Recommendations
- Future Work

Business Problem

- Composing a song is by no means an easy feat
- Music tastes nowadays are more diversified than ever
- Could an artist know if a song will be a hit (or flop) before releasing it?
- Minimise risk of failure

Data and Methods

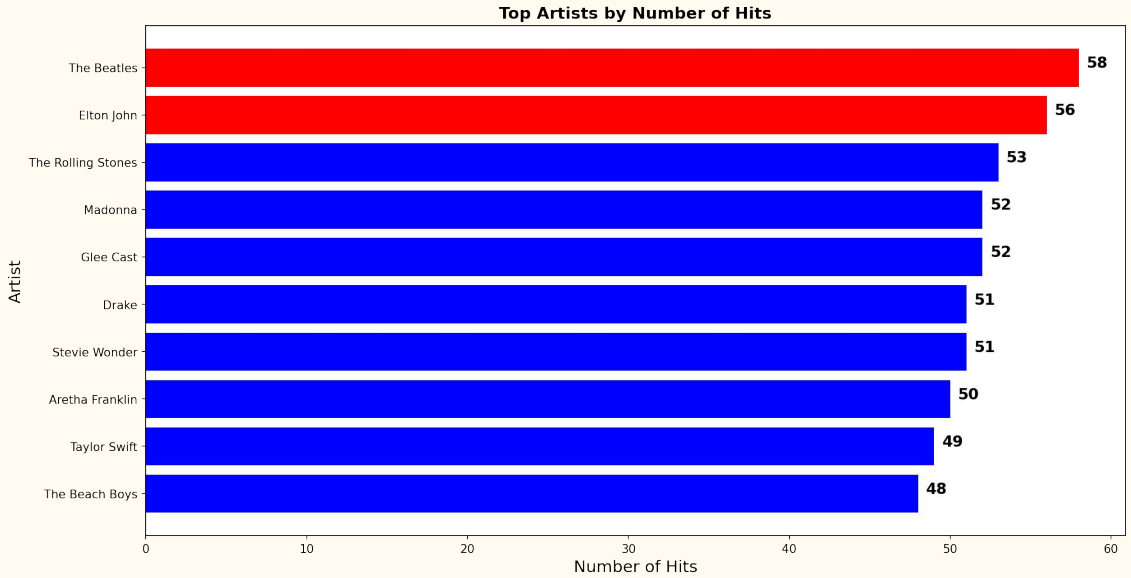


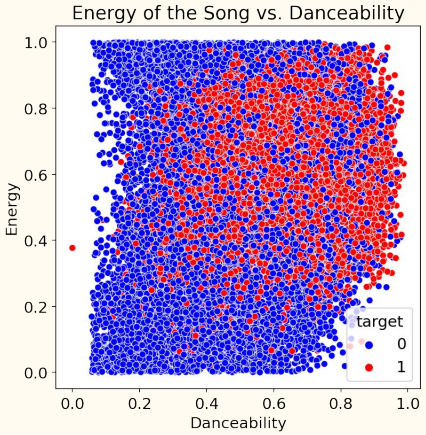


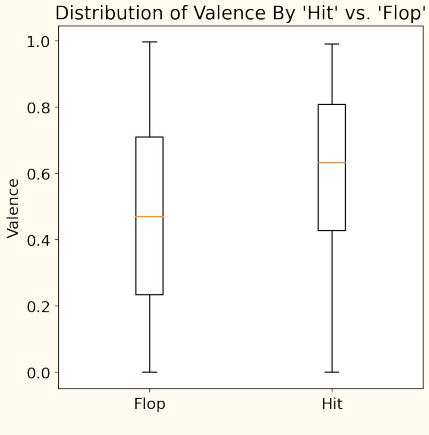
- Dataset fetched using Spotify's Web API contains information on over 40,000 songs from 1960-2019
- Features in this dataset include measures of a track's danceability, energy, loudness, etc.
- The 4 candidate machine learning (ML) algorithms used were Logistic Regression, K-nearest Neighbours, Random Forest, and Gradient Boosting

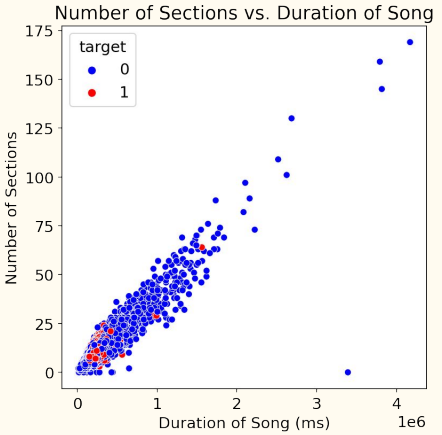
Results

- Unsurprisingly, *The Beatles* are the artist/band with the most number of hit songs in this dataset
- Data suggests higher danceability and energy have an influence on the track's success
- Hit songs in the dataset have a slightly higher valence measure than flops

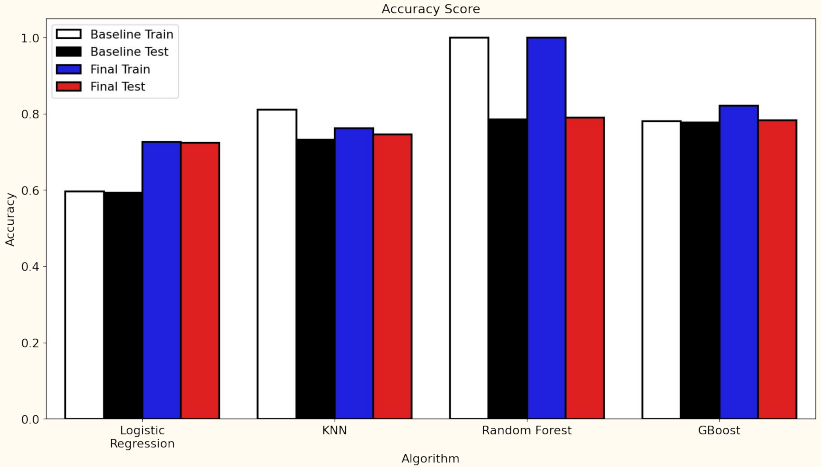


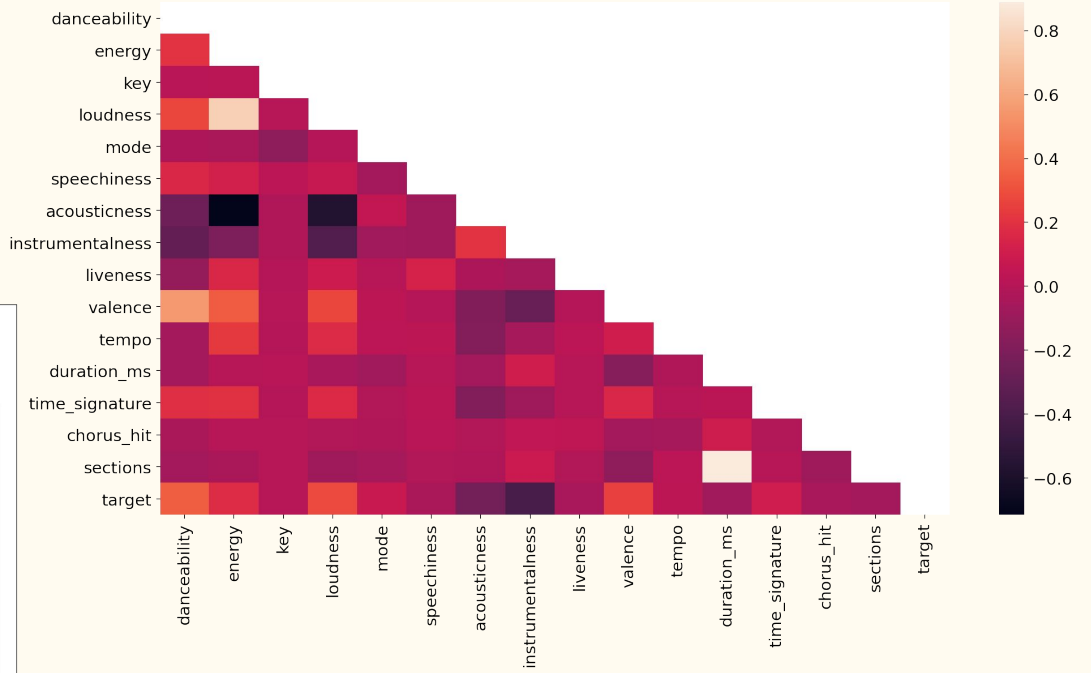






Results





- Random Forest and Gradient Boost classifiers perform the best in terms of accuracy
- Duration and number of sections were highly correlated - could cause multicollinearity problems

Conclusions and Recommendations

- Gradient Boost algorithm had the best overall predictive performance
- After model tuning, the algorithm that improved the most was Logistic Regression
- Gradient Boost classifier achieved an accuracy of 80%

Future Work

- Consider more features to predict a song's success: genre, language, etc.
- Further model tuning can be performed
- Choose other machine learning algorithms to see how they compare to the ones selected here

Thank You!

Email: pcontreras1797@gmail.com
Github: @p-contreras
LinkedIn: linkedin.com/in/pcontreras97/