The Billboard Hot 100 Chart remains one of the definitive ways to measure the success of a popular song. We investigated using machine learning techniques to predict which songs will become Billboard Hot 100 Hits.

We were able to predict the Billboard success of a song with ~75% accuracy using machine-learning algorithms including Logistic Regression, GDA, SVM, Decision Trees and Neural Networks.

Ten audio features were extracted from the Spotify API. Spotify assigns each song a value between 0 and 1 for these features, except loudness which is measured in decibels.

Audio Features

- Danceability
- Instrumentalness
- Acousticness
- Valence
- Energy
- Artist Score

Top Songs: December 2018

- “You Make It Feel Like Christmas” by Chris Norman
- “Thank You, Next” by Ariana Grande
- “Happy” by Pharrell
- “Without Me” by Halsey
- “Girls Like You” by Maroon 5 ft. Cardi B

Supervised Learning: data split 75/25 into training/validation. Logistic Regression and GDA yielded the strongest results.

Bagging using random forests corrected SVM from over-fitting.

Decision Tree performs poorly as it suffers from severe over-fitting.

Neural Network with regularization, using one hidden layer of six units with the sigmoid activation function. The L2 regularization function was applied to the cost function to avoid over-fitting.

Table 1: Audio features extracted from Spotify’s API. Spotify assigns each song a value between 0 and 1 for these features, except loudness which is measured in decibels.

Table 2: Error analysis for the two strongest-performing algorithms. The features at the end of the list decreased the accuracy of predictions.