Predicting the 59th GRAMMY (2017) Record of the Year Winner

Dylan Ong, Eric Hao, Max Schuman, Siyu Zhang
Northwestern University, EECS 349 Machine Learning 2016

Our task is to predict the winner of the 2017 GRAMMY Record of the Year award. Because our model can be easily applied to future GRAMMY seasons, these results will yield insights into subtle differences between popular expectations and actual winners, especially in upset years, and can help provide context for identifying shifts and trends in popular music. Our project output is a ranking of all songs that have made an appearance on the Billboard Top 100 since October 1, 2015 (the cutoff for eligibility for the 2017 GRAMMYs) by their likelihood of winning the Record of the Year award.

Our data set consisted of roughly 5400 songs from 1958-2015 that were on the year-end Billboard Top 100, giving us a set of popular and significant songs spread out over the time that the GRAMMYs have been contested. For each song, we scraped data from Spotify, Last.fm, and various song lyric websites to generate a list of 23 attributes and metrics. These covered general information about each song (year of release, genre, and song duration), measures of each song's popularity (Spotify popularity; Last.fm followers of the artist, listeners to the song, and play count of the song), metrics characterizing each song's lyrics (word count, Flesch reading ease, and the positivity and subjectivity of the lyrics), and metrics characterizing each song's music as calculated by Spotify (key, major/minor modality, and time signature; danceability, energy, tempo, and loudness; speechiness, acousticness, instrumentalness, and liveness; and the positivity of the music).

For testing potential models, we divided our set into a training set of songs from 1958-2010 and a validation set of songs from 2011-2015, as we felt our ability to predict the winners from the most recent years would be most predictive of our ability to predict future winners. To evaluate the ranking methods we tried, we looked at the predicted rankings of the Record of the Year winners and nominees relative to the other songs in their years in the validation set; methods that had winners and nominees closer to the top of their years' rankings were favored.

Two major choices in our approach improved our validation results greatly over time. The first was the decision to turn our task into one of numeric prediction instead of classification by computing the "winner score" (defined as 1 if a song won the Record of the Year and 0.2 if a song was nominated for the award, reflecting that there has almost always been five nominees for each year) for each song in our data set and creating a model that would predict the winner score of each song as a rating of that song's chance of winning the award. Prior to this, we examined logistic regression models that would compute the probability of a song's being classified as a winner in a binary classification problem and used this to generate rankings, but did very poorly in predicting the years 2011 through 2015. By incorporating the additional information about which songs were nominees, we were able to greatly improve our validation results while testing linear regression models, SVMs, and multilayer perceptron models. Ultimately, our best results were achieved using a multilayer perceptron model to predict song winner scores.

We saw another major improvement to our validation results when we began weighting more recent data more heavily in our training sets to account for the likelihood that more recent results are more indicative of current musical trends and therefore carry more predictive power moving forward. Through testing various weighting techniques, we settled on this logistic weighting equation:

$$W = \frac{2}{1 + e^{(Last\ Training\ Year - Year\ of\ Song)/5}}$$

Using these weights on our training data gave another boost to our validation performance, as our multilayer perceptron models went from ranking the winners in our validation years at roughly 10th on average to roughly 5th.

Our final model was a multilayer perceptron model consisting of two hidden layers of 23 and 4 nodes, respectively, using a learning rate of 0.1 and trained over 200 epochs. The following graph depicts the resulting validation rankings of the Record of the Year winners and nominees in their years using this model.
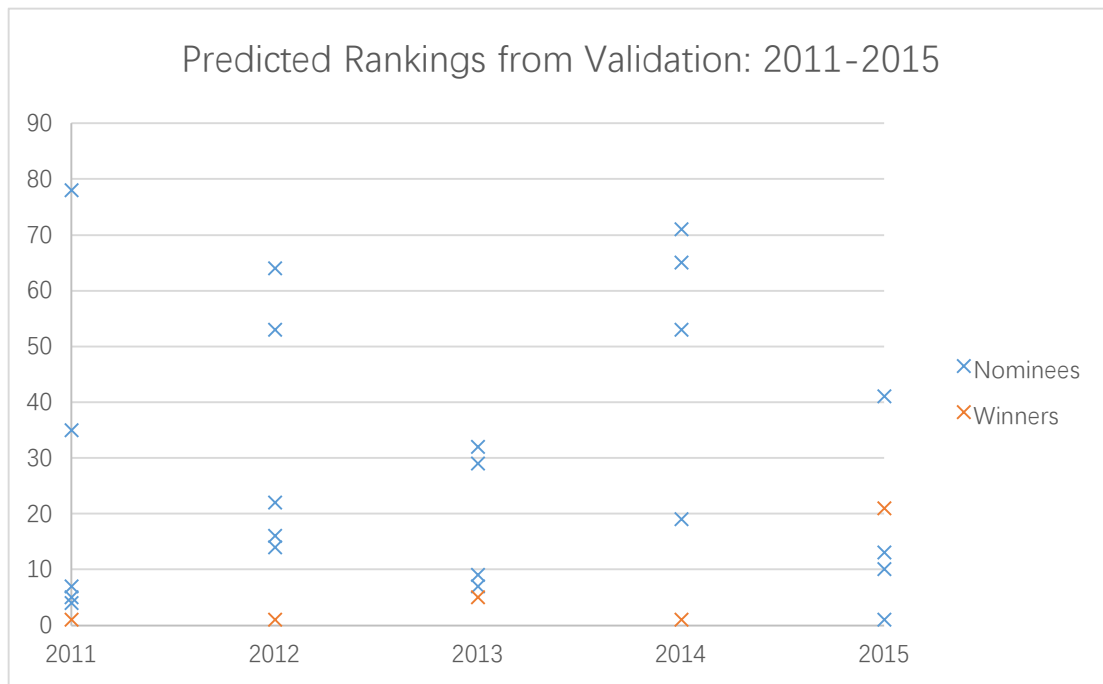
Notably, our model correctly predicted the winners in 2011, 2012, and 2014, while the winner in 2013 was ranked 5th. In all four of those years our model also ranked the eventual winners higher than any of the nominated songs. 2015 stands out as a bit of a miss, with our model ranking the eventual winner, Uptown Funk, 21st out of that year's songs, but given that the song was one of the few representatives of the funk genre in our data set and is the only song of that genre to ever win Record of the Year, it makes sense that we would have difficulties accurately predicting its performance relative to our performance in other years. Additionally, note that our performance in terms of ranking eventual nominees highly was much more inconsistent than our performance in ranking winners.

We then used our model to rank the songs that have appeared on the Billboard Top 100 in any capacity since October 1, 2015, by their predicted winner score to create our predictions. In total, we ranked 286 songs; our top-20 songs are listed below, with more listed on our project website.

Rank	Song	Artist	Predicted Winner_Score
1	Ultralight Beam	Kanye West	0.248
2	Sandcastles	Beyonce	0.151
3	Blackstar	David Bowie	0.113
4	Hold Up	Beyonce	0.083
5	Water Under The Bridge	Adele	0.076
6	Lazarus	David Bowie	0.049
7	Love Yourself	Justin Bieber	0.047
8	Zero	Chris Brown	0.046
9	Tennessee Whiskey	Chris Stapleton	0.044
10	Traveller	Chris Stapleton	0.044
11	Photograph	Ed Sheeran	0.035
12	Purpose	Justin Bieber	0.033
13	Dibs	Kelsea Ballerini	0.032
14	Die A Happy Man	Thomas Rhett	0.031
15	7 Years	Lukas Graham	0.030
16	All I Ask	Adele	0.029
17	Mark My Words	Justin Bieber	0.025
18	Life Is Worth Living	Justin Bieber	0.024
19	Lonesome Broken And Blue	Adam Wakefield	0.024
20	Blessings	Chance The Rapper	0.023

By our predictions, we feel that Ultralight Beam by Kanye West is the clear favorite for the 2017 Record of the Year award. Our top five songs, the projected nominees, are fairly separated from the pack (although it is worth nothing that no artist has ever had two songs nominated for Record of the Year in the same year, as we predict with Beyonce in our ranking). However, Ultralight Beam is the only song that clears the 0.2 threshold by projected score, indicating that our model does not rate this year's songs to date particularly highly.

Moving forward, it is possible that we could see better predictive results with a different formulation of our winner score metric or with a different weighting scheme with further iteration. Additionally, while our data set presently contains roughly 100 songs from each year, further work to expand that data set could be done to create a more representative sample of all music created in our time period of interest. Finally, as we've created these rankings roughly four months before the deadline for eligibility in the 2017 GRAMMYs, we ought to revisit these rankings in four months to incorporate newer songs and get a complete picture of the field's chances. We are certainly looking forward to next year's GRAMMYs as a test of our model's predictive capabilities.

Teamwork:

Data: Dylan Ong, Eric Hao, Max Schuman

Model: Max Schuman, Siyu Zhang

Report: Siyu Zhang, Max Schuman

Website: Eric Hao, Siyu Zhang