

Hit Song! Predicting Songs on the Billboard Top 100

Julien Kawawa-Beaudan and Gabriel Garza
CS 229: Machine Learning



Stanford University

Problem

To classify songs as Billboard Top 100 songs based on genre description and audio features

Dataset

10,000 song subset from 1 Million Song Dataset, Created by LabROSA and The Echo Nest

Labels

Positive: Song labeled positive if it appeared on any monthly Billboard Top 100 List since 1958.

Negative: Song labeled negative if its artist has never appeared on any monthly Billboard Top 100.

Features

- Pitches and timbres Sampled every 250ms
- Genre as indicated in 1 Million Song Dataset
- Metadata provided by The Echo Nest

Methods

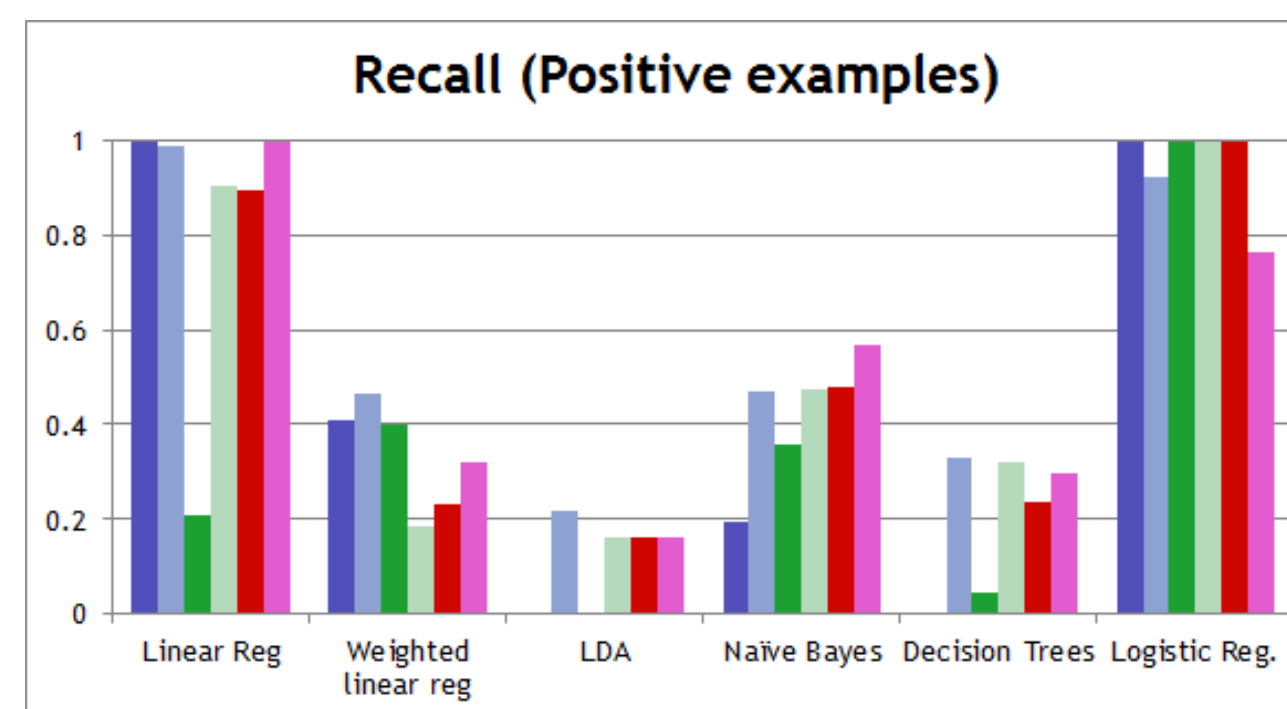
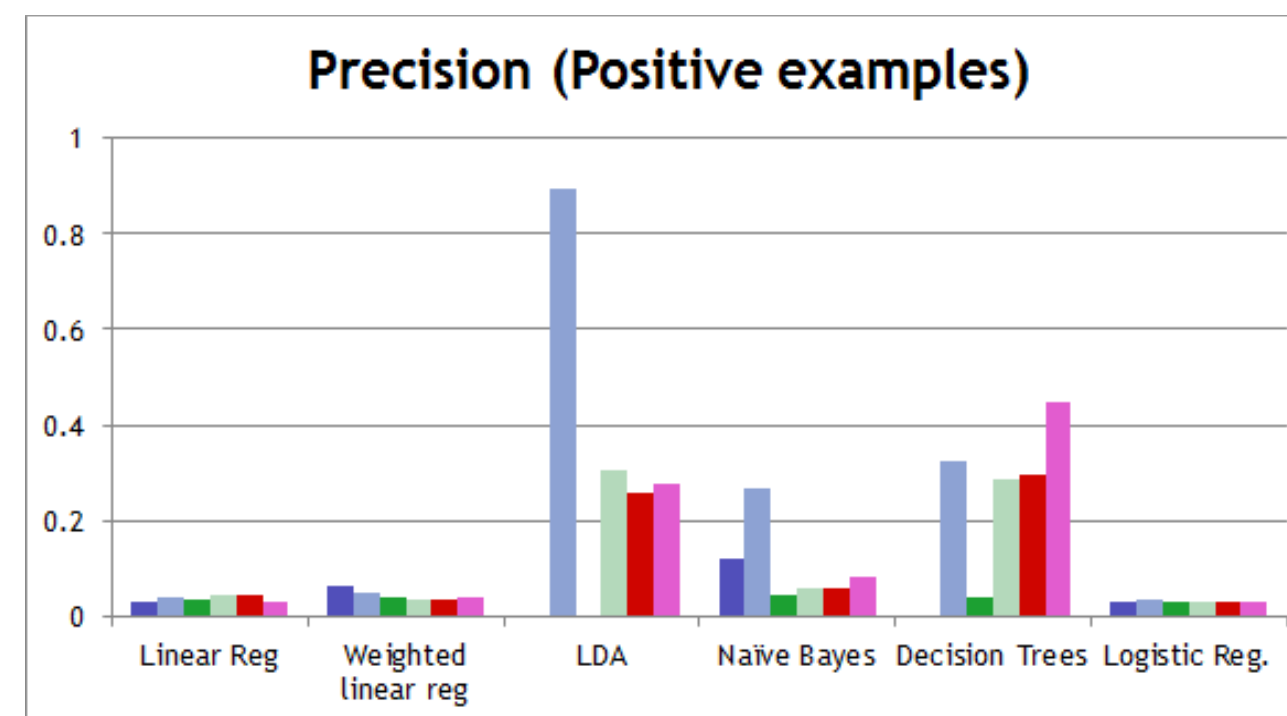
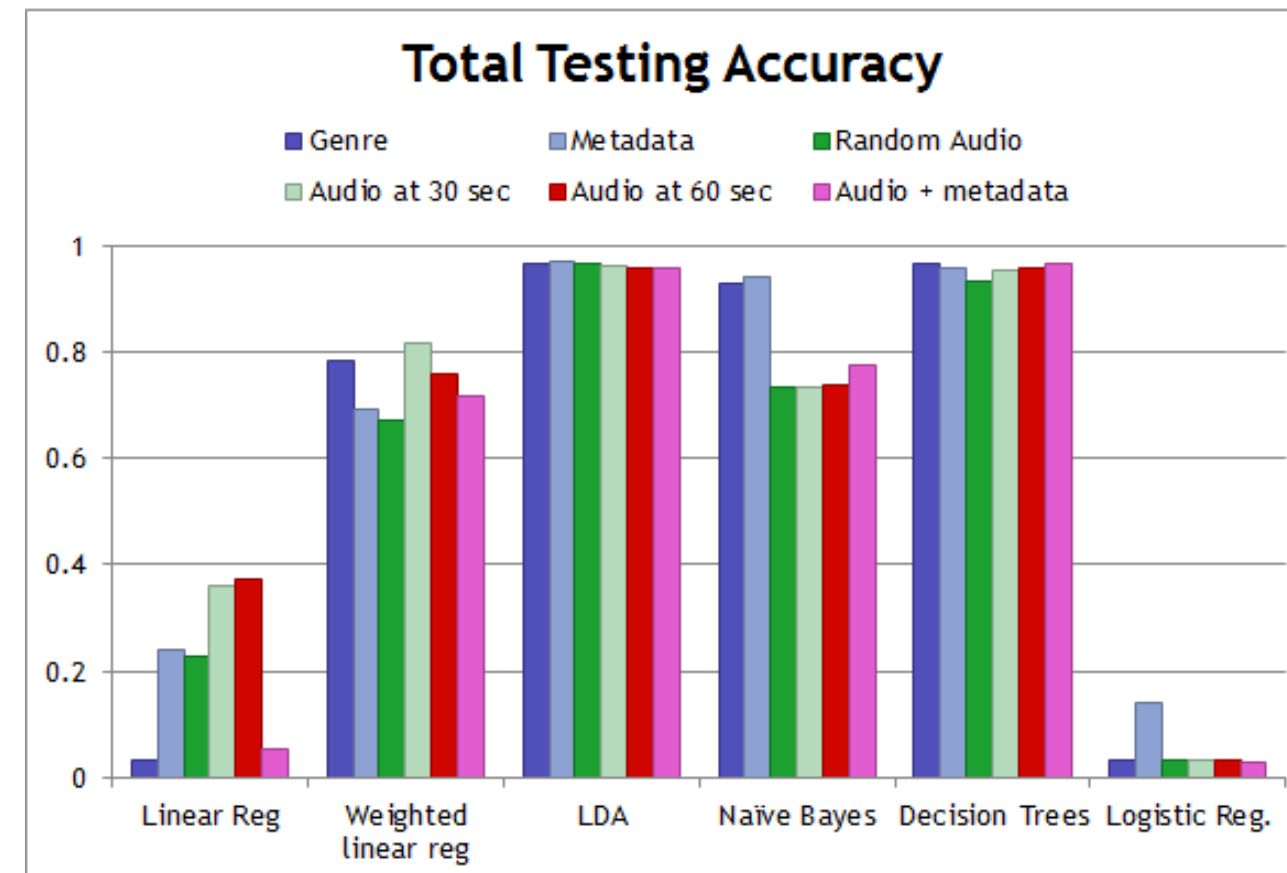
Linear and Logistic Regression
Naive Bayes
Support Vector Machine
Gaussian Discriminant Analysis
Decision Trees

References

Thierry Bertin-Mahieux, Daniel P.W. Ellis, Brian Whitman, and Paul Lamere. The Million Song Dataset. In Proceedings of the 12th International Society for Music Information Retrieval Conference (ISMIR 2011), 2011.

Provost, Foster. "Machine Learning From Imbalanced Data Sets." New York University. New York University. Web. Accessed Dec. 4, 2015.

Wang, Kedao. "Predicting Hit Songs with MIDI Musical Features." CS 229 Projects 2014, Dec. 2014. Web.



Results

Features	Best Classifier	Accuracy	Pos. Recall	Neg. Recall	Pos. Precision	Neg. Precision
Genre	Naïve Bayes	0.9287	0.9535	0.1919	0.9721	0.1226
Metadata	LDA	0.9684	0.993	0.0459	0.9688	0.9688
Random Audio	Naïve Bayes	0.7363	0.749	0.3586	0.9718	0.0461
Audio at 30 seconds	LDA	0.9605	0.981	0.1633	0.9722	0.3048
Audio at 60 seconds	Decision Tree	0.9517	0.7792	0.2479	0.9268	0.2864
Audio and Metadata	Decision Tree	0.9567	0.7855	0.185	0.9744	0.2975

- Challenges: Dataset is highly skewed, only 3.3% positive
 - Trivial to achieve high accuracy without predicting any songs to be positive
 - Recall and precision more important
- Neither genre nor metadata sufficient to classify with high accuracy (even with hotness feature)
- Experimented with selecting different audio segments
 - 4 segments equally spaced in song
 - 10 segment at 30 seconds and at 60 seconds
- Most successful classifiers were LDA and Decision trees

Future Directions

- Lyrics
 - Similarity to hit songs
- Additional Audio Processing
 - Time series
- MIDI features
 - Convert audio to midi for cleaner features