RAGE AGAINST THE MACHINE LEARNING: LEARNING TO PREDICT SONG POPULARITY

Andreas “King of the Keys” Garcia, Cody “C-Dawg” Kala, Gabriel “The Getaway Driver” Barajas

{andreas4, codykala, gbarajas}@stanford.edu

INTRODUCTION

- What about a song determines its popularity? What makes a song ‘good’ is subjective but there are certainly trends in how music is perceived and these trends are complex and interesting.
- We used machine learning to determine the features of songs that are most important for predicting song popularity.

DATASET (DS)

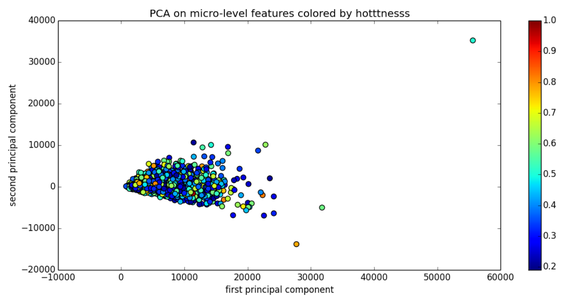
- The Million Song Dataset (MSD) contains almost 500 GB of song data and metadata from which we extract features for our learning models.
- We used a subset of the MSD containing 10,000 songs to train and develop our learning models.
- To measure popularity, we used “hotttnesss”, which is a metric developed by Echonest and used in the MSD.

Song Hotttnesss	Artist_Familiary	Artist_Hotttnesss	Artist Name Length	Song Title Length	Song Duration	End-of-Fade-In Time	Loudness	Tempo
0.7499	0.8289	0.6197	11.8431	17.7955	244.5903	0.8443	-6.6596	129.6999
0.6539	0.6279	0.4278	13.4236	16.5718	253.7144	0.8777	-9.8522	112.4614
0.5754	0.6903	0.4678	13.8771	17.4373	228.8378	0.5218	-7.5546	175.6721
0.537	0.781	0.5608	11.541	20.8525	236.4169	0.6664	-7.5523	107.2311
0.48	0.565	0.3856	12.6125	16.5646	238.0609	0.5959	-8.4368	113.5264
0.3843	0.5886	0.403	13.8471	17.8471	223.9007	1.5455	-22.489	100.8855
0.3208	0.5516	0.3802	12.2789	17.4917	234.8006	0.806	-9.3324	165.792
0.3111	0.7036	0.4839	12.8506	24.0482	248.6352	0.6782	-8.639	118.8304
0.2841	0.518	0.3598	12.7381	18.9416	253.0461	0.7468	-10.1417	98.8172
0.2754	0.3196	0.2096	15.9898	19.6837	234.3519	0.829	-11.6471	125.2274

K-means centroids on Macro-level features

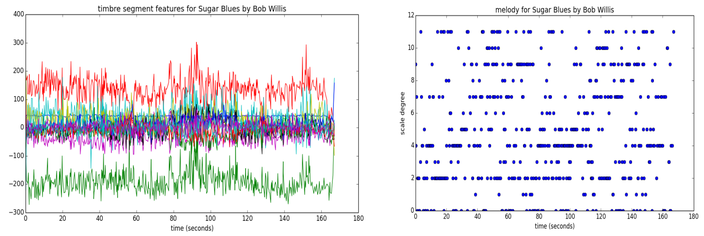
FEATURE EXTRACTION (FE)

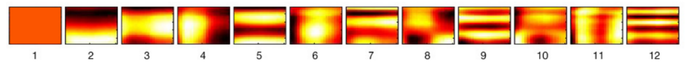
- **Macro-level features:** included song key, tempo, duration, length of song name, time signature, fadeout duration, loudness.
- **Micro-level features:** extracted features from the MSD “segment” feature arrays with various statistics.
- **Bag-of-words features**: used term frequencies of 100 most common words as features for our models.
- **Location features:** Artist origin is converted to indicator features for the top 10 most common locations.



To look for trends in the micro-level features, we used PCA to visualize a 177-dimension feature space.

FE (CON’T)

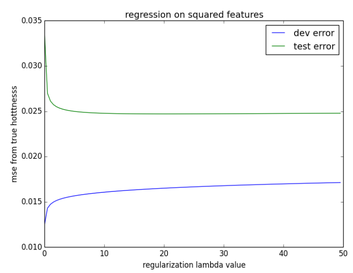


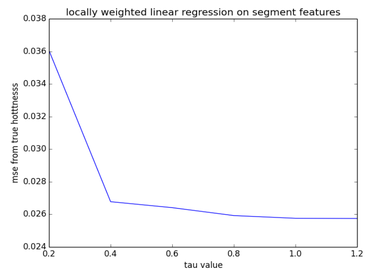


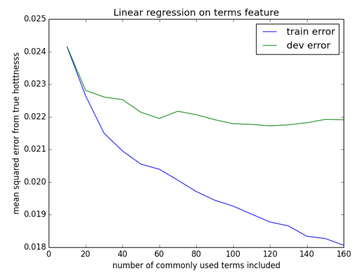
12 basis functions for the timbre vector: x = time, y = frequency, z = amplitude

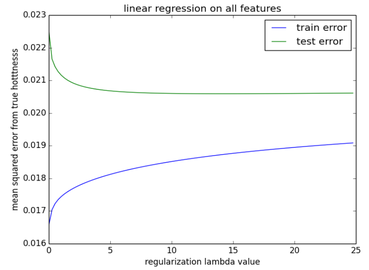
In the MSD, songs are split into hundreds of ‘segments’ that are roughly uniform in timbre and pitch. The timbre is analyzed for each of these segments, yielding a value corresponding to each of the 12 spectral basis functions above. Pitch information is also included for each segment. We manipulated this information to extract information about the song’s melodic and harmonic content. This yields incredibly descriptive information about what a song sounds like.

IMPLEMENTATION & EXPERIMENTS (I&E)



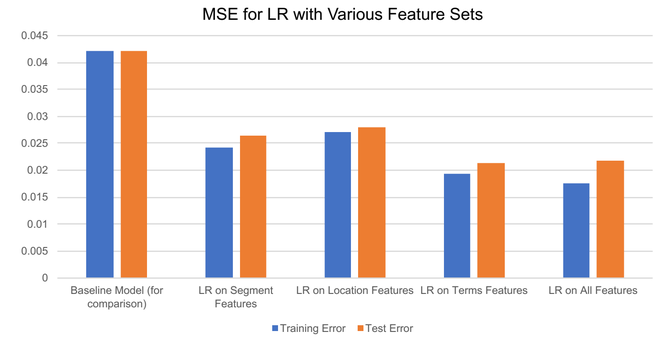






Due to the limited size of the workable dataset and the large number of features, certain models resulted in overfitting, including polynomial regression and locally weighted linear regression. Above are some plots used to tune parameters and choose models.

RESULTS & ANALYSIS (R&A)



Highest weighted terms	Lowest weighted terms
‘soundtrack’	‘r&b’
reggae’	‘techno’
‘acoustic’	‘70s’
‘electro’	‘indie rock’
‘guitar’	‘american’

- **Most common artist locations sorted by increasing hotttnesss**: 'United States’, 'Texas', 'Kingston, Jamaica’, 'London, England’, 'New York, NY’, 'Brooklyn, NY’, 'Los Angeles, CA’, 'Chicago, IL’, 'Atlanta, GA'

Highest weighted micro-level features	Most negative weighted features
Variance of segment max loudness	Average of timbre measure 5
Frequency of minor 3rd in melody (relative to key)	Frequency of minor 2nd in melody
Average density of major 7th	Average attack time of segments
Variance of change in timbre measure 8	Average of timbre measure 3
Variance of timbre measure 10	Average of timbre measure 4

FUTURE WORK (FW)

- We are considering implementing classifier models to make direct predictions about song popularity.
- Time permitting, we would like to implement a neural network for predicting song hotttnesss.

REFERENCES

[1] 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

[2] James Pham, and Edric Kyauk, and Edwin Park. ”Predicting Song Popularity”.

[3] Mussman, Stephen, and Moore, John, and Coventry, Brandon. “Using Machine Learning Principles to Understand Song Popularity”.