Dance Dance Devolution: Supervised Learning with two left feet

...

Spenser Johnson
Mentor: Koyuki Nakamori

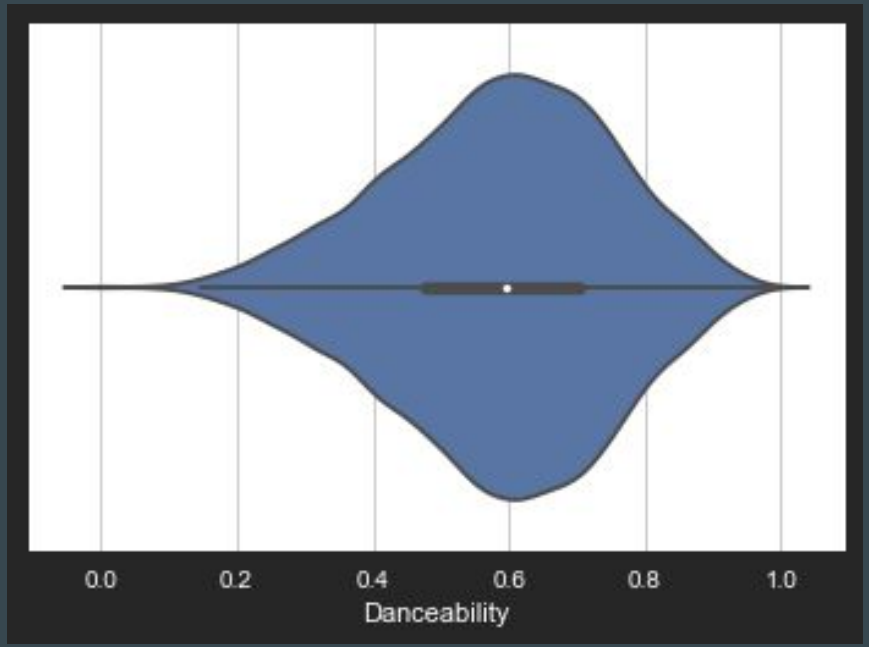
"Dancing! It is a primal art form used in ancient times to express yourself with the body and communicate!"
- Michael Gary Scott

Introduction

- Dataset currently contains 7,167 songs
- Contains top 100 most streamed Spotify songs in 2017, all time best-selling artists per Wikipedia, and a dataset from a Kaggle user
- Song data pulled from Spotify API with python script
- Song attributes calculated by neural network “EchoNest”
- Regression models scored poorly, so problem rephrased as classification
- Models will predict whether a song is danceable or not danceable

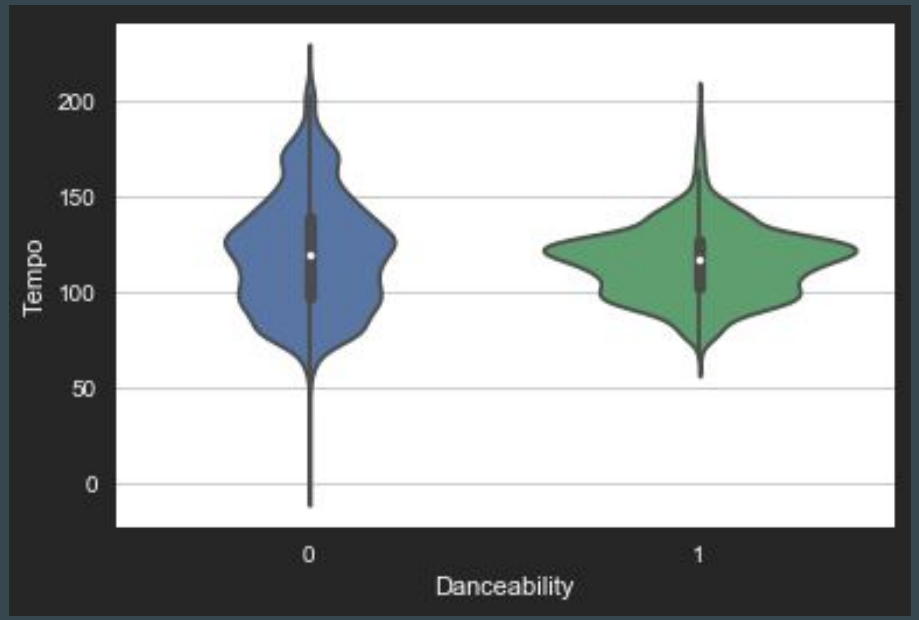
Problem Classification

- Danceability - Describes how suitable a track is for dancing based on a combination of musical elements. A song is considered “Danceable” if the danceability is greater than 0.7



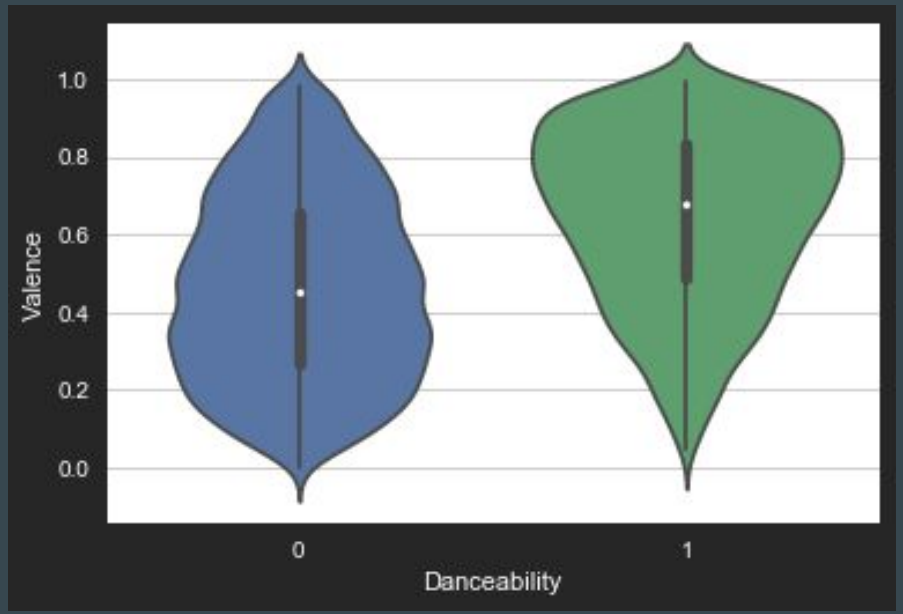
Exploratory Data Analysis

- Tempo - The overall estimated tempo of a track in beats per minute (BPM). Here 1 means the song has a danceability of greater than 0.7, 0 means less than 0.7



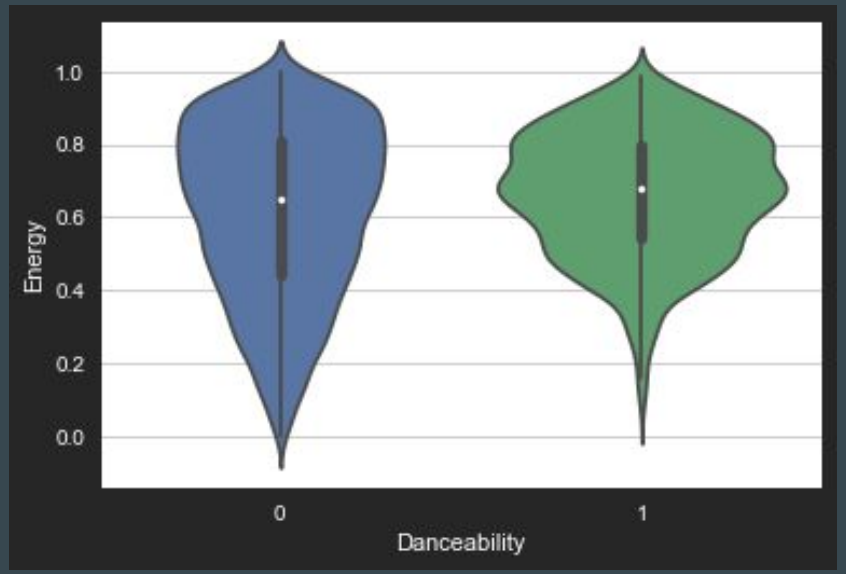
Exploratory Data Analysis, cont.

- Valence - A measure from 0.0 to 1.0 describing the musical positiveness conveyed by a track. Tracks with high valence sound more positive.



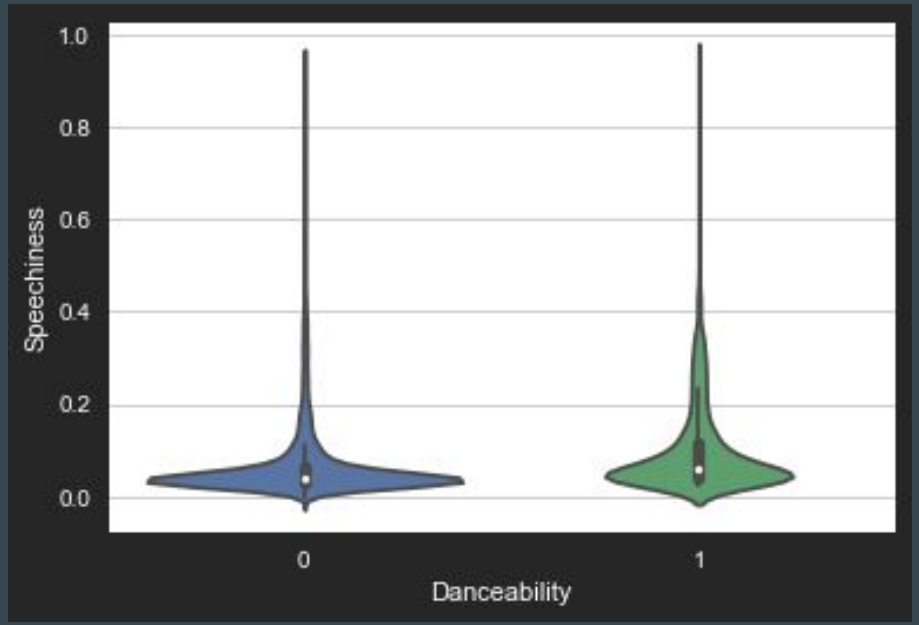
Exploratory Data Analysis, cont.

- Energy - Energy is a measure from 0.0 to 1.0 and represents a perceptual measure of intensity and activity. Perceptual features contributing to this attribute include dynamic range, perceived loudness, and timbre.



Exploratory Data Analysis, cont.

- Speechiness - Detects the presence of spoken words in a track. The more exclusively speech-like the recording (e.g. talk show, audio book, poetry), the closer to 1.0 the attribute value.



Exploratory Data Analysis, cont.

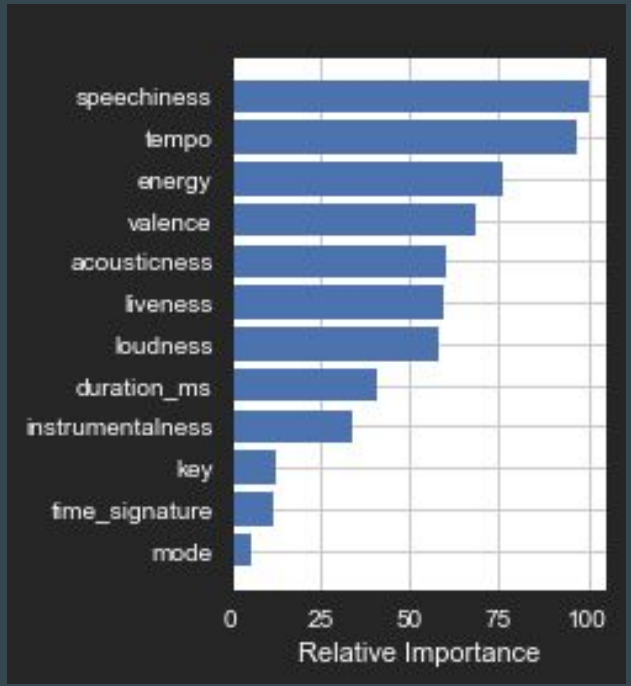
- Acousticness - A confidence measure from 0.0 to 1.0 of whether the track is acoustic. 1.0 represents high confidence the track is acoustic.
- Instrumentalness - Predicts whether a track contains no vocals. The closer the value is to 1.0, the greater likelihood the track contains no vocal content.
- Liveness - Detects the presence of an audience in the recording. Higher liveness values represent an increased probability that the track was performed live.
- Duration - Time of song
- Key - Key of composition
- Mode - Major / minor scale

Model Selection and Comparison

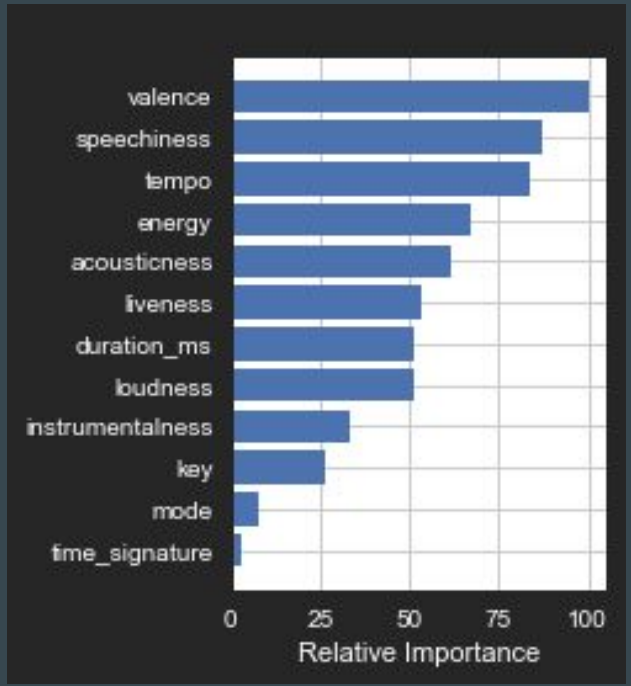
Model	Score	Runtime	F1 Score	Precision	Recall
KNN	.73	27 ms	.04	.34	.02
Logistic Regression	.73	19 ms	N/A	N/A	N/A
Random Forest	.81	48 ms	.61	.34	.51
Gradient Boost	.82	15 ms	.62	.71	.55
Ensemble Vote Classifier	.80	113 ms	.44	.83	.29

Feature Importance

Gradient Boosting Classifier



Random Forest Classifier



Further exploration

- Create a useful feature / transformation to help models
  - PCA did not improve model scores
  - NLP features (artist / title) were unimportant
  - Combining features was ineffective
  - Transforming features yet to make a difference
- Unsupervised learning methods, clustering, unsupervised feature generation
- Future projects:
  - UI where users can input audio features to see if a song is danceable, build out into app
  - Neural network DJ that uses tempo, energy, valence, etc. to create mixes