How well do we know our...

MUSIC

...And can machine learning improve it?

By Alex Cave

CAVE

How big is the Music Industry?

- Huge - $16.1bn in 2016, and up 7% YoY
- Digital music: 33%
- Streaming revenue - $5.4bn in 2016, up 57% YoY
- Spotify - 43% of 106.3m worldwide subscribers.
- Will increase by 40.3m by the end of 2017.

CAVE

Digitial music management

- Spotify and Apple Music – 30-40 million song catalogue!
- Classificaton of music increasingly important
  - A) Discovery
  - B) Subscriber attrition

How do we classify music?

- Imagine if this all had to be done by hand….

CAVE

Goals: Genre Classification

- strong to suggest that automatic genre classification is mistake free…
- …but can be more efficient than manual input
- Use machine learning to classify music based on the features of the track
- beneficial to any organization that needs to classify and group data within a large pool of observations

CAVE

Method of analysis

- Obtain Data from multiple sources and aggregate
- Clean the dataset (null values, unnecessary fields)
- Perform EDA
- Modelling
- Review

Potential datasets

- few datasets available - copyright restrictions
- million song dataset - 300gb, and no genres
- Sample sets with genres, but no features
- Solution! to create a new dataset using the spotify api
  - link the features to genres by combining datasets

CAVE

Million Song Dataset

- Freely-available collection of audio features and metadata for a million contemporary popular music tracks
- https://labrosa.ee.columbia.edu/millionsong/

```
# PATH TO Track Metadata from Million Song dataset
dbfile = '../../resource-datasets/msdextra/AdditionalFiles/track_metadata.db'

# connect to the SQLite database
conn = sqlite3.connect(dbfile)

# from that connection, get a cursor to do queries
c = conn.cursor()
q = "SELECT * FROM songs"
res = c.execute(q)
ids = res.fetchall()
```

CAVE

Acoustic Brainz

- Dataset of corresponding track ids from the million song dataset to other services such as Spotify (json format)
- http://labs.acousticbrainz.org/million-song-dataset-echonest-archive

```
In [353]: with open('../../../../Downloads/millionsongdataset_echonest2/SOAAADD12AB018A9DD.json')
          as json_data:

                  loaded_json = json.load(json_data)
                  tracks = loaded_json['response']['songs'][0]
                  song = tracks[u'tracks'][0]
                  print song
```


```
{u'album_type': u'unknown', u'release_image': u'http://artwork-cdn.7static.com/static/img/sle
eveart/00/006/594/0000659454_200.jpg', u'album_date': u'2000-08-15', u'foreign_release_id':
 u'7digital-UK:release:659454', u'preview_url': u'http://previews.7digital.com/clip/7307902',
u'catalog': u'7digital-UK', u'foreign_id': u'7digital-UK:track:7307902', u'album_name': u'The
Room', u'id': u'TRFXTSY12E5AC77165'}
```

Spotify

- Provides the API to extract music features
- https://developer.spotify.com/web-api/
- Output: List of list of dictionaries

```
import spotipy
from spotipy.oauth2 import SpotifyClientCredentials

client_credentials_manager = SpotifyClientCredentials(client_id='77d05ef2544d4bd9b0f6e5a6119f4d3
sp = spotipy.Spotify(client_credentials_manager=client_credentials_manager)

for x in sptable.spotify_uri:
    time.sleep(0.1)
    spdata.append(sp.audio_features(str(x)))
```

C Λ V E

Data Dictionary

- Acousticness
- Energy
- Danceability
- Liveness
- Tempo
- Speechiness
- Instrumentalness
- Loudness
- Key
- Mode
- Valence

How are these features derived?

- Analyse the audio of a track electronically, directly
  - Instruments
  - Vocals
  - Decibels

- User derived metadata
  - How does a song make you feel?
  - Mood -> Valence

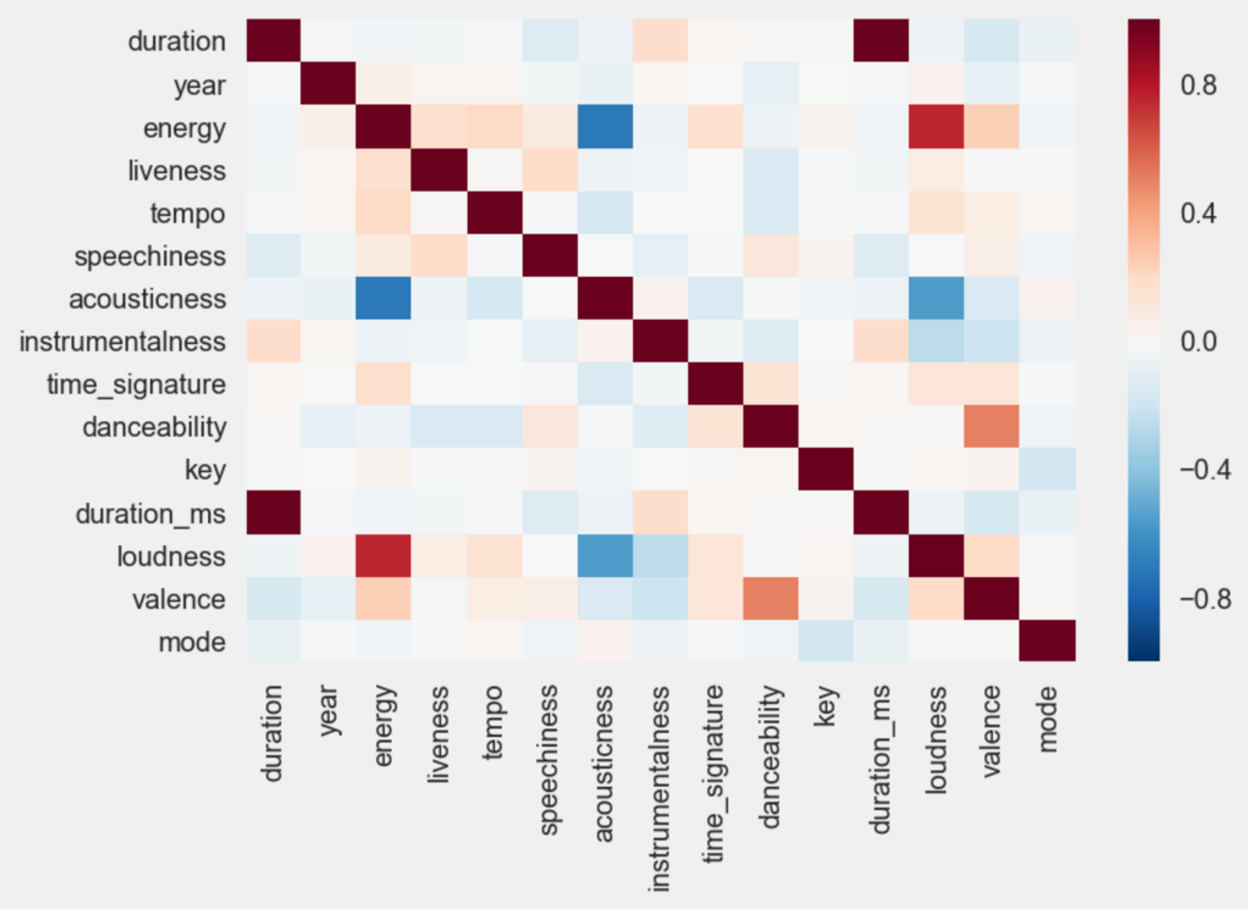
Final Dataset

Pop_Rock	61846
Electronic	9126
Rap	5375
Jazz	4099
Latin	3925
International	3606
RnB	3155
Country	2793
Blues	2167
Folk	1488
Vocal	1185
New Age	1168
Reggae	1039

- 100,972 Observations
- 13 Top-level Genres
- Baseline Accuracy: 0.6125
- A lot of Pop-Rock Songs!

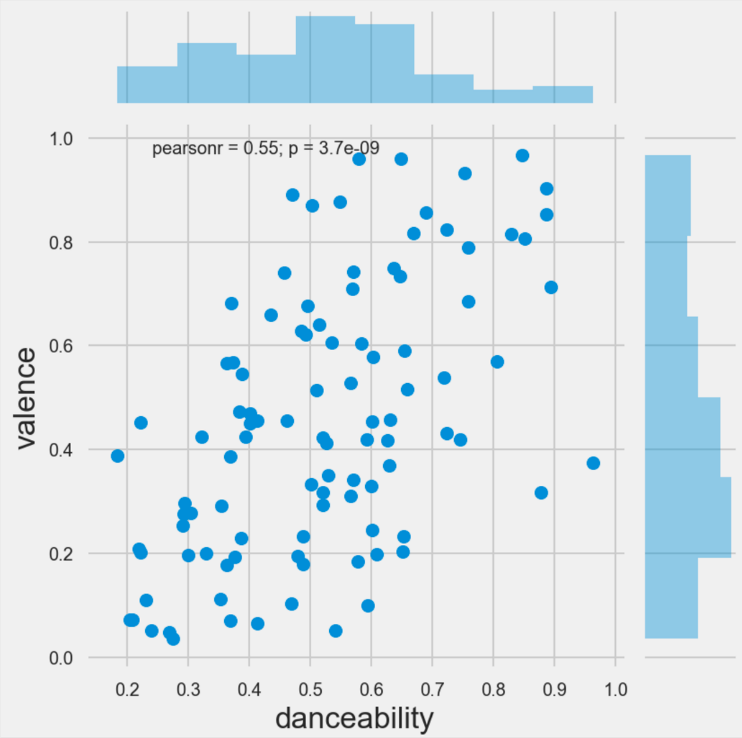
CAVE

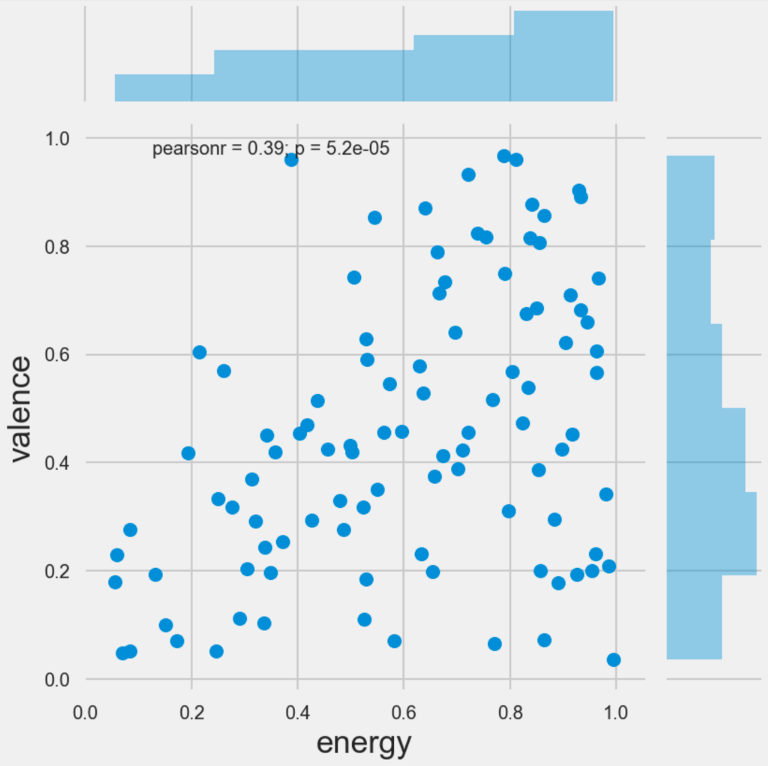
Correlation plots

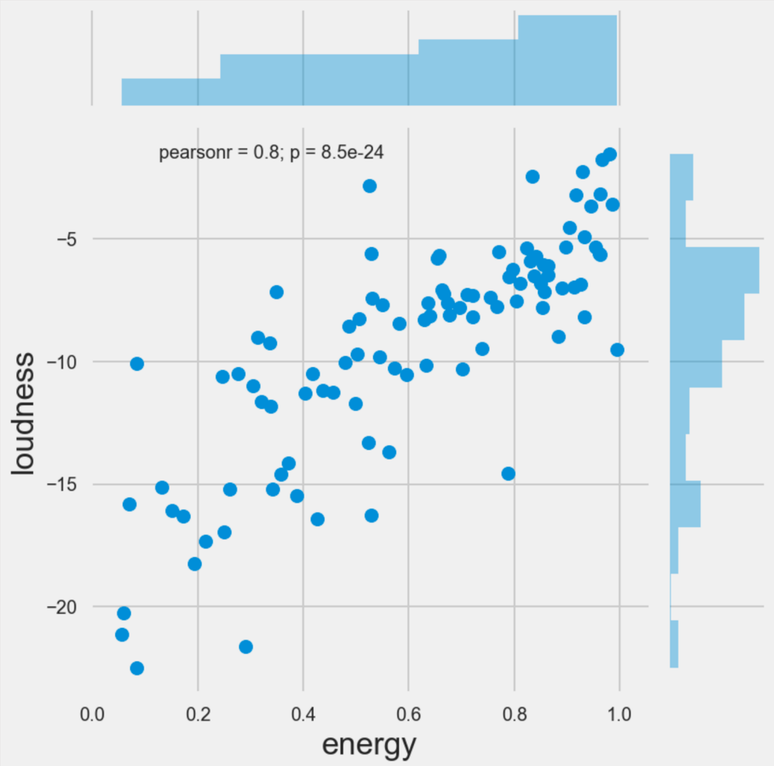


CAVE

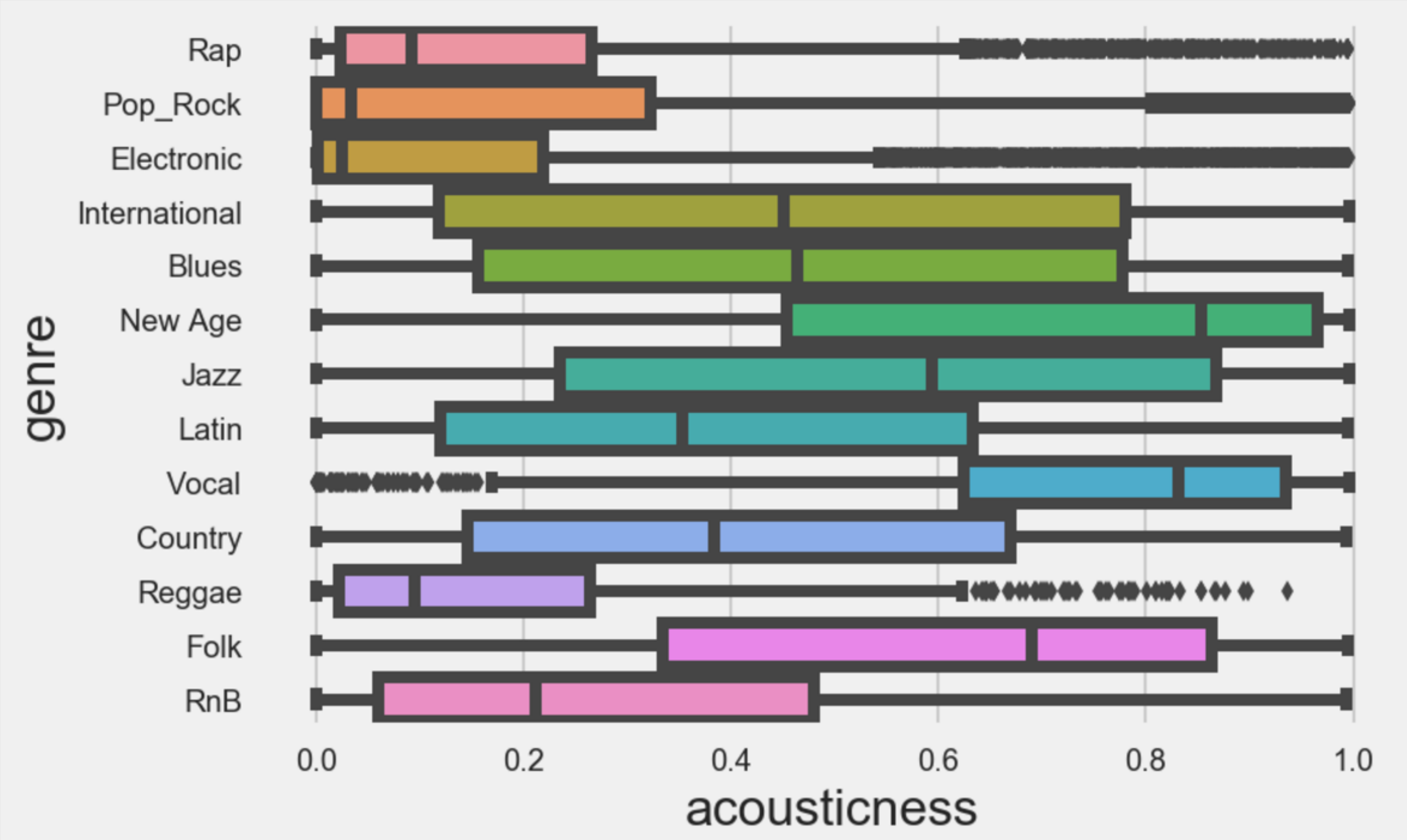
Correlations – A closer look



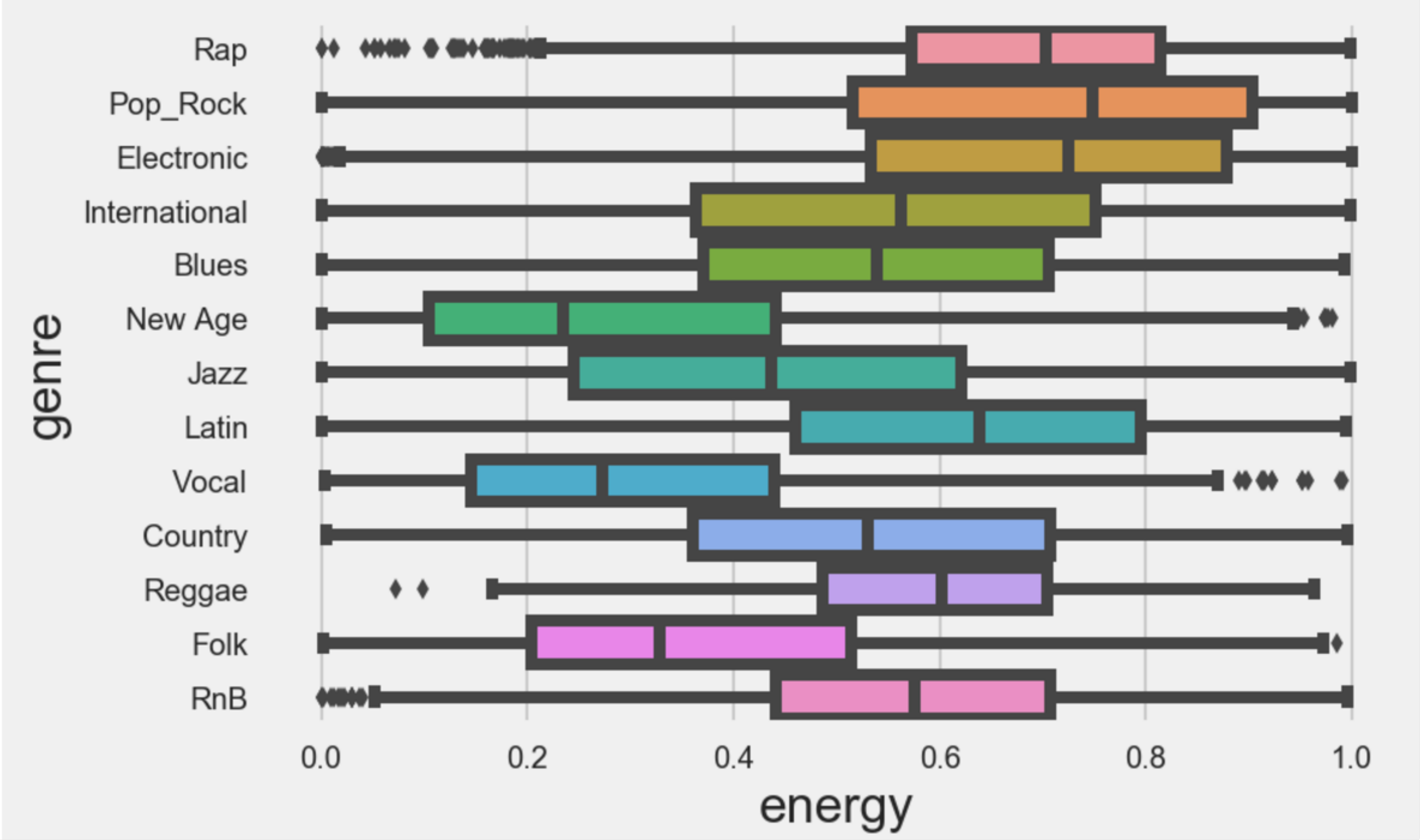




Acousticness

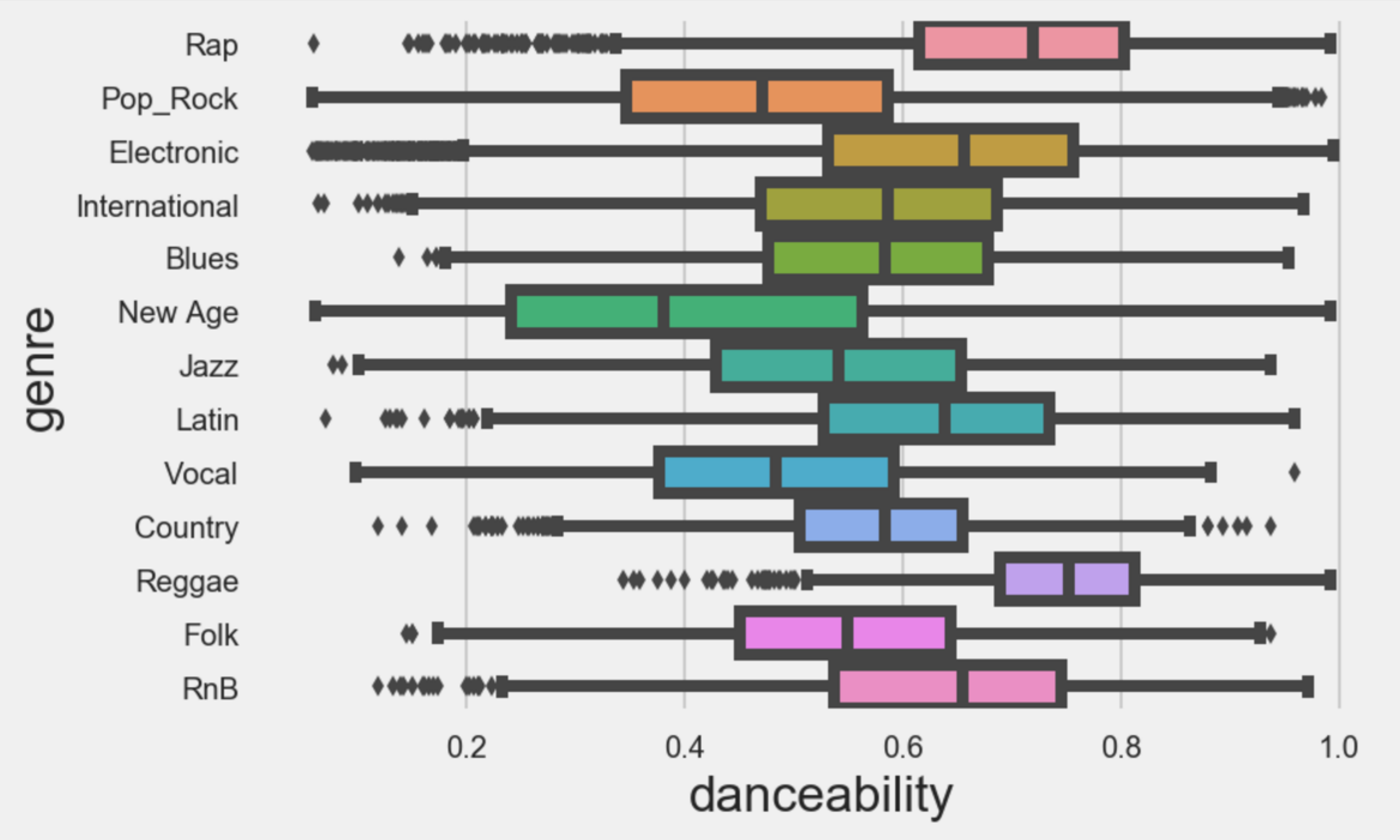


Energy



Rap
Pop_Rock
Electronic
International
Blues
New Age
Jazz
Latin
Vocal
Country
Reggae
Folk
RnB
genre
0.0
0.2
0.4
0.6
0.8
1.0
energy

Danceability



Types of Models

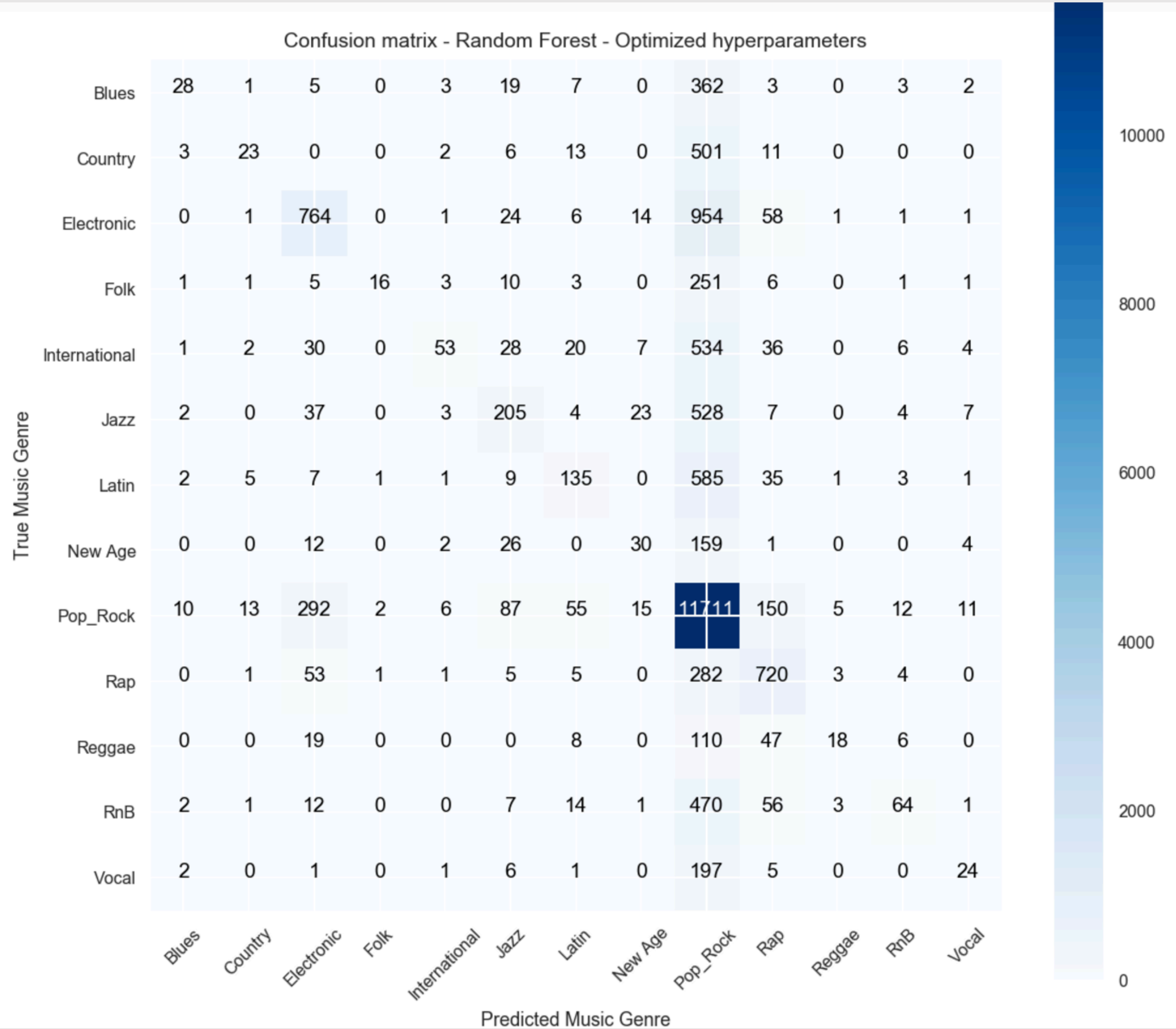
- Multinomial Logistic Regression
  - Similar to linear regression
  - Predicts to a particular class based on numerical variables

- Random Forest
  - Ensemble Method
  - Fit multiple decision trees
  - Returns the mean of the trees fitted for each class
  - Corrects for overfitting when using decision trees

Model Performance

	Logistic Regression	Random Forest
Accuracy Score pre optimization	0.647857521	0.65716737
Accuracy Score post optimization	0.647907201	0.682174381
Accuracy Score on unseen test data	0.647586036	0.682891805

CAVE




```
y_test.value_counts()

Pop_Rock         12369
Electronic        1825
Rap               1075
Jazz               820
Latin              785
International      721
RnB                631
Country            559
Blues              433
Folk               298
Vocal              237
New Age            234
Reggae             208
Name: genre, dtype: int64
```

- Pop: 94%
- Rap: 67%
- Electronic: 41%
- Jazz: 25%

But does it work in practice?

- Lets see how it looks in production:
  - Google a track
  - Check the Spotify API for the features
  - Enter into webpage
  - Analyze predictions

- http://localhost:4000/musicpage

CAVE

Misclassification – why?

CAVE

Test data – Class Imbalance

```
y_test.value_counts()
```

```
Pop_Rock         12369
Electronic        1825
Rap               1075
Jazz               820
Latin              785
International      721
RnB                631
Country            559
Blues              433
Folk               298
Vocal              237
New Age            234
Reggae             208
Name: genre, dtype: int64
```

How well do we know our...

MUSIC

...And can machine learning improve it?

Save time

- Classify up and coming Artists automatically

CAVE

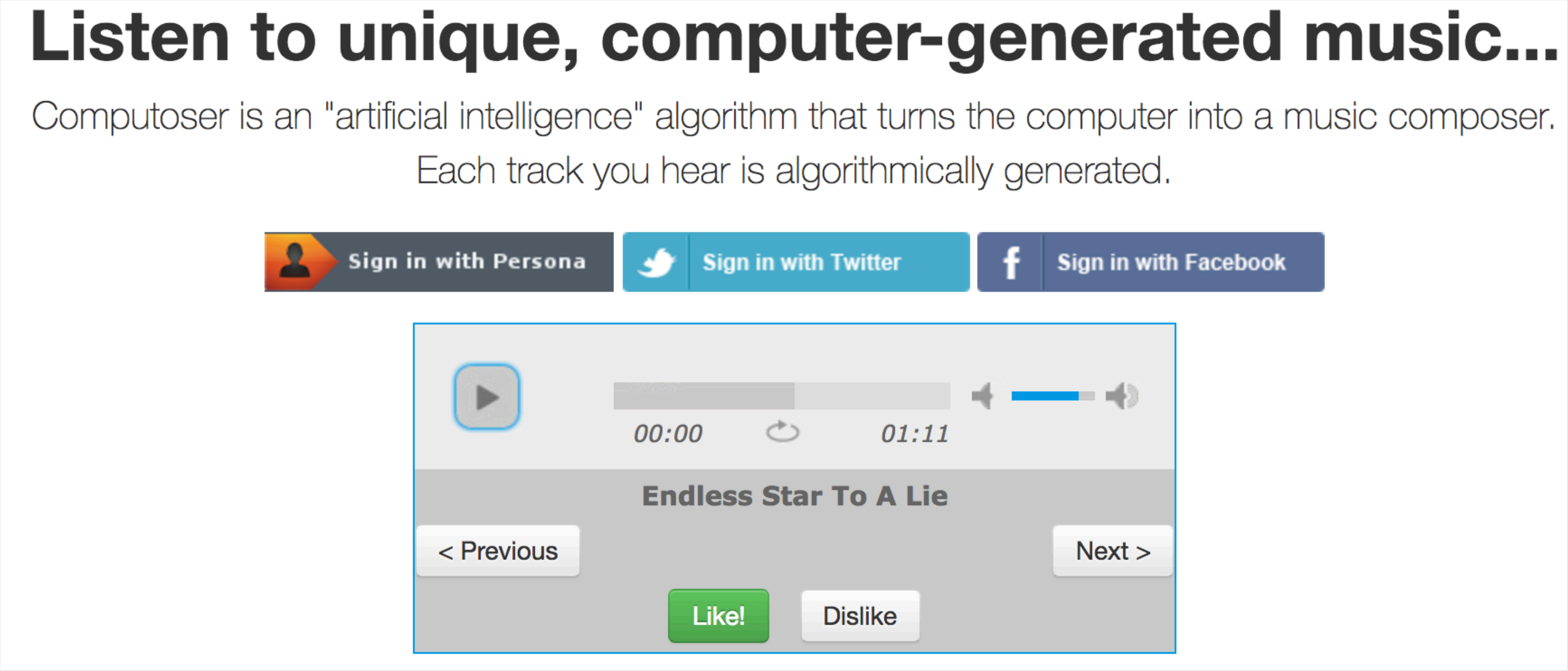
Save money

- Especially Man hours!

CAVE

Make money

- New Music generation



Recommendations

- Correct class imbalance for training the model
  - Under/oversampling
  - More Samples
- log the accuracy scores when predicting on new data
  - Monitor for degradation
  - Opportunity for retraining
- Combine with user metadata
  - Genres could be considered subjective at a song level
- Construct features from scratch using fingerprinting (Shazam)

CAVE

Check out my blog….

http://www.chi-squared.com

CAVE

Thank you!

MILLION SONG DATASET

Acoustic Brainz

Spotify

musixmatch

CAVE