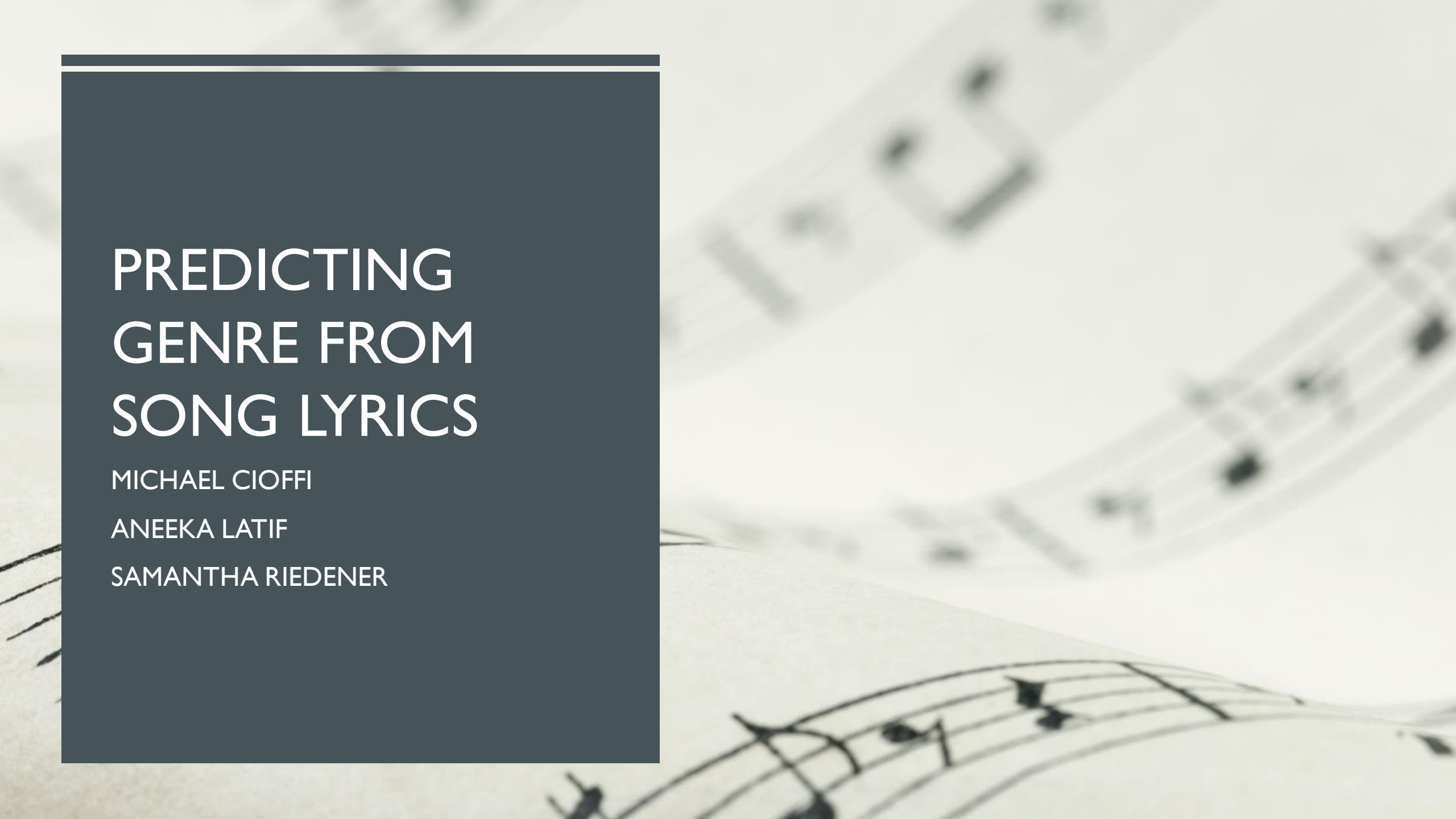
PREDICTING GENRE FROM SONG LYRICS

MICHAEL CIOFFI

ANEEKA LATIF

SAMANTHA RIEDENER



PURPOSE

- The Spotify API does all but include genre information. This is because genre can be subjective, and many songs fall into multiple category buckets.
  - Can we use song lyrics to predict genre?
  - Do certain artists use similar words in their songs?

ABOUT THE DATA

SPOTIFY API METRICS

Danceability
Based on a combination of musical elements: rhythm, beats, etc.

Energy
Represents a perceptual measure of intensity and activity.

Key
Mapped to pitches using standard Pitch Class notation.

Loudness
The overall loudness of a track in decibels (dB).

Mode
Indicates the modality (major or minor) of a track.

Speechiness
Detects the presence of spoken words.

Instrumentalness
Predicts whether a track contains no vocals.

Liveness
Detects the presence of an audience in the recording.

Valence
Describes the musical positiveness conveyed by a track.

Tempo
Estimated tempo of a track in beats per minute (BPM).

Duration and Time Signature
Duration in milliseconds and estimated overall time signature of a track.

DATA PREPROCESSING

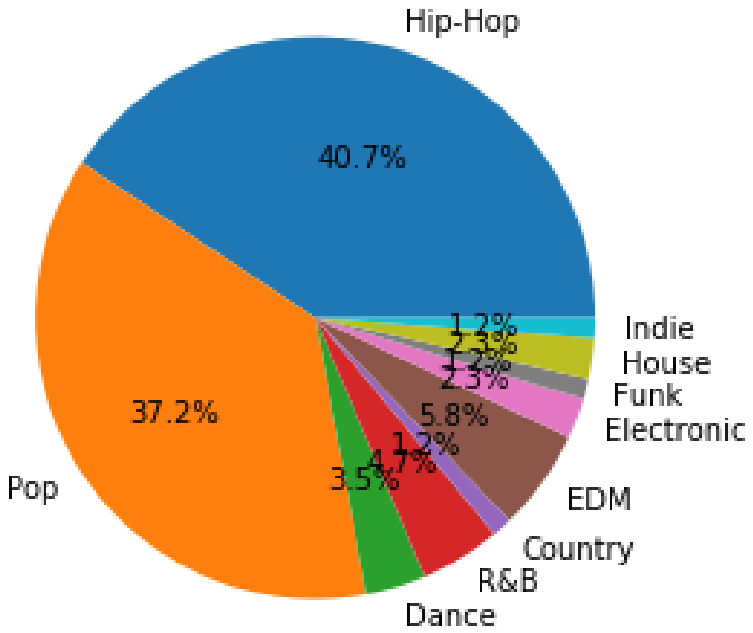
- Manually added lyric data to our CSV file
- 3x annotations for genre
- Label finalization
- Text normalization / formatting cleanup
  - Tokenization
  - Stemming
  - Stopwords removal

Artist	Song Title	Lyrics
Drake	God's Plan	Yeah they wishin' and wishin' and wishin' and wishin' They wishin' on me, yuh I bee
XXXTENTACION	SAD!	Yeah Who am I? Someone that's afraid to let go, uh You decide, if you're ever gonna
Post Malone	rockstar (feat. 21 Savage)	Ayy, I've been fuckin' hoes and poppin' pillies Man, I feel just like a rockstar (star) Ay
Post Malone	Psycho (feat. Ty Dolla $ign)	Damn, my AP goin' psycho, lil' mama bad like Michael Can't really trust nobody with
Drake	In My Feelings	Trap, TrapMoneyBenny This shit got me in my feelings Gotta be real with it, yup Kiki
Post Malone	Better Now	You probably think that you are better now, better now You only say that 'cause I'm
Cardi B	I Like It	Yeah baby, I like it like that You gotta believe me when I tell you I said I like it like tha
Calvin Harris	One Kiss (with Dua Lipa)	Let me take the night, I love real easy And I know that you'll still wanna see me On t
Dua Lipa	IDGAF	You call me all friendly Tellin' me how much you miss me That's funny, I guess you'v
Marshmello	FRIENDS	Ooh ooh, ooh ooh Ooh ooh, ooh ooh You say you love me, I say you crazy We're no
Camila Cabello	Havana	Hey Havana, ooh na-na (ay) Half of my heart is in Havana, ooh-na-na (ay, ay) He too
Juice WRLD	Lucid Dreams	No, no, no I still see your shadows in my room Can't take back the love that I gave y
Drake	Nice For What	I wanna know who mothafuckin' representin' in here tonight Hold on, hold on I kee
Maroon 5	Girls Like You (feat. Cardi B)	Spent 24 hours I need more hours with you You spent the weekend Getting even, o

ANNOTATING GENRE LABEL

Artist	Song Title	Genre (Aneeka)	Genre (Samantha)	Genre (Michael)
Drake	God's Plan	Pop	Hip-Hop	Hip-Hop
XXXTENTACION	SAD!	Hip-Hop	Rap	Hip-Hop
Post Malone	rockstar (feat. 21 Savage)	Hip-Hop	Hip-Hop	Hip-Hop
Post Malone	Psycho (feat. Ty Dolla $ign)	Hip-Hop	R&B	Hip-Hop
Drake	In My Feelings	Hip-Hop	Hip-Hop	Hip-Hop
Post Malone	Better Now	Hip-Hop	Hip-Hop	Hip-Hop
Cardi B	I Like It	Hip-Hop	Hip-Hop	Hip-Hop
Calvin Harris	One Kiss (with Dua Lipa)	Pop	Dance	Pop
Dua Lipa	IDGAF	Pop	Pop	Pop
Marshmello	FRIENDS	Pop	Pop	Pop

```
['Hip-Hop', 'Hip-Hop', 'Hip-Hop', 'Hip-Hop', 'Hip-Hop', 'Hip-Hop', 'Hip-Hop', 'Pop',
'Pop', 'Pop', 'Pop', 'Hip-Hop', 'Hip-Hop', 'Pop', 'Dance', 'Hip-Hop', 'Pop', 'Hip-
Hop', 'Hip-Hop', 'Dance', 'Pop', 'Pop', 'Pop', 'Pop', 'R&B', 'Country', 'Hip-Hop',
'Pop', 'Hip-Hop', 'EDM', 'R&B', 'Pop', 'Pop', 'Pop', 'EDM', 'Hip-Hop', 'Electronic',
'Hip-Hop', 'Hip-Hop', 'Pop', 'EDM', 'Funk', 'Pop', 'Hip-Hop', 'R&B', 'Hip-Hop',
'Electronic', 'Pop', 'Hip-Hop', 'Pop', 'Hip-Hop', 'EDM', 'House', 'Hip-Hop', 'Pop',
'Hip-Hop', 'Hip-Hop', 'Pop', 'Dance', 'Pop', 'Pop', 'Pop', 'Pop', 'Pop', 'Hip-Hop',
'Hip-Hop', 'Hip-Hop', 'Hip-Hop', 'Hip-Hop', 'Pop', 'Hip-Hop', 'Indie', 'Hip-Hop',
'Hip-Hop', 'Pop', 'R&B', 'House', 'Hip-Hop', 'Pop', 'Pop', 'Hip-Hop', 'Pop', 'EDM',
'Hip-Hop', 'Pop', 'Pop']
```



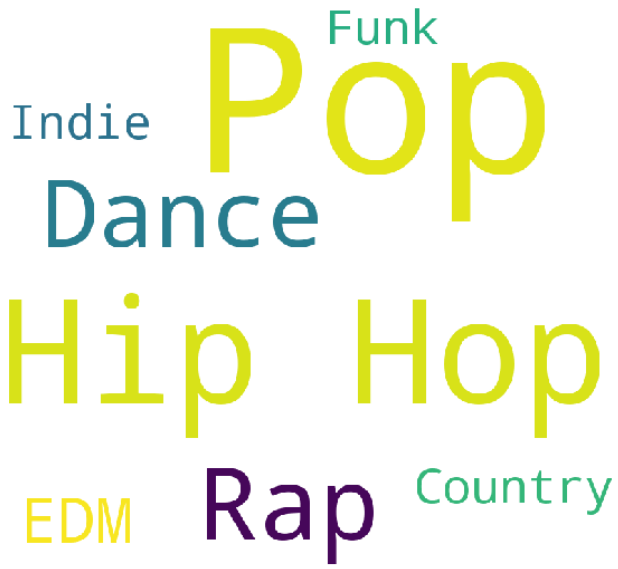
	Aneeka	Samantha	Michael
Aneeka		0.485	0.560
Samantha	0.485		0.435
Michael	0.560	0.435	

Overall Kappa score : 0.493

ANNOTATION DISTRIBUTION

Aneeka

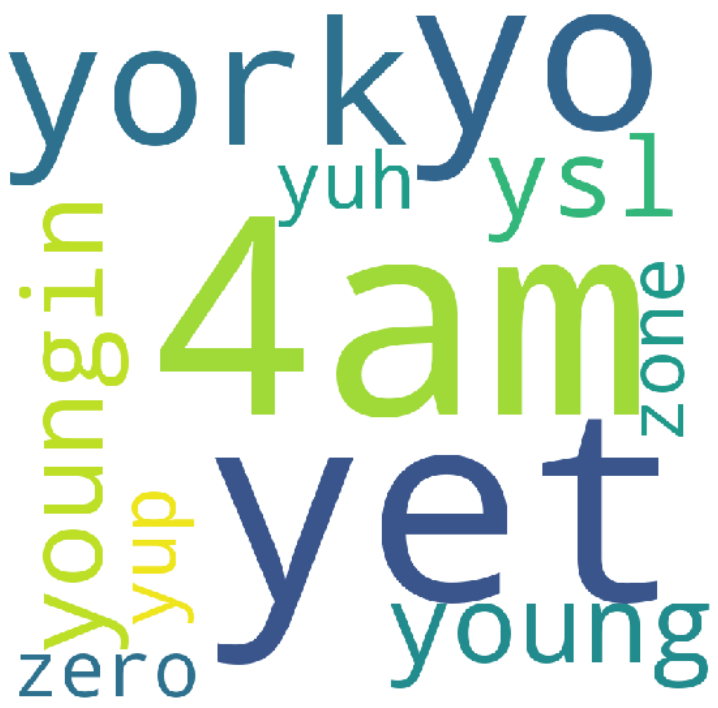
Samantha



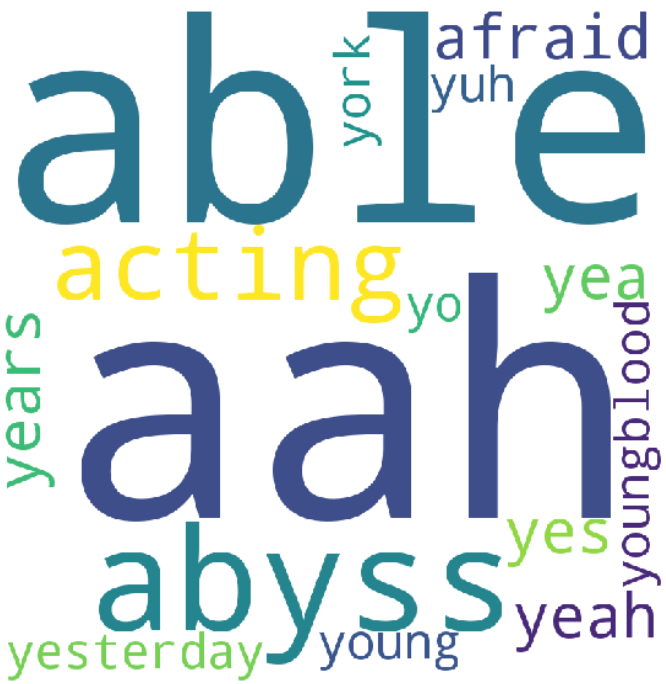
Michael

MOST COMMON WORDS BY GENRE

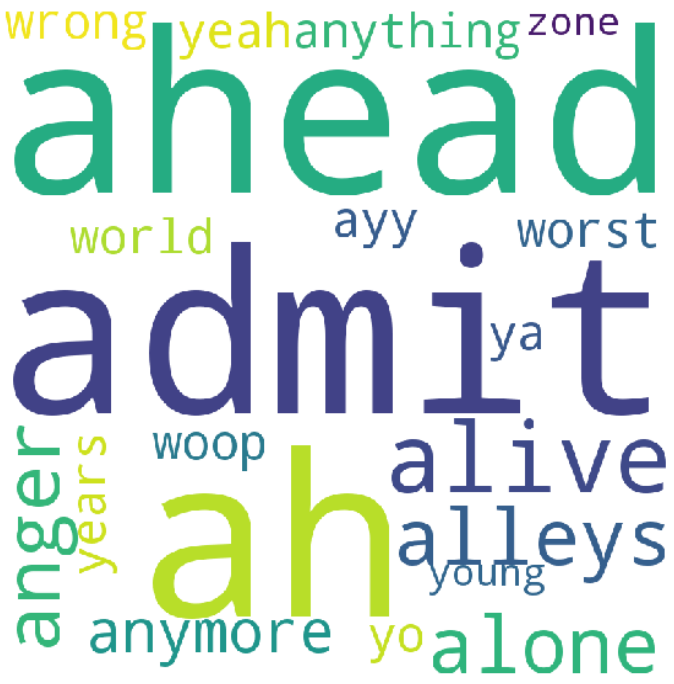
Hip hop



Pop



EDM



NAÏVE BAYES ANALYSIS

```
[[ 0  0  0  2  0]
 [ 0  0  1  2  0]
 [ 0  0 17  5  0]
 [ 0  0  1 11  0]
 [ 0  0  0  1  0]]
```

	precision	recall
Dance	0.00	0.00
EDM	0.00	0.00
Hip-Hop	0.89	0.77
Pop	0.52	0.92
R&B	0.00	0.00

Success Rate: 70%

- Half of the genres used due to too few in the data set
- Algorithm only sorted into Hip-Hop and Pop
- Successfully distinguished between the two
- Skewed data set led to these results – better representation would work more successfully

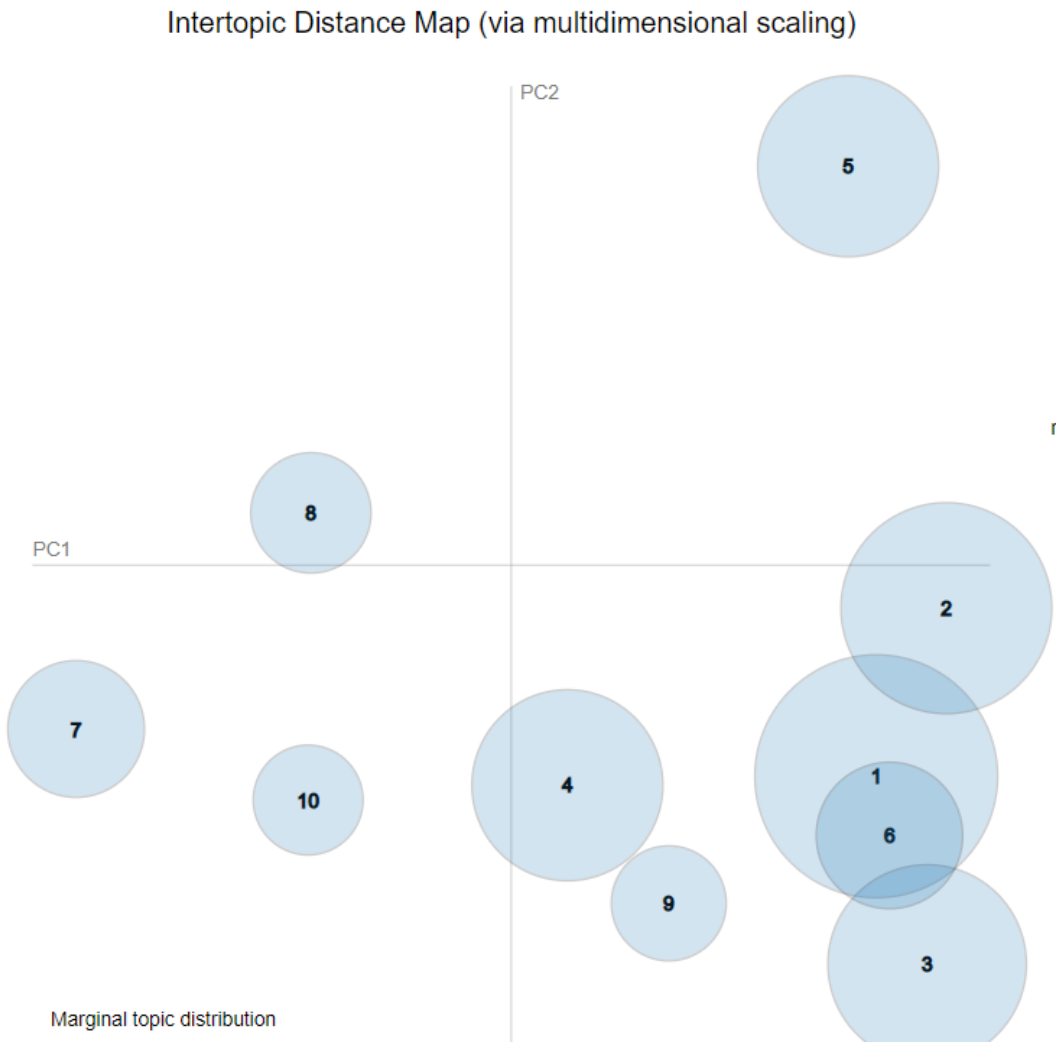
SUPPORT VECTOR MACHINE

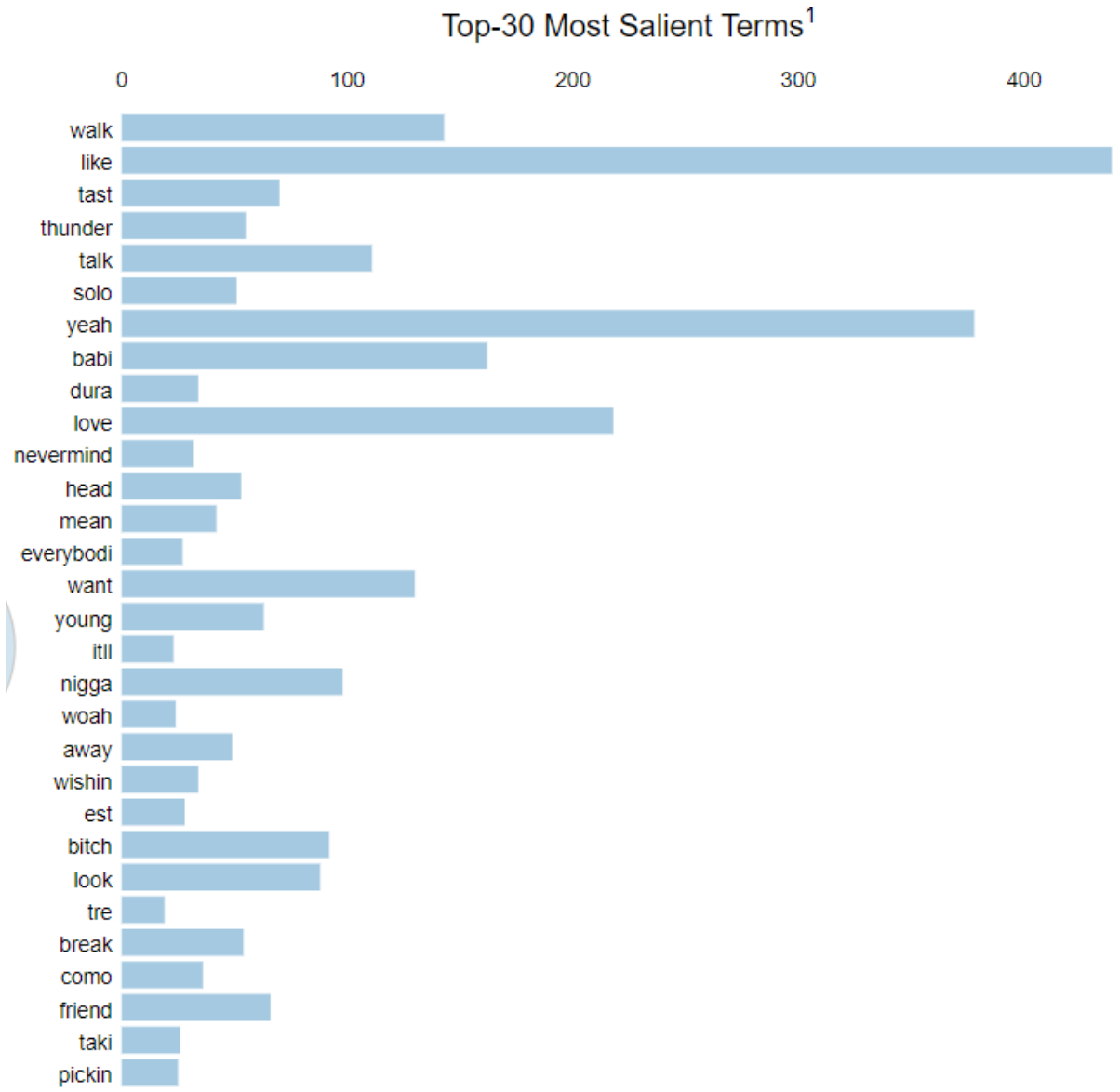
```
[0.5        0.66666667 0.75       ]
```

- Used cross fold validation to better accommodate data set
- Successive improvement with each fold
- Same subset of data used here as the Naïve Bayes

LDA

- 10 topics were chosen to match the 10 genre labels




Top-30 Most Salient Terms[1]
0
100
200
300
400
walk
like
tast
thunder
talk
solo
yeah
babi
dura
love
nevermind
head
mean
everybodi
want
young
itll
nigga
woah
away
wishin
est
bitch
look
tre
break
como
friend
taki
pickin

LDA CONT.