The Birth of a Chartbuster

Anisha Alluru
Khyathi Balusu
Manas Rai
Ravi Kiran Bobba
Vishal Ramachandran

AGENDA

- Project Objective
- Data Collection
- Exploratory Data Analysis
- Topic Modeling and Sentiment Analysis
- Modeling - Regression and Classification
- Song Recommendation Engine
- Key Takeaways

OBJECTIVE

Leverage our web scraping, modeling and text analysis skills to the best possible use and come up with a business product that would:

- Predict if a song would become a hit based on its lyrics and musical features
- Suggest the right ingredients for a banger hit
- Perform content based song recommendation

Getting blocked by our crush

Getting blocked by Web APIs

SERIAL SCRAPING

SONG FEATURES

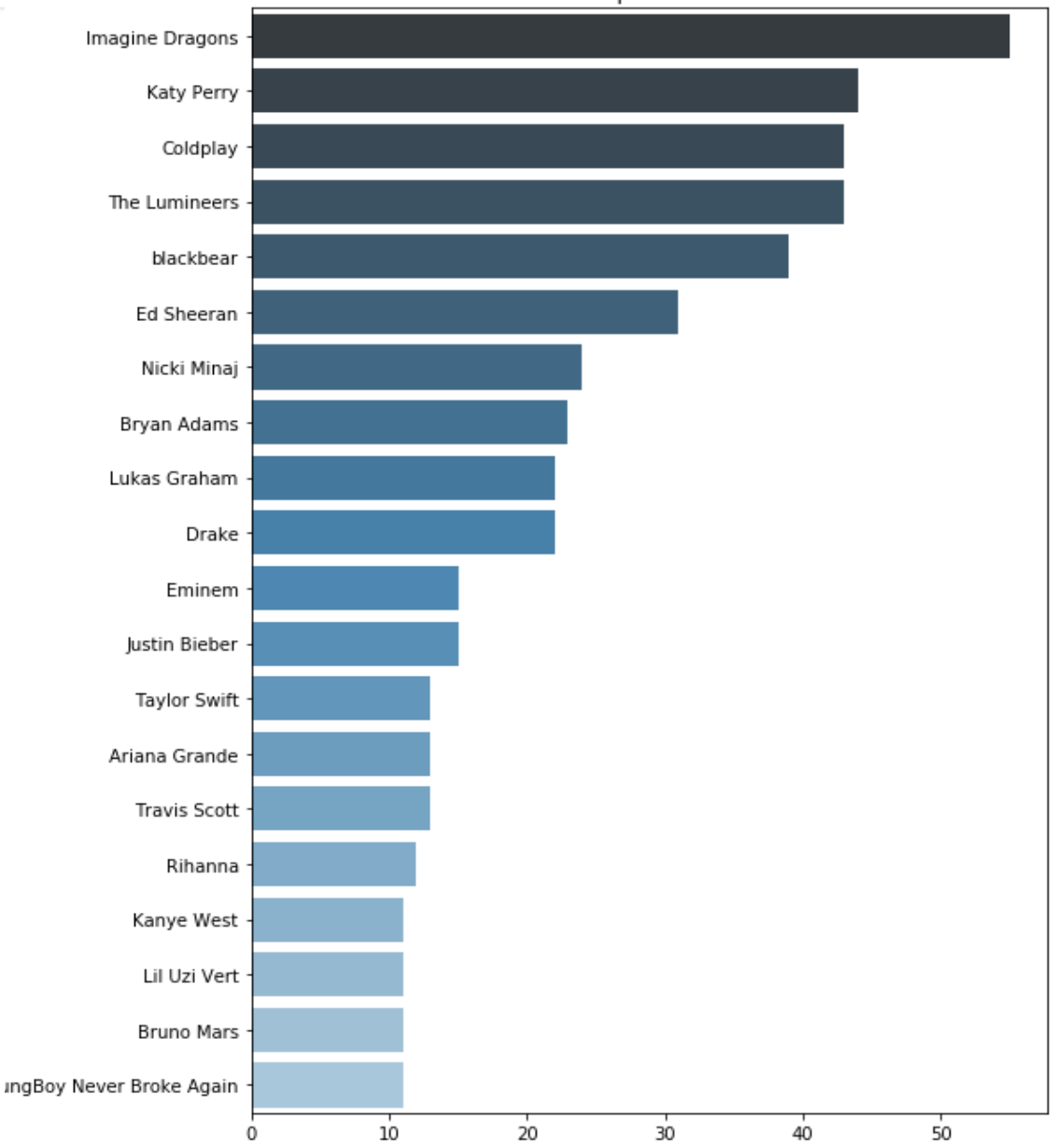
LISTENER COUNT

LYRICS

DATA OVERVIEW

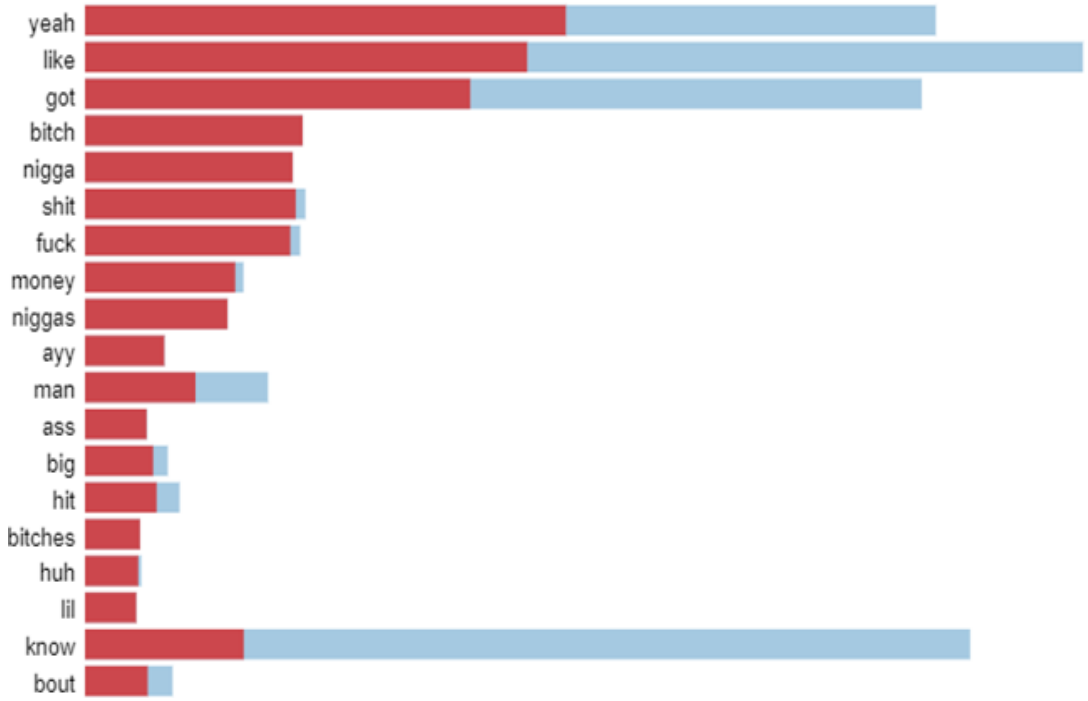
Most Frequent Words

Top Artists



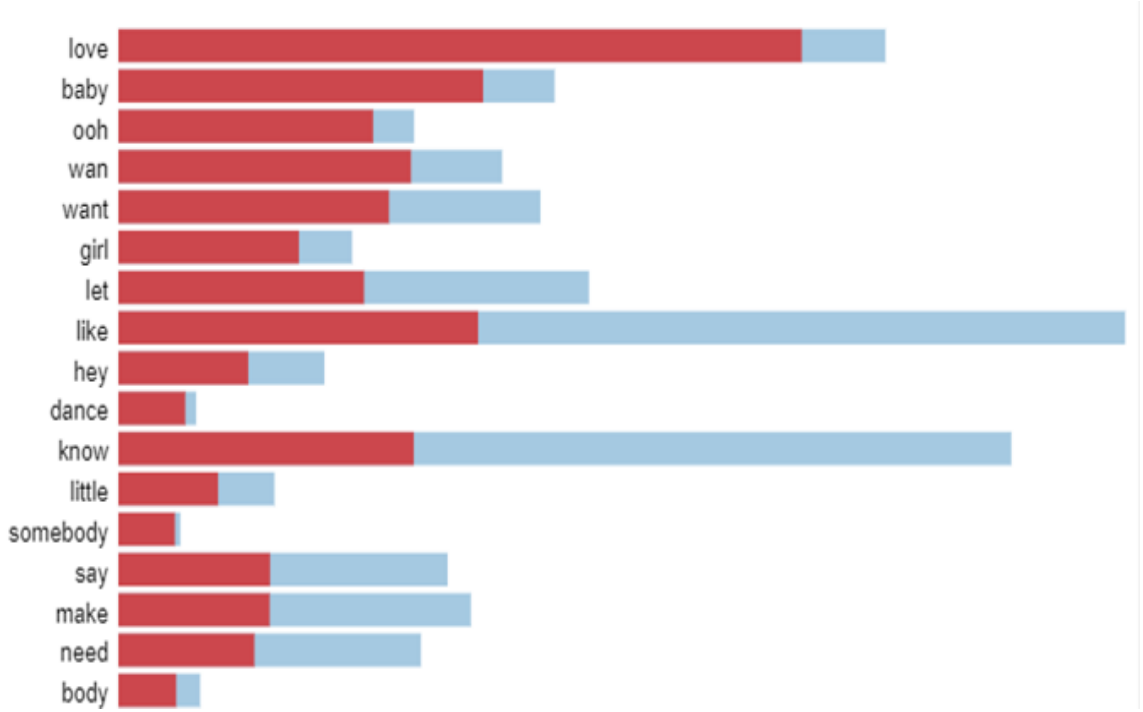
TOPIC MODELING

Topic 1: Explicit



- *B[redacted]h Better Have My Money*
- *Run the World (Girls)*
- *Truth Hurts*

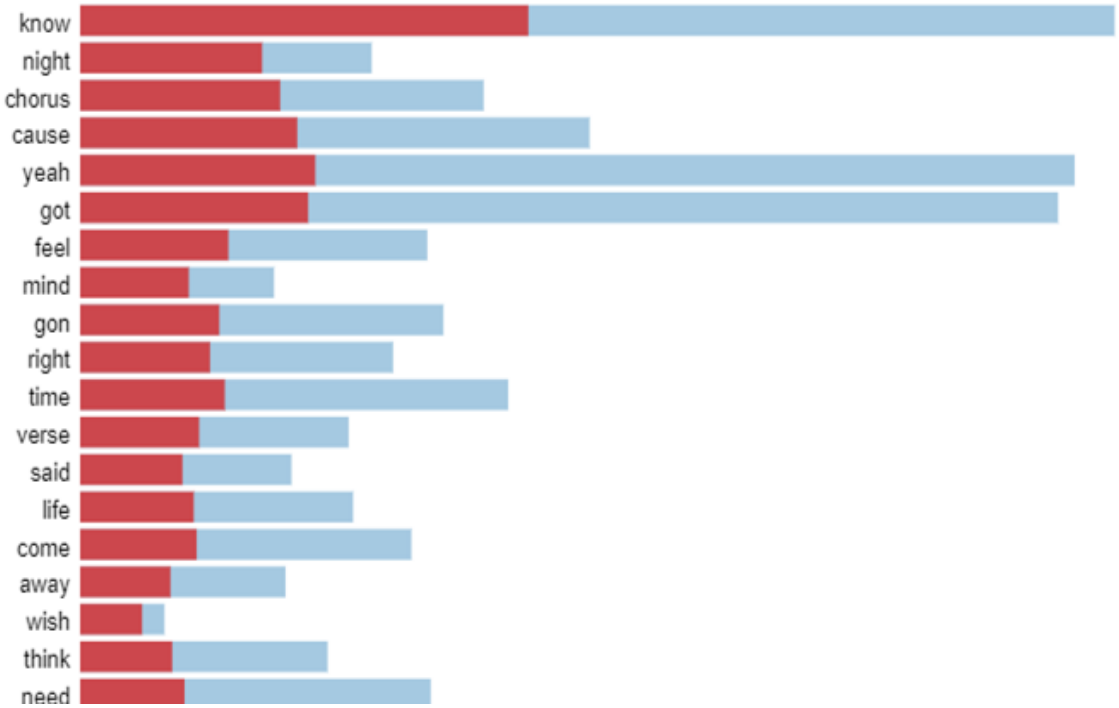
Topic 2: Love/Emotional



- *Faded - Alan Walker*
- *Hello - Adele*
- *Angela - The Lumineers*

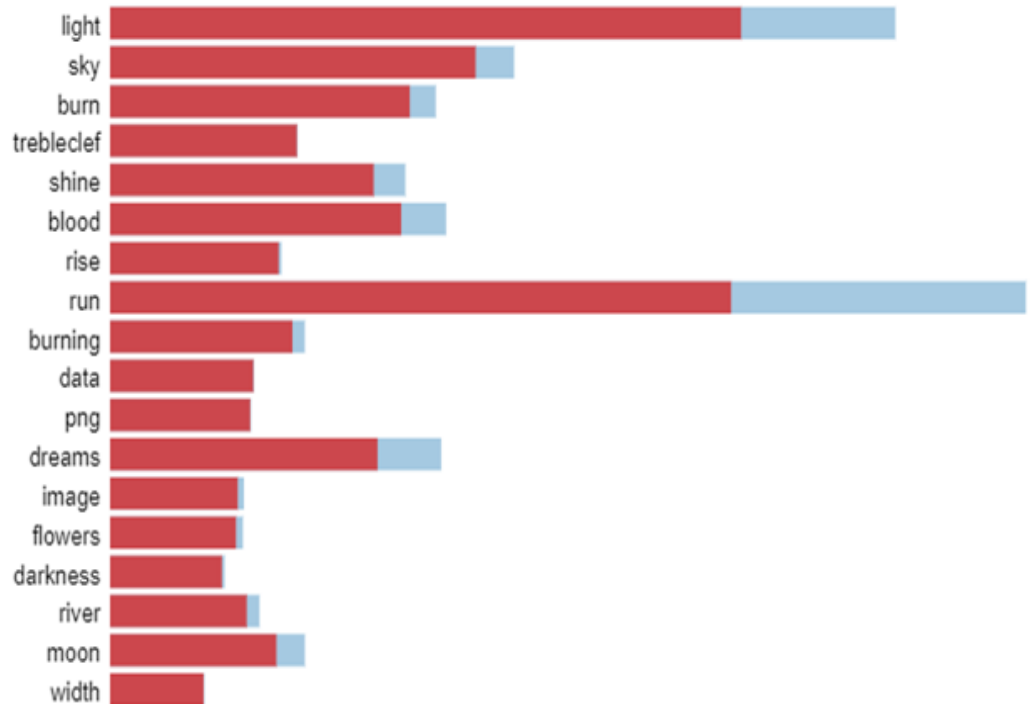
TOPIC MODELING

Topic 3: Reflective/Upbeat



- *Call Me Maybe*
- *Shape of You*
- *24K Magic*

Topic 4: Breakfree/Celebrating Yourself

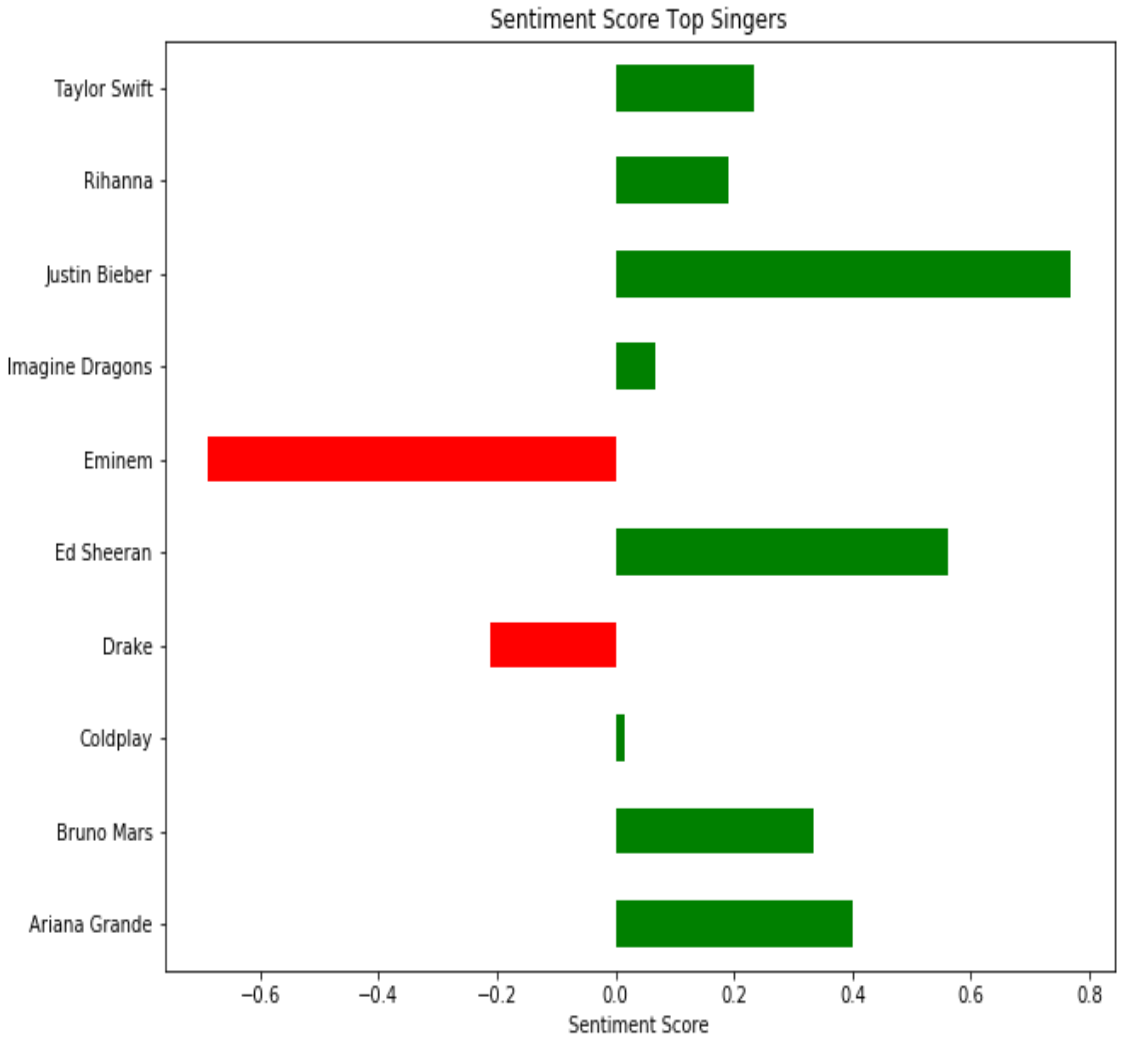


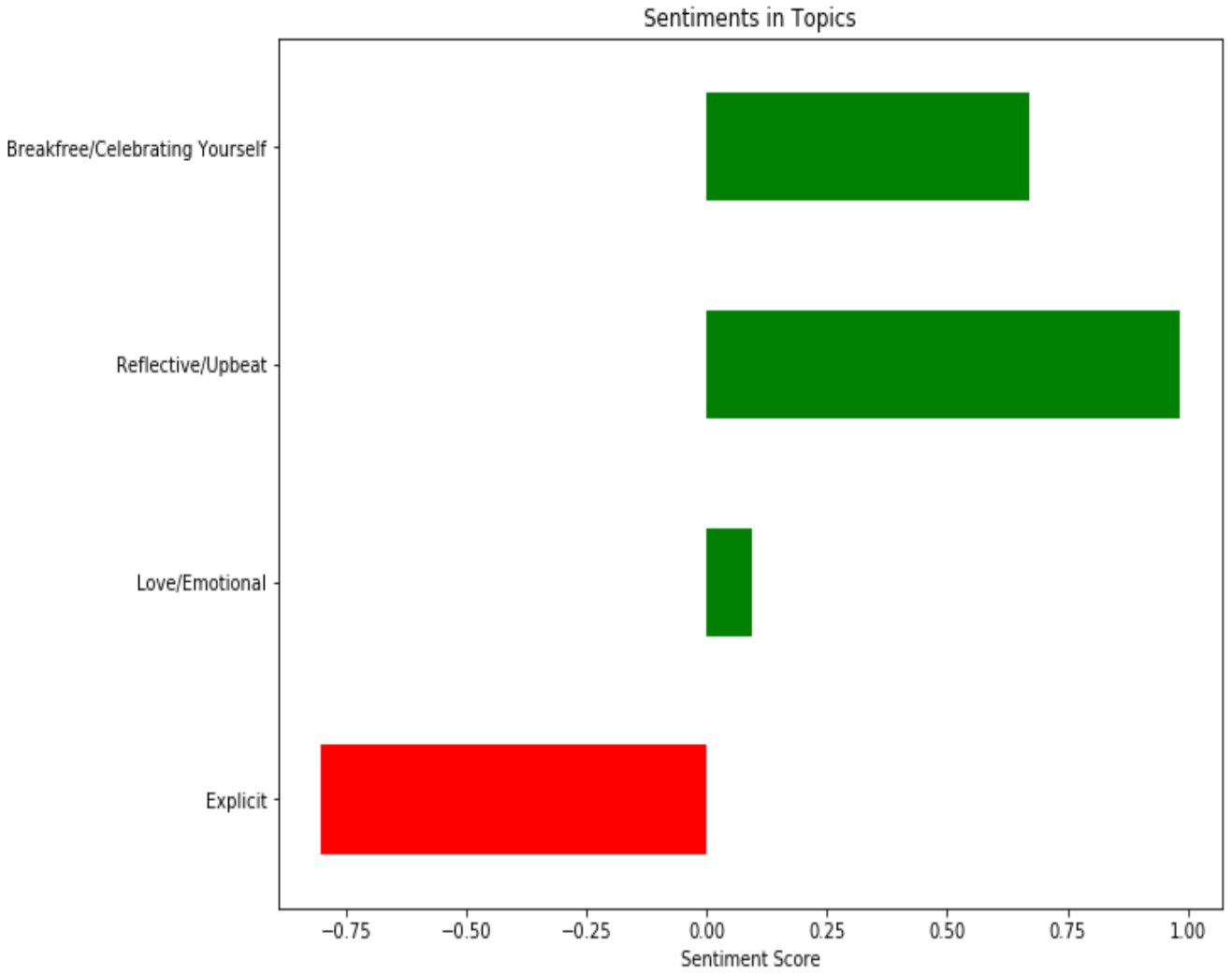
- *Shake It Off*
- *Bang Bang*
- *FRIENDS*

TOPICS AND TOP ARTISTS

Artist	Explicit	Love/Emotional	Reflective/Upbeat	Breakfree/Celebrating Yourself
Katy Perry	5%	27%	27%	41%
The Lumineers	2%	56%	19%	23%
Coldplay	7%	60%	19%	14%
Ed Sheeran	3%	29%	42%	26%
Nicki Minaj	67%	0%	17%	17%
Lukas Graham	0%	18%	14%	68%
Drake	73%	0%	9%	18%
Eminem	67%	7%	7%	20%
Taylor Swift	8%	31%	23%	38%
Ariana Grande	31%	0%	46%	23%
Rihanna	17%	8%	50%	25%
Imagine Dragons	2%	58%	15%	25%

SENTIMENT ANALYSIS





Predictive Modeling

MODELING - CLASSIFICATION

Hypothesis:

- Can Text Analysis Techniques like Topic Modelling, Predicting Sentiment, TF-IDF help in improving Prediction of Hit ?

Approach:

- Created a base model with different features of the song and then incorporated different metrics based on text analysis of Lyrics of the song into the model

Results:

Actual	Predicted	
	Miss	Hit
Miss	476	14
Hit	208	17

Actual	Predicted	
	Miss	Hit
Miss	472	18
Hit	192	33

94% Percent higher prediction in number of Hits compared to baseline Model with features of the song, with a slight improvement in accuracy (69% to 71%)

MODELING - REGRESSION

Hypothesis:

- Can Text Analysis Techniques, help in prediction of number of Views ? How much additional variation in number of views can be explained using Text ?

Approach:

- Created a base model with different features of the song and then incorporated different metrics based on text analysis of Lyrics of the song into the model

Results:

R^2 Value	Model using Song Features	Model using Topic Modeling and Sentiment
	0.66	0.74

10% Higher explanation in Variation in number of views using Text Analytics

Recommendations

RECOMMENDATION METHODOLOGY

Hypothesis:

- Can Text Analysis help in improving the Recommendations of Songs ?

Approach:

MODEL 1

- Based on Song Features like Danceability, Loudness, etc

MODEL 2

- Based on TF-IDF on Lyrics

MODEL 3

- Based on Song Features, TF-IDF, Sentiment and Topic Modelling

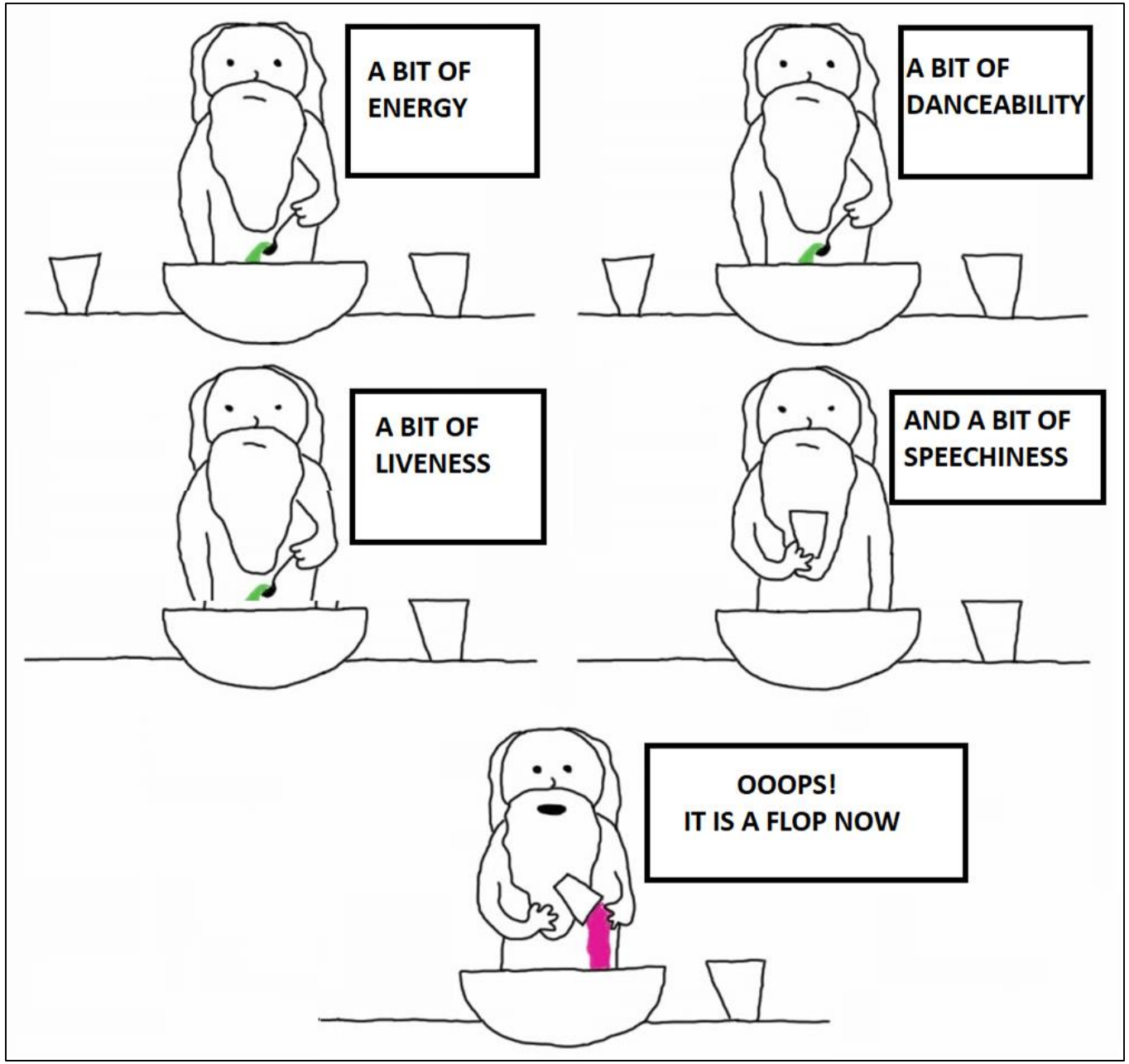
SONG RECOMMENDATION - 1

Model 1 (Features of Song)		Model 2 (TF-IDF)		Model 3 (Features, TF-IDF, Topic Modelling, Sentiment)	
Song	Artist	Song	Artist	Song	Artist
Bibia Be Ye Ye	Ed Sheeran	Me and Your Mama	Childish Gambino	Peacock	Katy Perry
Written in the Sand	Old Dominion	CoCo	O.T. Genasis	Mad Love	Mabel
Drown	Clairo	I Think I'm In Love	Kat Dahlia	Treasure	Bruno Mars
Englishman In New York	Sting	Can't Feel My Face	The Weeknd	What Do You Mean?	Justin Bieber
What Do You Mean?	Justin Bieber	Nothin to Somethin	Syd	Happy	Pharrell William

SONG RECOMMENDATION - 2

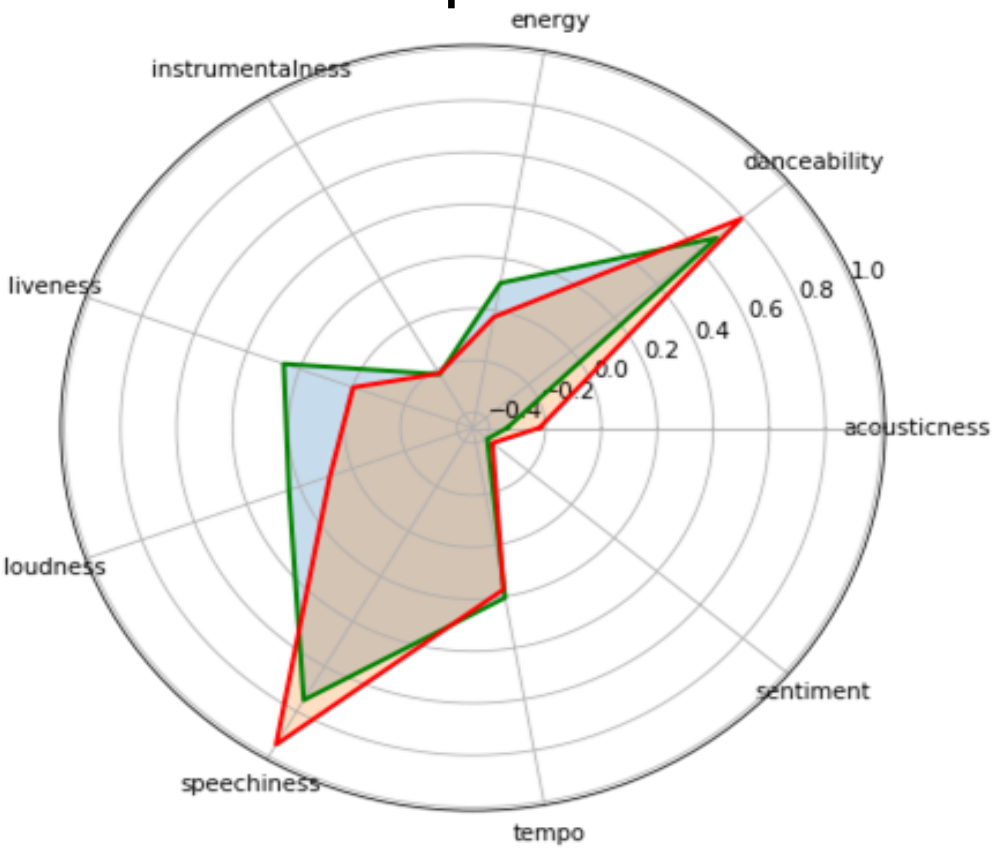
Model 1 (Features of Song)		Model 2 (TF-IDF)		Model 3 (Features, TF-IDF, Topic Modelling, Sentiment)	
Song	Artist	Song	Artist	Song	Artist
Heartbeat Song	Kelly Clarkson	Walk Thru (feat. Problem)	Rich Homie Quan	Uma Thurman	Fall Out Boy
In The Night	The Weeknd	Eyes Wide Open	Gotye	Exits	Foals
For the First Time	The Script	Your Own	SALES	Maps	Maroon 5
Blinding Lights	The Weeknd	My Heart Stood Still (Remastered)	Peggy Lee	Not Afraid	Eminem
Found You	Kane Brown	Boom Boom	John Lee Hooker	My Songs Know What You Did In The Dark	Fall Out Boy

WHEN THE ARTIST WAS MAKING A BREAK-FREE SONG

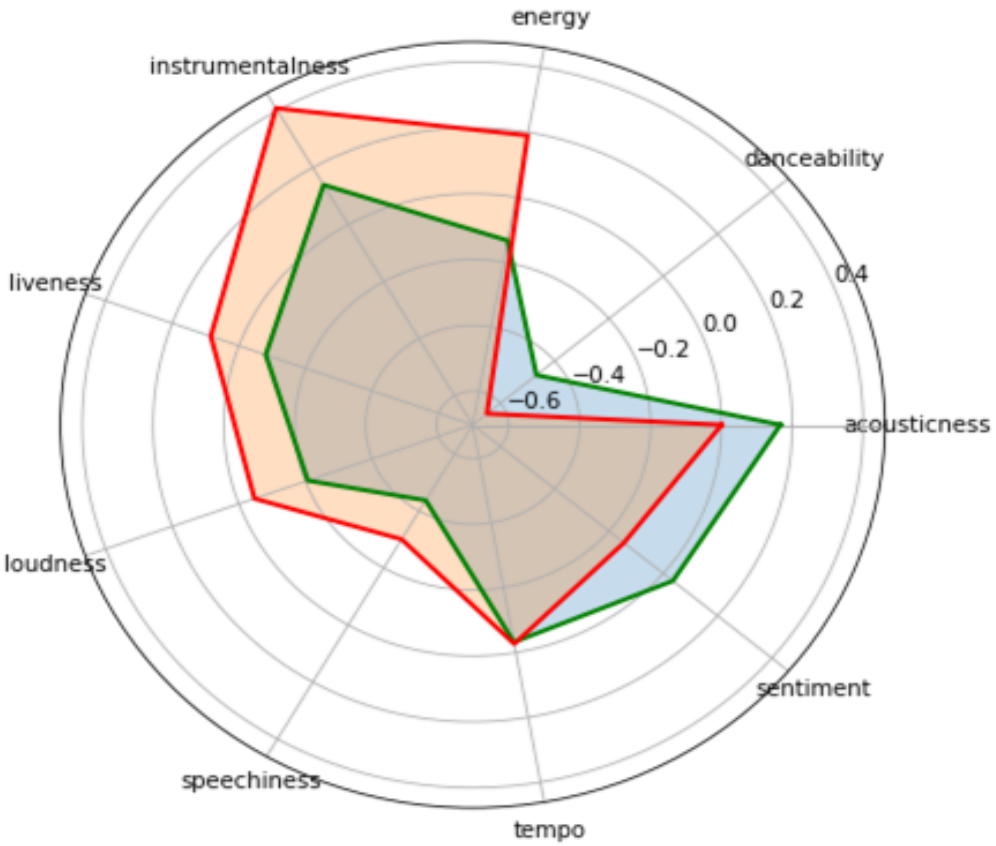


THE RIGHT INGREDIENTS!

Explicit



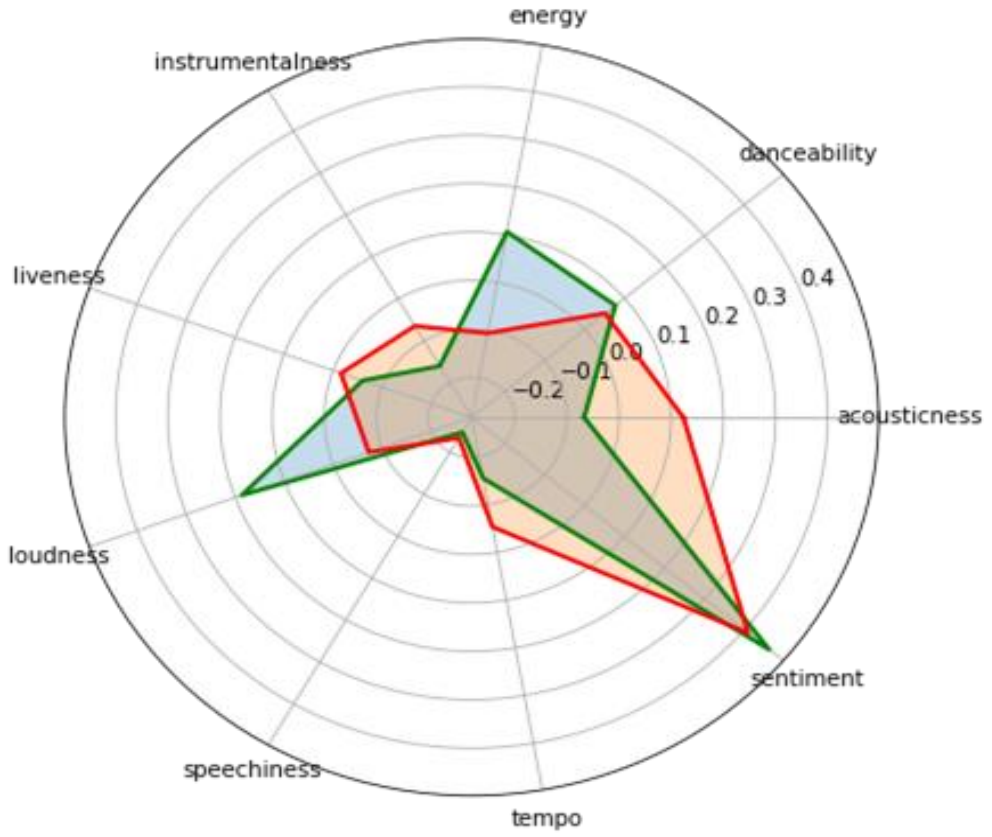
Love/Emotional



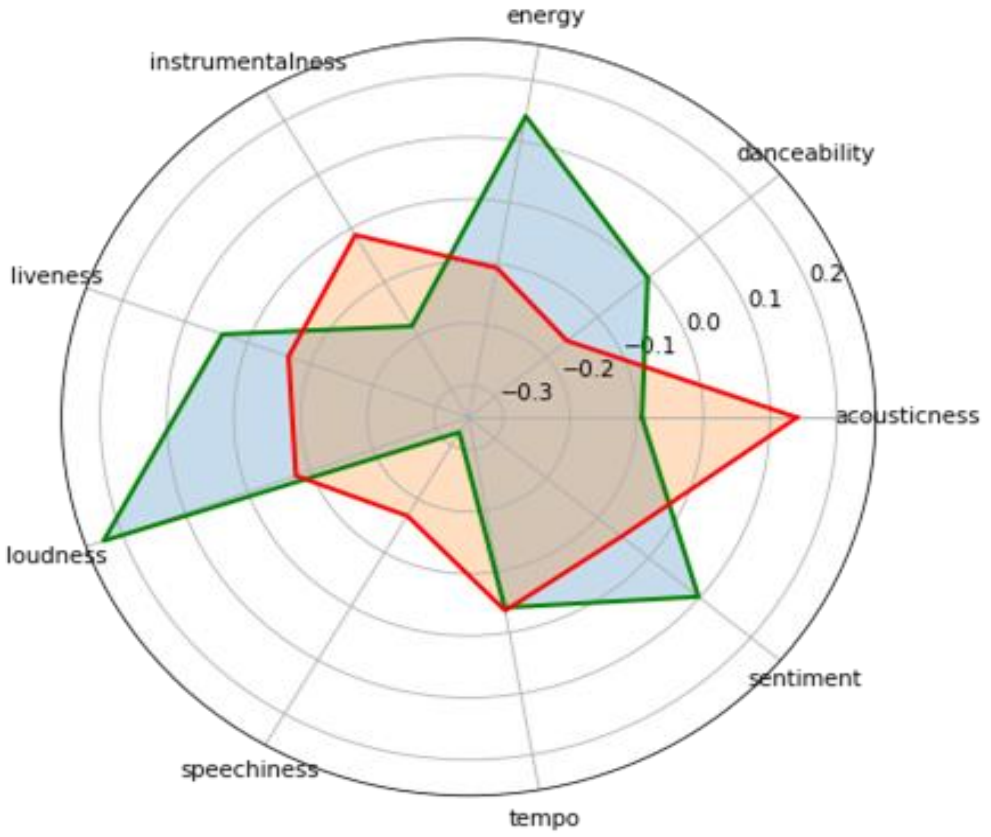
Liveliness & Energy	**Acousticness & Sentiment**
Speechiness & Danceability	**Instrumentalness & Energy**

THE RIGHT INGREDIENTS!

Reflective/Upbeat



Breakfree/Celebrating Yourself



Loudness & Energy	**Loudness & Energy**
Acousticness	**Instrumentalness & Acousticness**

RECIPE!

If you are planning to go pro making music, ditching the MSBA course, maybe you could try our recommendation recipe for a song?
Let us consider the breakfree Topic,
For this Topic,

- A healthy portion of loudness
- Generous serving of energy
- A pinch of sentiment to taste
- Avoid acousticness and instrumentalness, for the dish to come out in the best possible way!

Tada, your instant chartbuster is served piping hot!

Business Application

Content Based Recommendation of audio songs may not be as effective in Recommendation compared to Collaborative Recommendation; However it can be applied in following use cases.

- Launch of Upcoming Artists to identify the songs/Topics where marketing efforts for the launch should be directed
- Ad placement based on recommendations can be effectively handled

Also,

- These ingredients can be potential blockbuster material for budding artists vying for billboards/chartbusters

QUESTIONS?

APPENDIX?

Explicit songs are more verbose and high on energy