




● Audio music mood classification

Metis NYC DS9
Project McNulty
Ting Neo
October 28, 2016

- 
- Moods and emotions drive consumer decisions
 - 1. Association with music
 - a. Advertising and brand identity
 - b. Music for spaces and events
 2. Music industry
 - a. Music recommendations
 - b. Label/artist management
 - c. Assembling music metadata
 - d. Creating album, artist, or playlist mood profiles

Business objective: Create a mood profile for a user's Spotify playlists to identify their tastes

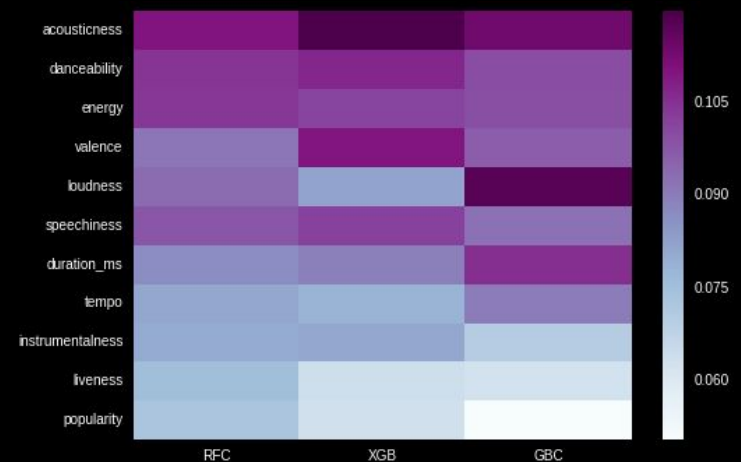
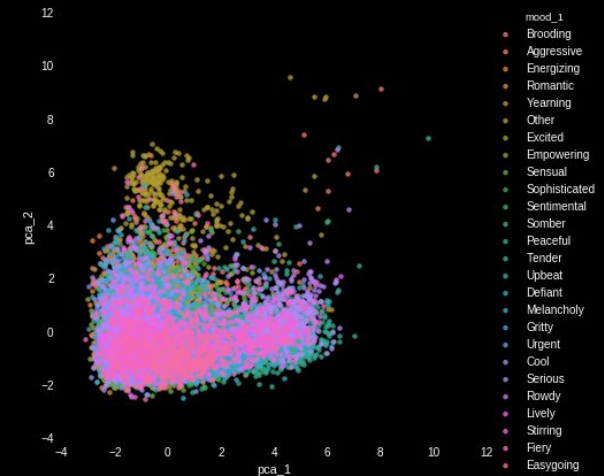
Identifying song features and labels

Compiled database of 21,059 songs with mood labels

Spotify API		Gracenote API	
Features		Labels	
Popularity	Aggressive	Romantic	Gracenote API
Energy	Brooding	Rowdy	
Liveness	Cool	Sensual	
Tempo	Defiant	Sentimental	
Speechiness	Easygoing	Serious	
Acousticness	Empowering	Somber	
Instrumentalness	Energizing	Sophisticated	
Time Signature	Excited	Stirring	
Danceability	Fiery	Tender	
Loudness	Gritty	Upbeat	
Valence	Lively	Urgent	
Explicit	Melancholy	Yearning	
Key (x11)	Peaceful		
Time Signature (x3)			
Mode			
Duration			
Genre (x9)			
Artist Era (x8)			
Artist Type (x7)			
Artist Origin (x18)			

Trial and error to determine the best process

- **Feature engineering:** Scaled Spotify audio features emerged as the best predictors
 - Explored interactions and PCA
- **Models:** Random Forests, Gradient Boosted Trees and XGBoost performed best
 - Random Forests consistently outperformed other models (w/ n estimators = 300)



Mood taxonomy is defined on an arousal-positivity scale

Deconstruct the problem into one of predicting level of arousal, level of positivity, then using coordinates to predict mood class

- Arousal and positivity defined on a scale of -2 to 2
- From one 25-class problem to two 5-class problems?



Evaluating two approaches to multinomial classification

Predict all moods at once, assuming no relationship

- One 25-class problem
- Model: Random Forest

100	12	10	84	0	5	32	35	0	0	0	1	0	0	1	5	0	2	0	0	0	0	0	7	2
14	129	4	39	0	47	7	32	0	3	0	10	0	2	2	26	4	1	3	4	0	2	2	4	24
4	6	84	8	3	9	37	18	2	0	2	0	0	0	1	63	1	1	1	1	0	1	7	4	9
70	33	43	247	0	16	13	31	0	1	0	1	1	0	5	14	0	0	0	0	0	1	0	6	8
0	0	6	1	25	11	1	1	1	5	8	7	1	0	0	12	11	1	0	4	1	7	1	0	6
1	43	4	29	3	247	2	18	0	3	2	18	1	3	2	23	17	4	0	5	0	7	2	43	
16	5	73	17	0	3	236	47	0	0	0	0	0	0	4	28	1	0	1	1	0	0	1	3	13
24	28	14	33	1	17	54	133	1	0	0	0	1	0	11	30	0	0	0	0	0	9	5	23	
1	6	12	2	1	3	23	14	14	3	3	2	1	0	5	12	10	1	0	0	0	0	2	0	7
0	5	15	1	6	34	8	7	1	22	7	9	0	0	1	28	13	0	0	6	0	3	2	2	17
0	4	12	0	5	4	3	3	0	5	16	1	0	0	2	11	7	3	0	1	0	8	10	3	6
0	11	1	0	4	44	1	6	0	1	0	80	4	6	0	10	38	7	1	4	1	13	0	0	3
0	1	0	0	2	0	1	2	0	1	0	5	37	4	0	4	10	23	12	1	0	41	0	0	1
2	7	2	3	3	29	3	4	0	3	3	13	7	25	0	12	9	14	3	6	0	14	3	0	10
8	11	9	34	0	13	14	27	1	1	0	0	0	0	17	12	1	1	1	0	0	3	4	7	9
1	26	75	5	2	26	30	28	1	1	4	14	0	0	0	244	9	3	0	4	0	1	1	5	20
0	5	2	0	1	32	0	1	0	3	1	30	4	6	0	11	30	6	1	8	0	19	1	0	4
1	3	1	0	1	5	1	0	0	1	0	6	12	7	0	4	7	30	3	0	1	34	0	0	0
1	4	2	0	0	3	0	0	0	0	0	4	24	0	0	6	3	21	17	0	0	15	0	1	1
0	3	4	1	6	42	0	1	0	4	6	12	0	4	0	21	12	1	0	23	0	7	1	1	12
2	1	1	1	1	3	2	1	0	0	1	1	2	0	0	1	4	17	0	0	2	29	1	1	2
1	1	0	0	1	4	0	0	2	2	1	4	9	3	0	1	17	31	2	2	3	125	1	0	2
1	11	11	0	0	29	11	22	1	2	5	1	0	0	8	15	0	0	0	1	0	26	7	13	
15	17	12	44	1	8	19	45	0	0	1	1	0	0	5	15	1	1	0	0	0	5	33	10	
2	26	15	15	3	52	7	19	1	6	3	6	0	0	3	30	3	0	0	0	4	6	10	105	

Precision

Recall

F-1

0.37

0.37

0.36

- Emergence of TP/TN diagonal in confusion matrix, as well as FP/FN in similar moods

Predict arousal and positivity separately, use intersection for mood

- Two 5-class problems
- Model: Random Forest

217	1	3	115	0	0	33	12	0	0	0	1	0	0	1	4	1	2	0	0	0	1	1	3		
37	48	3	61	0	25	8	11	2	1	2	2	1	1	2	28	8	36	6	4	0	2	2	12	57	
5	0	117	82	4	3	47	5	2	0	0	0	0	1	0	43	1	2	2	21	0	1	2	1	40	
88	4	16	303	0	6	13	13	0	0	0	1	0	0	1	8	0	5	1	2	0	1	0	10	18	
0	0	7	2	26	6	3	1	1	1	5	3	2	3	0	9	15	1	1	6	0	9	3	0	6	
7	12	2	43	6	176	2	3	0	1	0	11	0	8	1	46	22	15	3	11	1	12	8	10	76	
15	1	28	81	1	2	359	22	1	0	0	0	0	1	3	19	0	1	0	7	0	1	1	0	26	
32	4	3	103	2	7	66	82	0	0	1	0	0	1	2	19	1	6	0	4	0	1	7	3	30	
2	4	4	15	1	2	21	8	12	2	2	2	1	0	0	9	9	2	1	7	0	0	2	1	14	
0	2	9	9	13	22	3	3	1	9	1	3	0	7	0	22	13	5	1	25	0	6	6	2	25	
0	0	4	5	15	3	4	2	4	0	12	1	1	1	0	8	4	4	0	9	0	11	8	3	5	
2	3	1	3	2	21	1	2	1	0	0	46	3	5	0	13	51	10	7	10	1	38	1	0	15	
2	0	0	0	1	3	1	0	0	0	0	3	13	9	26	2	0	61	0	0	0	0	0	0	1	
3	2	1	5	4	26	3	1	1	2	0	5	3	12	0	7	10	11	19	7	1	35	2	2	13	
10	3	4	65	0	9	17	14	2	0	1	1	0	0	5	9	1	1	1	1	0	3	6	3	17	
6	3	30	43	1	17	31	30	7	1	0	4	0	0	0	58	15	19	5	33	0	2	4	0	71	
0	1	1	1	2	15	0	0	1	1	1	10	2	4	0	17	84	4	9	16	0	61	1	0	4	
2	0	0	0	2	5	1	0	0	1	0	2	3	1	0	1	5	56	37	3	0	57	0	0	1	
2	2	0	0	0	4	4	0	0	0	0	1	12	0	0	2	5	12	37	0	0	20	0	0	1	
0	0	5	3	9	24	0	1	0	0	0	6	1	7	0	20	13	2	0	40	0	13	0	0	17	
2	1	0	2	2	4	2	1	0	0	0	0	1	1	0	4	4	4	5	0	1	36	1	0	2	
1	0	0	0	2	5	0	0	0	3	0	1	2	5	3	0	1	13	8	17	1	0	145	1	0	3
3	2	4	16	3	13	14	14	1	0	4	0	0	1	0	12	1	2	0	5	1	0	27	3	38	
19	4	6	85	2	5	19	23	1	0	1	1	0	0	4	7	0	2	1	4	0	4	22	23		
5	3	1	47	5	34	6	7	0	1	3	2	0	2	0	31	5	11	0	11	0	3	4	1	135	

Precision

Recall

F-1

0.35

0.32

0.30

- Arousal and positivity models individually hit relatively high precision scores of 0.57 and 0.48



Playlist mood profile app demo



● Ideas for next steps and future work

● App

- Incorporate additional interactivity and data views
- Incorporating Spotify logins and authorizations so users can view and evaluate their own playlists
- Go live and make backend robust

Modeling

- Incorporating lyrics, topics and sentiment
- Approach mood identification as a clustering problem based on audio features

Thanks!

Any questions?

You can find me at

neo.kaiting@gmail.com

github.com/neokt