

Музыка в ответ на эмоции

подбираем музыку по эмоциям,
распознанным из поданного текста или звука

автор: Андрей Кукунов
ментор: Леонид Саночкин

Предпосылки

- Музыка -- это особый род языка, призванный
 - выражать и вызывать состояния  , эмоции  и мысли 
 - в концентрированном  виде
- Текст , интонации  и мелодические линии 
- Базовые состояния в нас  и в музыке  -- ядро работы
 - спокойное  -- энергичное 
 - грустное  -- радостное 

Маркетинговые идеи

1. Широкой аудитории:

- расскажи мне, что ты чувствуешь (в текстовом и/или звуковом сообщении)



- я подберу соответствующую музыку. Она отразит твоё состояние и поможет его выразить.



Маркетинговые идеи

2. Блогерам и подражающим им:

- пришли мне звуковую дорожку с голосом, голосовое сообщение или видео



- я подберу фоновую музыку по настроению присланного и подложу ее под твой голос.



Download YouTube as MP3

бот



/emo 17:25 ✓

👉 Нажми на кнопку под этим сообщением
👉 напиши мне всё, что ты чувствуешь



🎵 а я пришлю тебе музыку по твоему настроению! 🎵

P.S.: обязательно нажми на эту кнопку, а то случится
обычный поиск по YouTube!

17:25

Нажми, напиши текст и отправь мне! ↗

@Dnld_YT_mp3_bot /emo_text Трусишка зайка серенький
под елочкой плясал

17:26 ✓

🎵 Вот как я тебя услышал:



17:26

Download YouTube as MP3

бот



CORMAEL:

"David Bowie Space Oddity (1999 Remastered)"
17 2016-01-17 | 5 min | 20381

🔗 https://youtube.com/watch?v=c6WhqDIg1_o

YouTube

David Bowie || Space Oddity (1999 Remastered)
Album: David Bowie
Release: 1969
Label: Philips (UK) Mercury (US) RCA...



17:26

Нажми и отправь мне, чтобы скачать звук! ↗



South Park Mexican - I Must Be High.

04:45, 4.5 MB

YouTube | add. -3.3 dB
Listen YouTube anywhere

17:36

- Ввели: Трусишка зайка серенький под елочкой плясал
- Получили на панели приборов:

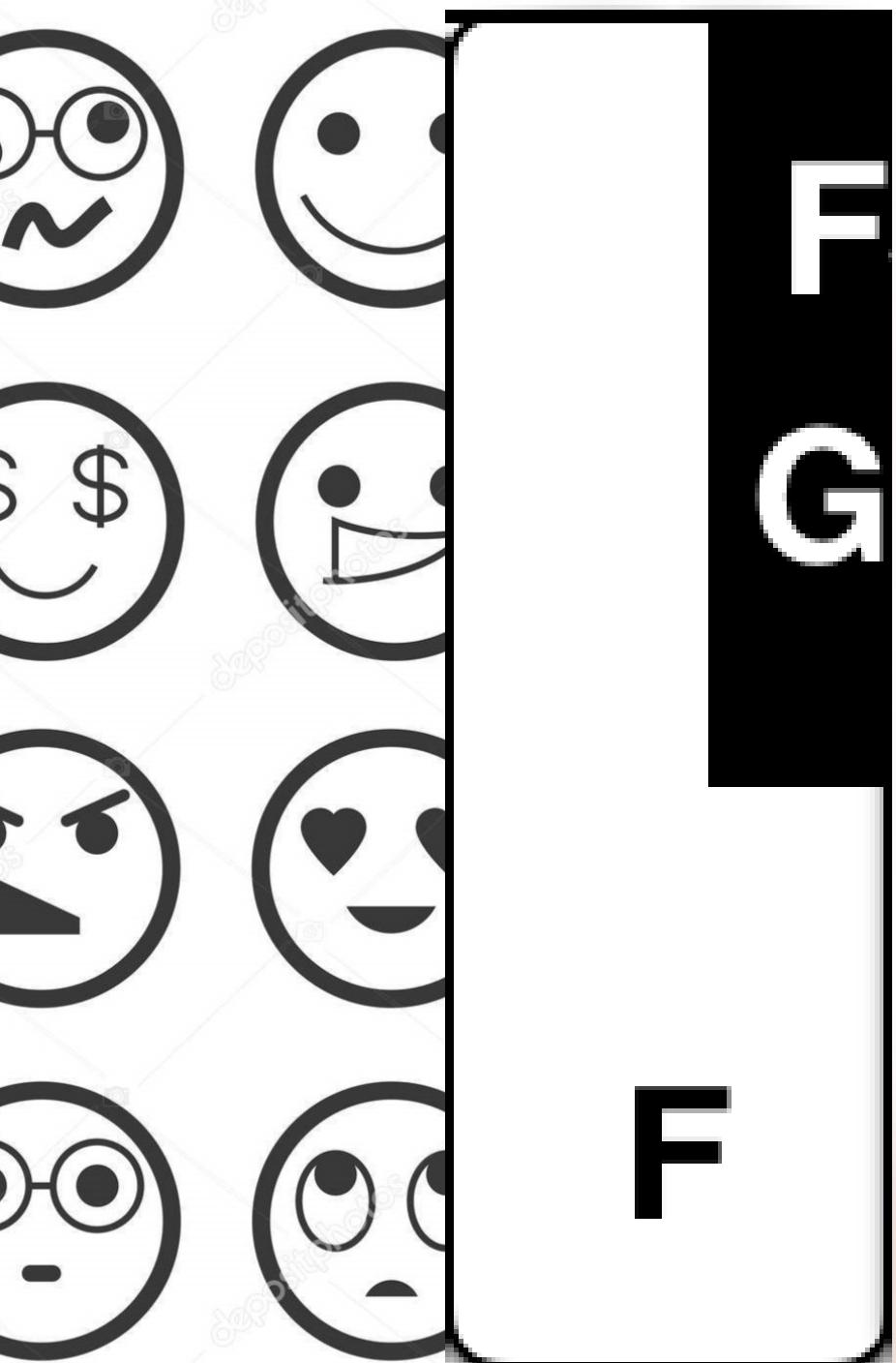
```
Panty bunny Grinky dancing under a Christmas tree
label    sadness        joy        love       anger       fear    surprise
0       0.024429  0.351417  0.058097  0.466345  0.092708  0.007003
text_emos_cluster: [3]
mus_feats_cb_pred:
    danceability   energy      mode  acousticness  valence      tempo
0       0.646914  0.699653  0.656441      0.142636  0.59337  0.458026
```

Data Science проект

Данные:

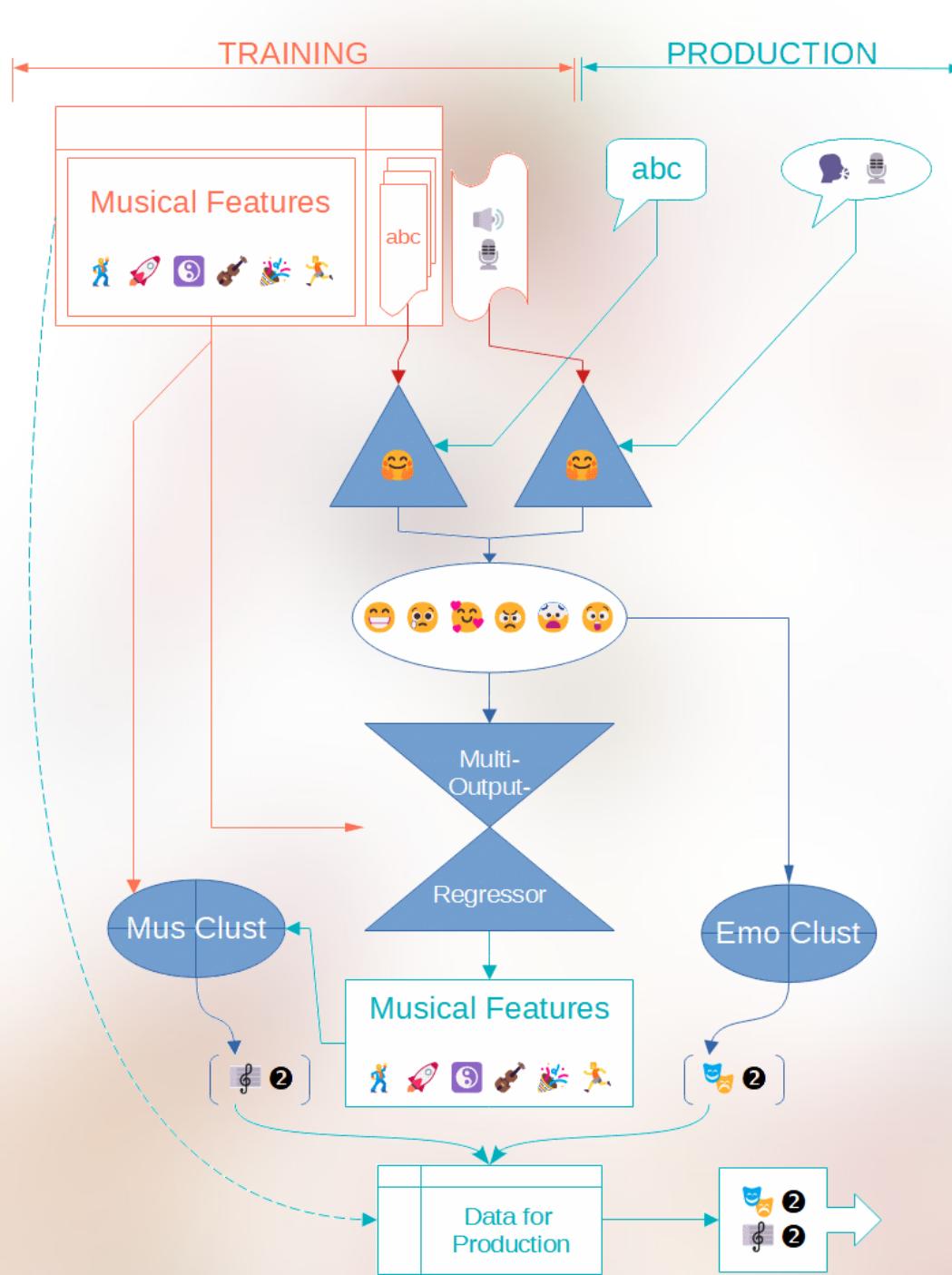
- датасет с открытыми данными **Spotify** и текстами песен 📄
 - более 17 тыс. популярных песен разных стилей
- отобранные музыкальные признаки 🎵:

danceability	energy	mode	acousticness	valence	tempo
танцевальность 	энергичность 	мажор/минор 	инструментальность 	позитивность 	темп 



Задачи:

1. Исследовать взаимосвязь ↔ между
 - эмоциями 😊, 😢, заложенными в музыке и
 - муз. признаками 🎵 от Spotify соответствующих песен
2. Получить предсказания эмоций 😊, 😢 пользователя
3. Предложить максимально соответствующую музыку 🎵



Pipeline

- Данные
 - Датасеты 📄, Звуковые файлы 🔊, Сообщения 💬
- Модели DL и ML
 - Распознавание эмоций 😊, 😢 из текста 📄 и звука 🔊
 - MultiOutput регрессия (6 эмоций на 6 муз. признаков)
 - Кластеризация эмоций 😊, 😢 и муз. признаков 🎵

Что сделано

1. Получены предсказания предобученной модели об эмоциях 😊, заложенных в музыке, исходя из соотв. текстов песен 📄.

```
bhadresh-savani/albert-base-v2-emotion      # best for "text --> emotions"
```



Hugging Face

2. Сделана кластеризация отдельно:

- предсказанных эмоций текста 😊😊📝 и
- музыкальных признаков 🎵 от Spotify
- цель -- естественная группировка материала

3. Создан сводный датасет с указанием обоих кластеров для каждого произведения.

	sadness	joy	love	anger	fear	surprise	danceability	energy	mode	acousticness	valence	tempo	label_y
label_X													
0	0.036563	0.056210	0.020408	0.838549	0.044406	0.003863	0.626428	0.693326	0.580516	0.146694	0.505406	0.479748	1.769473
1	0.028694	0.819773	0.071410	0.064584	0.010603	0.004936	0.597025	0.686263	0.602863	0.170475	0.537936	0.477611	1.884086
2	0.840218	0.040416	0.016180	0.080372	0.020068	0.002746	0.550774	0.680049	0.591208	0.183712	0.473982	0.482024	1.903992
3	0.1111623	0.297034	0.088393	0.411451	0.069034	0.022464	0.585141	0.693194	0.599259	0.166569	0.519356	0.481753	1.799461
4	0.018775	0.166712	0.731783	0.068545	0.009839	0.004346	0.600951	0.644405	0.581570	0.202845	0.549594	0.464131	2.086689
5	0.054763	0.041710	0.015420	0.136259	0.682751	0.069097	0.588499	0.665813	0.597305	0.187511	0.494243	0.471275	1.983533

4. Анализ сводного датасета показал:

- при усредняющей группировке по кластерам эмоций 😊 😢 ②
 - концентрацию большинства эмоций в разных кластерах
 - при этом все музыкальные признаки оказались равномерно распределены по кластерам эмоций.
 - среднее значение музыкального кластера 🎵 ② -- 2

	sadness	joy	love	anger	fear	surprise	danceability	energy	mode	acousticness	valence	tempo	label_y
label_X													
0	0.036563	0.056210	0.020408	0.838549	0.044406	0.003863	0.626428	0.693326	0.580516	0.146694	0.505406	0.479748	1.769473
1	0.028694	0.819773	0.071410	0.064584	0.010603	0.004936	0.597025	0.686263	0.602863	0.170475	0.537936	0.477611	1.884086
2	0.840218	0.040416	0.016180	0.080372	0.020068	0.002746	0.550774	0.680049	0.591208	0.183712	0.473982	0.482024	1.903992
3	0.1111623	0.297034	0.088393	0.411451	0.069034	0.022464	0.585141	0.693194	0.599259	0.166569	0.519356	0.481753	1.799461
4	0.018775	0.166712	0.731783	0.068545	0.009839	0.004346	0.600951	0.644405	0.581570	0.202845	0.549594	0.464131	2.086689
5	0.054763	0.041710	0.015420	0.136259	0.682751	0.069097	0.588499	0.665813	0.597305	0.187511	0.494243	0.471275	1.983533

4. Анализ сводного датасета показал:

- -> слабость связей между эмоциями 😊😊 и муз. признаками 🎵
- невысокая точность взаимных предсказаний 😊😊 <--> 🎵
- --> кл. 2 -- это центр 🎯 музыки в плоскости эмоций 😊😊
- некое "музыкальное коллективное бессознательное" 💭
- константа, которая нам пригодится

	sadness	joy	love	anger	fear	surprise	danceability	energy	mode	acousticness	valence	tempo	label_y
label_X													
0	0.036563	0.056210	0.020408	0.838549	0.044406	0.003863	0.626428	0.693326	0.580516	0.146694	0.505406	0.479748	1.769473
1	0.028694	0.819773	0.071410	0.064584	0.010603	0.004936	0.597025	0.686263	0.602863	0.170475	0.537936	0.477611	1.884086
2	0.840218	0.040416	0.016180	0.080372	0.020068	0.002746	0.550774	0.680049	0.591208	0.183712	0.473982	0.482024	1.903992
3	0.1111623	0.297034	0.088393	0.411451	0.069034	0.022464	0.585141	0.693194	0.599259	0.166569	0.519356	0.481753	1.799461
4	0.018775	0.166712	0.731783	0.068545	0.009839	0.004346	0.600951	0.644405	0.581570	0.202845	0.549594	0.464131	2.086689
5	0.054763	0.041710	0.015420	0.136259	0.682751	0.069097	0.588499	0.665813	0.597305	0.187511	0.494243	0.471275	1.983533

4. Анализ сводного датасета показал:

- при усредняющей группировке по муз. кластерам  ②
 - центральный кластер 2 отличается
 - негативностью , не-танцевальностью ,
 - энергичностью , быстрой темпом , не-инструм. 

	sadness	joy	love	anger	fear	surprise	label_X	danceability	energy	mode	acousticness	valence	tempo
label_y													
0	0.173743	0.268450	0.090176	0.391227	0.066205	0.010199	1.594841	0.522036	0.731254	1.000000	0.068585	0.348579	0.514518
1	0.120878	0.305917	0.126890	0.376710	0.059245	0.010359	1.676067	0.668431	0.728152	1.000000	0.112568	0.717097	0.465348
2	0.176822	0.274224	0.088898	0.384875	0.064570	0.010610	1.614410	0.520398	0.750635	0.000000	0.057721	0.351006	0.512018
3	0.118587	0.303052	0.120540	0.390496	0.057884	0.009442	1.620920	0.705732	0.714546	0.000000	0.113234	0.713233	0.453751
4	0.166446	0.262918	0.143803	0.353092	0.063636	0.010105	1.819795	0.597787	0.464397	0.000000	0.523116	0.407974	0.431264
5	0.183890	0.293944	0.130445	0.311818	0.066986	0.012917	1.851444	0.535838	0.434100	1.000000	0.602905	0.393260	0.439901

4. Анализ сводного датасета показал:

- при усредняющей группировке по муз. кластерам 🎵 ②
 - хорошая поляризация ☰ признаков по кластерам (6x6)
 - корреляции || эмоций 😊 и муз. признаков 🎵
 - гнев (решительность) 🔥 <--> энергичность 🚀
 - радость 😃 || позитивность 🎉 || грусть 😢

	sadness	joy	love	anger	fear	surprise	label_X	danceability	energy	mode	acousticness	valence	tempo
label_y													
0	0.173743	0.268450	0.090176	0.391227	0.066205	0.010199	1.594841	0.522036	0.731254	1.000000	0.068585	0.348579	0.514518
1	0.120878	0.305917	0.126890	0.376710	0.059245	0.010359	1.676067	0.668431	0.728152	1.000000	0.112568	0.717097	0.465348
2	0.176822	0.274224	0.088898	0.384875	0.064570	0.010610	1.614410	0.520398	0.750635	0.000000	0.057721	0.351006	0.512018
3	0.118587	0.303052	0.120540	0.390496	0.057884	0.009442	1.620920	0.705732	0.714546	0.000000	0.113234	0.713233	0.453751
4	0.166446	0.262918	0.143803	0.353092	0.063636	0.010105	1.819795	0.597787	0.464397	0.000000	0.523116	0.407974	0.431264
5	0.183890	0.293944	0.130445	0.311818	0.066986	0.012917	1.851444	0.535838	0.434100	1.000000	0.602905	0.393260	0.439901



4. Анализ сводного датасета показал:

- при усредняющей группировке по муз. кластерам 

 - Это и есть **базовые состояния**, кот. выражает музыка (   )
 - Свести к ним предсказания? -- Не хочется упрощать.

--> Ожидаем базовую сходимость  в предсказаниях муз. признаков по эмоциям   --> 

5. Педсказание муз. признаков (6) по эмоциям (6) 😊😊 --> 🎵

- модели `CatBoostRegressor` и `MultiOutputRegressor` (от Scikit-learn, на эстиматоре `GradientBoostingRegressor`)
- метрики : затрудненный подбор и интерпретация
- результаты предсказаний обеих моделей -- мало отличимы

model	(Multi)RMSE	R2
CatBoostRegressor	0.64	--
MultiOutputRegressor	0.23	0.02

6. Ручное тестирование

- берем песню из тестовой выборки

```
8204    Don't sweat the technique Don't sweat the technique Let's trace the hits and  
Name: lyrics, dtype: object
```

	track_id	track_artist	track_name	playlist_genre
8204	48ZUwXrEOhaXLCxvmRYhZv	Eric B. & Rakim	Don't Sweat The Technique	rap

text_emos:

	sadness	joy	love	anger	fear	surprise
8204	0.040200	0.088773	0.011515	0.760507	0.090509	0.008495

mus_feats:

	danceability	energy	mode	acousticness	valence	tempo
8204	0.719583	0.888132	1.000000	0.005603	0.816345	0.392714

	label_X	label_y	cluster_pred: [3]
8204	0	1	

6. Ручное тестирование

- реальные муз. признаки и кластер  / предсказанные

mus_feats:

	danceability	energy	mode	acousticness	valence	tempo
8204	0.719583	0.888132	1.000000	0.005603	0.816345	0.392714

label_X label_y

cluster_pred: [3]

8204 0 1

mus_feats_cb_pred:

	danceability	energy	mode	acousticness	valence	tempo
0	0.591261	0.705940	0.638632	0.118973	0.479353	0.474940

mus_feats_skl:

	danceability	energy	mode	acousticness	valence	tempo
0	0.594620	0.724891	0.579906	0.128167	0.483025	0.469346

Как улучшить работу моделей?

- --> поработать с предсказаниями эмоций 😊😊
- сделать предсказания по звуку 🎧 песен

```
Aniemore/Russian-Emotion-Recognition      # best for "sound --> emotions"
```



Hugging Face



7. Создана выборка в 300 примеров,
 - сбалансированная по музыкальным кластерам  2 (50 шт. из кластера)
8. По скачанным звуковым файлам песен получены предсказания эмоций   для
 - песен без обработки звука , с 40 по 100 сек.
 - выделенной голосовой партии 
9. Создан сводный датасет (300 песен) всех предсказаний и кластеров
 -    —   2 —  —  2

10. Педсказание напрямую муз. кластеров по эмоциям 🎤2

- Цель -- на уверенной метрике выбрать лучшие комбинации признаков эмоций 🎤2 (💻, 🔊, 🎤)
- Регрессия от CatBoost: для 95% песен предсказан
 - центральный 🎯 муз. кластер 2 🎤2, константа и опора
- Составляющие протестированных вариантов:
 - 💻 -- эмоции текста (😊, 😢, 😍, 😡, 😰, 😮)
 - 🔊 -- эмоции звука без обработки (😊, 😢, 😍, 😡, 😰, 😕, ✌, 😮)
 - 🎤 -- эмоции звука выделенной голосовой партии
 - ✂️ 💻 -- не конкурирующие эмоции текста (😊, 😮)



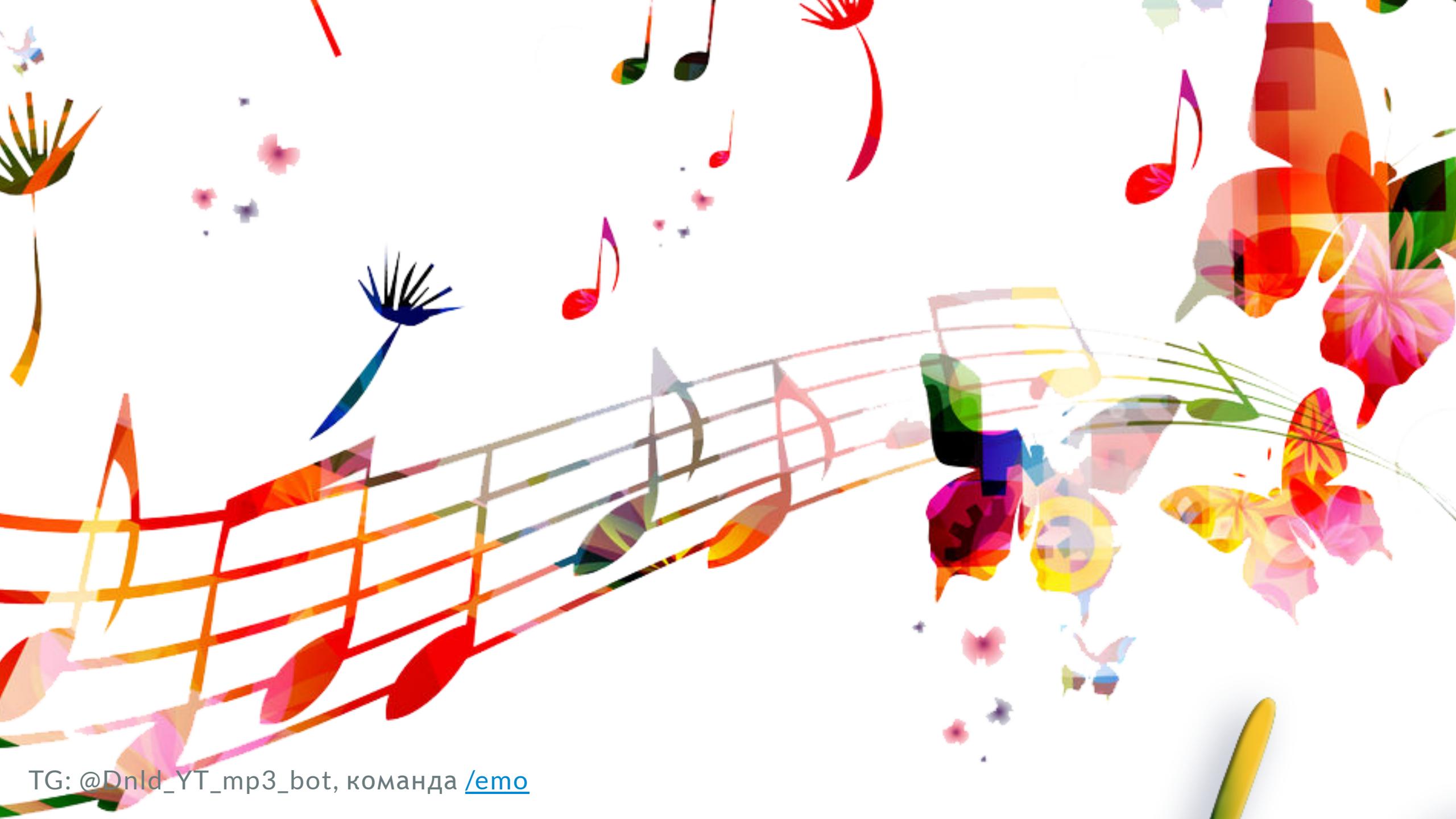
Комбинация	MAPE
Speaker icon	0.242
(Speaker + Microphone) / 2 — Scissors icon	0.242
(Speaker + Microphone) / 2 — Chat icon	0.248
Microphone icon — Scissors icon	0.251
Microphone icon	0.253
Speaker icon — Scissors icon	0.255
Chat icon (15000)	0.263
Chat icon (300)	0.279

Выводы для production

1. Музыкальные признаки  произведений, предлагаемые Spotify, не могут служить надежной опорой  для определения эмоций  произведения.
2. Чтобы предложить "городу и миру" сервис подбора песен по эмоциям, необходимо:
 - кластеризовать эмоции  из их текста  и звука 
 - кластеризовать музыкальные признаки 
 - предлагать имеющиеся в сети песни на пересечении кластеров эмоций ② и музыкальных признаков ②

3. Для подбора фоновой инструментальной 🎻 музыки можно пробовать запрашивать в Spotify рекомендации, отправляя:

- 3-5 примеров из кластера эмоций 😊HSV
- предсказания музыкальных признаков 🎵 (как целевые параметры)
- желаемый уровень инструментальности 🎻



TG: @Dnld_YT_mp3_bot, команда [/emo](#)