Recomendação de Músicas em Playlists no Spotify

Integrantes

Breno Martins
Marina Moreira
Ronald Pereira

1

Uma breve lembrança

O Problema

- Recomendação de músicas

- Centenas de serviços de streaming

- Milhões de usuários

- Dezenas de milhões de músicas

O Problema

- Spotify RecSys Challenge 2018

- Participantes deveriam recomendar músicas para continuar uma playlist

2

Configuração Experimental

Coleção de Treinamento/Teste

- Spotify Playlist Dataset
  - ▷ Conjunto de dados privado
  - ▷ 11.000 playlists
    - ▷ Colaborativas/Individuais
  - ▷ 731.360 músicas ao total
    - ▷ 180.409 músicas únicas
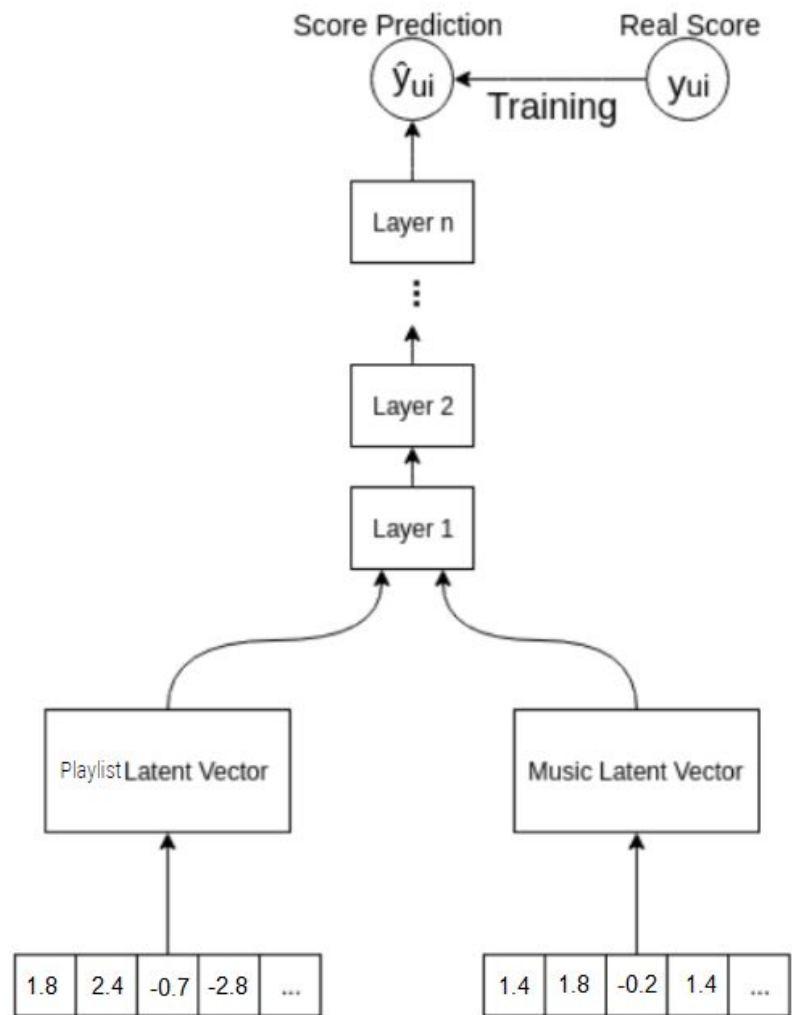 - ▷ 37.476 artistas únicos

Coleção de Treinamento/Teste

	trackid	artist_name	track_name	pid	trackindex
0	spotify:track:0UaMYEvWZi0ZqiDOoHU3YI	Missy Elliott	Lose Control (feat. Ciara & Fat Man Scoop)	0	11647
1	spotify:track:6I9VzXrHxO9rA9A5euc8Ak	Britney Spears	Toxic	0	145808
2	spotify:track:0WqIKmW4BTrj3eJFmnCKMv	Beyoncé	Crazy In Love	0	12511
3	spotify:track:1AWQoqb9bSvzTjaLralEkT	Justin Timberlake	Rock Your Body	0	27379
4	spotify:track:1lzr43nnXAijIGYnCT8M8H	Shaggy	It Wasn't Me	0	41255
5	spotify:track:0XUfyU2QviPAs6bxSpXYG4	Usher	Yeah!	0	12745
6	spotify:track:68vgtRHr7iZHpzGpon6Jlo	Usher	My Boo	0	142288
7	spotify:track:3BxWKCI06eQ5Od8TY2JBeA	The Pussycat Dolls	Buttons	0	73997
8	spotify:track:7H6ev70Weq6DdpZyyTmUXk	Destiny's Child	Say My Name	0	168512
9	spotify:track:2PpruBYCo4H7WOBJ7Q2EwM	OutKast	Hey Ya! - Radio Mix / Club Mix	0	56109

Coleção de Treinamento/Teste

	acousticness	danceability	duration_ms	energy	instrumentalness	liveness	loudness	speechiness	tempo	uri	valence
0	0.03110	0.904	226864	0.813	0.006970	0.0471	-7.105	0.1210	125.461	spotify:track:0UaMYEvWZi0ZqiDOoHU3YI	0.810
1	0.02490	0.774	198800	0.838	0.025000	0.2420	-3.914	0.1140	143.040	spotify:track:6I9VzXrHxO9rA9A5euc8Ak	0.924
2	0.00238	0.664	235933	0.758	0.000000	0.0598	-6.583	0.2100	99.259	spotify:track:0WqIKmW4BTrj3eJFmnCKMv	0.701
3	0.20200	0.891	267267	0.714	0.000234	0.0521	-6.055	0.1400	100.972	spotify:track:1AWQoqb9bSvzTjaLraIEkT	0.818
4	0.05610	0.853	227600	0.606	0.000000	0.3130	-4.596	0.0713	94.759	spotify:track:1Izr43nnXAijIGYnCT8M8H	0.654
5	0.02120	0.881	250373	0.788	0.000000	0.0377	-4.669	0.1680	104.997	spotify:track:0XUfyU2QviPAs6bxSpXYG4	0.592
6	0.25700	0.662	223440	0.507	0.000000	0.0465	-8.238	0.1180	86.412	spotify:track:68vgtRHr7iZHpzGpon6Jlo	0.676
7	0.15800	0.544	225560	0.823	0.000000	0.2680	-4.318	0.3200	210.750	spotify:track:3BxWKCI06eQ5Od8TY2JBeA	0.434
8	0.27300	0.713	271333	0.678	0.000000	0.1490	-3.525	0.1020	138.009	spotify:track:7H6ev70Weq6DdpZyyTmUXk	0.734
9	0.10300	0.728	235213	0.974	0.000532	0.1750	-2.261	0.0665	79.526	spotify:track:2PpruBYCo4H7WOBJ7Q2EwM	0.965

Score Prediction — $\hat{y}_{ui}$ Training   Real Score — y_{ui}

Layer n

Layer 2

Layer 1

Playlist Latent Vector

Music Latent Vector

| 1.8 | 2.4 | -0.7 | -2.8 | ... |

| 1.4 | 1.8 | -0.2 | 1.4 | ... |

- R-Precision

$$\text{R-precision} = \frac{|G \cap R_{1:|G|}|}{|G|}$$

- Normalized Discounted Cumulative Gain (NDCG)

$$\text{DCG} = rel_1 + \sum_{i=2}^{|R|} \frac{rel_i}{\log_2(i+1)} \quad \text{IDCG} = 1 + \sum_{i=2}^{|G|} \frac{1}{\log_2(i+1)} \quad \text{nDCG} = \frac{\text{DCG}}{\text{IDCG}}$$

- Recommended Songs Clicks

$$\text{clicks} = \left\lfloor \frac{\text{argmin}_i\{R_i : R_1 \in |G|\} - 1}{10} \right\rfloor$$

G → Relevant Musics Set R → Returned Musics Set rel_i → Returned Music Position

Spotify® RecSys Challenge 2018

Final ranking for the main track

This leaderboard was computed on July 18, 2018 and shows the scores of final leaderboard for the **main** track. It takes into account the final submission scores evaluated against 100% of the challenge set, and rule compliance. Congratuations to the winners and to all that participated.

#	TEAM	RPREC	RPREC RANK	NDCG	NDCG RANK	CLICKS	CLICKS RANK	BORDA	CODE
1	vl6	0.22413802335	1	0.394622225334	1	1.7839	2	329	Code
2	hello world!	0.223353092072	2	0.393233606092	2	1.8952	6	323	Code
3	Avito	0.215316047292	6	0.384550795868	4	1.7818	1	322	Code

3

Resultados

Resultados dos Baselines

Nome	R-Precision	nDCG	Clicks
Vl6	0.22413802335	0.394622225334	1.7839
hello world!	0.223353092072	0.393233606092	1.8952
Avito	0.215316047292	0.384550795868	1.7818

Tabela 1: Resultados finais dos algoritmos de baseline

Acurácia dos Modelos Utilizados

Metric/Algorithm	Accuracy for Non-Colaborative Playlists	Accuracy for Colaborative Playlists
Original Recommendation	0.513	0.149
Recommendation as Treatments	0.665	**0.467**
Causal Embed	**0.783**	0.132

Tabela 2: Acurácia dos diferentes modelos implementados

Metric/Algorithm	R-Precision	R-Precision Rank	nDCG	nDCG Rank	Clicks	Clicks Rank	Borda	Ranking Position
Original Recommendation	0.0458	89	0.16388	78	4.2016	47	119	72
Recommendation as Treatments	0.20777	14	0.385688	3	2.864	29	287	13
Causal Embed	0.1852	33	0.36231	17	1.9519	9	274	15

Tabela 3: Resultados dos modelos implementados nas métricas propostas pela competição

4

Conclusão

- Para playlists não-colaborativas
  - ▷ Estilo da playlist entre as músicas é evidente
  - ▷ Ambos algoritmos melhoraram as predições
    - ▷ Acoplamento causal → +0.27 de acurácia
    - ▷ "Desenviesamento" → +0.152 de acurácia

Playlists Não-Colaborativas

Recommended Songs ▲
Based on the songs in this playlist

REFRESH

ADD	Falling Away from Me	EXPLICIT	Korn	Issues	4:31
ADD	Killpop	EXPLICIT	Slipknot	.5: The Gray Chapter (Special Edition)	3:45
ADD	Chop Suey!		System Of A Down	Toxicity	3:30
ADD	Break Stuff		Limp Bizkit	Significant Other	2:47
ADD	Down with the Sickness	EXPLICIT	Disturbed	The Sickness	4:39
ADD	Bodies		Drowning Pool	Sinner	3:22
ADD	One Step Closer		Linkin Park	Hybrid Theory (Bonus Edition)	2:37
ADD	Between Angels And Insects	EXPLICIT	Papa Roach	Infest	3:54
ADD	Absolute Zero	EXPLICIT	Stone Sour	House of Gold & Bones Part 1	3:50
ADD	My Own Summer (Shove It)		Deftones	Around The Fur	3:35

- Para playlists colaborativas
  - Estilo da playlist é uma "bagunça"
    - Sem evidências causais que o modelo consegue capturar
  - "Desenviesamento" causa, ainda assim, uma melhora na acurácia → +0.318
  - Acoplamento Causal causa uma leve piora na acurácia → -0.017

Playlists Colaborativas

Recommended Songs ▲
Based on the songs in this playlist

REFRESH

ADD	Shiroyama	Sabaton	The Last Stand	3:36
ADD	Sanctified with Dynamite	Powerwolf	Blood of the Saints	4:26
ADD	Rasputin	Turisas	Rasputin	3:54
ADD	Toda Forma De Poder	Engenheiros Do Hawaii	Longe Demais Das Capitais	3:15
ADD	Choram As Rosas	Hardneja Sertacore	Hardneja Sertacore	3:37
ADD	Batendo Na Porta Do Céu (Knockin' On Heaven'S	Zé Ramalho	Zé Ramalho Ao Vivo 2005	3:58
ADD	Eu	Pato Fu	Ruido Rosa	3:01
ADD	Last Man Standing	Hammerfall	Steel Meets Steel - 10 Years Of Glory	4:30
ADD	How High	Welshly Arms	No Place Is Home	4:07
ADD	A Escolhida	Mc Pedrinho	A Escolhida	3:31

OBRIGADO

Dúvidas?

github.com/ronaldpereira/spotify-playlist-song-recommendation/