

Music Information Retrieval

Relatório

Feito por:

- Marco Pais - Nº 2019218
- Tiago Oliveira - Nº 2019219068

1. Preparação

- 1.1. Ganhar familiaridade com um sistema de recomendação real, e.g., Jango.com, Spotify, Last.fm, Torch ou outro.
- 1.2. Descarregar a base de dados a utilizar (4Q audio emotion dataset) do seguinte URL:
<http://mir.dei.uc.pt/downloads.html>.
- 1.3. Analisar a base de dados.
 - 1.3.1. Excertos áudio: ficheiros mp3 (em particular as 4 queries).
 - 1.3.2. Metadados: ficheiro panda_dataset_taffc_metadata.csv, colunas Song, Artist, Quadrant, MoodsStrSplit e GenresStr.
 - 1.3.3. Features: ficheiros top100_features.csv (features extraídas) e features.csv (descrição das features).
- 1.4. Estudar a framework de processamento áudio librosa
 - 1.4.1. Instalar a framework a partir do seguinte URL: <https://librosa.org/>. Será também necessário instalar a framework ffmpeg para leitura de ficheiros .mp3: <https://ffmpeg.org/>.
 - 1.4.2. Analisar o código fonte de base fornecido no InforEstudante (ficheiro mrs.py).
 - 1.4.3. Estudar a documentação da framework: <https://librosa.org/doc/>.

2. Extracção de Features

- 2.1. Processar as features do ficheiro top100_features.csv.
 - 2.1.1. Ler o ficheiro e criar um array numpy com as features disponibilizadas.
 - 2.1.2. Normalizar as features no intervalo [0, 1].
 - 2.1.3. Criar e gravar em ficheiro um array numpy com as features extraídas (linhas = músicas; colunas = valores das features).
- 2.2. Extrair features da framework librosa.
 - 2.2.1. Para os 900 ficheiros da BD, extrair as seguintes features (sugestão: guardar todas as músicas na mesma pasta):
 - Features Espectrais: mfcc, spectral centroid, spectral bandwidth, spectral contrast, spectral flatness e spectral rolloff.
 - Features Temporais: F0, rms e zero crossing rate.
 - Outras features: tempo.

- Utilize os parâmetros por omissão do librosa (sr = 22050 Hz, mono, window length = frame length = 92.88 ms e hop length = 23.22 ms).
- Guarde as features num array numpy 2D, com **número de linhas = número de músicas e número de colunas = número de features**
- 2.2.2. Calcular as 7 estatísticas típicas sobre as features anteriores: média, desvio padrão, assimetria (skewness), curtose (kurtosis), mediana, máximo e mínimo. Para o efeito, utilizar a biblioteca scipy.stats (e.g., scipy.stats.skew).
- 2.2.3. Normalizar as features no intervalo [0, 1].
- 2.2.4. Criar e gravar em ficheiro o array numpy com as features extraídas.
- 2.3. **Alínea com bonificação de 10% na nota final! (Sugestão: desenvolver nas semanas 5 e 6).** Implementar features de raiz. Neste ponto, não é permitido utilizar o librosa nem qualquer outra biblioteca, à excepção do scipy e numpy, e.g., scipy.fftpack.
 - 2.3.1. Desenvolver o código Python/numpy para extrair as features anteriores (à excepção de spectral contrast, F0 e tempo), usando a mesma parametrização. Comparar os resultados obtidos com os resultados do librosa.
 - 2.3.2. Criar e gravar em ficheiro um array numpy com as features extraídas (i.e, as estatísticas correspondentes, com normalização).

3. Implementação de métricas de similaridade

- 3.1. Desenvolver o código Python/numpy para calcular as seguintes métricas de similaridade:
 - 3.1.1. Distância Euclidiana
 - 3.1.2. Distância de Manhattan
 - 3.1.3. Distância do Coseno
- 3.2. Criar e gravar em ficheiro 6 (ou 9, se bonificação) matrizes de similaridade (900x900), uma para cada conjunto de features e métrica de distância utilizada.
- 3.3. Criar os 6 (ou 9) rankings de similaridade (para as 4 queries fornecidos). Considere apenas recomendações de 20 músicas.
- **3.4. Apresentar, comparar e discutir os resultados.**

Como podemos verificar pela análise das tabelas seguintes, para várias distâncias são obtidas recomendações diferentes para a mesma música.

Euclidiana: top100 - MT0000202045

MT0014878397	MT0011821215	MT0011380622	MT0005129157	MT0033415296
MT0040033011	MT0005091539	MT0012534566	MT0006001707	MT0029874624
MT0011413068	MT0004896738	MT0032235381	MT0000011975	MT0013313448
MT0009016829	MT0004958762	MT0009010830	MT0003311798	MT0008733057

Manhattan: top100 - MT0000202045

MT0011821215	MT0014878397	MT0005129157	MT0011380622	MT0040033011
--------------	--------------	--------------	--------------	--------------

MT0033415296	MT0005091539	MT0012534566	MT0029874624	MT0004896738
MT0032235381	MT0011413068	MT0006001707	MT0009016829	MT0004958762
MT0009010830	MT0013313448	MT0000011975	MT0005752234	MT0001236649

Cosseno: top100 - MT0000202045

MT0010344415	MT0027048677	MT0011821215	MT0000092267	MT0017797643
MT0001217651	MT0029874624	MT0000821772	MT0001419145	MT0009845275
MT0034125967	MT0008575372	MT0004537445	MT0036368550	MT0040033011
MT0005261375	MT0004958762	MT0005129157	MT0000044741	MT0004896738

Euclidiana: Librosa - MT0000202045

MT0005129157	MT0011899302	MT0012001409	MT0002233402	MT0007043504
MT0007799677	MT0004428604	MT0002161109	MT0011975274	MT0010624346
MT0009217411	MT0000092267	MT0007766156	MT0001676671	MT0005270263
MT0003787478	MT0011376343	MT0005213723	MT0000732821	MT0002634024

Manhattan: Librosa - MT0000202045

MT0005129157	MT0011899302	MT0012001409	MT0007799677	MT0002233402
MT0007043504	MT0010624346	MT0000732821	MT0011975274	MT0004428604
MT0000092267	MT0003787478	MT0002161109	MT0009217411	MT0002379222
MT0002634024	MT0005213723	MT0011376343	MT0036368550	MT0000446304

Cosseno: Librosa - MT0000202045

MT0005129157	MT0012001409	MT0011899302	MT0002233402	MT0007043504
MT0007799677	MT0004428604	MT0011975274	MT0002161109	MT0000092267
MT0010624346	MT0009217411	MT0007766156	MT0003787478	MT0010085729
MT0005270263	MT0001676671	MT0000732821	MT0005213723	MT0026158301

Ao analisar os resultados dos diversos top20 obtidos pode-se verificar que as músicas primeiras músicas são iguais para as 3 distâncias e o com afastamento dessas primeiras músicas é verificado que as músicas são cada vez mais distintas.

Visto que os resultados não foram os mais satisfatórios, uma forma de os melhorar passaria pela escolha de outro conjunto de *Features* e de distâncias para o cálculo da similaridade entre as músicas.

4. Avaliação

- 4.1. Avaliação objectiva.

- 4.1.1. Para cada uma das 4 queries, obter o ranking das 20 músicas recomendadas com base na correspondência com os **metadados** seguintes: artista, género, quadrante e emoção. Por cada item coincidente, adicionar um ponto à qualidade da música alvo, e.g., se tanto a música de referência como o alvo tiverem género = jazz e emoção = happy, a qualidade do alvo será 2.
- 4.1.2. Criar e gravar a matriz de similaridade baseada em contexto (i.e., nos metadados).
- 4.1.3. Para cada um dos rankings determinados em 3.3, calcular a métrica precision, assumindo como relevantes as músicas devolvidas em 4.1.1 (metadados).
- **4.1.4. Apresentar, comparar e discutir os resultados.**

Metadados - MT0000202045

MT0014475915	MT0012862507	MT0000888329	MT0007556029	MT0031898123
MT0004867564	MT0001494812	MT0003022328	MT0011922905	MT0030369896
MT0007453719	MT0034186620	MT0004850690	MT0011938737	MT0034577404
MT0003025046	MT0005285696	MT0002846256	MT0001058887	MT0007766156

	Precision of MT0000202045.mp3	Precision of MT0000379144.mp3	Precision of MT0000414517.mp3	Precision of MT0000956340.mp3
euclidiana_top100.csv	0.0	0.0	0.05	0.05
manhattan_top100.csv	0.0	0.0	0.05	0.05
coseno_top100.csv	0.0	0.0	0.05	0.05
euclidiana_normalizada.csv	0.05	0.0	0.15	0.0
manhattan_normalizada.csv	0.0	0.0	0.1	0.0
coseno_normalizada.csv	0.05	0.0	0.1	0.0

Analisando os resultados obtidos no cálculo das precisões é possível verificar que as precisões são todas muito baixas, sendo o valor mais alto 0.15.

Os valores mais altos de precisão podem ser encontrados na música **MT0000414517.mp3**. Em contraste com a música MT0000414517.mp3, a música **MT0000379144.mp3** apresenta sempre valores de precisão nulos, ou seja, não existe qualquer *match* entre o top20 gerado pelos metadados e os top20 gerados pelas várias distâncias.

- 4.2. Avaliação subjectiva.
 - 4.2.1. Para cada uma das 4 queries, conjunto de 100 features e distância do coseno, avaliar a qualidade de cada uma das 20 recomendações, com base na seguinte escala de Likert ": 1 – Muito Má; 2 – Má; 3 – Aceitável; 4 – Boa; 5 – Muito Boa(para as 4 queries). Cada elemento do grupo deverá efectuar individualmente a avaliação da recomendação.
 - a) Calcular a média e o desvio-padrão de todos os membros por query, assim como a média e o desvio-padrão global para as 4 queries.
 - b) Definindo um score mínimo de 2.5 para "recomendação relevante", calcular a precision resultante.
 - 4.2.2. Para cada uma das 4 queries e similaridade com base nos metadados, avaliar a qualidade de cada uma das 20 recomendações, com base na escala de Likert anterior. Cada elemento do grupo deverá efectuar individualmente a avaliação da recomendação.

a) Calcular a média e o desvio-padrão de todos os membros por query, assim como a média e o desvio-padrão global para as 4 queries.

b) Definindo um score mínimo de 2.5 para “recomendação relevante”, calcular a precision resultante e actualizar o valor da F-measure.

◦ **4.2.3. Apresentar, comparar e discutir os resultados.**

MT0000202045.mp3

	Marco	Tiago	Média	Desvio Padrão
MT0010344415	1,5	1	1,25	0,25
MT0027048677	2,5	2	2,25	0,25
MT0011821215	1,5	2	1,75	0,25
MT0000092267	1,5	2	1,75	0,25
MT0017797643	2	1	1,5	0,5
MT0001217651	3	2,5	2,75	0,25
MT0029874624	2,5	3	2,75	0,25
MT0000821772	1,5	1	1,25	0,25
MT0001419145	2	1,5	1,75	0,25
MT0009845275	3	2	2,5	0,5
MT0034125967	2	1	1,5	0,5
MT0008575372	2	2,5	2,25	0,25
MT0004537445	1,5	2	1,75	0,25
MT0036368550	1,5	1	1,25	0,25
MT0040033011	1,5	2	1,75	0,25
MT0005261375	2	1,5	1,75	0,25
MT0004958762	3	3,5	3,25	0,25
MT0005129157	3	4	3,5	0,5
MT0000044741	2,5	2	2,25	0,25
MT0004896738	1,5	1	1,25	0,25
Média Total		2		
Desvio Padrão Total		0,7245688373		
Precisão		5/20		

MT0000379144.mp3

	Marco	Tiago	Média	Desvio Padrão
MT0030214520	2	1,5	1,75	0,25
MT0003272751	2	1	1,5	0,5

	Marco	Tiago	Média	Desvio Padrão
MT0003262589	1,5	1	1,25	0,25
MT0003213835	1	2	1,5	0,5
MT0002132088	3	2,5	2,75	0,25
MT0005387539	1	1	1	0
MT0009904717	1,5	1,5	1,5	0
MT0028167155	2	1	1,5	0,5
MT0006367176	3	3,5	3,25	0,25
MT0000729701	1,5	1	1,25	0,25
MT0027002641	3,5	3	3,25	0,25
MT0012124855	1,5	1,5	1,5	0
MT0009188643	2,5	2,5	2,5	0
MT0006742179	1,5	1	1,25	0,25
MT0012396528	2,5	3	2,75	0,25
MT0007067293	2	1	1,5	0,5
MT0008716237	3,5	3,5	3,5	0
MT0015005100	3	2,5	2,75	0,25
MT0008570712	1	1	1	0
MT0008222676	2	1,5	1,75	0,25
Média Total		1,95		
Desvio Padrão Total		0,8426149773		
Precisão		6/20		

MT0000414517.mp3

	Marco	Tiago	Média	Desvio Padrão
MT0006939177	3	2,5	2,75	0,25
MT0028335228	2	2	2	0
MT0009169626	2	2	2	0
MT0027976714	1	1	1	0
MT0030319227	1	1	1	0
MT0001676671	2	1	1,5	0,5
MT0003949060	3	2,5	2,75	0,25
MT0004131058	1	1	1	0
MT0007477802	3	4	3,5	0,5
MT0001703346	4	2	3	1

	Marco	Tiago	Média	Desvio Padrão
MT0002448379	2,5	2,5	2,5	0
MT0009437817	3	2	2,5	0,5
MT0007349999	1	2	1,5	0,5
MT0002595792	2	1	1,5	0,5
MT0001193971	2	2	2	0
MT0001624303	2	1,5	1,75	0,25
MT0006410285	1	2	1,5	0,5
MT0005270263	2	2	2	0
MT0026776967	3	3	3	0
MT0004840819	1	1	1	0
Média Total		1,9875		
Desvio Padrão Total		0,825284042		
Precisão		6/20		

MT0000956340.mp3

	Marco	Tiago	Média	Desvio Padrão
MT0012168286	0,5	0,5	0,5	0
MT0002385077	0,5	1	0,75	0,25
MT0004645468	1	1	1	0
MT0002399275	2	1	1,5	0,5
MT0006798861	1	2	1,5	0,5
MT0011836290	1	1	1	0
MT0009800907	1	1	1	0
MT0000956340	4	3	3,5	0,5
MT0007535042	3	1	2	1
MT0006552038	2	1	1,5	0,5
MT0005751512	1	2	1,5	0,5
MT0028345470	1	2	1,5	0,5
MT0028560561	1	1	1	0
MT0007413949	2	2	2	0
MT0028750297	2	2	2	0
MT0028633715	1	1	1	0
MT0003968586	2	2	2	0
MT0002372242	3	3,5	3,25	0,25

	Marco	Tiago	Média	Desvio Padrão
MT0029772184	1	4	2,5	1,5
MT0005515169	1,5	2,5	2	0,5
Média Total		1,65		
Desvio Padrão Total		0,916515139		
Precisão		3/20		

Através da análise objetiva seria de esperar que na análise subjetiva as músicas **MT0000202045** e **MT0000379144** apresentassem valores de precisão mais baixos que as restantes músicas. No entanto, isso não se verificou, tendo até valores de precisão equiparáveis à música **MT0000414517**.

Essa diferença nos resultados deve-se ao facto de termos sido nós a escolher as músicas relevantes ao contrário da alínea anterior em que estas eram as músicas com os metadados mais semelhantes aos das queries.