



UNIVERSIDADE FEDERAL DA FRONTEIRA SUL
CIÊNCIA DA COMPUTAÇÃO

DANIELE KAROLINE CARVALHO ROSA

Atividade 6 - Clustering

CHAPECÓ

2024

K-means: Agrupa dados em um número fixo de clusters (grupos) ao encontrar pontos centrais e atribuir cada ponto ao centróide mais próximo.

HDBSCAN: Agrupa dados com base na densidade, identificando áreas densas de pontos e permitindo clusters de formas e tamanhos variados. Ele também consegue identificar pontos que não pertencem a nenhum cluster.

Spectral Clustering: Usa grafos para agrupar dados, identificando conexões fortes entre pontos próximos. Ele transforma o problema em uma questão de álgebra linear, facilitando a detecção de clusters.

Para o agrupamento e análise dos dados, utilizou-se os algoritmos:

```
from sklearn.cluster import KMeans
from sklearn.cluster import HDBSCAN
from sklearn.cluster import SpectralClustering
```

Adicionando os parâmetros e treinando o algoritmo:

```
# Crie o modelo usando K-Means e defina parâmetros essenciais (número de grupos/clusters)
modelo_ia_1 = KMeans(n_clusters=3)

# Crie o modelo usando HDBSCAN e defina parâmetros essenciais (quantidade mínima de
elementos para formar um grupo)
modelo_ia_2 = HDBSCAN(min_cluster_size=4)

# Crie o modelo usando SpectralClustering e defina parâmetros essenciais (número de
grupos/clusters)
modelo_ia_3 = SpectralClustering(n_clusters=2)

# Treina o modelo
modelo_ia_1.fit(x)
modelo_ia_2.fit(x)
modelo_ia_3.fit(x)
```

Resultado (parcial) após execução:

	Nome	Pop	Sertanejo	...	Grupo KMeans	Grupo HDBSCAN	Grupo SpectralClustering
	Paulo Sergio Nunes	1	0	...	0	-1	0
	Diogo Maciel da Cunha	1	1	...	0	-1	0
	Heitor Machado	1	0	...	0	0	0
	João Marcos Soares	1	1	...	0	-1	0
	Eduardo Folle Miotto	1	0	...	0	0	0
	Marco Antonio Balestrin	1	0	...	0	0	0
	Jacquet Leme	1	1	...	0	-1	1
	Lucas Smaniotto	1	0	...	0	0	0
	Gabrieli Loeblein Brutscher	1	1	...	0	-1	0
	Valtemir Gomes	1	0	...	0	-1	0
	Gabriel Schmitz Vieira	1	0	...	0	-1	0
	Robert Drey	1	0	...	0	-1	0
	Larissa Mones Bedin	1	0	...	0	-1	0
	João Lamaison	1	1	...	0	-1	0
	Richard Facin Souza	1	1	...	0	-1	0
	Victor de Conto	1	1	...	0	-1	0
	Daniele Karoline	1	1	...	0	-1	0
	João Bettu	1	1	...	0	-1	0
	Bruno Neckel	1	1	...	0	0	0
	Matheus Slama	1	0	...	0	0	0
	Diogo Kemmerich	1	0	...	0	0	0
	Mateus Azor Frutuoso	1	0	...	0	0	0
	Diego S. Minichiello	1	1	...	0	-1	0
	Yuri Lanzini	1	1	...	0	-1	0
	Leonardo Nakamichi	1	1	...	0	-1	0
	Henrique Alberto de Andrade	1	0	...	0	-1	0
	Carlos Thomas	1	1	...	1	2	1
	Gustavo Zampieron	1	0	...	1	-1	0
	Pedro Dutra	1	0	...	1	-1	0
	Letícia da Rosa	1	1	...	1	-1	0
	Ashley Saint-Louis	1	0	...	1	-1	1
	Felipe Grando	1	0	...	1	1	0
	Jeffley Garçon	1	0	...	1	1	1
	Gabriel Menoncin	1	1	...	1	-1	1
	Bianca Gabriela Fritsch	1	0	...	1	-1	0
	Linka Sofia	1	1	...	1	-1	1
	Gabriel Kolling da Silva	1	1	...	1	-1	1
	Matheus Vieira Santos	1	0	...	1	1	1
	Luiz Fernando Klein	1	0	...	1	-1	0
	Thaila Schmidt	1	0	...	1	-1	0
	João Gomes	1	1	...	1	-1	0
	Nicole Moritz schneider	1	1	...	1	-1	1
	Ana Clara Brusamarello	1	0	...	1	-1	0
	Aline Moraes	1	0	...	1	-1	0
	Isabeli Reik	1	0	...	1	1	1
	Luiz Paulo	0	0	...	1	-1	0
	Natalia Ferrandin	1	0	...	1	1	1
	Patrícia Trevisan	1	1	...	1	-1	1
	Kellane	1	0	...	1	-1	0
	MATHEUS HENRIQUE RODRIGUES DA COSTA	1	0	...	1	-1	1
	Axel Aviloff	1	0	...	1	-1	0

Pop	Sertanejo	Funk	Rock	Bandinha	Pisadinha	Eletrônica	K-Pop	Trap	Hip-Hop	House	Folk	MPB	Lo-Fi	Indie	Bossa Nova	Samba	Pagode	Gaúchesca	Dubstep	Brega	Reggae	Metal	Reggaeton	Clássica	FPS	Mundo Aberto	Guerra	MOBA	Estratégia		
Corrida	Battle Royale	Survival	Simulação	Esporte	Arcade	RPG	Plataforma	Cartas	Tabuleiro	Luta	Azar	MMORPG	Puzzle	Hack and slash	Rímico	Point and click	Furtivo	Realidade Virtual	Nome												
0	1.00	0.50	0.65	0.96	0.38	0.19	0.77	0.31	0.58	0.81	0.54	0.58	0.81	0.65	0.62	0.77	0.77	0.81	0.46	0.46	0.27	0.81	0.69	0.58	0.85	0.85	0.96	0.81	0.50	0.85	
0	0.69	0.69	0.83	0.81	0.46	0.92	0.85	0.73	0.92	0.81	0.92	0.46	0.73	0.88	0.88	0.73	0.58	0.58	0.58	0.58	0.58	0.69	0.62	Centro 0	0.62	0.52	0.62	0.66	0.24	0.31	0.62
1	0.97	0.38	0.83	0.86	0.38	0.31	0.41	0.21	0.34	0.66	0.14	0.31	0.62	0.38	0.66	0.62	0.38	0.45	0.62	0.28	-0.80	0.17	0.34	0.41	0.41	0.52	0.62	0.66	0.24	0.31	0.62
0	0.34	0.34	0.66	0.41	0.31	0.28	0.41	0.28	0.76	0.79	0.14	0.34	0.28	0.62	0.62	0.10	0.24	0.14	0.10	0.14	0.10	0.10	0.07	Centro 1	0.38	0.46	0.67	0.75	0.54	0.25	0.67
2	0.71	0.42	0.12	0.83	0.58	0.12	0.83	0.04	0.25	0.38	0.33	0.33	0.17	0.42	0.33	0.08	0.08	0.12	0.54	0.08	0.08	0.17	0.50	0.38	0.46	0.67	0.75	0.54	0.25	0.67	
0	0.67	0.42	0.54	0.58	0.46	0.54	0.58	0.33	0.28	0.33	0.38	0.21	0.33	0.38	0.38	0.25	0.04	0.54	0.54	0.54	0.54	0.54	0.54	Centro 2	0.54	0.54	0.54	0.54	0.54	0.54	0.54

K-Means: Os valores de agrupamento variam entre 0 e 1, sugerindo dois clusters principais nesse método.

HDBSCAN: Utiliza um valor -1 para indicar pontos considerados como outliers ou que não pertencem a nenhum cluster. Clusters reais têm valores 0, 1, e 2, o que indica

a formação de alguns grupos significativos e a presença de muitos pontos fora dos clusters (rotulados como -1).

SpectralClustering: Indica a formação de dois clusters, mas não trata explicitamente outliers como HDBSCAN. Utiliza 0 e 1 como agrupamento.

K-Means:

- **Centro 0:** Focado em músicas populares (Pop, Rock) e jogos de ação (FPS, Battle Royale). Perfil dinâmico e mainstream.
- **Centro 1:** Preferências por músicas suaves (MPB, Samba) e jogos estratégicos ou calmos (Puzzle, Tabuleiro).
- **Centro 2:** Estilos alternativos (Indie, Metal) e jogos de imersão e cooperação (RPG, MMORPG).

HDBSCAN:

- **Cluster Denso 0:** Músicas rápidas (Eletrônica, Trap) e jogos de ação intensa (FPS, Corrida). Perfil ativo e estimulante.
- **Cluster Denso 1:** Músicas calmas (Folk, Clássica) e jogos estratégicos (Puzzle, Tabuleiro). Perfil mais tranquilo.
- **Ruído:** Usuários com preferências muito variadas, sem cluster específico.

Spectral Clustering:

- **Cluster 0:** Interesse em músicas populares (Pop, Rock, Eletrônica) e jogos de ação (Battle Royale, Arcade).
- **Cluster 1:** Preferências por músicas complexas (MPB, Clássica) e jogos analíticos (Estratégia, RPG).

1) Qual estilo de jogo seria indicado para alguém que gosta de sertanejo?

Para recomendar um estilo de jogo para quem gosta de sertanejo, olhamos o cluster onde a pontuação de Sertanejo é mais alta. No mesmo cluster, buscamos os estilos de jogo com valores altos. Para quem gosta de sertanejo, os estilos de jogo recomendados são **RPG, Simulação** ou **Esporte**.

2) Qual estilo de música seria indicado para alguém que gosta de jogos de estratégia?

Para recomendar um estilo de música para quem gosta de jogos de estratégia, olhamos o cluster onde Estratégia tem a pontuação mais alta. Nesse cluster, vemos quais estilos musicais também têm valores altos. Para quem gosta de jogos de estratégia, os estilos de música recomendados são **Clássica, Folk** ou **Indie**.