

SISTEMA DE ANÁLISE, RECOMENDAÇÃO E CLASSIFICAÇÃO DE JOGOS

MÉTODOS PROBABILÍSTICOS PARA ENGENHARIA INFORMÁTICA

Eduardo Alves nº104179

Joaquim Martins nº115931



UNIVERSIDADE DE AVEIRO, ENGENHARIA DE COMPUTADORES E INFORMÁTICA



Ideia de aplicação conjunta e exemplos de utilização

Ideia da Aplicação

Sistema capaz de:

- Classificação de Jogos usando Naïve Bayes
- Análise de Jogos usando Filtro de Bloom
- Recomendação de Jogos usando Minhash

Exemplos de Utilização

- Recomendação de prendas de Natal
- Auxilia pais a identificar jogos adequados para os filhos

DataSets

Apocalypse	Stealth	Dark Atmosphere	Crafting	Class
0	0	0	1	-18
0	0	0	0	-18
1	0	1	0	+18
0	0	0	1	-18
0	0	0	0	-18
0	0	0	0	-18
0	0	0	0	-18
1	0	1	0	+18
1	0	0	0	+18

naivebayes_data.xlsx

Dados de treino e teste para o
Naïve Bayes.

Apocalypse	Stealth	Dark Atmosphere	Crafting	Class
0	0	1	1	The Witcher 3
1	0	0	1	Fortnite
0	0	0	1	Minecraft
0	0	1	0	Dark Souls III
0	1	0	1	Read Dead Redemption 2
0	0	0	0	League of Legends
0	0	0	0	Overwatch 2
0	0	0	0	Among Us
0	0	1	1	God of War (2018)

jogos_normais.xlsx

Jogos PC/PlayStation com tags.
(Usados no Bloom Filter e MinHash)

Stealth	Dark Atmosphere	Crafting	Class
0	1	1	The Elder Scrolls V: Skyrim VR
0	0	1	No Man's Sky VR
0	0	0	Beat Saber
0	1	0	Half-Life: Alyx
0	0	0	Star Wars: Squadrons
0	1	0	Boneworks

Jogos_VR.xlsx

Jogos VR com tags associadas.
(Usados no minHash)

Naïve Bayes

Utilização

Classificar jogos como apropriados para maiores ou menores de 18 anos, com base nas tags (“features”).

Implementação

- Previsão de idade recomendada (+18 ou -18), baseada nas tags ‘features’ que o utilizador indica que gosta.
- Ajuste de pesos de tags (“features”) para personalizar recomendações.

Filtro de Bloom

Utilização

Verifica se um jogo escolhido pelo utilizador tem uma determinada tag escolhida pelo mesmo

Implementação

- Cada jogo possui um Bloom Filter dedicado às suas tags associadas.
- Permite identificar se uma tag está vinculada a um jogo escolhido pelo utilizador.
- Pode haver falsos positivos, mas nunca falsos negativos.

Minhash

Utilização

Recomenda os jogos VR mais semelhantes a um jogo normal escolhido pelo utilizador.

Implementação

- Cria assinaturas usando funções hash para os conjuntos de dados dos jogos normais e dos jogos VR .
- Usa a similaridade de Jaccard para comparar as assinaturas geradas.
- São apresentados os jogos dos mais aos menos semelhantes. Até 5 recomendações diferentes (limite que colocamos).

Naïve Bayes – Testes

true positive	false positive
false negative	true negative

Matriz de Confusão

$$\text{Accuracy} = \frac{tp+tn}{tp+fp+tn+fn}$$

Fórmula da Accuracy

$$\text{recall} = \frac{tp}{tp+fn}$$

Fórmula do Recall

$$F_1 = \frac{2PR}{P+R}$$

Fórmula de F1-Score

```
>> naivebayes_teste
```

```
Após dividir o modelo em 70% de treino e 30% de testes:
```

```
A precisão do modelo é de 92.46%
```

```
Matriz de Confusão:
```

```
129    20
 3    153
```

```
Verdadeiros Positivos (jogos para +18 e o modelo classificou como +18): 153
```

```
Falsos Positivos (jogos para +18 mas o modelo classificou mal como -18): 20
```

```
Falsos Negativos (jogos para -18 mas o modelo classificou mal como +18): 3
```

```
Verdadeiros Negativos (jogos para -18 e foram classificados como -18): 129
```

```
Accuracy: 0.92459
```

```
Recall: 0.97727
```

```
F1-score: 0.9502
```

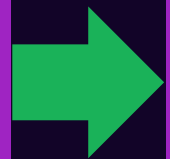
```
>>
```

Resultados Obtidos:

Naïve Bayes

Filtro de Bloom – Testes

No nosso programa
são inseridas 407
tags no filtro de
bloom



Colocamos o
tamanho do nosso
filtro 8x Maior, logo
3256

$$k_{\text{ótimo}} \approx \frac{n \ln(2)}{m}$$

Fórmula de K (número de funções dispersão) ótimo

$$K_{\text{ótimo}} = (\ln(2) * 3256) / 407$$

$$K_{\text{ótimo}} = 5,544 \text{ logo } 6$$

$$p_{fp} = \left[1 - \left(1 - \frac{1}{n} \right)^{km} \right]^k$$

Fórmula de probabilidade de ocorrer falsos positivos

```
Resultado do Teste Bloom Filter:  
Número total de verificações: 1175  
Número total de tags inseridas no Bloom Filter (ou seja com o valor 1): 407  
Número de Falsos positivos: 36  
Percentagem de falsos positivos: 3.06%  
Probabilidade de ocorrer falsos positivos: 100.00%  
>>
```

```
Novo Falso positivo detectado: Jogo "Horizon Zero Dawn", Tag "MOBA"  
Novo Falso positivo detectado: Jogo "Gof of War Ragnarok", Tag "MOBA"  
Novo Falso positivo detectado: Jogo "Hitman 1", Tag "MOBA"  
Novo Falso positivo detectado: Jogo "Hitman 2", Tag "Historical Fiction"  
Novo Falso positivo detectado: Jogo "Hitman 3", Tag "Historical Fiction"
```

```
Resultado do Teste Bloom Filter:  
Número total de verificações: 1175  
Número total de tags inseridas no Bloom Filter (ou seja com o valor 1): 407  
Número de Falsos positivos: 27  
Percentagem de falsos positivos: 2.30%  
Probabilidade de ocorrer falsos positivos: 2.16%  
>>
```

Resultados Obtidos:
Filtro de Bloom

Minhash – Testes

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|}$$

Fórmula da semelhança de Jaccard Real

```
>> minhash_teste
Jogo normal escolhido: Persona 5
Comparação entre Similaridade de MinHash e Jaccard Real do jogo escolhido com todos os jogos VR disponíveis:
Jogo VR: The Elder Scrolls V: Skyrim VR
Similaridade de Jaccard (MinHash): 0.52
Similaridade de Jaccard Real: 0.53846
Diferença Absoluta: 0.018462
Percentagem de Diferença: 3.4286%
-----
```

```
Ordem de recomendações de jogos VR baseadas na similaridade com o jogo escolhido pelo utilizador:
Recomendação nº1: Moss (Similaridade MinHash: 0.71)
Recomendação nº2: Dark Souls VR (Fan-Made) (Similaridade MinHash: 0.68)
Recomendação nº3: Chronos (Similaridade MinHash: 0.64)
Recomendação nº4: Westworld Awakening (Similaridade MinHash: 0.62)
Recomendação nº5: The Climb (Similaridade MinHash: 0.54)
Recomendação nº6: The Walking Dead: Saints & Sinners (Similaridade MinHash: 0.54)
Recomendação nº7: Blaston (Similaridade MinHash: 0.53)
Recomendação nº8: The Elder Scrolls V: Skyrim VR (Similaridade MinHash: 0.52)
Recomendação nº9: Blade & Sorcery (Similaridade MinHash: 0.52)
Recomendação nº10: Vader Immortal: A Star Wars VR Series (Similaridade MinHash: 0.51)
```

Resultados Obtidos:

Minhash

Artigos Relacionados

Menu: <https://www.mathworks.com/help/matlab/ref/menu.html>

Listdlg: <https://www.mathworks.com/help/matlab/ref/listdlg.html>

Matriz de Confusão: <https://www.mathworks.com/help/stats/confusionmat.html>

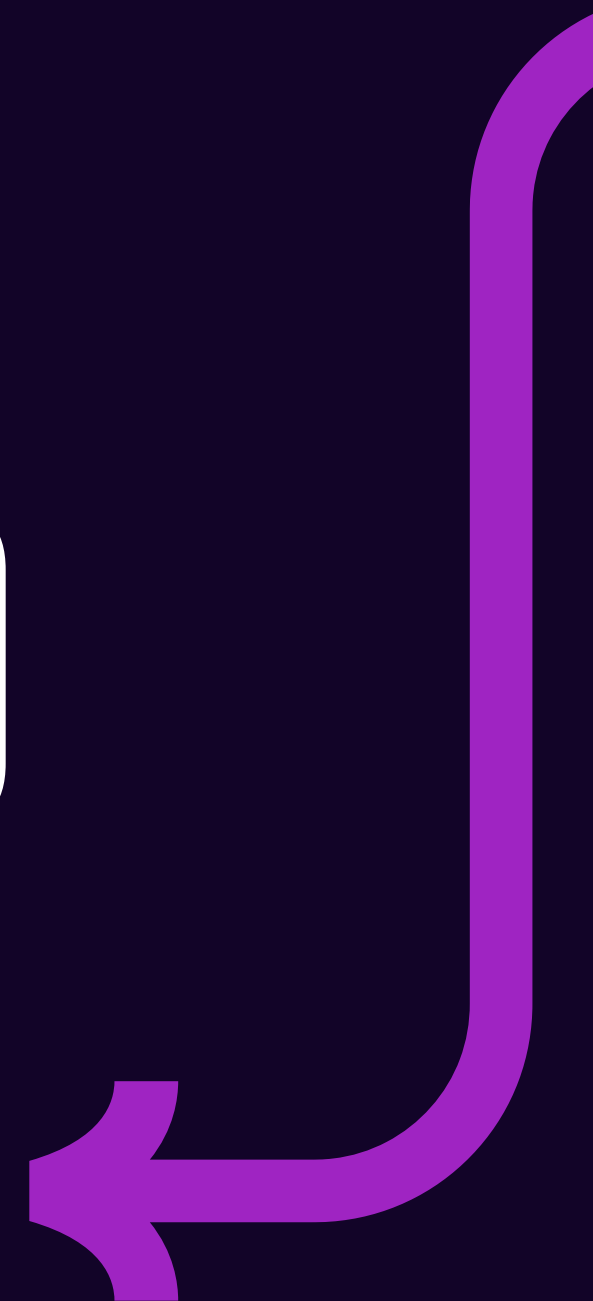
Exist: <https://www.mathworks.com/help/matlab/ref/exist.html>

Naive Bayes-I: <https://www.mathworks.com/help/stats/naive-bayes-classification.html>

Bloom Filter-II: <https://systemdesign.one/bloom-filters-explained/>

Minhash: <https://medium.com/@chnwsw01/minhash-algorithm-54381a39ad62>

DEMONSTRAÇÃO LIVE



OBRIGADO PELA ATENÇÃO!

ALGUMA QUESTÃO?

