CURSO DE SISTEMAS DE INFORMAÇÃO TÓPICOS AVANÇADOS EM SISTEMAS DE INFORMAÇÃO 2

Análise Preditiva de Tendências e Sucesso de Jogos na Steam





Agenda



- Sobre o dataset
- Dados brutos
- Pré-processamento
- Análise de tópicos (LDA)
- Agrupamento (Kmeans)
- Conclusão

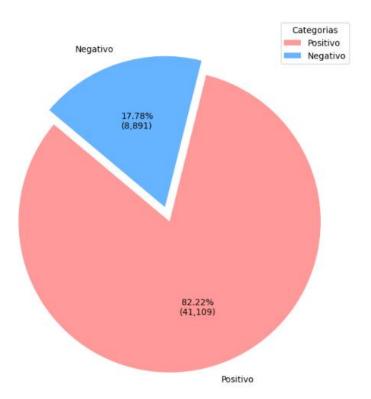
Dataset



- Conteúdo do Conjunto de Dados:
- Mais de 6,4 milhões de avaliações públicas (em inglês).
- Origem: Seção de Avaliações da Steam (operada pela Valve).
- Inclui: Texto da avaliação, ID do jogo, sentimento da avaliação (positivo ou negativo), número de usuários que acharam a avaliação útil
- Utilização do Conjunto de Dados:
- Análise de sentimentos para prever a pontuação com base no texto da avaliação
- Explorar os sentimentos e temas das avaliações entre diferentes jogos e gêneros

Distribuição dos Valores de Score







Nuvem de palavras

```
seem
                                      easytake
                              everything
                                          far
thing
hard
see
```

Nuvem de palavras para reviews positivas e negativas



```
actually
bad lovenice make back interesting good far good game though great feelgive though great free really thing buzzle run level many keep everything level many keep gameplay great game friend gameplay great game friend even graphic people steam wgot say graphic people played

Think review graphic people played

Think review graphic people steam wgot say great lot best made amazing combat much playing look well much playing look well much go multiplayer character will worth little hard fan map new player point
```

```
review right played friend boring end worth people trying run thingissue recommend part want lot work bad nothing a actually want lot makego hour to gameplay be people trying feel weapon everything feel weapon everything feel weapon free need still minute try good see got much player bug made mission money fun made mission playing map playing maybe control playing maybe c
```

"great", "amazing", "fun", "nice".

"bad", "problem", "terrible", "issue".



Pré-processamento

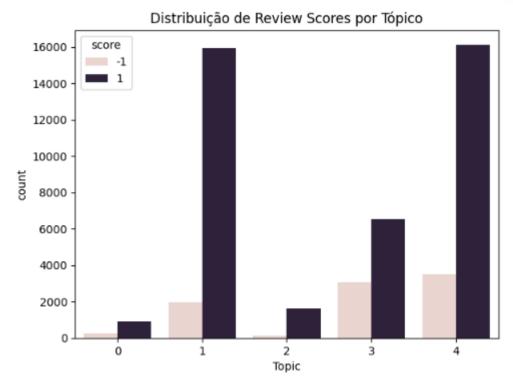
- LowerCase
- Tokenização
- Remoção de Caracteres Especiais
- Remoção de Stop Words e Pontuação
 - Stemming

Análise de tópicos com LDA

0

 Modelagem de tópicos para identificar temas comuns e os temas predominantes nas reviews.

- Tecnologia e Ficção Científica
- Narrativa e Enredo
- Jogabilidade e Experiência de Jogo
- Ação e Combate
- Diversão e Qualidade



Tópicos de destaque



- Tópico 1: Tecnologia e Ficção Científica

 Termos como "portal" e "black" sugerem discussões sobre elementos tecnológicos ou de ficção científica nos jogos. Títulos como "Portal" e "Saints Row" ou que reforçam essa temática.

Tópico 4: Ação e Combate

- Este tópico inclui palavras como "enemi", "weapon", "kill", e "get", indicando uma ênfase em ação e combate. "Counter-Strike" e "Left 4 Dead 2" se encaixam bem aqui.



Jogos relacionados a cada tópico

```
Jogos relacionados ao Tópico 1:
['Assetto Corsa' 'BattleBlock Theater' 'Arma 3' 'Saints Row IV'
 'NEKOPARA Vol. 1' 'Cry of Fear' 'Chivalry: Medieval Warfare' 'DETOUR'
 "Garry's Mod" 'Dota 2'1
Jogos relacionados ao Tópico 2:
['Sonic Lost World' 'Furi' "Five Nights at Freddy's" 'Intrusion 2'
 'The Binding of Isaac' 'Undertale' 'Bastion'
 '3089 -- Futuristic Action RPG' 'Insanity Clicker' 'The Shivah']
Jogos relacionados ao Tópico 3:
['The Elder Scrolls V: Skyrim' 'Fallout 3 - Game of the Year Edition'
 'Grand Theft Auto V' 'Super Meat Boy' 'Take On Helicopters'
 'The Bureau: XCOM Declassified' 'Viridi' 'Blood and Bacon'
 'Sleeping Dogs: Definitive Edition' 'Watch Dogs']
Jogos relacionados ao Tópico 4:
['DARK SOULS": Prepare To Die Edition' 'Poly Bridge'
 'Alien Breed 2: Assault' 'Counter-Strike Nexon: Studio' 'Dead Space 2'
 'Emily is Away' 'The Binding of Isaac' 'Divinity: Dragon Commander'
 'Stellaris' 'Middle-earth™: Shadow of Mordor™']
Jogos relacionados ao Tópico 5:
['Call of Duty: Black Ops' "Dragon's Dogma: Dark Arisen"
 'Serious Sam 3: BFE' 'PAC-MAN Championship Edition DX+' 'Riven'
 'Red Faction: Guerrilla Steam Edition' 'Terraria' 'Hacknet' 'I Am Bread'
 "Sid Meier's Civilization VI"]
```

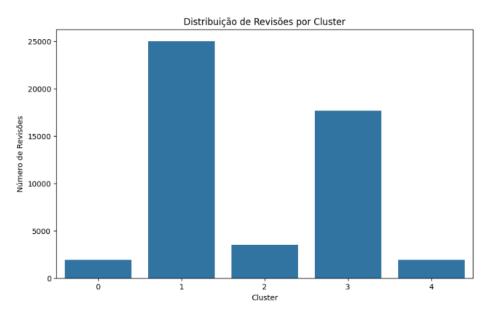


Agrupamento com KMeans



Agrupar as revisões em clusters baseados em semelhanças.

 Identificar padrões em revisões positivas e negativas, bem como temas comuns entre jogos diferentes.



```
Termos mais representativos por cluster:
Cluster 0:
best, game, play, ve, rpg, time, fp, probabl, steam, stori

Cluster 1:
game, play, like, buy, love, good, time, awesom, better, worth

Cluster 2:
fun, game, play, friend, realli, lot, good, addict, like, great

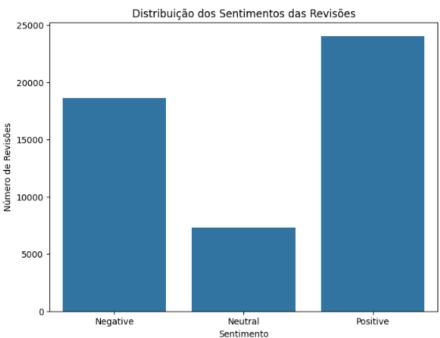
Cluster 3:
game, play, good, like, realli, time, love, stori, make, recommend

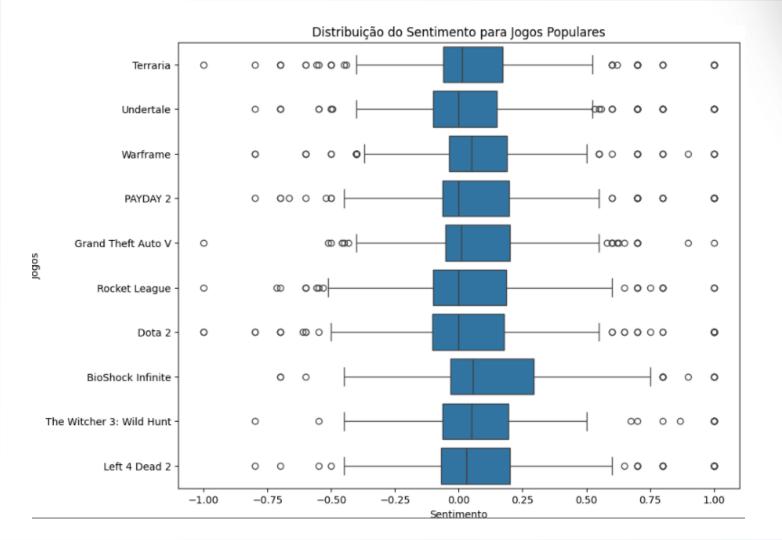
Cluster 4:
great, game, stori, play, fun, love, gameplay, graphic, realli, good
```





Para esta análise mais detalhada, examinamos o sentimento subjacente em cada revisão através do peso e contexto de palavras específicas:









Conclusões

Engajamento Comunitário: "Terraria", "PAYDAY 2" e "Dota 2" marcam alta presença em todos os clusters, indicando forte engajamento comunitário. Estes jogos são consistentemente ligados a termos positivos como "fun", "great" e "good".

Insights para Desenvolvedores: Jogos com temas de shooter, ficção científica ou futuristas frequentemente recebem reviews positivas, apontando para uma recepção entusiástica desses gêneros.

Obrigado!