

ANALISE DE DADOS

Aplicada a uma Plataforma de Notas de Obras Japonesas

Ciência dos Dados | Algoritmos e Estrutura de Dados

Docentes:

Higor Alexandre Duarte Mascarenhas
Michel Pires da Silva

Discente:

Arthur Santana de Mesquita



SUMÁRIO

1 Sobre o Projeto

1.1 Objetivo

1.2 Importância

2 Extração

2.1 Mangás

2.2 Usuários

2.3 Animes

3 Aplicabilidade

3.1 Vantagens e Desvantagens

3.2 Exemplos de aplicações

4 Análise Comportamental

4.1 Ratings x Timestamp

4.2 Lista Estática

4.3 Lista Dinâmica

4.4 Pilha

4.5 Fila

5 Conclusão

1 SOBRE O PROJETO

1.1 Objetivo

- Objetivo: Analisar os padrões de uso do público consumidor de obras de animação japonesa (animes) de forma a criar agrupamentos que auxiliem na recomendação de obras para adaptação para esta mídia
- Para atingir tal objetivo, foram extraídos, tratados e analisados dados da plataforma de notas de obras japonesas MyAnimeList.

1 SOBRE O PROJETO

1.2 Importância

Como o streaming pode se beneficiar

- Sistemas de streaming que investirem cedo neste mercado podem ter maior controle sobre propriedade intelectual existente.
- Aquisição e criação de estúdios de animação japoneses para criação de obras que atendam às demandas do mercado.
- Muitos shows criados em cada temporada, poucos bem sucedidos.
- O conhecimento de quais gêneros e elementos são mais populares dentro de cada demografia pode prover uma vantagem.

2 EXTRAÇÃO

2.1 Sobre o MyAnimeList

- Notas de 1-10 para cada obra (26K animes 68K mangás)
- Gênero, Tipo de Mídia e Fonte Original das Obras
- Apesar de limitado em termos de requisições, pode ser extraído facilmente.

Organise, Discuss, Discover

MyAnimeList

HOW MUCH ANIME HAVE YOU SEEN?

2 EXTRACÃO

2.1 Mangá e Literatura

All Manga Rankings are updated twice a day. How do we rank shows?					Next 50 >
Rank	Title	Score	Your Score	Status	
1	Berserk Manga (? vols) Aug 1989 - 776,542 members	★ 9.47	★ N/A	Add to My List	
2	JoJo no Kimyou na Bouken Part 7: Steel Ball Run Manga (24 vols) Jan 2004 - Apr 2011 307,910 members  MANGA MIRAI Vol. 1 \$16.99 > Read Sample	★ 9.33	★ N/A	Add to My List	
3	Vagabond Manga (37 vols) Sep 1998 - May 2015 443,823 members	★ 9.27	★ N/A	Add to My List	
4	One Piece Manga (? vols) Jul 1997 - 682,443 members  MANGA MIRAI Vol. 1 \$6.99 > Read Sample	★ 9.22	★ N/A	Add to My List	
5	Monster Manga (18 vols) Dec 1994 - Dec 2001 281,523 members	★ 9.16	★ N/A	Add to My List	
6	Guimi Zhi Zhu Novel (18 vols) May 2020 - Aug 2023 20,033 members	★ 9.15	★ N/A	Add to My List	

- Ignorou-se obras com nota inferior a 7.0
- Ignorou-se obras que já tenham recebido adaptação para anime.
- O próprio sistema do MyAnimeList já elimina obras +18 do sistema de pesquisa.
- Ainda com estas limitações, a base de dados final ainda totalizou 4060 mangás

2 EXTRACÃO

2.2 Usuários

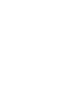
Advanced Search

Location: Brazil
Ex: California

Age: > 0 and < 0
Ex: > 18 and < 24

Gender: Don't care

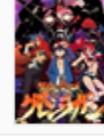
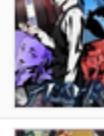
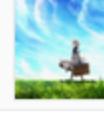
User Search Results

danieros	carlahime28br	bodim2	mikeneko	Lyzette	[1] 2 3 ... 20 Lawland
					
Aug 18, 2018 11:06 AM	Feb 19, 2021 12:23 PM	Sep 24, 2011 10:36 AM	Sep 14, 2022 4:33 AM	Jan 25, 2012 6:31 PM	Nov 27, 10:58 AM
Mamo-kun	thalitamassae	ichigofujiwara	MissYaoiBL	_blake_	biotmessiah
					
Jan 14, 4:45 AM	Dec 27, 2021 4:54 AM	Jul 30, 2020 7:16 AM	Sep 4, 6:17 AM	Apr 29, 2024 12:23 PM	Sep 21, 2023 6:50 PM
Haddad	alumina	narusenpai	PancaLoca	neomauii	Schenker
					
Oct 17, 2015 9:25 AM	Nov 7, 2016 7:49 AM	Dec 6, 12:29 PM	Dec 7, 11:09 AM	Jul 5, 2008 6:44 PM	Nov 16, 2007 5:38 PM
Pop-chi	GaranhaoItaliano	Nagash	OverkilledRed	hunterbr	boris_kato
					
Apr 24, 2023 11:20 AM	Jul 10, 2008 10:09 PM	Aug 27, 2010 11:08 AM	Nov 12, 4:47 PM	Nov 2, 5:32 PM	Dec 27, 2009 4:34 PM

[1] 2 3 ... 20

- Para a criação de páginas de busca de usuários necessita-se de inserir parâmetros de filtro: **gênero, localização ou nome**.
- Optou-se pelo parâmetro localização = Brazil.
- Para evitar contas inativas há muito tempo, verifica-se se houve acesso desde de 2017.

2 EXTRACÃO

ALL ANIME						 Stats	 Filters
#	Image	Anime Title	Score	Type	Progress	Tags	
1		Princess Tutu	Add - More	10	TV	38	
2		Gurren Lagann	Add - More	10	TV	27	
3		Clannad: After Story	Add - More	10	TV	24	
4		Puella Magi Madoka Magica	Add - More	10	TV	12	
5		Into the Forest of Fireflies' Light	Add - More	10	Movie	1	
6		Death Parade	Add - More	10	TV	12	
7		Mob Psycho 100	Add - More	10	TV	12	
8		Violet Evergarden	Add - More	10	TV	13	

2.3 Animes

- Os animes foram extraídos conforme apareciam nas listas dos usuários, formando um cachê de 5430 animes.
- Animes da lista com nota de usuário inferior a 7.0 também foram excluídos da coleta.
- Animes já existentes no cachê tem sua informação verificada localmente, evitando requisições e acelerando o processo.

3 MANIPULAÇÃO

3.1 Perfis de Consumo

CALCULO EFETUADO

- Afinidade com cada gênero e mídia baseado nas notas dos animes
- Valores normalizados por proporção
- Pesos que permitam maior aparecimento de Light Novels

TRATAMENTO DE DADOS

- Remoção de usuários com menos de 5 animes válidos.
- União dos conjuntos Manhwa e Manhua
- Exclusão de gêneros pouco diferenciáveis (“Award Winning”)

3 MANIPULAÇÃO

3.2 Geração de comunidades e grafo

PERDA AMOSTRAL

- Grande número de usuários perdido no passo anterior, 281 restantes.

TFIDF

- Aplicado para aumentar a influência de gêneros **pouco comuns** e diminuir a relevância de *Ação*, *Aventura* e *Comédia*

BANCO DE DADOS

- Aumento do peso das fontes na geração das arestas para garantir mídias além de mangá na recomendação.

NETWORKX

- Layout *spring* na representação gráfica.
Similaridade de cosseno

NOMES

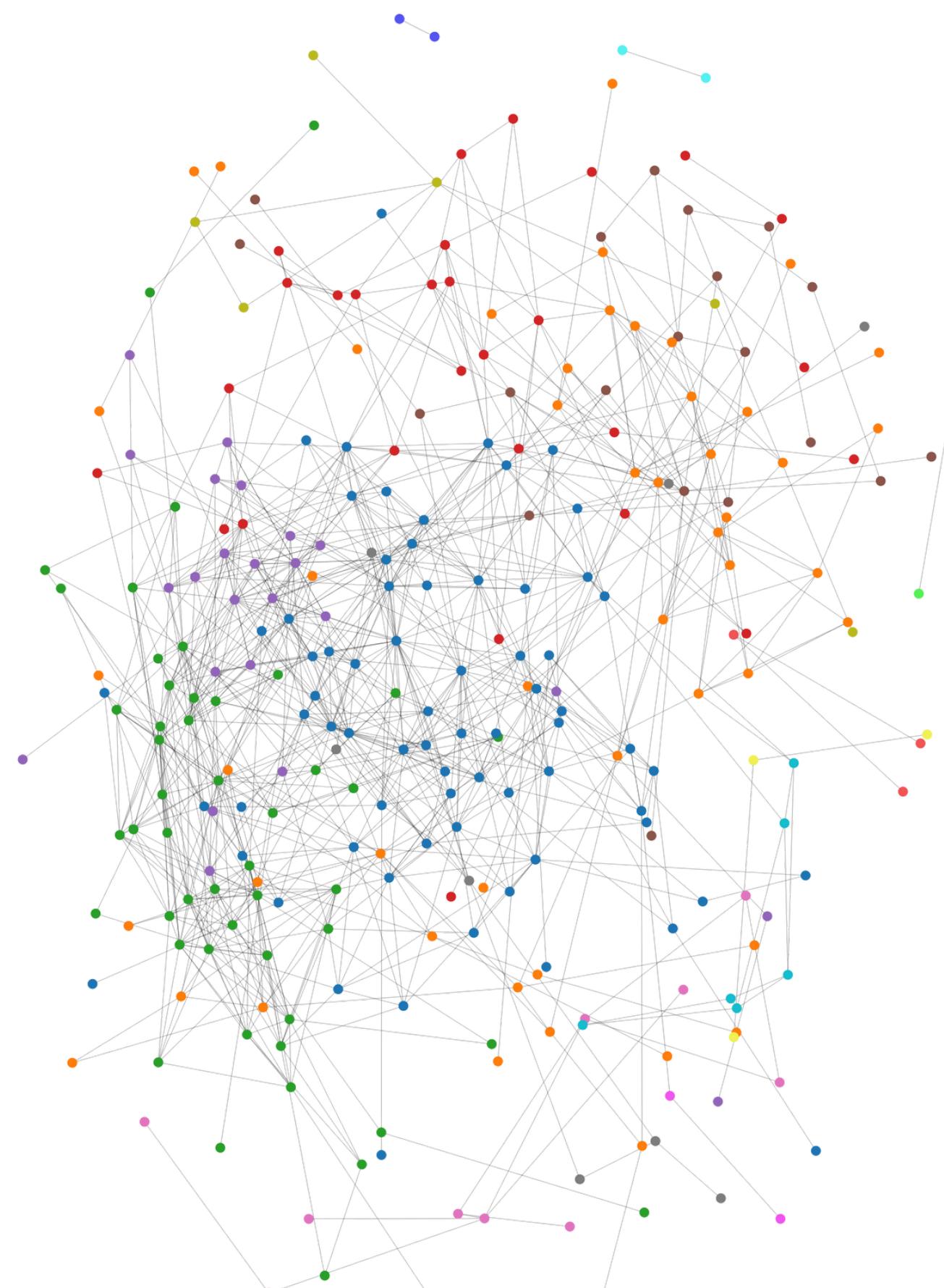
- Gerados com base nos 2 gêneros principais, com um terceiro sendo adicionado conforme necessário.

3 RESULTADOS

3.3 Grafo gerado:

Comunidades (gêneros dominantes)	
Action / Fantasy	70 usuários
Action / Adventure	51 usuários
Romance / Comedy	50 usuários
Romance / Comedy / Supernatural	28 usuários
Comedy / Romance	24 usuários
Action / Adventure / Fantasy	20 usuários
Drama / Sci-Fi	10 usuários
Romance / Drama	8 usuários
Comedy / Sci-Fi	6 usuários
Action / Supernatural	6 usuários
Suspense / Drama	3 usuários
Slice of Life / Sports	3 usuários
Drama / Suspense	2 usuários
Romance / Supernatural	2 usuários
Adventure / Sci-Fi	2 usuários
Action / Adventure / Suspense	2 usuários

GRAFO DE SIMILARIDADE:



3 RESULTADOS

3.3 Grafo gerado:

Comunidades (gêneros dominantes)	
Action / Fantasy	70 usuários
Action / Adventure	51 usuários
Romance / Comedy	50 usuários
Romance / Comedy / Supernatural	28 usuários
Comedy / Romance	24 usuários
Action / Adventure / Fantasy	20 usuários
Drama / Sci-Fi	10 usuários
Romance / Drama	8 usuários
Comedy / Sci-Fi	6 usuários
Action / Supernatural	6 usuários
Suspense / Drama	3 usuários
Slice of Life / Sports	3 usuários
Drama / Suspense	2 usuários
Romance / Supernatural	2 usuários
Adventure / Sci-Fi	2 usuários
Action / Adventure / Suspense	2 usuários

GRAFO SOBREPOSTO POR ZONAS DE INFLUÊNCIA:



3 RESULTADOS

3.4 Recomendações Geradas

==== RESULTADO DA CLUSTERIZAÇÃO E RECOMENDAÇÃO ===

Total de comunidades: 16

[COMUNIDADE 1] (70 usuários)

Foco Principal: Origem - Light_Novel. Gêneros - Action, Fantasy

--- Top 5 Mangás Não Adaptados Recomendados ---

[0.7878] Yu☆Gi☆Oh!

- Score MAL: 7.33 | Tipo: Light Novel | Gêneros: Action, Adventure, Comedy, Drama, Fantasy, Mystery, Supernatural

[0.7782] Code Name wa Sailor V (Codename: Sailor V)

- Score MAL: 7.50 | Tipo: Manga | Gêneros: Action, Adventure, Comedy, Drama, Fantasy, Romance

[0.7782] Watashi no Messiah-sama

- Score MAL: 7.62 | Tipo: Manga | Gêneros: Action, Adventure, Comedy, Drama, Fantasy, Romance, Ecchi

[0.7727] One Piece Novel: A (One Piece Novel: Ace's Story)

- Score MAL: 7.93 | Tipo: Light Novel | Gêneros: Action, Adventure, Comedy, Fantasy

[0.7727] Fairy Tail

- Score MAL: 7.54 | Tipo: Light Novel | Gêneros: Action, Adventure, Comedy, Fantasy

[COMUNIDADE 2] (51 usuários)

4 CONCLUSÃO

- O algoritmo desenvolvido conseguiu extrair e tratar os dados da plataforma MyAnimeList
- As recomendações finais também podem ser consideradas válidas para o conjunto de fontes e gêneros
- Reitera-se a necessidade da criação de novos mecanismos que permitam a identificação de estratégias comerciais para a criação de entretenimento.



R E F E R Ê N C I A S

- Python Software Foundation. Python Documentation.
[https://docs.python.org/3/.](https://docs.python.org/3/)
- CORMEN, T. H. Algoritmos - Teoria e Pratica. MITPress, 1989.
- Leonard Richardson. Beautiful soup documentation, 2024. URL
[https://www.crummy.com/software/BeautifulSoup/bs4/doc/.](https://www.crummy.com/software/BeautifulSoup/bs4/doc/)
Accessed: 2025-12-08.
- NetworkX Developers. Networkx documentation, 2025. URL
[https://networkx.org/documentation/stable/.](https://networkx.org/documentation/stable/)
Accessed: 2025-12-08.

Obrigado !!!

Dúvidas ?