

Contexto

Correlatos

Proposta

Referências

Término



MÉTODOS DE RECOMENDAÇÃO DE FILMES na NETFLIX



Começar Filme! Ou
melhor, apresentação..



I TÓPICOS

01

CONTEXTO

Recomendação de
Filmes e Netflix
Prize

02

TRABALHOS CORRELATOS

Trabalhos utilizados como base
para a orientação do trabalho

03

O QUE SERÁ FEITO?

Objetivos do
trabalho e os seus
processos

04

REFERÊNCIAS BIBLIOGRÁFICAS

Alguns dos materiais
utilizados





01

CONTEXTO

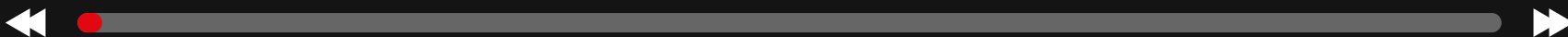
Método de Recomendação de Filmes e Netflix Prize

- Abordagem segundo Xavier Amatriain, chefe científico na Netflix e membro do setor de desenvolvimento dos sistemas de recomendação





Pense na seguinte situação
DEPOIS DE UM LONGO DIA...



Contexto

Correlatos

Proposta

Referências

Término



Contexto

Correlatos

Proposta

Referências

Término



Contexto

Correlatos

Proposta

Referências

Término



Contexto

Correlatos

Proposta

Referências

Término



Contexto

Correlatos

Proposta

Referências

Término



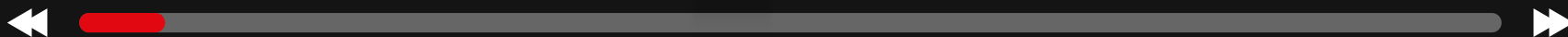
Contexto

Correlatos

Proposta

Referências

Término



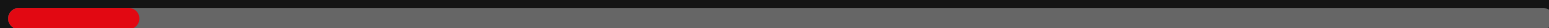
Contexto

Correlatos

Proposta

Referências

Término



Contexto

Correlatos

Proposta

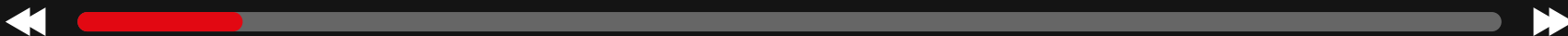
Referências

Término



COMO?

Como a Netflix faz isso?





Diversas Técnicas de Análise de Dados





Diversas Técnicas de Análise de Dados

- 01 Dados sobre o que assiste:
- Gênero de Filme
 - Data
 - Tempo Assistido

- 02 Como você navega:
- Termos na barra de pesquisa
 - Locais da tela que mais clica
 - Tempo procurando filme

- 03 Quem você é:
- Local
 - Dados Demográficos
 - Email



Home

About me



Diversas

HOUSE of CARDS

de Dados

01 Dados
assiste:

- Go
- Da
- Te

0



no você navega:

Termos na barra de pesquisa
Locais da tela que mais clica
Tempo procurando filme





I LINHA DO TEMPO (DA NETFLIX)

1997

ALUGUEL DE DVD

Oferecia serviço de aluguel de DVD's por correio, sem sair de casa

2006

CINEMATCH

Projetado para realizar análises de preferências e recomendar filmes

2006 - 2009

NETFLIX PRIZE

Desafio proposto pela Netflix, com o objetivo de melhorar o algoritmo de recomendação do Cinematch

2023

HOJE EM DIA

Potência no mundo do streaming, com 232,5 milhões de assinantes

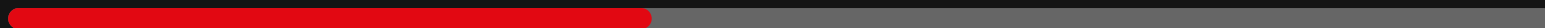




TRABALHOS CORRELATOS

02

Mecanismos de busca, seleção
e métodos empregados



Contexto

Correlatos

Proposta

Referências

Término



RECOMMENDER SYSTEM

TECHNIQUES APPLIED TO NETFLIX

MOVIE DATA





I ORIGEM DOS DADOS



NETFLIX PRIZE

Conjunto disponibilizado pela Netflix para os participantes



IMDb

Conjunto de dados público de filmes internacionais



Contexto

Correlatos

Proposta

Referências

Término



I CONJUNTO DE DADOS INICIAL

17.7K

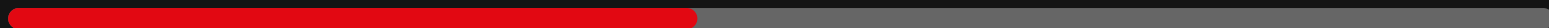
FILMES

480K

USUÁRIOS

100M

AVALIAÇÕES



Contexto

Correlatos

Proposta

Referências

Término



I CONJUNTO DE DADOS FINAL

3.5K

FILMES

236K

USUÁRIOS

34.8M

AVALIAÇÕES





I PRÉ-PROCESSAMENTO DOS DADOS



VETORIZAÇÃO

Vetorização do conjunto de dados não estruturados



MEDIDAS DE CORRELAÇÃO

Correlação de Pearson ou Similaridade por Cosseno





| ABORDAGENS



COLABORATIVO

Recomendação com base no comportamento de usuários semelhantes



BASEADO EM CONTEÚDO

Recomendação com base nas características dos itens



HIBRIDO

Combinação do colaborativo e baseado em conteúdo para sugestões mais precisas e personalizadas





I COLABORATIVO DE USUÁRIOS



JURASSIC PARK

USUÁRIO 1



JURASSIC PARK

USUÁRIO 2



JURASSIC WORLD

USUÁRIO 1



JURASSIC WORLD

USUÁRIO 2

?????





I BASEADO EM CONTEÚDO



DIRETOR

Steven Spielberg

FILME 1

ANO DE LANÇAMENTO

1993

GÊNEROS

Ação, Aventura e Ficção Científica

ATORES

Sam Neill, Laura Dern e Jeff Goldblum

FILME 2

ANO DE LANÇAMENTO

2015

GÊNEROS

Ação, Aventura e Ficção Científica

ATORES

Chris Pratt, Bryce Dallas e Irrfan Khan



DIRETOR

Colin Trevorrow





I COMBINAÇÃO DE ABORDAGENS



MODELO A

Colaborativo
baseado em
usuário



MODELO B

Baseado em
conteúdo





I RESULTADOS FINAIS



	MODELO A	MODELO B	COMBINAÇÃO
1	Jurassic Park	Homem-Aranha	Jurassic Park
2	Vingadores	Jurassic Park	Homem-Aranha
3	Homem-Aranha	Vingadores	Vingadores
4	Carros	Carros	Carros



Contexto

Correlatos

Proposta

Referências

Término



03

O QUE FAREMOS?

Qual a nossa proposta?



EXPLORANDO ESTRATÉGIAS DE RECOMENDAÇÃO

**Uma Análise de Redes Sociais na Reprodução de
Experimentos para o Desenvolvimento de Sistemas de
Recomendação de Filmes na Netflix**



04

REFERENCIAS BIBLIOGRÁFICAS

Material principal e
auxiliares





I REFERÊNCIAS BIBLIOGRÁFICAS

- <<https://vu-business-analytics.github.io/internship-office/papers/paper-postmus.pdf>>
- <https://www.asc.ohio-state.edu/statistics/statgen/journal_aut2009/BigChaos.pdf>
- <<https://www.maxwell.vrac.puc-rio.br/57546/57546.PDF>>
- <<https://xamat.medium.com/on-the-usefulness-of-the-netflix-prize-403d360aaf2>>



Contexto

Correlatos

Proposta

Referências

Término



OBRIGADO!

PELA ATENÇÃO





I INTEGRANTES

- Bruno Friedrich Raquel 13672450
- Ezequiel Park 5172519
- Felipe Mateos Castro de Souza 11796909
- Rodrigo Dorneles Ferreira de Souza 11295831



Introdução

Alternativas

Resultados

Referências

Término



Sobre os Dados para Sistemas de Recomendações na NETFLIX



Começar Filme! Ou
melhor, apresentação..



| TÓPICOS

01

DESAFIOS

O tamanho

02

ALTERNATIVAS

Datasets que cabem no bolso

03

RESULTADOS

Alguns resultados

04

REFERÊNCIAS BIBLIOGRÁFICAS

Alguns dos materiais
utilizados





I O DESAFIO



O TAMANHO DO NETFLIX PRIZE

100 milhões de linhas
4 gb de dados



Desvantagens

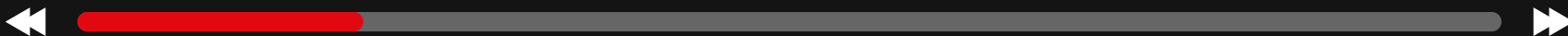
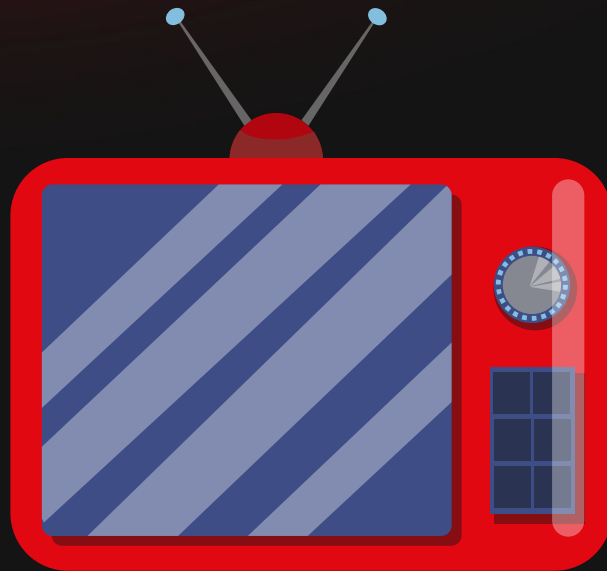
Atributos inconsistentes
Impraticável cruzar com outros Bancos de
Dados abertos ao público





Alternativas Movie Lens

Movie Lens é uma base de dados colaborativa e aberta.





| Vantagens



Variedade

Tamanho para todos os gostos - possui datasets que variam entre 100k e 1B



IDs

Fornece os IDs dos filmes em outras bases de dados facilitando o cruzamento de dados



Atributos

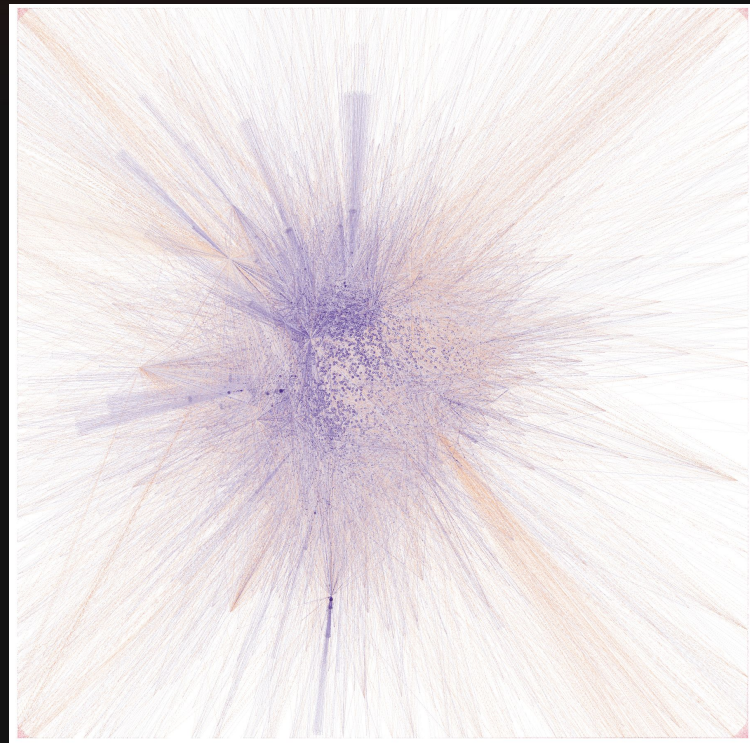
Possui maior variedade de atributos a serem utilizados





| Resultados Preliminares

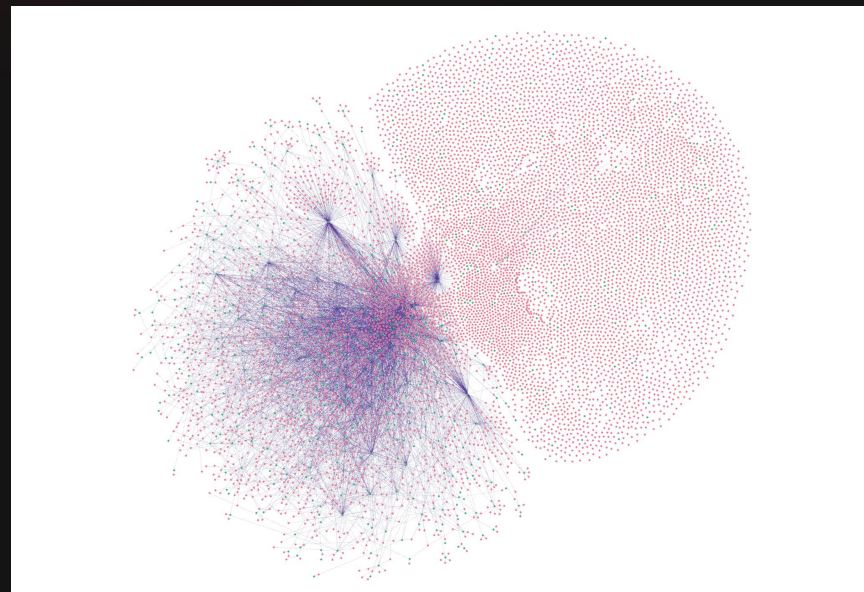
- Grafo Bipartido
- Arestas representam a nota dado pelo usuário
- Aresta Roxas/Violetas são notas mais altas
- Arestas Vermelhas/Laranjas são notas mais baixa
- Possível notar alguns focos isolados de interesse (conjunto de feixes)





I Resultados Preliminares

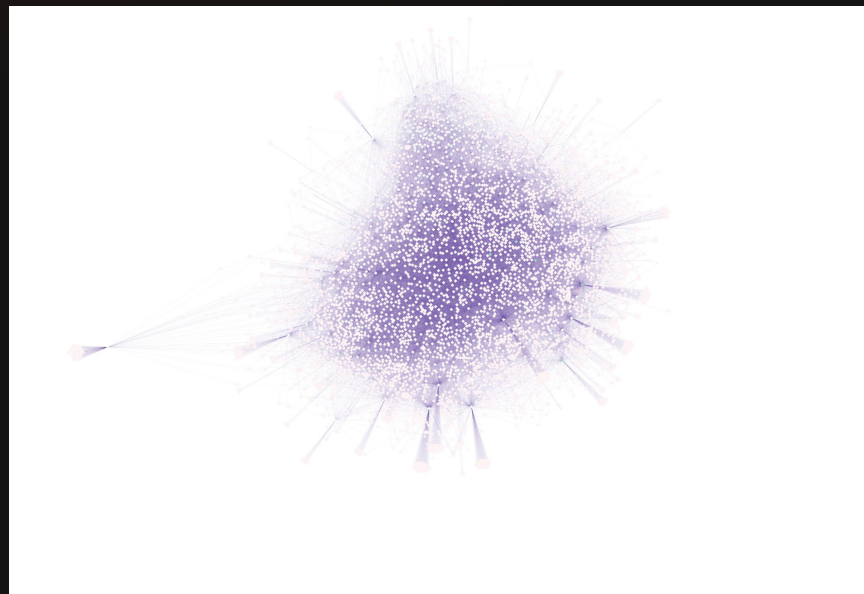
- Grafo Bipartido
- Arestas representa notas baixas (entre 0,5 e 1,5)
- É possível visualizar filmes pouco desvalorizados e alguns usuários que não costumam dar nota baixa (melhor seleção do filme ou gostam de quase tudo)
- Concentração de alguns filmes muito odiados e alguns usuários que costumam dar notas baixas.





| Resultados Preliminares

- Grafo Bipartido
- Arestas representa notas altas (entre 4 e 5)
- Possível ver vários nichos e possivelmente alguns filmes de sucesso.





I REFERÊNCIAS BIBLIOGRÁFICAS

- Movie Lens:
<https://grouplens.org/datasets/movielens/>
- Netflix Prize:
<https://www.kaggle.com/datasets/netflix-inc/netflix-prize-data>
- Gephi: <https://gephi.org/>





I INTEGRANTES

- Bruno Friedrich Raquel 13672450
- Ezequiel Park 5172519
- Felipe Mateos Castro de Souza 11796909
- Rodrigo Dorneles Ferreira de Souza 11295831

