

Universidade Presbiteriana Mackenzie

Projeto Aplicado III - Recomendação de Livros

CESAR VALENTIM SILVA - 10416087
JOAO PEDRO SANTOS OLIVEIRA - 10423752
WILLIANS CARVALHO DA SILVA – 10416087

São Paulo, 2025

Sumário

1. Introdução
 - 1.1 Contextualização
 - 1.2 Motivação
 - 1.3 Justificativa
 - 1.4 Objetivo geral
 - 1.5 Objetivos específicos
2. Referencial Teórico
 - 2.1 Referências compatíveis conforme normas ABNT
3. Metodologia
 - 3.1 Coleta e Pré-processamento dos Dados
 - 3.2 Técnicas e Algoritmos de Recomendação Utilizados
 - 3.3. Avaliação do Modelo
4. Resultados
5. Conclusão

1 INTRODUÇÃO

Os sistemas de recomendação desempenham um papel fundamental na personalização da experiência do usuário em diversas plataformas digitais. No setor literário, plataformas como Amazon, Google Books e Goodreads utilizam algoritmos avançados para sugerir livros aos leitores com base em preferências e interações passadas. A crescente quantidade de títulos disponíveis torna essencial o uso de técnicas de inteligência artificial para otimizar o processo de recomendação e facilitar a descoberta de novos conteúdos.

1.1 Contexto do trabalho

O presente trabalho propõe o desenvolvimento de um sistema de recomendação de livros com base em dados reais de avaliações de usuários. Utilizando um conjunto de dados extraído da plataforma Kaggle, o projeto tem como objetivo principal identificar padrões de preferências e oferecer recomendações personalizadas com base no histórico de avaliações.

1.2 Motivação

A baixa taxa de leitura no Brasil e em diversos países em desenvolvimento é um fator preocupante. Estudos apontam que muitos leitores deixam de consumir livros devido à dificuldade em encontrar obras alinhadas aos seus interesses. Além disso, estudantes que se preparam para vestibulares e concursos enfrentam desafios para selecionar materiais adequados à sua jornada acadêmica. Um sistema de recomendação eficiente pode contribuir significativamente para incentivar a leitura e melhorar a qualidade da preparação educacional.

1.3 Justificativa que demonstre a importância do projeto.

O desenvolvimento de um sistema de recomendação de livros fundamentado em Big Data e Machine Learning visa criar uma ferramenta capaz de fornecer sugestões personalizadas e de alta relevância para cada usuário. Diferente dos métodos tradicionais de curadoria, o uso de filtragem colaborativa, análise de conteúdo e redes neurais permite recomendações mais precisas, baseadas não apenas no histórico de leitura, mas também em avaliações de outros leitores e padrões de consumo.

1.4 Objetivo Geral

- *O principal objetivo deste estudo é analisar como os sistemas de recomendação de livros podem ser aplicados no contexto educacional, auxiliando na personalização do aprendizado e no cumprimento do ODS 4, garantindo uma educação justa, abrangente e de qualidade para todos*

1.5 Objetivos Específicos:

- *Investigar as abordagens principais utilizadas em plataformas de recomendação (como filtragem colaborativa, baseada em conteúdo e híbrida) e suas aplicações no contexto educacional.*
- *Analisar o impacto dos sistemas de recomendação na promoção da excelência educacional, com foco na recomendação de livros digitais, e como essa abordagem pode contribuir para reduzir as desigualdades no ensino.*
- *Avaliar a viabilidade da implementação de sistemas de recomendação personalizados para alunos com diferentes perfis e estágios de aprendizado, considerando variáveis como custo, acessibilidade e infraestrutura tecnológica.*
- *Propor alternativas para adaptar os sistemas de recomendação às especificidades do cenário educacional brasileiro, levando em consideração aspectos culturais, econômicos e sociais, com o objetivo de promover um ensino mais inclusivo e equitativo.*

2 Referencial teórico compatível com o problema e objetivo do projeto

Os sistemas de recomendação (SR) são ferramentas computacionais que auxiliam na filtragem e seleção de itens relevantes para um usuário, com base em suas preferências e comportamentos (AGGARWAL, 2016). Esses sistemas são amplamente utilizados em plataformas de comércio eletrônico (como Amazon e Netflix), redes sociais e, mais recentemente, na educação (RICCI et al., 2015).

No contexto educacional, os SRs desempenham um papel crucial na personalização do aprendizado, sugerindo materiais didáticos, livros e atividades que se alinhem ao perfil e às necessidades individuais dos alunos (GÓMEZ-URIBE; HUNT, 2016). Essa abordagem está alinhada com o Objetivo de

Desenvolvimento Sustentável (ODS) 4 da ONU, que visa garantir uma educação inclusiva, equitativa e de qualidade para todos.

2.1. Referências compatíveis com o tema estudado e citadas conforme normas ABNT

- AGGARWAL, Charu C. *Recommender Systems: The Textbook*. Springer, 2016.
- *Este livro oferece uma introdução detalhada aos sistemas de recomendação, abordando desde os fundamentos e algoritmos clássicos até as técnicas mais avançadas aplicadas em diversos domínios, como comércio eletrônico, redes sociais e educação.*
- GÓMEZ-URIBE, Carlos A.; HUNT, Neil. "The Netflix Recommender System: Algorithms, Business Value, and Innovation". *ACM Transactions on Management Information Systems*, 2016.
- *Este artigo examina o sistema de recomendação da Netflix, detalhando os algoritmos usados para personalizar o conteúdo para os usuários e como esses algoritmos agregam valor comercial à empresa.*
- KAGGLE. *Amazon Books Reviews Dataset*. Disponível em: <https://www.kaggle.com/datasets/mohamedbakhet/amazon-books-reviews>. Acesso em: 27 mar. 2025.
- *Este dataset contém resenhas de livros da Amazon, oferecendo uma fonte rica de dados para a análise de sentimentos, comportamento do consumidor e desenvolvimento de sistemas de recomendação de livros.*
- MCKINSEY & COMPANY. "How Amazon's Recommendation Algorithm Drives 35% of Sales", 2018.
- *Este estudo analisa o impacto do algoritmo de recomendação da Amazon, destacando como ele contribui para 35% das vendas totais da empresa e a importância da personalização na experiência de compra.*
- RICCI, Francesco; ROKACH, Lior; SHAPIRA, Bracha. *Recommender Systems Handbook*. Springer, 2015.
- *Este manual é uma referência abrangente sobre sistemas de recomendação, explorando diversas técnicas como filtragem colaborativa, análise de conteúdo e algoritmos híbridos, além de discutir desafios práticos e questões éticas relacionadas.*

3 Metodologia

3.1. Coleta e Pré-processamento dos Dados

Na presente atividade, a lógica de recomendação foi construída a partir da biblioteca *pandas*, utilizando agregações simples para prever a preferência do usuário com base na média das avaliações históricas por livro. Para melhorar a personalização, foi incluído um ajuste com base na média de avaliações do próprio usuário. A avaliação da performance foi feita com a biblioteca *sklearn.metrics*, com cálculo de acurácia e matriz de confusão, e a visualização dos resultados foi feita com *matplotlib*

A base de dados foi composta por dois arquivos principais:

- *Books_rating.csv*: contendo avaliações dos usuários sobre os livros, incluindo colunas como *User_id*, *Title*, *review/score* e *Price*.
- *books_data.csv*: contendo informações bibliográficas dos livros, como *Title*, *authors* e *categories*.

O processo de pré-processamento incluiu:

- Remoção de dados ausentes (nulos) e duplicados.
- Padronização dos títulos.
- Extração da primeira categoria de cada livro para fins analíticos.
- Conversão das notas para formato binário (recomendo ou não recomendo), considerando:
 - Recomendo (nota ≥ 4)
 - Não recomendo (nota ≤ 3)

3.2. Técnicas e Algoritmos de Recomendação Utilizados

Foram implementadas duas abordagens distintas:

A. Filtragem Baseada em Conteúdo Simples (Baseline)

- Técnica: Previsão com base na média histórica de avaliações por título.
- Implementação: Para cada livro no conjunto de teste, foi atribuída a média das notas que ele recebeu no conjunto de treino.
- Objetivo: Servir como baseline para comparação futura com métodos mais sofisticados.

B. Ajuste Personalizado por Usuário

- Técnica: A média do livro foi ajustada pela tendência do usuário (se ele costuma dar notas mais altas ou mais baixas).
- Implementação: Foi aplicado um desvio da média global com base na média

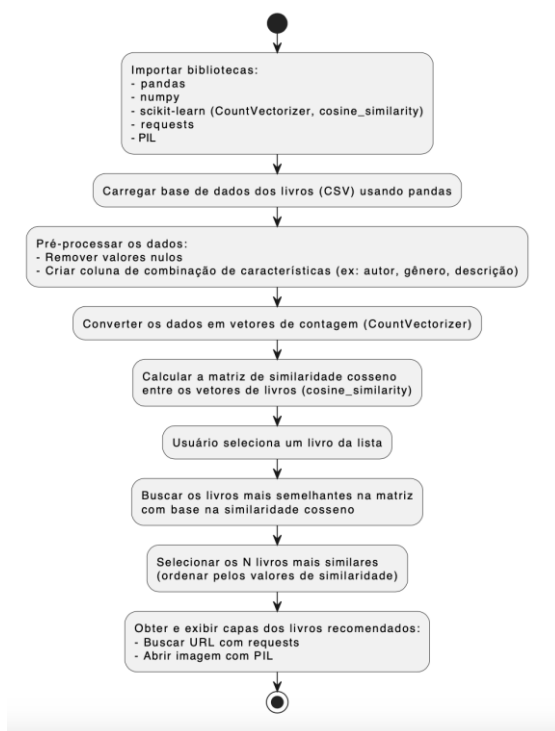
- **Objetivo:** Aumentar a personalização da recomendação.

3.3. Avaliação do Modelo

A avaliação do desempenho do sistema foi feita com base nas seguintes etapas:

- Divisão dos dados em treino (80%) e teste (20%), com amostra de 5.000 registros na fase de desenvolvimento e toda a base na fase final.
- Conversão para classificação binária (Recomendo / Não recomendo).
- Cálculo da acurácia e visualização por meio de matriz de confusão.
- Repetição do experimento com validação cruzada para obter média e desvio padrão da acurácia.

4 Fluxo da Análise



5 Resultados

A abordagem metodológica empregada permitiu a construção de um sistema de recomendação funcional que alcançou boa acurácia binária, superior a 75%. Embora técnicas mais sofisticadas - como Singular Value Decomposition (SVD), algoritmos K-Nearest Neighbors (KNN) e redes neurais - possam proporcionar maior precisão e níveis mais elevados de personalização, a simplicidade e a alta

interpretabilidade inerentes ao modelo baseado em médias mostraram-se adequadas para os propósitos didáticos e demonstrativos deste trabalho.

6 Conclusão

Futuramente, o sistema pode ser aprimorado com:

- *Técnicas de filtragem colaborativa avançada.*
- *Integração com dados textuais (descrição do livro, resenhas).*
- *Interface interativa com Streamlit ou Gradio.*

[Link para o Repositório no Github](#)

[Link para o Python – Jupyter \(Google Colab\)](#)