

Análise de Sentimentos em Avaliações de Livros Utilizando a API Gemini para Recomendação Personalizada

Fagner Silva¹, Sebastiao Alves Filho¹, Raul Paradedda¹, Lenardo Silva¹

¹Programa de Pós-Graduação em Ciência da Computação
Universidade do Estado do Rio Grande do Norte (UERN) e
Universidade Federal Rural do Semi-Árido (UFERSA)
v. Francisco Mota, s/n – Bairro Costa e Silva, Mossoró/RN, Brasil

{fagnersilva,sebastiaoalves,raulparadedda}@uern.br, lenardo@ufersa.edu.br

Abstract. *This study investigates the use of the Gemini API for sentiment analysis in book reviews, focusing on genre-categorized evaluations. The methodology involved extracting predominant emotions for use in personalized recommendation systems. Results showed effectiveness in detecting positive sentiments, especially in fiction. Emotion-based recommendations closely matched user preferences which demonstrates how sentiment analysis can enhance personalization in book recommendation systems. Challenges in less popular genres were also identified, highlighting the need for more adaptive approaches.*

Resumo. *Este trabalho investiga o uso da API Gemini na análise de sentimentos em avaliações de livros, com foco em resenhas categorizadas por gênero. A metodologia envolveu a extração de emoções predominantes para aplicação em sistemas de recomendação personalizados. Os resultados mostraram eficácia na detecção de sentimentos positivos, especialmente em obras de ficção. As recomendações baseadas nas emoções apresentaram alta correspondência com as preferências dos usuários, o que mostra como a análise de sentimentos pode melhorar a personalização em sistemas de recomendação de livros. Também foram identificados desafios em gêneros menos populares, indicando a necessidade de abordagens adaptativas.*

1. Introdução

A análise de sentimentos, subárea do Processamento de Linguagem Natural (PLN), vem ganhando destaque na personalização de recomendações em plataformas digitais de leitura. Nesse contexto, ferramentas de aprendizado profundo surgem como alternativas promissoras para a extração de informações emocionais de resenhas literárias.

Apesar dos avanços, a análise de sentimentos ainda enfrenta desafios significativos. Destacam-se a dificuldade em capturar nuances emocionais complexas, como sarcasmo e ironia, e a variabilidade observada em categorias literárias com menor volume de resenhas em grandes *datasets*, o que dificulta a generalização de conclusões [Devika and Milton 2025]. Adicionalmente, a aplicação de APIs especializadas no contexto literário permanece um campo com exploração limitada [Gogula et al. 2023].

Este estudo investiga a aplicação da API Gemini para análise de sentimentos em um volume expressivo de resenhas de livros. O objetivo principal é avaliar a viabilidade e eficácia desta ferramenta na identificação de variações emocionais entre diferentes

gêneros literários, visando integrar essas informações emocionais em sistemas de recomendação personalizados. A contribuição central deste trabalho reside na análise empírica da API Gemini em um cenário de larga escala, explorando seus desafios e potencialidades para o domínio literário, um campo onde sua utilização ainda é incipiente [Gogula et al. 2023]. Adicionalmente, busca-se compreender os padrões emocionais no consumo de literatura, dada sua relevância prática [Sahoo et al. 2023].

2. Fundamentação Teórica

A análise de sentimentos consolidou-se no campo do PLN, evoluindo a partir de abordagens baseadas em léxicos de palavras, que visavam identificar polaridades emocionais. Contudo, tais abordagens frequentemente demonstravam limitações na captura de emoções complexas, como sarcasmo e ironia, devido à sua dependência de significados fixos de palavras, desconsiderando o contexto frasal e discursivo [Chen 2015]. O advento do aprendizado profundo impulsionou avanços significativos: técnicas como redes neurais convolucionais (CNNs) [Severyn and Moschitti 2015] e, mais notavelmente, a arquitetura *transformer* [Ashish et al. 2017], proporcionaram maior capacidade de interpretação de nuances contextuais e, conseqüentemente, ganhos substanciais de precisão.

Posteriormente, redes neurais recorrentes (RNNs) profundas, como explorado por [Graves et al. 2013], e os mecanismos de atenção, introduzidos por [Bahdanau et al. 2014], desempenharam papéis cruciais. As RNNs permitiram modelar sequências e capturar dependências temporais, enquanto os mecanismos de atenção refinaram essa capacidade, permitindo que os modelos focassem em segmentos textuais mais relevantes para a inferência de sentimentos, especialmente em textos extensos. Em contraste, para textos mais concisos, como *tweets*, modelos baseados em convolução, a exemplo dos propostos por [Dos Santos and Gatti 2014], mostraram-se eficazes ao capturar padrões locais importantes. Paralelamente, a investigação da semântica composicional em modelos profundos, como a realizada por [Socher et al. 2013], evidenciou a importância de compreender como o significado de sentenças complexas é construído a partir de seus componentes, uma consideração particularmente relevante para resenhas literárias, que constituem o foco deste estudo.

A API Gemini, ferramenta central nesta investigação, situa-se neste panorama ao integrar modelos de classificação supervisionada com algoritmos avançados para detecção emocional [Kheiri and Karimi 2023]. Sua arquitetura, que pode incorporar elementos de redes neurais avançadas como *Long Short-Term Memory* (LSTM) [Computation 2016], é projetada para capturar dependências de longo alcance em textos como resenhas literárias. Esta capacidade é fundamental para uma interpretação precisa dos sentimentos expressos, que muitas vezes se desdobram ao longo de várias sentenças. Dessa forma, este estudo se fundamenta na aplicação prática dessas técnicas modernas de PLN e aprendizado profundo, utilizando a API Gemini para explorar padrões emocionais em resenhas de livros e avaliar sua integração em sistemas de recomendação.

3. Trabalhos Relacionados

A aplicação da análise de sentimentos para enriquecer sistemas de recomendação de livros tem sido um campo de investigação ativa. Por exemplo, [Mounika and Saraswathi 2021] desenvolveram um sistema de recomendação de livros baseado em sentimentos, demonstrando o potencial da análise emocional para aprimorar a personalização das sugestões

aos leitores. Essa abordagem sublinha a importância de ir além das métricas tradicionais de popularidade ou similaridade de conteúdo, incorporando a dimensão afetiva da experiência de leitura.

A utilização de ferramentas de PLN prontas para uso, como APIs para análise emocional, foi discutida por Kheiri e Karimi [Kheiri and Karimi 2023]. Os autores analisaram as vantagens, como a rápida implementação e o acesso a modelos pré-treinados robustos, e as limitações, incluindo a menor flexibilidade e o potencial desalinhamento com domínios específicos. Embora tais ferramentas ofereçam uma solução prática, sua adaptação e validação no contexto específico da análise de resenhas literárias, com suas nuances e vocabulário próprios, permanecem como um desafio e uma área que demanda investigação adicional, lacuna que o presente estudo busca endereçar com a API Gemini.

Em paralelo à utilização de APIs, modelos de linguagem pré-treinados de grande escala, como o BERT (*Bidirectional Encoder Representations from Transformers*) [Devlin et al. 2019] e suas variantes (e.g., RoBERTa [Liu et al. 2021]), revolucionaram o PLN, representando o estado da arte em diversas tarefas, incluindo a análise de sentimentos. Trabalhos como o de [Gogula et al. 2023] exploraram a aplicação desses modelos em plataformas de nuvem para análise de sentimentos em larga escala, destacando sua eficácia e escalabilidade. A principal diferença em relação à abordagem deste estudo reside no nível de abstração: enquanto o uso direto de modelos como BERT exige *fine-tuning* e um gerenciamento mais complexo da infraestrutura de treinamento e inferência, APIs como a Gemini oferecem uma solução mais acessível e de rápida implementação, embora com menor controle sobre os detalhes do modelo subjacente e menor especialização em uma temática. Este estudo opta pela API Gemini para avaliar a viabilidade de uma solução de alto nível para personalização em larga escala, contrastando com abordagens que demandam maior especialização em aprendizado profundo.

A otimização de arquiteturas para lidar com grandes volumes de dados e para análise contínua, como discutido em trabalhos fundamentais sobre mecanismos de atenção [Bahdanau et al. 2014] e redes recorrentes [Graves et al. 2013], embora não diretamente comparáveis a sistemas de recomendação, informam a infraestrutura necessária para a implementação eficiente de sistemas de recomendação baseados em sentimentos em larga escala, como o proposto.

Os trabalhos supracitados fornecem o embasamento teórico e prático para a aplicação proposta neste artigo. A investigação da API Gemini para extrair nuances emocionais de resenhas de livros, com o objetivo de personalizar sistemas de recomendação, e a análise de sua aplicabilidade em comparação com abordagens que utilizam modelos como BERT, constituem contribuições centrais do presente estudo ao domínio literário.

4. Metodologia

A metodologia deste estudo foi estruturada para analisar sentimentos em um grande volume de resenhas de livros e, subsequentemente, utilizar esses sentimentos para gerar recomendações personalizadas, avaliando a aplicabilidade da API Gemini nesse contexto.

4.1. Descrição da API Gemini

A API Gemini, desenvolvida pelo Google, é uma família de modelos de linguagem multimodais de grande escala (LLMs). Para este estudo, foram testadas suas capacidades de

processamento de linguagem natural para realizar a análise de sentimentos. O modelo subjacente é treinado em vastos volumes de dados textuais e é capaz de interpretar o conteúdo de um texto para classificar os sentimentos expressos em categorias como positivo, negativo, neutro ou, em alguns casos, misto, além de identificar emoções específicas. No contexto deste trabalho, a API foi empregada para extrair os sentimentos predominantes das resenhas de livros. Para os experimentos conduzidos neste estudo, foi utilizada a camada gratuita da API Gemini, o que implicou em considerações sobre limites de taxa de requisição, gerenciados conforme descrito na Seção 4.5.

4.2. Organização dos Dados

Foi escolhido o *dataset Amazon Books Reviews*, disponível publicamente na plataforma Kaggle¹. O dataset é composto por dois arquivos principais: `Books_rating.csv` e `books_data.csv`. Após a junção desses arquivos, o dataset consolidado incluiu as seguintes variáveis principais para este estudo: o texto das resenhas (`review/text`), que representou a variável essencial para a análise de sentimentos; as avaliações numéricas (`review/score`), expressas em notas de uma a cinco estrelas, utilizadas como referência complementar e para validação inicial; as categorias literárias (`categories`), que correspondem aos gêneros dos livros definidos após a integração dos dados, possibilitando a análise comparativa dos sentimentos entre diferentes gêneros; o identificador do usuário (`User_ID`), código único fundamental para a criação de perfis e personalização das recomendações; e o título dos livros (`Title`), utilizado para identificar as obras nas resenhas e nas sugestões geradas. Após a consolidação, os dados foram filtrados para garantir a inclusão apenas de resenhas que continham texto significativo e categorias literárias bem definidas, assegurando a qualidade dos dados de entrada.

4.3. Pré-processamento dos Dados

Dado o volume expressivo de dados (mais de 2,8 milhões de registros no *dataset* original), o pré-processamento foi realizado da seguinte forma: inicialmente, foi feita a filtragem de dados irrelevantes, na qual foram removidas resenhas com campos vazios em variáveis críticas para a análise, como `review/score`, `review/text` e `categories`. Em seguida, realizou-se a segmentação do dataset, dividindo o conjunto completo de dados em arquivos menores, cada um contendo 100.000 resenhas. Essa segmentação facilitou a construção de um *pipeline* de processamento, especialmente das requisições à API, além de permitir a análise incremental dos dados. Por fim, procedeu-se à preparação para análise, onde, para cada segmento, foram selecionadas apenas as colunas necessárias para a análise de sentimentos e criação de perfis. Cada lote processado foi salvo separadamente, garantindo a possibilidade de retomada do processo em caso de interrupções.

4.4. Limitações da API Gemini e Mitigação

A utilização de uma API como a Gemini, apesar de sua eficácia, apresenta certas limitações intrínsecas e observadas. A primeira delas é a natureza de “caixa-preta”: sendo um modelo proprietário, os mecanismos internos detalhados da API não são totalmente transparentes, o que limita a capacidade de realizar um *fine-tuning* específico para o domínio literário ou de entender completamente as razões por trás de certas classificações. Nosso estudo reconhece essa limitação, focando na avaliação da aplicabilidade e desempenho

¹<https://www.kaggle.com/datasets/mohamedbakhmet/amazon-books-reviews>

prático da ferramenta “como está” para a tarefa proposta. Outra limitação observada refere-se à interpretação de sentimentos: houve uma tendência da API em classificar alguns sentimentos que poderiam ser considerados neutros como “mistos”, possivelmente devido à complexidade ou ambiguidade de algumas resenhas. Além disso, a sensibilidade a categorias menos representadas também se mostrou um desafio, já que a interpretação de nuances em categorias literárias com menor volume de dados ou com linguagem muito específica pode ser limitada em modelos generalistas. Para mitigar parcialmente esses desafios, especialmente os relacionados à interpretação de sentimentos, foram realizadas revisões manuais de amostras das classificações da API para calibrar a compreensão dos resultados e, em análises exploratórias, foram considerados ajustes nos limiares de confiança para a classificação final dos sentimentos, embora a classificação primária da API tenha sido mantida para consistência.

4.5. Pipeline de Análise de Sentimentos com a API Gemini

Devido à limitação na taxa de requisições da camada gratuita da API Gemini e ao grande volume de resenhas, o processo de análise de sentimentos foi cuidadosamente gerenciado: **Envio em Lotes Menores:** Cada segmento de 100.000 resenhas (conforme Seção 4.3) foi subdividido, e as resenhas foram enviadas à API em fragmentos de 10 textos por vez para otimizar o tempo de resposta e o manejo de erros. **Gerenciamento de Chaves de API:** Implementou-se um sistema para alternar entre múltiplas chaves de API (obtidas dentro dos termos de uso da plataforma) para garantir a continuidade do processamento e minimizar interrupções devido a limites de requisição por chave. **Tratamento de Erros e Validação da Resposta:** Foram implementadas rotinas para verificar as respostas da API, tratando possíveis erros de conexão, respostas vazias ou erros de decodificação JSON. Em caso de falha em um lote específico, o sistema tentava reenviar o lote após um intervalo ou após alternar a chave da API. **Extração e Classificação dos Sentimentos:** Para cada resenha processada, a API retornou informações detalhadas, incluindo as principais emoções detectadas (e.g., alegria, tristeza, raiva), uma pontuação de sentimento (*sentimentscore*) e uma classificação geral do sentimento do texto (positivo, negativo, neutro ou misto). Estes dados foram armazenados para cada resenha. Os resultados da análise de sentimentos de cada lote foram salvos periodicamente. Ao final do processo, foi gerado um dataset consolidado contendo todas as resenhas originais enriquecidas com as classificações de sentimento e emoção fornecidas pela API.

4.6. Critérios de Seleção dos Usuários

Para a análise detalhada e a geração de recomendações personalizadas, foram selecionados os 10 usuários mais ativos do dataset, identificados pelo maior número de resenhas publicadas. Esta escolha metodológica visou trabalhar com perfis que possuem um volume substancial de dados textuais e de avaliações, garantindo uma base rica e representativa para a análise da capacidade da API em capturar nuances emocionais. O limite também foi estabelecido para equilibrar a profundidade da análise individual de cada perfil com a viabilidade computacional do processamento via API, buscando otimizar a precisão das recomendações geradas e evitar o viés provocado por dados escassos.

4.7. Criação do Perfil Emocional do Usuário

Com base nos resultados da análise de sentimentos, foram criados perfis emocionais detalhados para os 10 usuários selecionados (conforme Seção 4.6). A formalização do perfil

de cada usuário (`User_ID`) consistiu na agregação das seguintes informações: **Histórico de Leitura e Sentimentos Associados:** Uma lista de todos os livros avaliados pelo usuário, contendo para cada livro: o título, a classificação de sentimento geral (positiva, negativa, neutra, mista) e as principais emoções detectadas pela API na resenha do usuário para aquele livro. **Preferências de Gênero e Sentimentos Predominantes por Gênero:** As categorias literárias (`categories`) mais frequentes nas resenhas do usuário, e um resumo dos sentimentos (e.g., percentual de resenhas positivas, negativas) associados a cada uma dessas categorias preferenciais. **Vetor Emocional Agregado:** Um sumário dos sentimentos e emoções mais frequentemente expressos pelo usuário em todas as suas resenhas, formando um panorama geral de suas reações emocionais à leitura. Esses perfis estruturados, contendo tanto dados granulares por livro quanto agregados por usuário e gênero, serviram como a entrada para a etapa de geração de recomendações personalizadas.

4.8. Geração de Recomendações Personalizadas de Livros com a API Gemini

A geração de recomendações personalizadas utilizou a API Gemini em uma segunda etapa, desta vez para inferir sugestões de novas leituras com base nos perfis emocionais dos usuários. Para cada um dos 10 usuários selecionados, foi construído um *prompt* detalhado. Este *prompt* incluía uma solicitação explícita para que a API atuasse como um sistema de recomendação de livros, o perfil emocional consolidado do usuário (conforme descrito na Seção 4.7), destacando seus gêneros literários favoritos e os sentimentos/emoções que ele mais valoriza ou busca em suas leituras (inferidos a partir das resenhas positivas), além de uma amostra de aproximadamente 50% dos títulos de livros que o usuário já leu e avaliou positivamente, com o objetivo de fornecer exemplos concretos de suas preferências. Essa amostra não “treina” a API no sentido tradicional, mas serve como um forte condicionamento para a geração de sugestões relevantes. O *prompt* construído foi então enviado à API Gemini, solicitando uma lista de pelo menos 20 novas sugestões de livros. As recomendações recebidas da API foram processadas com o intuito de remover quaisquer títulos duplicados, filtrar sugestões para garantir que os livros recomendados, sempre que possível, pertencessem a categorias literárias de interesse do usuário (identificadas em seu perfil) e estivessem alinhados com os padrões emocionais positivos observados, além de excluir livros que o usuário já havia lido (presentes no seu histórico completo). As recomendações finais para cada usuário foram armazenadas para a etapa de validação e análise. Esta abordagem permitiu que a API Gemini utilizasse o conhecimento contextual fornecido pelo perfil do usuário para gerar recomendações que fossem não apenas baseadas em similaridade de conteúdo, mas também em alinhamento emocional.

4.9. Validação das Recomendações

A validação inicial das recomendações personalizadas concentrou-se em uma métrica de “coincidência”, comparando os livros recomendados pela API com a porção do histórico de leitura do usuário que não foi fornecida à API durante a geração das recomendações. Para cada usuário, seu histórico de livros lidos foi dividido: aproximadamente 50% foi usado para construir o *prompt* de recomendação, e os 50% restantes foram retidos como um conjunto de “controle” ou “validação cega”. Após receber as recomendações da API, estas foram comparadas com os títulos presentes no conjunto de controle. Uma “coincidência” foi registrada se um livro recomendado pela API já tivesse sido lido pelo usuário

(e estivesse neste conjunto de controle). Para cada coincidência identificada, os sentimentos originalmente associados à resenha daquele livro pelo usuário foram analisados para verificar o alinhamento. É importante ressaltar as limitações desta abordagem de validação. A métrica de “coincidência” avalia primariamente a capacidade do sistema de redescobrir preferências conhecidas, mas não mede diretamente a qualidade ou a novidade das recomendações de livros que o usuário ainda não leu. O fato de um livro recomendado não coincidir com o histórico de leitura não implica que a recomendação seja ruim; pode ser uma sugestão nova e relevante. Além disso, esta validação não incorpora feedback direto do usuário sobre as novas recomendações. Apesar dessas limitações, essa análise de coincidência fornece um primeiro indicador quantificável do alinhamento do sistema com as preferências estabelecidas dos usuários.

5. Resultados e Discussão

5.1. Análise Descritiva do Conjunto de Dados

O conjunto de dados utilizado para a análise de sentimentos compreende um total de **1.947.532 resenhas**. Este volume reflete uma quantidade relevante de *feedback* de usuários sobre livros, fornecendo uma base diversificada para a extração de opiniões, sentimentos e emoções dos leitores, e, conseqüentemente, para as análises subsequentes.

A análise da distribuição das resenhas por categoria literária revela que o gênero mais proeminente é **ficção**, com 647.855 resenhas. Outras categorias com representatividade significativa incluem **biografia e autobiografia** (89.116 resenhas) e **religião** (85.671 resenhas), seguidas por **história** e **negócios e economia**. A **Figura 1a** ilustra a distribuição das 10 principais categorias, evidenciando a predominância da ficção. A variedade de gêneros sugere uma ampla gama de interesses entre os leitores da plataforma, um aspecto crucial para a personalização de recomendações literárias.

As classificações gerais atribuídas às resenhas demonstram um viés predominante para **avaliações positivas** (1.447.194). Embora **avaliações negativas** (340.510) também representem um volume considerável, seguidas por classificações **neutras** e **mistas**, a tendência positiva é clara, como ilustrado na **Figura 1b**. Este padrão pode indicar que há uma maior propensão dos leitores satisfeitos a registrar suas experiências de leitura.

A análise de sentimentos das resenhas revelou uma predominância de **emoções positivas** (553.889 ocorrências), que englobam sentimentos como **alegria**, **amor** e **apreciação**, entre outros. A **Figura 1c** apresenta a distribuição das principais emoções identificadas. A prevalência de emoções positivas sugere que os livros avaliados frequentemente proporcionam experiências de leitura agradáveis, o que reforça a relevância de considerar o componente emocional nas recomendações.

5.2. Resultados das Recomendações Personalizadas

Para a avaliação, foram selecionados 10 usuários, cujos IDs originais foram mapeados para identificadores numéricos simplificados (User_01 a User_10). Foram realizadas três rodadas de recomendações para esses usuários.

As Tabelas 1, 2 (Primeira Rodada), 3, 4 (Segunda Rodada) e 5, 6 (Terceira Rodada) detalham o número de recomendações feitas, o número de coincidências com livros já lidos pelos usuários (e não utilizados na construção do perfil para aquela rodada) e os títulos coincidentes.

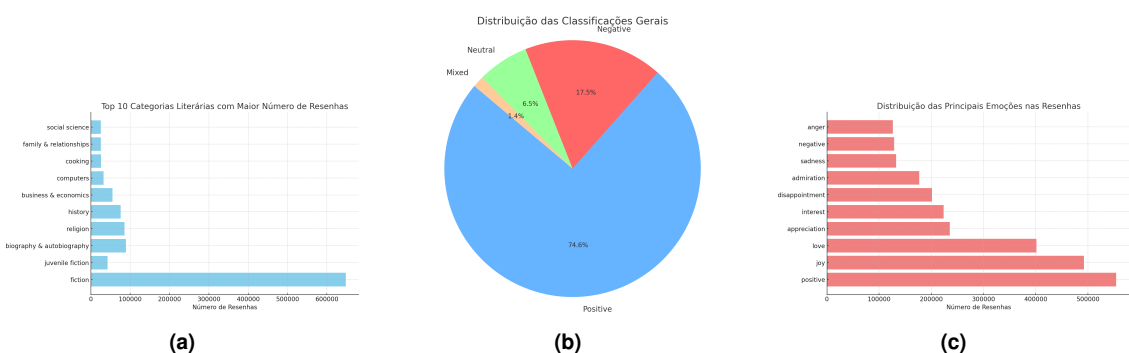


Figura 1. (a) categorias literárias, (b) classificações gerais, (c) principais emoções.

Tabela 1. Primeira Rodada de Recomendações e Coincidências

User_ID	Recomendações	Coincidências
User_01	30	1
User_02	30	1
User_03	31	6
User_04	30	1
User_05	31	4
User_06	30	0
User_07	31	6
User_08	31	0
User_09	30	1
User_10	31	0

Tabela 2. Recomendações e Sentimentos - Primeira Rodada

User_ID	Livro Coincidente	Sentimento Geral	Categoria
User_01	The Secret Garden	positive	['juvenile fiction']
User_02	Persuasion	positive	['fiction']
User_03	Persuasion	positive	['fiction']
User_03	Wuthering Heights	positive	['audiobooks']
User_03	A Tale of Two Cities	positive	['executions and executioners']
User_03	David Copperfield	positive	['fiction']
User_03	A Room with a View	positive	['fiction']
User_03	A Room with a View	positive	['fiction']
User_04	The Giver	positive	['juvenile fiction']
User_05	Wuthering Heights	positive	['audiobooks']
User_05	A Tree Grows in Brooklyn	positive	['fiction']
User_05	David Copperfield	positive	['fiction']
User_05	A Tale of Two Cities	positive	['executions and executioners']
User_07	And Then There Were None	positive	['fiction']
User_07	And Then There Were None	positive	['fiction']
User_07	Twenty Thousand Leagues Under the Sea	positive	['fiction']
User_07	Twenty Thousand Leagues Under the Sea	positive	['fiction']
User_07	Twenty Thousand Leagues Under the Sea	positive	['fiction']
User_07	Around the World in Eighty Days	positive	['juvenile fiction']
User_09	The Alchemist	positive	['fiction']

Tabela 3. Segunda Rodada de Recomendações e Coincidências

User_ID	Recomendações Recebidas	Coincidências
User_01	30	2
User_02	31	0
User_03	30	3
User_04	31	1
User_05	31	2
User_06	31	0
User_07	31	0
User_08	30	0
User_09	30	0
User_10	30	0

Tabela 4. Recomendações e Sentimentos - Segunda Rodada

User_ID	Livro Coincidente	Sentimento Geral	Categoria
User_01	The Secret Garden	positive	['juvenile fiction']
User_01	The Notebook	positive	['computers']
User_03	Persuasion	positive	['fiction']
User_03	Wuthering Heights	positive	['audiobooks']
User_03	A Tree Grows in Brooklyn	positive	['fiction']
User_04	The Giver	positive	['juvenile fiction']
User_05	Wuthering Heights	positive	['audiobooks']
User_05	A Tree Grows in Brooklyn	positive	['fiction']

Tabela 5. Terceira Rodada de Recomendações e Coincidências

User_ID	Recomendações Recebidas	Coincidências
User_01	20	0
User_02	20	1
User_03	20	2
User_04	20	0
User_05	20	1
User_06	20	0
User_07	20	0
User_08	20	0
User_09	20	0
User_10	20	0

Observa-se uma variação no número de coincidências entre os usuários e entre as rodadas (Tabela 7). Usuários como User_03 e User_07 apresentaram um número maior de coincidências na primeira rodada, enquanto outros, como User_06, User_08 e User_10, não tiveram coincidências em diversas rodadas. É importante ressaltar que a métrica de “coincidência com livros já lidos” é uma avaliação indireta da qualidade da recomendação. Embora útil, ela possui limitações: um livro recomendado pode ser relevante para o usuário mesmo que ele não o tenha lido anteriormente, e a ausência de coincidência não implica necessariamente uma recomendação inadequada, mas sim uma sugestão de algo novo. Esta métrica foi adotada como uma forma pragmática de validação offline, dada a impossibilidade de realizar testes com usuários em tempo real para este estudo. A interpretação da eficácia do sistema deve, portanto, considerar essas nuances.

5.3. Avaliação Quantitativa das Recomendações

Para mensurar o desempenho das recomendações personalizadas geradas a partir dos perfis emocionais dos usuários, foram calculadas as métricas de *Precision@k*, *Recall@k* e

Tabela 6. Recomendações e Sentimentos - Terceira Rodada

User_ID	Livro Coincidente	Sentimento Geral	Categoria
User_02	Persuasion	positive	['fiction']
User_03	Wuthering Heights	positive	['audiobooks']
User_03	A Tree Grows in Brooklyn	positive	['fiction']
User_05	Wuthering Heights	positive	['audiobooks']

Tabela 7. Métricas de Avaliação das Recomendações

User_ID	Rodada	Recs Feitas (k)	Coincidências	Precision@k	Recall@k	F1-Score@k
User_01	1 ^a	30	1	0,033	0,067	0,044
User_02	1 ^a	30	1	0,033	0,050	0,040
User_03	1 ^a	31	6	0,194	0,400	0,266
User_04	1 ^a	30	1	0,033	0,083	0,048
User_05	1 ^a	31	4	0,129	0,267	0,175
User_06	1 ^a	30	0	0,000	0,000	0,000
User_07	1 ^a	31	6	0,194	0,333	0,244
User_08	1 ^a	31	0	0,000	0,000	0,000
User_09	1 ^a	30	1	0,033	0,050	0,040
User_10	1 ^a	31	0	0,000	0,000	0,000

F1-Score@k. Estas métricas foram aplicadas considerando como “relevantes” os livros que constavam no histórico de leitura dos usuários (e que não foram utilizados na criação do perfil para evitar *overfitting*), e que coincidiram com os livros recomendados pelo sistema. A Tabela 7 detalha esses resultados para a primeira rodada de recomendações, onde *k* representa o número de recomendações feitas a cada usuário (variando entre 30 e 31).

A *Precision@k* média observada entre os 10 usuários na primeira rodada foi de aproximadamente 6,5%, enquanto a *Recall@k* média foi de aproximadamente 12,5%. O *F1-Score@k* médio, que combina harmonicamente precisão e revocação, situou-se em torno de 8,5%. Estes valores indicam que, em média, uma pequena proporção das recomendações coincidiu com livros previamente lidos e avaliados positivamente pelos usuários e que não fizeram parte da construção do perfil.

É importante notar uma variabilidade considerável no desempenho entre os usuários. Por exemplo, o *User_03* (*Precision@31* = 19,4%, *Recall@31* = 40,0%) e o *User_07* (*Precision@31* = 19,4%, *Recall@31* = 33,3%) apresentaram métricas muito superiores à média. Em contraste, usuários como *User_06*, *User_08* e *User_10* não tiveram nenhuma coincidência (*Precision@k* = 0%, *Recall@k* = 0%) na última rodada. Esta disparidade sugere que a eficácia a recomendação baseada em perfis emocionais pode ser mais acurada para usuários com históricos de leitura e perfis de sentimento mais consistentes ou claramente definidos. Perfis mais dispersos emocionalmente ou com uma gama de interesses literários muito ampla podem representar um desafio maior.

É fundamental contextualizar estes resultados. A métrica de “coincidências com livros já lidos” (e não usados no treino) serve como um proxy para avaliar se o sistema consegue identificar preferências conhecidas, mas possui limitações. Ela não mede diretamente a capacidade do sistema de promover a descoberta de novos livros que o usuário apreciaria, nem a satisfação geral do usuário com as recomendações que não são coincidências. Portanto, embora os valores de precisão e revocação sejam baixos, eles devem ser interpretados como um indicador inicial da capacidade do sistema em alinhar-se com alguns aspectos do gosto do leitor, dentro do escopo da metodologia de validação ado-

tada. Trabalhos futuros contemplarão avaliações com usuários reais para coletar *feedback* outras métricas como novidade, diversidade e serendipidade, bem como a comparação com métodos de referência (*baselines*) para melhor contextualizar o desempenho.

5.4. Discussão da Abordagem de Recomendação Baseada em Emoções

Este trabalho consistiu em usar perfis emocionais, extraídos de resenhas via análise de sentimentos, para guiar a API Gemini na geração de recomendações literárias personalizadas. Os *prompts* utilizados continham os sentimentos predominantes e os gêneros preferidos dos usuários, buscando dar sugestões alinhadas às suas afinidades emocionais.

A análise qualitativa dos resultados (Tabela 7) revelou uma variabilidade significativa. Usuários como o *User_03*, com um histórico de leitura de clássicos emocionalmente ricos, como “*Persuasion*” e “*Wuthering Heights*”, e um perfil emocional mais coeso, apresentaram um número maior de coincidências (6 em 31). Isso sugere que a abordagem capturou sinais relevantes do seu perfil. Em contraste, a ausência de coincidências para outros usuários, como *User_06* e *User_08*, pode indicar perfis emocionais mais complexos ou multifacetados. A dificuldade do modelo em identificar padrões emocionais menos evidentes no conjunto de teste, ressalta a necessidade de uma avaliação mais aprofundada.

Um ponto crítico observado (Tabelas 2–6) é que todos os livros coincidentes foram associados a um “sentimento geral positivo”. Embora esperado, dado o viés do dataset de resenhas para avaliações positivas (Figura 1b) e o fato de que os livros no histórico dos usuários foram presumivelmente apreciados, isso levanta a questão se o sistema está verdadeiramente alinhando as sugestões com as nuances do perfil emocional ou apenas sugerindo obras populares com sentimentos positivos. Para uma avaliação mais precisa, seria necessário analisar o conteúdo emocional de todas as recomendações geradas, não apenas das coincidentes, e compará-las com o perfil detalhado do usuário.

A variabilidade nos resultados indica que a segmentação por categoria emocional e gênero literário é um ponto de partida promissor, mas que se beneficiaria de refinamentos. Para usuários com perfis emocionais menos definidos ou com interesse em uma gama mais ampla de emoções, estratégias híbridas que combinem a análise emocional com outros fatores (como popularidade, similaridade de itens, ou a exploração de diferentes “estados de espírito” pelo usuário) poderiam enriquecer a experiência de recomendação.

5.5. Implicações Práticas

Os resultados deste estudo, embora exploratórios, oferecem implicações promissoras para o desenvolvimento de sistemas de recomendação de livros. A principal delas é a **Perssonalização Baseada em Emoções**: a identificação de sentimentos extraídos de resenhas (Tabela 2) demonstra potencial para a construção de perfis de usuários mais ricos e detalhados. Plataformas de recomendação podem integrar análises emocionais para refinar suas sugestões, buscando um alinhamento não apenas com o gênero literário, mas também com o impacto emocional desejado pelo usuário. A observação de que a API Gemini, mesmo guiada por *prompts* emocionais, conseguiu capturar algumas preferências (Tabelas 1–5, 7) reforça a viabilidade de integrar ferramentas de PLN para este fim, abrindo caminho para experiências de leitura mais personalizadas e ressonantes.

6. Conclusão

Este estudo explorou a aplicação da análise de sentimentos, utilizando a API *Gemini*, como uma ferramenta para personalizar recomendações de livros, com o objetivo de alinhar as sugestões às preferências emocionais inferidas dos usuários. Os resultados indicaram que, para certos usuários e em determinados contextos, foi possível observar uma correspondência entre os livros recomendados e aqueles previamente lidos e avaliados, particularmente em gêneros como ficção, com grande volume de dados. Isso sugere um potencial para que ferramentas de PLN contribuam para a criação de sistemas de recomendação com maior sensibilidade ao perfil emocional do leitor, sendo a segmentação por emoções predominantes um passo inicial nessa direção.

Contudo, é essencial reconhecer as limitações inerentes a esta investigação exploratória. A metodologia de validação adotada, embora tenha fornecido percepções iniciais, é restrita, e as métricas quantitativas (Tabelas 1–5, 7) apontaram para um desempenho modesto e variável, indicando espaço significativo para aprimoramento no alinhamento emocional. Persistem desafios como a interpretação de sarcasmo e nuances culturais pela API, além das limitações dos modelos generativos em nichos literários específicos ou menos representados, onde a diversidade de interesses e a complexidade emocional exigem modelos mais robustos. Adicionalmente, o possível viés para avaliações positivas no conjunto de dados (Figura 1b) pode direcionar as recomendações, ressaltando a necessidade de estratégias para promover diversidade e evitar "bolhas de filtro" emocionais. As capacidades da ferramenta de PLN utilizada e as características do conjunto de dados também impuseram restrições à generalização e à precisão da análise.

Diante desse panorama, trabalhos futuros devem concentrar-se no fortalecimento das metodologias de avaliação e na superação das limitações identificadas. Recomenda-se a implementação e comparação com métodos de referência, como o TF-IDF e KNN, para contextualizar o desempenho da abordagem proposta. Outra possibilidade é utilizar sentimentos e emoções, detectadas nas resenhas, como *features* para enriquecimento do *dataset*, permitindo assim seu uso indireto com abordagens tradicionais de sistemas de recomendação, como a baseada em conteúdo e em filtragem colaborativa.

É fundamental o desenvolvimento de métricas de avaliação mais abrangentes, que considerem aspectos como diversidade, serendipidade e a aceitação do usuário, obtida por meio de validação direta, como testes A/B ou estudos qualitativos e quantitativos. Pesquisas futuras devem também explorar abordagens híbridas que combinem perfis emocionais com outras fontes de dados e algoritmos de recomendação, assim como investigar o impacto de perfis emocionais claros versus dispersos na qualidade das sugestões. A exploração de técnicas baseadas em *embeddings* contextuais ou outras abordagens adaptativas para lidar com a complexidade emocional em categorias menos representadas é igualmente promissora. O desenvolvimento contínuo nessa área é essencial para o avanço na criação de sistemas de recomendação que não apenas busquem precisão, mas que sejam genuinamente adaptáveis à rica diversidade emocional e de interesses dos usuários.

Declarações

Estudo realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) – Brasil – Código de Financiamento 001.

Referências

- Ashish, V., Noam, S., Niki, P., Jakob, U., Llion, J., Aidan, N., Łukasz, K., and Illia, P. (2017). "attention is all you need", advances in neural information processing systems. *NeurIPS Proceedings*.
- Bahdanau, D., Cho, K., and Bengio, Y. (2014). Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473*.
- Chen, Y. (2015). Convolutional neural network for sentence classification. Master's thesis, University of Waterloo.
- Computation, N. (2016). Long short-term memory. *Neural Comput*, 9:1735–1780.
- Devika, P. and Milton, A. (2025). Book recommendation using sentiment analysis and ensembling hybrid deep learning models. *Knowledge and Information Systems*, 67(2):1131–1168.
- Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2019). Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers)*, pages 4171–4186.
- Dos Santos, C. and Gatti, M. (2014). Deep convolutional neural networks for sentiment analysis of short texts. In *Proceedings of COLING 2014, the 25th international conference on computational linguistics: technical papers*, pages 69–78.
- Gogula, S. D., Rahouti, M., Gogula, S. K., Jalamuri, A., and Jagatheesaperumal, S. K. (2023). An emotion-based rating system for books using sentiment analysis and machine learning in the cloud. *Applied Sciences*, 13(2):773.
- Graves, A., Mohamed, A.-r., and Hinton, G. (2013). Speech recognition with deep recurrent neural networks. In *2013 IEEE international conference on acoustics, speech and signal processing*, pages 6645–6649. Ieee.
- Kheiri, K. and Karimi, H. (2023). Sentimentgpt: Exploiting gpt for advanced sentiment analysis and its departure from current machine learning. *arXiv preprint arXiv:2307.10234*.
- Liu, Z., Lin, W., Shi, Y., and Zhao, J. (2021). A robustly optimized bert pre-training approach with post-training. In *China national conference on Chinese computational linguistics*, pages 471–484. Springer.
- Mounika, A. and Saraswathi, S. (2021). Design of book recommendation system using sentiment analysis. In *Evolutionary Computing and Mobile Sustainable Networks: Proceedings of ICECMSN 2020*, pages 95–101. Springer.
- Sahoo, C., Wankhade, M., and Singh, B. K. (2023). Sentiment analysis using deep learning techniques: a comprehensive review. *International Journal of Multimedia Information Retrieval*, 12(2):41.
- Severyn, A. and Moschitti, A. (2015). Twitter sentiment analysis with deep convolutional neural networks. In *Proceedings of the 38th international ACM SIGIR conference on research and development in information retrieval*, pages 959–962.

Socher, R., Perelygin, A., Wu, J., Chuang, J., Manning, C. D., Ng, A. Y., and Potts, C. (2013). Recursive deep models for semantic compositionality over a sentiment treebank. In *Proceedings of the 2013 conference on empirical methods in natural language processing*, pages 1631–1642.