

Recomendação de Psicólogos por meio de algoritmos de filtragem colaborativa, conteúdo e híbrida.

[Augusto Rosário]¹, [Silvana Bocanegra]¹

¹Departamento de Estatística e Informática – Universidade Federal Rural de Pernambuco
Rua Dom Manuel de Medeiros, s/n, - CEP: 52171-900 – Recife – PE – Brasil

augusto.gomesjunior@ufrpe.br, silvana.bocanegra@ufrpe.br

Resumo. *[É crescente o número de pessoas diagnosticadas com transtornos de saúde mental como depressão e ansiedade, transtornos esses que foram por muito tempo negligenciados pela ciência e sociedade. Ainda assim, cada vez mais surgem avanços nas formas de tratamento para essas pessoas, como é o caso das plataformas que oferecem atendimento psicológico de forma remota. Entretanto, escolher um psicólogo ou terapeuta nem sempre é uma tarefa fácil, dada a grande quantidade de informação que envolve o processo de escolha. Partindo desse princípio, o objetivo deste trabalho foi desenvolver um sistema de recomendação de psicólogos baseado em modelo híbrido, que fosse capaz de recomendar psicólogos com especialidades que supram as necessidades de diferentes tipos de pacientes. O modelo apresentou resultados promissores, onde a similaridade entre os psicólogos recomendados foi consistente e as métricas de avaliação MAE e RMSE atingiram bons resultados ($<0,5$). Além disso, foi possível mitigar pontos fracos de ambos os tipos de recomendação utilizados.]*

Abstract. *[The number of people diagnosed with mental health disorders such as depression and anxiety is growing, disorders that have were long neglected by science and society. Even so, more and more advances are being made in the ways of treating these people, such as platforms that offer psychological care remotely. However, choosing a psychologist or therapist is not always an easy task, given the large amount of information involved in the choosing process. Based on that, the goal of this article was to develop a psychologist recommender system based on a hybrid model, which should be able to recommend psychologists with expertise skills that meet the needs of different types of patients. The model showed promising results, where the similarity between the recommended psychologists was consistent and the MAE and RMSE evaluation metrics achieved good results (<0.5). It was also possible to mitigate weaknesses from both types of recommendations used.]*

1. Introdução

Durante muito tempo condições psicológicas como depressão, ansiedade e outros transtornos específicos, oriundos de experiências próprias de cada pessoa, foram banalizados e negligenciados, tanto pela ciência, como pela sociedade. Devido ao preconceito e ignorância sobre o tema, esses transtornos e/ou condições psicológicas foram por muitas vezes considerados como dramaticidade, ingratidão ou ainda falta de religiosidade por parte das pessoas que passavam por momentos ou até mesmo uma vida inteira enfrentando esses problemas, afetando a busca de ajuda e a adesão ao tratamento psicológico [Nascimento and Souza 2021].

É muito recente o movimento de se importar, e principalmente, estudar as pessoas que sofrem com essas condições para que se possa oferecer um tratamento adequado e com evidências científicas de um impacto positivo na melhoria da saúde mental e qualidade de vida dos pacientes [Nascimento and Souza 2021]. Um dos motivos para a popularização e a crescente do investimento nessa área da medicina é o aumento, nas últimas décadas, do número de pessoas diagnosticadas com depressão, ansiedade e outras patologias psicológicas. Segundo dados da OMS, somente entre os anos de 2005 e 2015, esse número cresceu em 18,4%, chegando a representar 4,4% da população mundial. No Brasil, esse número chega a 5,8% da população, e com previsões para que esse número continue crescendo. [OMS 2015].

Somada essa previsão ao recente momento de pandemia do Covid-19 vivenciado, houve uma aceleração no crescimento do número de pessoas que sofrem com essas condições [Barros 2020, Faro 2020]. Além disso, a pandemia também foi responsável por estabelecer novos desafios na forma como o atendimento psicológico é prestado. Foi necessário realizar o atendimento por meios que cumprissem com as restrições sanitárias, como o distanciamento social e *lockdown*. Portanto, houve um aumento na já crescente popularização de plataformas que oferecem serviços de cuidado com a saúde mental de forma remota [Fonseca 2020, Furlan 2021]. Aumento esse que tende a continuar crescendo mesmo após a suavização das restrições sanitárias da pandemia, principalmente devido a problemas como a locomoção nas grandes cidades e a falta de tempo livre nas rotinas contemporâneas.

Entretanto, mesmo com a popularização dessas plataformas, o ato de escolher um psicólogo que atenda às suas necessidades não é tão simples, e pode acabar se tornando um problema devido a quantidade de informação envolvida. Além da escolha da plataforma, o paciente deve se atentar a uma grande variedade de informações a respeito de cada psicólogo.

Dentro desse contexto, surge a necessidade de um modelo que possibilite uma melhor filtragem desse conteúdo, melhorando a experiência do paciente na procura por um profissional adequado ao seu perfil. Atualmente, uma das melhores formas de gerar esse tipo de filtragem de informação é por meio de sistemas de recomendação, que são capazes de fazer um recorte dos itens mais interessantes, utilizando dados sobre o usuário e o conteúdo que ele consome ou avalia positivamente, além de também utilizarem os dados de outros usuários semelhantes [Sharma and Mann 2013].

1.1. Objetivo geral

Desenvolver um sistema de recomendação que possa recomendar psicólogos com especialidades que satisfaçam as necessidades específicas de diferentes tipos de pacientes.

1.2. Objetivos específicos

- Implementar um modelo híbrido, que utiliza algoritmos de recomendação colaborativa e baseada em conteúdo, visando conseguir mitigar os pontos fracos de ambas;
- Atingir parâmetros positivos das métricas avaliativas MAE e RMSE, assim como os de coeficiente de similaridade entre os psicólogos, garantindo a precisão das recomendações geradas pelo sistema, sejam essas baseadas em conteúdo ou colaborativas.

1.3. Organização do trabalho

O presente trabalho está disposto de acordo com as seguintes seções: Referencial Teórico, abordando a teoria e conceitos utilizados no desenvolvimento do projeto. Trabalhos Relacionados, descrevendo e comparando outros trabalhos relevantes. Abordagem Proposta, discorrendo a metodologia aplicada e desenvolvimento dos algoritmos de recomendação. Avaliação, onde os resultados são analisados. Por fim, as Conclusões e Trabalhos Futuros.

2. Referencial teórico

Nessa seção serão definidos conceitos fundamentais para a compreensão do funcionamento dos sistemas de recomendação, incluindo as características específicas dos diferentes modelos de recomendação que serão desenvolvidos. Também estão inclusos os conceitos utilizados diretamente na implementação desses modelos, e por fim, uma abordagem sobre as métricas avaliativas que serão utilizadas na análise dos resultados obtidos.

2.1. Sistemas de Recomendação

Essa subseção visa não apenas definir o conceito de um sistema de recomendação, mas também apresentar os principais modelos que podem ser utilizados dentro dessa área, abordando suas características, vantagens e desvantagens.

A princípio, um sistema de recomendação pode ser entendido como uma forma de filtrar uma grande quantidade de informação visando recomendar itens de interesse para cada usuário em específico. O sistema faz isso com base em algumas informações prévias sobre o usuário, como: últimos itens visitados, itens melhor avaliados, ou até mesmo itens melhor avaliados por usuários semelhantes. Além disso, também é possível gerar recomendações em cima de informações obtidas diretamente do próprio usuário, através de um questionário ou formulário. Formalmente, pode ser definido como: “Sistema de recomendação é um sistema inteligente que faz sugestões para usuários sobre itens que podem interessá-los” [Sharma and Mann 2013].

Portanto, dentro da ideia proposta pelo projeto, um sistema de recomendação buscaria indicar psicólogos adequados às necessidades de cada usuário. Isso seria feito ao levar em consideração as especialidades que estão sendo buscadas recentemente pelo usuário dentro do sistema, além das características do perfil dos psicólogos que foram melhor avaliados pelo usuário, e também por usuários semelhantes.

Por fim, é importante ressaltar que as especialidades de cada psicólogo estão relacionadas a diferentes tipos de transtornos psicológicos cujo o psicólogo possui expertise para tratar (ex: depressão, bipolaridade, ansiedade, fobias, síndrome do pânico e transtorno obsessivo compulsivo) e também diversas temáticas que podem ser abordadas dentro desses transtornos (ex: separação, luto, adolescência, procrastinação, *burnout* e identidade de gênero). A seguir, serão vistas as abordagens mais comuns para a implementação de sistemas de recomendação.

2.1.1. Recomendação Baseada em Conteúdo

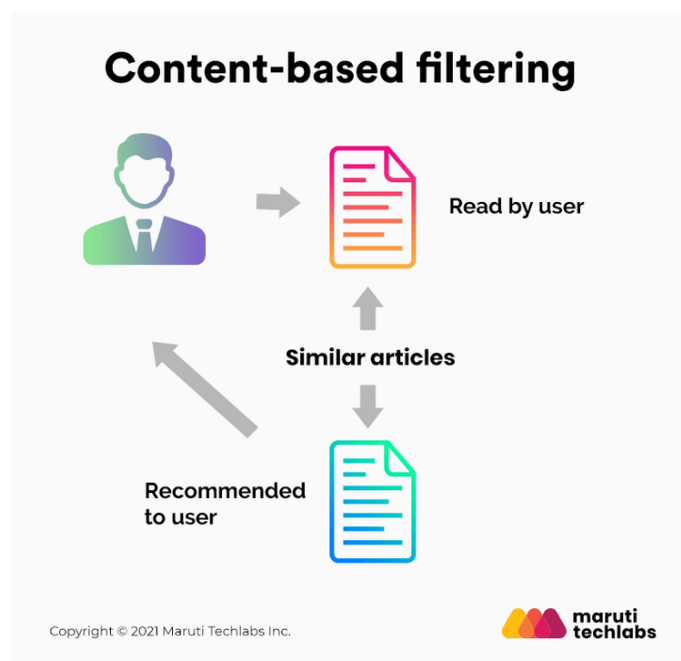


Figura 1. Esquema da Recomendação por Conteúdo.
Fonte: Maruti Techlabs [Ariwala 2021].

Como pode ser observado na Figura 1, os sistemas de recomendação baseados em conteúdo funcionam com base na recomendação de itens semelhantes àqueles que o usuário avaliou bem no passado. Ou seja, partem do princípio que os usuários tendem a continuar buscando por itens parecidos aos que já gostaram anteriormente. Portanto, os sistemas de recomendação baseados em conteúdo podem ser melhor entendidos por meio de dois conceitos fundamentais: perfil de interesse do usuário e similaridade [Sharma and Mann 2013].

O primeiro a ser abordado é o de perfil de interesse do usuário, afinal, antes de estabelecer o grau de semelhança entre os itens, é necessário identificar quais itens podem ser considerados interessantes pelo usuário. Esse perfil pode ser traçado a partir de dois métodos diferentes. As informações sobre o perfil de interesse podem ser coletadas diretamente do usuário de maneira consciente, a partir das respostas inseridas por ele em um questionário ou formulário. Outra forma de obter essas informações, é através do uso contínuo do próprio sistema, ou seja, conforme o usuário procura por itens específicos, e os avalia, o seu perfil de interesse vai sendo construído [Sharma and Mann 2013].

Outro conceito fundamental para a compreensão de sistemas de recomendação baseados em conteúdo, é o de similaridade. Uma vez definido o perfil de interesse do usuário, cabe ao sistema definir o nível de similaridade entre os itens que o usuário já demonstrou interesse e avaliou positivamente, e os demais itens presentes no sistema.

A similaridade entre os itens é calculada a partir do grau de semelhança entre as suas características, geralmente definidas por palavras-chaves. A partir daí, pode-se calcular e tentar prever o quanto um item é recomendável para um dado usuário. Esse cálculo pode ser feito de diversas formas, e a maioria está ligada a técnicas de *machine learning*, como: algoritmos de classificação e/ou de recomendação, *clustering* e redes neurais [Sharma and Mann 2013].

Dentro da ideia do sistema proposto, um algoritmo de recomendação baseado em conteúdo definiria a similaridade entre os psicólogos por meio de suas características. Essas características estariam dispostas no citado padrão de palavras-chaves, e como abordado anteriormente, se tratam das temáticas e transtornos psicológicos cujo o psicólogo possui expertise para tratar. Os detalhes das técnicas utilizadas estão descritos na Seção 2.2.1.

2.1.2. Recomendação Baseada em Filtragem Colaborativa.



Figura 2. Esquema da Recomendação Colaborativa.
Fonte: Maruti Techlabs [Ariwala 2021].

Os sistemas de recomendação baseados em filtragem colaborativa, assim como os sistemas baseados em conteúdo, também utilizam o conceito de similaridade, entretanto, como pode ser observado na Figura 2, essa similaridade também é calculada entre os usuários do sistema, e não mais somente entre os itens. Outra diferença é que a filtragem colaborativa pode gerar dois tipos de resultados diferentes: predição ou recomendação.

A predição se trata de um valor numérico que representa a avaliação que o usuário daria para um item específico, de acordo com a previsão do algoritmo. Já a recomendação, é uma lista composta pelos melhores itens para um usuário em questão, baseando-se na predição da avaliação do usuário para diversos itens do sistema [Sharma and Mann 2013].

Algoritmos de recomendação baseados em filtragem colaborativa podem ser divididos em dois grupos: baseado em memória e baseado em modelo. O esquema baseado em memória, por sua vez, também pode ser dividido em dois tipos: baseado em usuários ou itens. Quando baseado em usuários, o sistema estabelece a similaridade entre os usuários ao comparar as avaliações fornecidas por eles a determinados itens. Só então, é utilizada uma média ponderada entre as avaliações dadas a um determinado item por usuários semelhantes, para prever a nota que seria dada ao mesmo item pelo usuário que se deseja prover uma recomendação. Já quando baseada em itens, a similaridade é calculada entre os itens por meio das avaliações fornecidas para um grupo de itens por um mesmo usuário. Nesse momento a avaliação de um determinado item é calculada por meio da média ponderada entre as avaliações dadas por um determinado usuário, para um grupo de itens [Sharma and Mann 2013].

Já o esquema baseado em modelo faz uso de modelos estatísticos para gerar recomendações, dessa forma, sendo menos dependente de informação sobre os usuários ou itens para fornecer uma recomendação. Nesse cenário, técnicas de aprendizado de máquina e mineração de dados são usados para criar um modelo que consiga identificar padrões para simular o comportamento dos usuários e fazer predições. Todavia, esse tipo de modelagem precisa de muito treinamento para funcionar de maneira adequada. Alguns modelos estatísticos que podem ser usados para gerar previsões baseadas em probabilidades são os modelos de *cluster* e redes bayesianas [Sharma and Mann 2013].

Um dos benefícios de utilizar a filtragem colaborativa, é a capacidade de inovação das recomendações. Os algoritmos de recomendação baseados em conteúdo podem sofrer de problemas relacionados à falta de variedade nas recomendações, não conseguindo recomendar algo novo ou diferente, mas que ainda assim pode ser de interesse do usuário. Esse problema é menos presente num algoritmo de filtragem colaborativa por conta de sua capacidade de gerar recomendações inesperadas baseadas no interesse de outros usuários considerados semelhantes.

Entretanto, a filtragem colaborativa pode apresentar alguns problemas devido a falta de informação sobre os itens ou usuários, como a escassez de dados de avaliações. Além disso, itens recém inseridos no sistema podem ser recomendados por possuírem uma boa avaliação, ainda que não devessem, tendo em vista o baixo número de usuários que avaliaram o item em questão. Esses problemas podem levar os novos usuários do sistema a sofrerem com recomendações imprecisas, devido à falta de informação sobre os seus interesses, problema esse que é conhecido como *cold start*.

Por fim, pode-se concluir que a recomendação colaborativa também pode ser aplicada dentro da proposta do projeto, visando gerar variedade e inovação dentre os psicólogos recomendados. A abordagem colaborativa utilizada será a baseada em modelo, onde os dados de avaliações serão utilizados para gerar predições de notas dadas por um usuário a um psicólogo. Mais detalhes sobre o modelo utilizado estão descritos na Seção 2.2.2.

2.1.3. Recomendação Baseada em Filtragem Híbrida

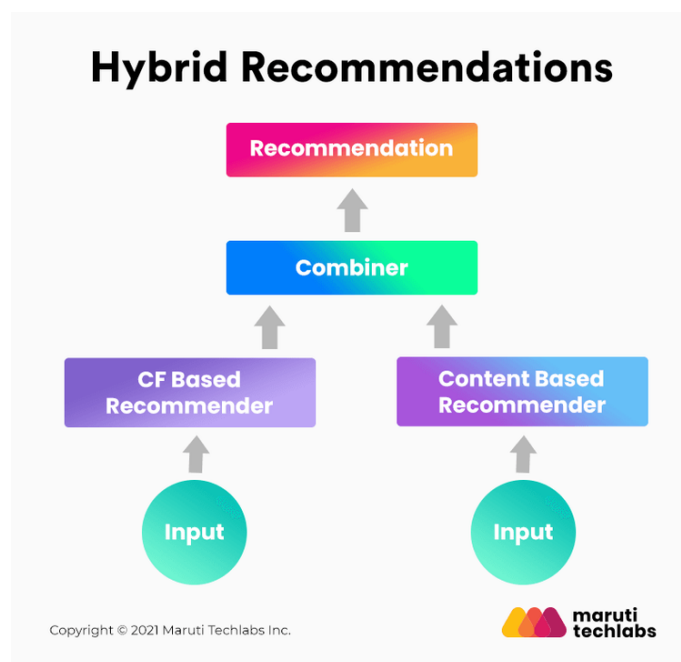


Figura 3. Esquema da Recomendação Híbrida.

Fonte: Maruti Techlabs [Ariwala 2021].

Como pode ser observado na Figura 3, a recomendação baseada em filtragem híbrida é um método que visa mesclar as técnicas dos modelos de recomendação baseada em conteúdo e filtragem colaborativa, buscando balancear os pontos fracos de ambas. Dessa forma, a falta de inovação da recomendação por conteúdo poderia ser minimizada por meio da recomendação colaborativa com base em usuários semelhantes, e os problemas de *cold start* da recomendação colaborativa poderiam ser minimizados através de técnicas da recomendação por conteúdo, como questionários e formulários [Sharma and Mann 2013].

A recomendação híbrida pode ser implementada de diversas formas. Os sistemas de recomendação colaborativo e baseado em conteúdo podem ser implementados de maneira independente, e terem os seus resultados combinados. Outra forma, é incorporar características de um método no outro. Por fim, uma forma mais complexa seria implementar um modelo único que contenha técnicas de ambos os sistemas de recomendação. Esta será a abordagem utilizada neste projeto.

2.2. Modelos usados para a implementação do Sistema de Recomendação

Nessa seção serão abordadas especificamente os modelos, algoritmos e/ou técnicas utilizadas na implementação dos diferentes tipos de sistemas de recomendação abordados.

2.2.1. Similaridade em Conteúdo Textual: TF-IDF e Similaridade Cosseno

Para a implementação da recomendação baseada em conteúdo, especificamente para gerar a matriz por onde a similaridade é calculada, será utilizado o TF-IDF (*Term Frequency - Inverse Document Frequency*), uma medida estatística que pode ser usada para avaliar a importância de uma palavra em um documento. Essa medida se baseia na frequência de um termo dentro de um documento em relação com a frequência desse mesmo termo em todos os outros documentos de um conjunto. Formalmente, o cálculo do TF-IDF é feito pelo produto da frequência de um termo (TF) pela frequência inversa do documento (IDF), onde quanto maior for o valor obtido, maior é a relevância do termo para o documento em questão [Pedregosa et al. 2011].

Medindo o número de vezes que um termo aparece em um documento específico, a frequência de um termo (TF) é definida pelo número de vezes (n) que um termo (i) aparece em um documento (j), dividido pelo número total de termos do documento.

$$TF_{i,j} = \frac{n_{i,j}}{\sum_{i'} n_{i',j}}, \quad (1)$$

onde: (i') são os demais termos do documento.

Já a frequência inversa do documento (IDF), que mede a raridade de um termo na coleção de documentos, é calculada pelo logaritmo do número total de documentos (N) dividido pelo número de documentos (df) que contêm o termo (i) em questão.

$$IDF_i = \log\left(\frac{N}{df_i}\right), \quad (2)$$

Como dito anteriormente, o TF-IDF é calculado ao fazer o produto dos dois valores acima, sendo assim:

$$TFIDF_{i,j} = \frac{n_{i,j}}{\sum_{i'} n_{i',j}} \log\left(\frac{N}{df_i}\right). \quad (3)$$

Existem diversas aplicabilidades para o TF-IDF, como: processamento de linguagem natural, recuperação de informações e documentos, agrupamento de documentos por similaridade, extração de palavras-chave e até mesmo para gerar resumos automatizados de documentos [Pedregosa et al. 2011].

Dentro desse projeto, o TF-IDF foi utilizado em conjunto com a técnica de vetorização (*TFIDF-Vectorizer*), para gerar a matriz por onde a similaridade será calculada. Esse processo consiste na avaliação textual da descrição das especialidades, ou seja, transformar as informações à respeito das especialidades dos psicólogos em vetores. Dessa forma é construída uma matriz esparsa contendo os vetores dos psicólogos nas linhas, e as especialidades que compõem os vetores nas colunas. Essa matriz é preenchida apenas com 0 e 1, sendo utilizado o 0 para uma especialidade de tratamento que o psicólogo não possui, e 1 para uma especialidade de tratamento que o psicólogo possui. A Tabela 1 representa a matriz criada pelo processo descrito acima.

Tabela 1. Matriz TF-IDF.

	Depressão	Ansiedade	Estresse	Relações	Luto	Fobia
Psicólogo x	0	0	1	0	1	0
Psicólogo y	1	0	1	1	0	1
Psicólogo z	1	1	0	0	0	0

A partir dessa matriz é possível calcular a similaridade entre os psicólogos por meio da similaridade cosseno, que é uma métrica matemática empregada para avaliar a semelhança entre dois vetores não-nulos. Essa métrica calcula o cosseno do ângulo formado por esses vetores e fornece um resultado dentro da faixa de -1 a 1. Nessa escala, o valor 1 indica que os dois vetores são perfeitamente iguais, enquanto o valor 0 indica que são completamente diferentes (ortogonais), e -1 indica que são opostos em direção diametral.

$$CosineSimilarity = \cos(\theta) = \frac{A \cdot B}{||A|| \cdot ||B||} = \frac{\sum_{i=1}^n A_i \cdot B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \cdot \sqrt{\sum_{i=1}^n (B_i)^2}}, \quad (4)$$

onde: A e B são os vetores cujo a similaridade é avaliada.

Ao aplicar o cálculo da similaridade cosseno nos vetores da Matriz TF-IDF, o resultado obtido é a matriz de similaridade entre os psicólogos, representada na Tabela 2.

Tabela 2. Matriz de Similaridade.

	Psicólogo x	Psicólogo y	Psicólogo z
Psicólogo x	1	0.37	0.86
Psicólogo y	0.37	1	0.42
Psicólogo z	0.86	0.42	1

2.2.2. Fatoração Matricial: SVD

Para o cálculo de notas previstas da recomendação colaborativa, será utilizado o SVD (*Singular Value Decomposition*), uma técnica de fatoração matricial que está relacionada à diversas finalidades, como: processamento de imagens, reconhecimento de padrões, regressão linear e sistemas de recomendação. A principal vantagem da utilização do SVD é a sua capacidade de redução de dados, que possibilita o cálculo das notas previstas com um tempo e custo computacional consideravelmente menores, quando comparados a percorrer a matriz de avaliações por completo enquanto calcula individualmente as previsões de notas de cada usuário para todos os itens do sistema [Hug 2017]. A ideia por trás de um SVD clássico consiste na decomposição de uma matriz original R em uma combinação linear de três matrizes fundamentais: matriz de valores singulares Σ , matriz de autovetores à esquerda U e matriz de autovetores à direita V .

$$R_{[m \times n]} = U_{[m \times r]} \Sigma_{[r \times r]} V_{[n \times r]}^t$$

A matriz Σ é uma matriz diagonal que ordena os fatores latentes em ordem de poder de predição, que seria a capacidade de um fator latente de gerar recomendações mais precisas. Os fatores latentes podem ser entendidos como características ou recursos que os itens compartilham de forma implícita, e que podem ser obtidos por meio do SVD. Dessa forma, a matriz original pode ser aproximada por uma matriz de menor dimensão considerando apenas os k fatores latentes mais significativos ($k \ll R$), ou seja, a matriz original pode ser mapeada em um espaço de dimensão k .

As matrizes de autovetores U e V^t obtidas à partir da matriz original são capazes mapear tanto os usuários (pacientes) como os itens (psicólogos) em um conjunto de fatores latentes de dimensão k . Dentro desse projeto, um exemplo de fator latente que os psicólogos poderiam compartilhar entre si seria a especialidade de tratamento em transtornos relacionados a identidade de gênero, ou simplesmente a especialidade de tratamento de quaisquer transtornos específicos, como depressão ou ansiedade.

Na matriz V^t cada item i está associado a um vetor q_i^t , que retrata as características do item por meio de uma sequência de valores que mostram o quanto o item é representado por cada fator latente. Análogamente, na matriz U , cada usuário u está associado a um vetor p_u , que retrata as preferências do usuário por meio de uma sequência de valores que representam o seu nível de interesse para cada fator latente.

Uma estimativa para a nota do usuário u para o item i pode ser obtida pelo produto escalar $\hat{r}_{(u,i)}$ entre os vetores de usuário p_u e de item q_i^t . Esses fatores latentes não possuem significado semântico, sendo apenas compreendidos matematicamente pelo SVD, entretanto, por motivos didáticos, foram representados como as especialidades de tratamento dos psicólogos como explicado anteriormente. A Tabela 3 exemplifica o produto escalar entre os vetores de um psicólogo e um paciente pelo SVD.

Tabela 3. Produto Escalar entre Vetores de Fatores Latentes.

Fatores Latentes	Depressão	Ansiedade	Estresse	Relações	Luto	Fobia
vetor-psicólogo (i)	3.7	0	4.7	4.3	1.4	2.9
vetor-paciente (u)	4.1	0.6	-3.9	4.4	0	-3.8

$$\hat{r}_{(u,i)} = p_u \cdot q_i^t = 3.7 \cdot 4.1 + 0 \cdot 0.6 + 4.7 \cdot (-3.9) + 4.3 \cdot 4.4 + 1.4 \cdot 0 + 2.9 \cdot (-3.8)$$

$$\text{nota prevista: } \hat{r}_{(u,i)} = 4.74$$

O mapeamento de cada item e cada usuário no espaço k pode ser feito usando o SVD. No entanto, considerando a esparsidade dos dados na matriz de avaliações (em geral os usuários avaliam poucos itens), o SVD clássico pode ser impraticável. Nesse cenário, algoritmos de aprendizado de máquina são usados para minimizar o erro quadrático entre o valor estimado para a nota do usuário ao item e o valor real considerando apenas os dados conhecidos na matriz de avaliação usuário item. O problema pode ser formulado matematicamente como:

$$\min_{q^*, p^*} \sum_{(u,i) \in k} (r_{ui} - q_i^t p_u)^2 + \lambda(||q_i^t||^2 + ||p_u||^2)$$

onde: x é o conjunto dos pares (u, i) em que as notas $\hat{r}_{(u,i)}$ são conhecidas, ou seja, a avaliação realmente existiu.

Para solucionar o problema pode ser aplicado o algoritmo do gradiente descendente. A cada iteração serão avaliadas os erros obtidos com os dados do treino e os parâmetros p_u e q_i^t serão corrigidos. A constante λ controla a regularização e usualmente é determinada pela técnica de Cross-Validation (mais detalhes sobre o uso dessa técnica estão presentes na Seção de Modelagem 4.3).

Especificamente para a versão do SVD utilizada nesse projeto (Funk SVD), além do produto escalar entre os vetores de usuários e itens, o cálculo da predição de notas envolve outras variáveis que representam tendências no comportamento dos usuários e características observadas nas avaliações dos itens. A avaliação média geral é denotada por μ , enquanto b_u representa o comportamento ou padrão avaliativo do usuário, por exemplo, se um usuário é mais crítico e tende a avaliar os psicólogos com uma nota 0.5 abaixo da média de avaliação de outros usuários, esse seria o seu valor. Para os itens, b_i representa o desvio observado com relação à média de avaliações, isto é, para um psicólogo bem avaliado cujo a nota média está 0.3 acima dos demais psicólogos, esse seria o seu valor. Dessa forma, o cálculo final é dado por:

$$\hat{r}_{(u,i)} = \mu + b_u + b_i + p_u \cdot q_i$$

Esses parâmetros são acrescentados ao modelo de aprendizado.

2.3. Métricas Avaliativas

Para garantir a precisão das recomendações geradas pelos modelos, se faz necessário o uso de algumas métricas avaliativas. Essas métricas são cálculos matemáticos responsáveis por analisar o desempenho das recomendações por meio de um valor numérico.

Para a análise das recomendações baseadas em conteúdo serão realizadas observações sobre a variação do grau de similaridade de um psicólogos enquanto suas especialidades sofrem alterações. Já para a avaliação do algoritmo de recomendação colaborativa, serão utilizadas as métricas MAE (*Mean Absolut Error*) e RMSE (*Root Mean Squared Error*).

- **MAE:** Se trata de uma métrica que calcula o "Erro Absoluto Médio", ou seja, calcula o grau dos erros entre os valores reais e previstos. Na prática, o MAE tende a evidenciar os pequenos desvios dos valores previstos ao longo de toda a matriz. Quanto mais próximo de zero, significa que melhor é o resultado gerado pelo modelo, sendo valores abaixo de 0,75 considerados satisfatórios.

$$MAE = \frac{1}{n} \sum_{i=1}^{100} |y_i - \hat{y}_i|, \quad (5)$$

onde: y_i = valor real e \hat{y}_i = valor previsto.

- **RMSE:** Calculando a "Raiz do Erro Quadrático Médio" e sendo uma variação da métrica MSE, sua principal diferença se dá por elevar ao quadrado a diferença entre o valor real e o valor previsto. O impacto dessa alteração se dá principalmente em cima dos outliers da matriz (pontos fora da curva, onde a diferença é discrepante) tornando os desvios maiores mais evidenciados e causando maior impacto no valor final. Quanto mais próximo de zero, significa que melhor é o resultado gerado pelo modelo, sendo valores abaixo de 1 considerados satisfatórios.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{100} |y_i - \hat{y}_i|^2}. \quad (6)$$

onde: y_i = valor real e \hat{y}_i = valor previsto.

- **Variedade:** Um dos problemas da recomendação baseada em conteúdo é a falta de inovação nas recomendações geradas, tendo em vista que são sempre baseadas nos conteúdos que o usuário consumiu no passado. Portanto, serão realizadas comparações entre os resultados do modelo baseado em conteúdo e do modelo híbrido, visando validar a capacidade do modelo híbrido de mitigar esse problema. Especificamente essas comparações irão funcionar contando o número de psicólogos alterados entre uma recomendação e outra, além das alterações nos valores de similaridade.
- **Cold Start:** Já para a recomendação colaborativa, como citado anteriormente, o *cold start* se trata de um problema na falta de precisão nas recomendações geradas para usuários recentes dentro do sistema, que ainda não possuem dados suficientes para garantir uma recomendação assertiva. Também como forma de verificar a capacidade do modelo híbrido de mitigar esse problema, serão realizadas comparações entre os resultados dos dois modelos, isto é, entre os valores das notas previstas pelos algoritmos para um usuário com poucos dados de avaliação no sistema.

3. Trabalhos Relacionados

Plataformas de cuidado com a saúde mental de forma remota ainda se encontram numa fase recente da sua popularização e crescimento, portanto, não foram encontrados trabalhos que relacionem diretamente sistemas de recomendação e psicólogos. Entretanto, como forma de diversificar o campo de estudo e enriquecer o presente artigo, foram estudados trabalhos que relacionam os sistemas de recomendações com outras áreas.

Mesmo que não relacionados diretamente com a psicologia, existem cada vez mais artigos que abordam o uso de sistemas de recomendação na área de saúde, principalmente devido ao aumento do número de dados disponíveis a respeito da saúde da população. Dados esses que muitas vezes são coletados a partir de *weareble devices* ou "dispositivos vestíveis" como os Smart Watches, relógios inteligentes capazes de coletar inúmeras informações a respeito da saúde do usuário, informações essas que podem ser utilizadas junto à técnicas de sistemas de recomendação baseados em conteúdo para gerar previsões de possíveis problemas de saúde [Ko et al. 2022].

Ainda na área de saúde, mas agora sobre o uso de modelos de recomendação colaborativos, existem estudos que mostram a possibilidade de, com base nos sintomas, formar grupo de pacientes que sofram de doenças em comum. Thong e Son [Thong and Son 2015] fizeram uso de técnicas de *fuzzy clustering* ou "agrupamento difuso" com base em dados médicos sobre os pacientes e foram capazes de classificá-los em grupos de acordo com as características de suas doenças. O uso desse tipo de recurso pode facilitar o tratamento médico em grupo, como em sessões oncológicas, ou até mesmo na formação de grupos de apoio para esses pacientes.

Além disso, existem estudos que se aproximam ainda mais da proposta desse artigo, como é o caso de Yongjie Yan, Guang Yu e Xiangbin Yan [Yan et al. 2020] que usaram um modelo híbrido de recomendação baseado nas técnicas CNN (*Convolutional Neural Network*) e PMF (*Probabilistic Matrix Factorization*) visando gerar recomendações de médicos para pacientes. O estudo mostrou resultados superiores ao modelo não-híbrido com que foi comparado, ainda que apresentasse problemas de *cold start* devido à alta dependência do modelo aos dados de avaliações.

Fora da área de saúde temos o famoso caso da Netflix, que em 2006 promoveu um campeonato onde os vencedores seriam os desenvolvedores do melhor algoritmo de recomendação baseado em filtragem colaborativa [Netflix 2006]. O valor de 1 milhão de dólares da premiação do concurso chamou atenção, mostrando a relevância dos sistemas de recomendação que começavam a se popularizar na época. Um fato interessante relacionado com esse caso que acabada sendo ofuscado pelo alto valor da premiação, é o critério levado em consideração para determinar o tal melhor algoritmo, que foi uma das métricas avaliativas abordadas anteriormente. Os ganhadores do concurso conseguiram melhorar consideravelmente os resultados, saindo de um RMSE inicial de 1.0540 para um RMSE de 0.8563 [Netflix 2009]. Esse caso também foi um dos responsáveis pela consolidação das técnicas de fatorização de matrizes, como o SVD, como as melhores técnicas para a geração de recomendações colaborativas.

Através da análise dos trabalhos relacionados foi possível reafirmar e consolidar a importância da implementação de um modelo híbrido, visando mitigar os pontos fracos das diferentes técnicas de recomendação. Além disso, a análise também mostrou a importância do uso das métricas avaliativas como forma de garantir a assertividade dos resultados gerados pelo modelo implementado. Por fim, a principal forma de diferenciação do presente artigo é aplicar as técnicas e conceitos estudados sobre sistemas de recomendação numa base de dados pertencente à uma subárea da saúde ainda pouco explorada por esses trabalhos; a saúde mental e psicologia.

4. Abordagem proposta

Nessa seção será abordada a metodologia utilizada para o desenvolvimento do sistema de recomendação proposto, o qual visa indicar psicólogos adequados para as necessidades específicas de cada usuário/paciente. A metodologia adotada foi o Cross Industry Standard for Data Mining (CRISP-DM) [Shearer 2000], que fornece uma divisão em 6 fases para o desenvolvimento de projeto relacionados com a área de dados. A Figura 4, representa as fases da metodologia e as possíveis interações entre elas:

- **Entendimento dos negócios:** Representa a fase inicial do projeto onde são definidas as atividades a serem desenvolvidas e os objetivos a serem alcançados, principalmente do ponto de vista de negócios e mercado;
- **Entendimento dos dados:** Essa fase começa com a extração dos dados necessários para a realização do projeto. Também é nesse primeiro contato com os dados que algumas informações adicionais são adquiridas, como a qualidade e completude dos dados;

- **Preparação dos dados:** Essa etapa está relacionada principalmente com o pré-processamento dos dados, isto é, limpeza, organização e estruturação dos dados;
- **Modelagem:** Nessa etapa são selecionados os modelos e algoritmos que serão aplicados ao conjunto final de dados que foram obtidos a partir de todo o pré-processamento realizado. Também é fundamental nesse momento que os modelos e algoritmos sofram os ajustes que forem observados necessários;
- **Avaliação:** Nessa etapa são avaliados os resultados obtidos a partir da aplicação dos modelos e algoritmos da etapa anterior. É nesse momento que deve ocorrer o uso de algumas métricas para avaliar os resultados, e novamente avaliar necessidade de ajustes;
- **Implantação:** Depois de construído, avaliado e revisado, o modelo finalmente pode ser disponibilizado para aplicação, monitoramento e criação de relatórios.

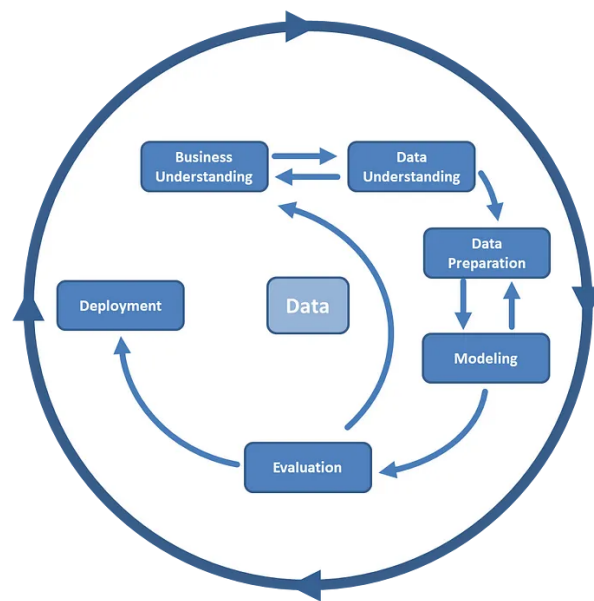


Figura 4. Diagrama da Metodologia CRISP-DM.
Fonte: Towards Data Science [Heinrichs 2020].

A seguir serão detalhadas cada fase empregada no desenvolvimento do projeto, com a avaliação do sistema implementado ocorrendo posteriormente na Seção 5.

4.1. Entendimento dos negócios

Nos últimos anos houve um aumento no número de pessoas que sofrem com transtornos psicológicos como ansiedade e depressão, com previsões para continuar crescendo [OMS 2015]. Somado isso à aceleração desse crescimento devido à pandemia do Covid-19 [Barros 2020, Faro 2020], e também à popularização e investimento nas plataformas que oferecem atendimento psicológico remoto [Fonseca 2020, Furlan 2021], pode-se afirmar que a área de implantação do sistema proposto possui espaço e longevidade no mercado atual.

Além disso, por meio da análise de algumas dessas plataformas como: Zenklub, Psicologia Viva, Telavita e Vittude, foi possível observar que nenhuma das plataformas possuem um sistema de recomendação, sendo disponibilizada apenas a busca por filtros. Essa lacuna mostra que aderir à um sistema de recomendação pode se mostrar uma forma de inovar e se diferenciar da concorrência, aprimorando a experiência do usuário.

4.2. Entendimento e Preparação dos dados

Antes de iniciar o desenvolvimento dos algoritmos de recomendação, foi necessário encontrar uma base de dados que tornasse isso possível. Para gerar recomendações adequadas às necessidades de cada usuário, é fundamental que a base de dados que contenha as informações de especialidades dos psicólogos. Isto é, quais os transtornos e temáticas que o psicólogo possui expertise para oferecer tratamento. Além disso, para gerar também as recomendações colaborativas (baseada em usuários semelhantes) também são necessários dados à respeito das avaliações que esses psicólogos receberam de seus pacientes.

Devido à popularização das plataformas de atendimento psicológico remoto se tratar de um fenômeno recente, não foram encontradas bases de dados dedicadas com esse tipo de informação crucial para o projeto. Portanto, foi necessário a implementação de um *"scraper/crawler"* [Rosário 2023a], isto é, um *bot* (robô) automatizado para coletar as informações presentes no site e convertê-las em um formato estruturado.

Além das plataformas citadas na Seção 4.1, também foram ponderadas plataformas estrangeiras, como: *PsychOn*, *Psychology Today*, *Talkspace* e *BetterHelp*. Entretanto, considerando não apenas a completude dos dados mas também o possível futuro impacto do projeto para a sociedade, a plataforma escolhida para realização da extração de dados foi a Psicologia Viva [Psicologia Viva 2023].

Ao todo foram extraídos os dados de 13.848 pacientes e 1.038 psicólogos, incluindo informações à respeito de suas especialidades, formação e biografia, além das quase cem mil entradas de avaliações. A preparação dos dados foi realizada em conjunto com a extração, por meio de testes de validação das informações que eram obtidas em cada perfil e cada avaliação, garantindo assim a completude e qualidade dos dados obtidos. O sistema de avaliações da plataforma escolhida funciona com base em consultas, dessa forma, um paciente pode avaliar diversas vezes o mesmo psicólogo. Portanto, se mostrou necessário remover as avaliações duplicadas para garantir a fidelidade dos resultados, restando 22.245 avaliações após a limpeza dos dados.

Como pode ser observado na Tabela 4, as informações estão divididas em: *name* (nome do psicólogo), *bio* (biografia escrita pelo psicólogo), *expertises* (transtornos e temáticas cujo o psicólogo possui expertise para oferecer tratamento), *total ratings* (número total de avaliações recebidas pelo psicólogo), *average rating* (nota média do psicólogo na plataforma).

Tabela 4. Representação da Base de Dados

name	bio	expertises	total ratings	average rating
Eliane Aguiar	Olá...	Estresse, Conflitos...	269	5.00
Rafael Cristiano	Olá...	Compulsões, Casamento...	111	4.78
Jhennifer Campos	Olá...	Gravidez, Ansiedade...	94	4.96

4.3. Modelagem

Essa seção tem como objetivo abordar a implementação dos algoritmos de recomendação que foram utilizados na modelagem do problema [Rosário 2023b]. Implementados na linguagem de programação Python [Python 2023], os algoritmos utilizados foram: TF-IDF, para a recomendação por conteúdo, e SVD, para a recomendação colaborativa, através das bibliotecas Scikit-Learn [Scikit-Learn 2023] e Surprise [Surprise 2023], respectivamente. Essas são bibliotecas que contêm diversos algoritmos de aprendizado de máquina e se destacam pela facilidade de uso, servindo como base para projetos de especialistas e também para não especialistas.

Para a recomendação baseada em conteúdo, como abordado na Seção 2.2.1, foi utilizado o TF-IDF em conjunto com a similaridade cosseno para calcular a similaridade entre os psicólogos. Cálculo esse que pode ser feito de duas formas, à depender do tipo de entrada do usuário. Na primeira, a similaridade é calculada diretamente com base em um psicólogo, em relação aos outros, bastando apenas o usuário inserir o nome do psicólogo em questão. Nesse cenário as expertises do psicólogo serão os termos utilizados para calcular a similaridade. Dentro de um sistema real, esse tipo de recomendação se aplicaria dentro do formato "por ter avaliado bem esse psicólogo, recomendamos esses outros especialistas", que pode ser exibida ao final do perfil de um psicólogo, por exemplo.

Na segunda forma, não há necessidade de um psicólogo, mas apenas que o usuário insira alguns problemas ou transtornos pelos quais está buscando tratamento. Nesse cenário, os dados especificados pelo usuário serão os termos utilizados para calcular a similaridade. Dentro de um sistema real, esse tipo de recomendação se aplicaria com base nas últimas pesquisas feitas pelo usuário, ou até mesmo com base em informações coletadas por meio de um questionário durante o cadastro na plataforma. O resultado de uma recomendação baseada em conteúdo está ilustrado na Tabela 5, que apresenta os dados originais do psicólogo novamente acrescidos da coluna *similarity*, representando o nível de similaridade dos psicólogos em relação ao psicólogo que foi utilizado como entrada nesse cenário.

Tabela 5. Representação da Recomendação por Conteúdo.

name	expertises	total ratings	average rating	similarity
Isabela Rodrigues	Adoção, Ansiedade...	44	5.00	0.67
Herika Blumer	Dependência, Luto...	1810	4.96	0.64
Rafael Franco	Fobia, Sexualidade...	18	5.00	0.63

Já para a recomendação baseada em filtragem colaborativa, o algoritmo utilizado foi o SVD, o mesmo utilizado pelos ganhadores do prêmio da Netflix, sendo um dos principais motivos da escolha o fato dele ter se mostrado superior a outras técnicas como KNN e Co-clustering em diversos trabalhos.

Com o objetivo de aprimorar os resultados e avaliar as previsões geradas pelos algoritmos preditivos, podem ser utilizadas técnicas de treinamento durante o desenvolvimento de sistemas de recomendação. Dentro desse projeto, o treinamento do modelo foi realizado por meio da técnica de Cross-Validation, que é utilizado para avaliação do desempenho de modelos de aprendizado. Essa técnica funciona dividindo a base de dados em campos chamados de *folds*, e em cada rotação/iteração o algoritmo de predição (nesse

caso, o SVD) é treinado sobre uma divisão dos campos em um conjunto de treino e teste diferentes [Hug 2017]. Na prática, o algoritmo funciona tentando usar um recorte dos dados de avaliação para observar parâmetros e aprender padrões, e então tenta prever as notas que os usuários dariam para um dos *folds* do conjunto. No final de cada iteração, a validação é feita se baseando na variância entre os valores reais e os valores previstos pelo algoritmo, onde foram utilizadas as métricas avaliativas descritas na Seção 2.3.

A Figura 5, elucida esse processo através de um exemplo em que a base de dados é dividida em 5 *folds*. Nesse cenário os campos de cor verde representam o conjunto de treino, onde algoritmo é aplicado buscando aprender os padrões dos dados de avaliações. Só então, o algoritmo tenta prever as notas dos campos de cor azul, que representam o conjunto de teste e vão se alterando ao longo das iterações do algoritmo, visando o não envasamento dos resultados por exclusão de alguma parte dos dados. O mesmo padrão de *splits* e *folds* da figura foi utilizado nesse projeto.

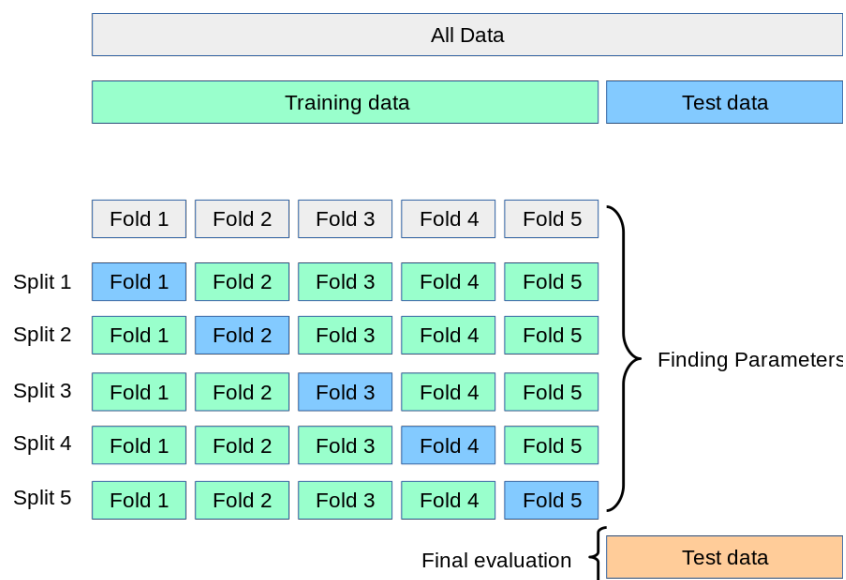


Figura 5. Exemplo de funcionamento do Cross-Validation.

Fonte: Scikit-Learn [Scikit-Learn 2023].

Com o treinamento concluído e a matriz reconstruída em definitivo, é possível localizar um usuário e então ordenar e exibir a recomendação com base nos psicólogos que receberam as melhores notas previstas por ele, segundo o algoritmo. A Tabela 6 ilustra o resultado de uma recomendação colaborativa gerada ao inserir um usuário do sistema como entrada, exibindo os dados dos psicólogos recomendados com o acréscimo da coluna *predict*, que contém as notas que o algoritmo previu que o usuário daria para cada um dos psicólogos.

Tabela 6. Representação da Recomendação Colaborativa.

name	expertises	total ratings	average rating	predict
Monica Freitas	Depressão, Relações...	96	5.00	4.79
Gustavo Brehmer	Infância, Adolescência...	384	5.00	4.77
Tainá Kurtz	Pânico, Hiperatividade...	251	5.00	4.74

Por fim, para a recomendação baseada em filtragem híbrida, foram implementados dois modelos que mesclam técnicas de ambos os algoritmos de formas distintas. O primeiro modelo consiste em, inicialmente, utilizar o algoritmo TF-IDF em conjunto com a similaridade cosseno para definir uma lista com os psicólogos mais similares de acordo com a entrada do usuário (problemas/transtornos que busca tratamento, ou outro psicólogo). Em seguida, o algoritmo SVD é utilizado para prever a nota que o usuário atribuiria aos psicólogos da lista obtida como resultado do TF-IDF e similaridade cosseno. Enfim, a recomendação híbrida é gerada por meio da ordenação da lista inicial com base nas notas previstas.

O segundo modelo funciona de maneira inversa, inicialmente é formada uma lista com os psicólogos com as maiores notas previstas pelo algoritmo SVD, e só então a lista de psicólogos é ordenada por meio da similaridade calculada pelo TF-IDF e a similaridade cosseno com base nos dados inseridos pelo usuário (novamente, problemas/transtornos ou outro psicólogo). A Figura 6 ilustra o funcionamento dos dois modelos de recomendação híbrida descritos acima.

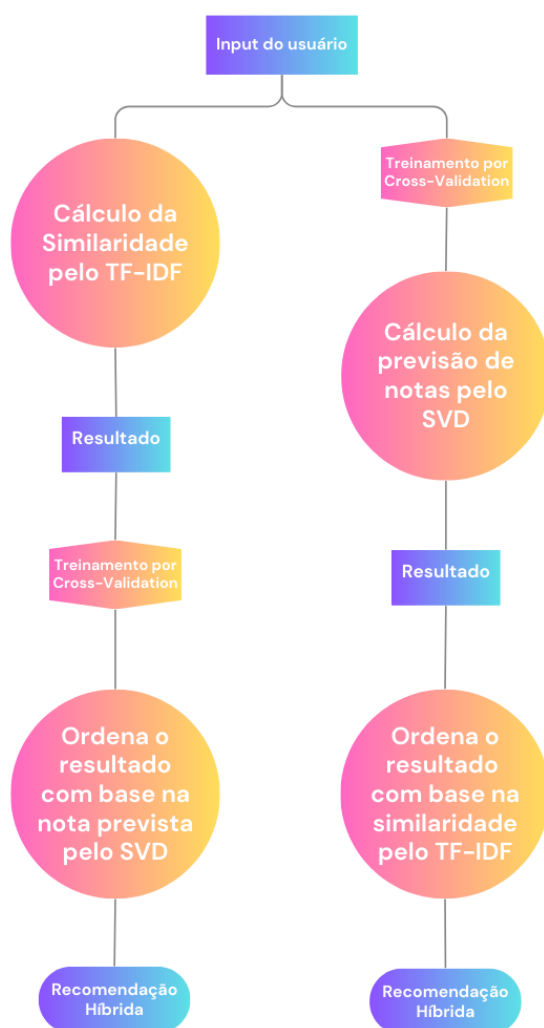


Figura 6. Diagrama da Recomendação Híbrida.
Fonte: Autoria Própria.

5. Avaliação

Nessa seção serão avaliados os resultados obtidos pelos modelos de recomendação cujo o desenvolvimento foi descrito na subseção anterior. Para cada modelo implementado, foram utilizados alguns dos diferentes critérios de avaliação presentes na Seção 2.3. A estratégia utilizada para avaliar o grau de coerência dos resultados da recomendação baseada em conteúdo consiste em calcular o grau de similaridade entre dois psicólogos que possuam as mesmas especialidades de tratamento, a partir daí, continuar recalculando a similaridade enquanto são realizadas alterações nas especialidades de apenas um dos psicólogos, para verificar se os resultados permanecem coerentes.

O resultado desse experimento pode ser consultado na Tabela 7, onde *expertises* são as especialidades do primeiro psicólogo, *expertises 2* são as especialidades do segundo psicólogo, e *n* é o número de especialidades que são compartilhadas por ambos. A princípio os dois psicólogos possuem exatamente as mesmas especialidades de tratamento, portanto, o valor de *n* é máximo, ou seja, o número do total de especialidades dos psicólogos. Em seguida são realizadas suscetivas alterações nas especialidades do segundo psicólogo, que são representadas pela diminuição do valor de *n*, significando que os psicólogos deixaram de possuir *n* especialidades em comum e vão passando a possuir *n-1*, *n-2*, *n-3* e assim suscetivamente. Como pode ser observado na última coluna da Tabela, o coeficiente de similaridade (*similarity*) acompanhou coerentemente as modificações que foram realizadas.

Tabela 7. Avaliação da Recomendação por Conteúdo.

n	expertises	expertises 2	similarity
n-0	Estresse, Depressão, Ansiedade...	Estresse, Depressão, Ansiedade...	1.00
n-1	Estresse, Depressão, Ansiedade...	Estresse, Depressão, Luto...	0.95
n-2	Estresse, Depressão, Ansiedade...	Estresse, Traumas, Luto...	0.93
n-3	Estresse, Depressão, Ansiedade...	Fobias, Traumas, Luto...	0.91

Já para o algoritmo de recomendação baseado em filtragem colaborativa, foram utilizados as métricas avaliativas MAE e RMSE, também descritas na Seção 2.3. Como explicado anteriormente durante a fase de modelagem do projeto, a base de dados foi dividida entre 5 campos chamados de *folds*, e para cada rodada/iteração foi utilizado um conjunto diferente de treino e teste entre os *folds* para treinar e avaliar o modelo.

A Tabela 8 representa o resultado das métricas aplicadas em cada um dos campos de teste utilizados, além de mostrar a margem de valores considerados Bons e Ideais. Como pode ser observado, o algoritmo apresentou um resultado positivo, ficando dentro dos parâmetros considerados como ideais.

Tabela 8. Avaliação da Recomendação Colaborativa.

	fold 1	fold 2	fold 3	fold 4	fold 5	Média	Bom	Ideal
MAE	0.12	0.13	0.12	0.12	0.11	0.12	0.5 - 0.75	< 0.5
RMSE	0.37	0.40	0.39	0.36	0.33	0.37	0.75 - 1.00	< 0.75

Por padrão, os resultados obtidos são comparados com diferentes algoritmos desenvolvidos em outros trabalhos, mas que fazem uso da mesma base de dados. Porém, dada as circunstâncias do presente projeto, onde não foram encontrados trabalhos diretamente relacionados à recomendação de psicólogos, e a base de dados foi extraída unicamente para estes devidos fins, foram realizadas comparações com os trabalhos relacionados da Seção 3, que associam sistemas de recomendação com diferentes áreas. Essas comparações podem ser observadas na Tabela 9, que mostra que o sistema implementado conseguiu obter resultados superiores aos dos trabalhos relacionados.

Tabela 9. Comparação de Resultados da Recomendação Colaborativa.

	Augusto	Thong and Son	Ko et al.	Netflix	Bom	Ideal
MAE	0.12	0.39	0.66	-	0.5 - 0.75	< 0.5
RMSE	0.37	-	-	0.85	0.75 - 1.00	< 0.75

Para a avaliação do algoritmo de recomendação baseado em filtragem híbrida, o critério analisado foi a capacidades do modelo de mitigar os pontos fracos de ambos os tipos de recomendação. A falta de variedade é considerada como o problema mais recorrente da recomendação baseada em conteúdo. Para realizar essa etapa da avaliação foram geradas recomendações de ambos os tipos com base em uma mesma entrada, e os resultados comparados. Como pode ser observado ao comparar as Tabelas 10 e 11, a recomendação híbrida foi capaz de alterar 4 dos 5 psicólogos exibidos como recomendação quando comparada ao resultado da recomendação baseada em conteúdo, apresentando um bom grau de variação, além de não sofrer quedas drásticas nos valores de similaridade, mostrando que apesar da variedade, a recomendação ainda permanece fiel à entrada feita pelo usuário.

Tabela 10. R Conteúdo.

name	expertise	similarity
Amanda	Ansiedade...	0.66
Karine	Depressão...	0.64
Karen	Compulsões...	0.63
Monique	Casamento...	0.60
Tatiane	Ansiedade...	0.59

Tabela 11. R Híbrida.

name	expertise	similarity
Amanda	Ansiedade...	0.66
Jurece	Casamento...	0.46
Lorrayne	Gravidez...	0.53
Eloize	Estresse...	0.47
Camilla	Sexualidade...	0.52

Uma estratégia semelhante foi utilizada para comparar os resultados da recomendação híbrida com a colaborativa. Nesse cenário o maior problema enfrentado pela recomendação colaborativa é o *cold start*, e para isso foram geradas recomendações de ambos os tipos, feitas a partir de um mesmo usuário novo no sistema, isto é, um paciente com poucos dados de avaliação. O resultado pode ser observado ao comparar as Tabelas 12 e 13, onde apesar dos resultados da recomendação colaborativa serem positivos mesmo para um usuário com poucas avaliações no sistema, ainda assim as notas previstas aumentaram na recomendação híbrida. Outro fator a ser levado em consideração é o fato de todos os psicólogos terem sido alterados entre uma recomendação e outra, mostrando que a recomendação híbrida também é capaz de gerar variedade mesmo em comparação com o modelo colaborativo.

Tabela 12. R Colaborativa.

name	expertise	predict
Vinicius	Depressão...	4.78
Lucca	Suicídio...	4.70
Guilherme	Adolescência...	4.70
Juliana	Ansiedade...	4.69
Mariana	Obesidade...	4.69

Tabela 13. R Híbrida.

name	expertise	predict
Pamela	Casamento...	4.87
Gizely	Sexualidade...	4.79
Elisa	Ansiedade...	4.77
Ariadne	Dependência...	4.77
Beatriz	Maternidade...	4.75

Por fim, foram realizadas comparações entre os dois modelos de recomendação híbrida implementados. As Tabelas 14 e 15 foram nomeadas de TFIDF-SVD e SVD-TFIDF respectivamente, representando a ordem em que os algoritmos são aplicados em cada um dos modelos. Diferente da comparação anterior, dessa vez foi utilizado como base um usuário real do sistema, que possui diversos dados de avaliações. Nesse cenário, ambos os modelos apresentaram resultados positivos, sendo capazes de retornar psicólogos com nota prevista máxima, entretanto, o modelo SVD-TFIDF conseguiu apresentar resultados superiores nos níveis de similaridade (coluna 'sim').

Tabela 14. TFIDF-SVD.

name	expertise	predict	sim
Amanda	Relações...	5.00	0.66
Jurece	Depressão...	5.00	0.46
Lorrayne	Gravidez...	5.00	0.53
Eloize	Casal...	5.00	0.47
Camilla	Fobias...	5.00	0.52

Tabela 15. SVD-TFIDF.

name	expertise	predict	sim
Isabelle	Ansiedade...	5.00	0.77
Geiziane	Fobias...	5.00	0.70
Katiussia	Estresse...	5.00	0.64
Henrique	Luto...	5.00	0.57
Larissa	Conflitos...	5.00	0.54

6. Conclusões

Este artigo apresenta uma abordagem de estudo e implementação a respeito de três modelos de sistemas de recomendação diferentes: baseada em conteúdo, filtragem colaborativa e filtragem híbrida, feitas em cima de uma base de dados obtida a partir de um *scraping* da plataforma Psicologia Viva. Os objetivos incluíam implementar um modelo híbrido que fosse capaz de mitigar os pontos fracos de ambos os outros modelos de recomendação, além de atingir níveis satisfatórios para as métricas avaliativas escolhidas, como MAE e RMSE. Para a implementação dos modelos foram utilizados os algoritmos TF-IDF em conjunto com a similaridade cosseno, para o cálculo da similaridade entre os psicólogos, e Cross-Validation em conjunto com SVD, como forma de treinamento e cálculo das previsões por meio da fatorização de matrizes.

Analisando os resultados obtidos pelos modelos implementados, pode-se afirmar que os objetivos inicialmente determinados foram alcançados, com o grau de similaridade gerado pelo modelo de recomendação baseada em conteúdo se comportando de maneira coerente de acordo com os testes realizados, e com o modelo de recomendação colaborativa se encaixando dentro dos parâmetros ideais das métricas MAE e RMSE, além do modelo híbrido ter se mostrado capaz de mitigar problemas de falta de variedade e *cold start* dos outros modelos implementados.

Entretanto, ao analisar os significados desses resultados de forma mais profunda, é possível observar alguns pontos interessantes a serem comentados. Quando comparado até mesmo com outros trabalhos renomados, os resultados das métricas MAE e RMSE do presente artigo ficaram consideravelmente acima da média. Entretanto, causas relacionadas a diferenças técnicas e de implementação foram descartadas, tendo em visto que para isso foram utilizadas versões dos algoritmos das bibliotecas Scikit-Learn e Surprise. As causas mais prováveis apontam para a baixa diversidade das avaliações dos pacientes, que normalmente costumam se consultar várias vezes com apenas um ou poucos psicólogos, e para a uniformidade da base de dados, isto é, a alta ocorrência de avaliações com nota máxima.

Essas causas explicam não somente o resultado das métricas avaliativas mas também podem explicar as altas notas previstas pelo filtro colaborativo, até mesmo para usuários novos no sistema. Por si só, essa uniformidade pode ter algumas causas possíveis, como uma manipulação dos dados de avaliações exibidos, como forma de atrair mais pacientes, ou até mesmo causas humanas que envolvem esse tipo de avaliação. Em outras palavras, as relações de avaliações entre humanos são mais complexas e é comum que pessoas se sintam mais confortáveis em avaliar de maneira negativa um produto do que uma pessoa. Isso se dá devido a relação de empatia que existe quando a avaliação é feita a respeito de um outro ser humano, que pode depender daquele serviço prestado para sua sobrevivência, e que no caso desse projeto, ainda pode ser considerado um ótimo psicólogo para um outro paciente com um perfil diferente. Nesse caso, é mais provável que o paciente dê uma avaliação positiva por questões empáticas, e simplesmente não volte a consumir o serviço daquele psicólogo.

Todavia, uma real avaliação do sistema de recomendação desenvolvido só poderia ocorrer por meio da sua implantação em um sistema real, que está inclusa dentre os trabalhos futuros deste projeto. Além disso, os trabalhos futuros também incluem outros possíveis caminhos, como a extração de uma nova base de dados para a realização de testes e validação dos modelos implementados, assim como a própria modificação desses modelos. Por exemplo, o uso de diferentes técnicas de mesclagem das recomendações por conteúdo e colaborativa, visando gerar outros modelos de recomendação híbrida e compará-los entre si, e até mesmo fazendo alterações e inclusão de outras técnicas mais avançadas relacionadas com a área de sistemas de recomendação e machine learning, como as citadas nos trabalhos relacionados: *Fuzzy Clustering* e *Convolutional Neural Network*. Em acréscimo a isso, fazer uso de outras métricas para a avaliação dos resultados do modelo, como: *Mean Reciprocal Rank (MRR)*, *Mean Average Precision (MAP)*, *Discounted Cumulative Gain (DCG)*, Correlação de Spearman, Serendipidade ou até mesmo métricas que envolvem performance temporal, isto é, levam conta o momento em que as avaliações foram realizadas.

Referências

- Ariwala, P. (2021). How do recommendation engines work? what are the benefits? Maruti Techlabs. Acesso disponível em: <https://towardsdatascience.com/using-crisp-dm-to-grow-as-data-scientist-a07ce3fd9d56/>.
- Barros, M. B. A. (2020). Relato de tristeza/depressão, nervosismo/ansiedade e problemas de sono na população adulta brasileira durante a pandemia de covid-19. *Epidemiologia e Serviços de Saúde*, 29(4).
- Faro, A. (2020). Covid-19 e saúde mental: a emergência do cuidado. *Estudos de Psicologia (Campinas)*, 37.
- Fonseca, M. (2020). Startup de terapia online recebe aporte de R\$ 6 milhões. Revista Pequenas Empresas & Grandes Negócios. Acesso disponível em: <https://revistapegn.globo.com/Startups/noticia/2020/05/startup-de-terapia-online-recebe-aporte-de-r-6-milhoes>. Data de acesso: 18/05/2022.
- Furlan, L. (2021). Cresce o número de psicólogos que atendem online: como se adaptar? Revista Você RH Abril. Acesso disponível em: <https://vocerh.abril.com.br/politicasepraticas/cresce-o-numero-de-psicologos-que-atendem-online-como-se-adaptar>. Data de acesso: 18/05/2022.
- Heinrichs, F. (2020). Using crisp-dm to grow as data scientist. Towards Data Science. Acesso disponível em: <https://towardsdatascience.com/using-crisp-dm-to-grow-as-data-scientist-a07ce3fd9d56/>.
- Hug, N. (2017). Contributions to the use of analogical proportions for machine learning: Theoretical properties and application to recommendation. *Université de Toulouse*.
- Ko, H., Lee, S., Park, Y., and Choi, A. (2022). A survey of recommendation systems: Recommendation models, techniques, and application fields. *Electronics*, 11(1):141.
- Nascimento, V. A. and Souza, I. D. (2021). Transtornos mentais e sociedade: vãos e desvãos do sofrimento psíquico em perspectiva multidisciplinar. *Londrina: Editora Científica*.
- Netflix (2006). Acesso disponível em: <https://web.archive.org/web/20070821194257/http://www.netflixprize.com/faq>.
- Netflix (2009). Acesso disponível em: <https://web.archive.org/web/20090924184639/http://www.netflixprize.com/community/viewtopic.php?id=1537>.
- OMS (2015). Depressão cresce no mundo, segundo oms; brasil tem maior prevalência da américa latina. Portal G1. Acesso disponível em: <https://g1.globo.com/bemestar/noticia/depressao-cresce-no-mundo-segundo-oms-brasil-tem-maior-prevalencia-da-america-latina.ghtml>.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.
- Psicologia Viva (2023). Acesso disponível em: <https://www.psicologiaviva.com.br/>.
- Python (2023). Acesso disponível em: <https://www.python.org/>.

- Rosário, A. (2023a). Acesso disponível em: <https://github.com/augustoljr/PsyCare-Scraping>.
- Rosário, A. (2023b). Acesso disponível em: <https://github.com/augustoljr/PsyCare-Backend>.
- Scikit-Learn (2023). Acesso disponível em: <https://scikit-learn.org/stable/>.
- Sharma, M. and Mann, S. (2013). A survey of recommender systems: Approaches and limitations. *International Journal of Innovations in Engineering and Technology*.
- Shearer, C. (2000). The crisp-dm model: The new blueprint for data mining. *Journal of data warehousing, THE DATA WAREHOUSE INSTITUTE*, pages 13–22.
- Surprise (2023). Acesso disponível em: <https://surpriselib.com/>.
- Thong, N. T. and Son, L. H. (2015). Hifcf: An effective hybrid model between picture fuzzy clustering and intuitionistic fuzzy recommender systems for medical diagnosis. *Expert Systems with Applications*, 42(7):3682–3701.
- Yan, Y., Yu, G., and Yan, X. (2020). Online doctor recommendation with convolutional neural network and sparse inputs. *Computational Intelligence and Neuroscience.*, 2020:10.