

Relatório de criação do sistema de recomendação de medicamentos com base em sintomas utilizando a Similaridade do Cosseno

Júlia Perez Umbelino

Maio de 2025

Objetivo

O objetivo deste trabalho foi desenvolver um algoritmo de recomendação baseado no conceito de similaridade do cosseno, conforme estudado nas aulas de Álgebra Linear. Essa métrica é utilizada para medir o grau de semelhança entre dois vetores que representam frases e, nesse trabalho, listas de sintomas.

Neste projeto, a similaridade do cosseno foi aplicada na criação de um sistema capaz de recomendar medicamentos com base nos sintomas fornecidos pelo usuário. Essa abordagem auxilia usuários em farmácias, plataformas de saúde digital ou atendimentos remotos, oferecendo sugestões baseadas em dados. Contudo, ela não substitui uma consulta médica.

Tratamento de dados

O dataset utilizado foi retirado da plataforma Kaggle e convertido para o formato `.xlsx`, facilitando a leitura para bibliotecas como `pandas`. A planilha contém as seguintes colunas principais:

- **Medicine Name:** nome do medicamento;
- **Uses:** informações sobre os sintomas.

Como etapa de pré-processamento, os dados textuais foram padronizados para permitir a vetorização textual e o cálculo preciso da similaridade.

Algoritmos desenvolvidos

O sistema de recomendação foi construído utilizando o cálculo da similaridade do cosseno entre os sintomas fornecidos pelo usuário e os sintomas registrados no dataset. Algumas das etapas principais do algoritmo são:

- **Cálculo da Similaridade:** para cada entrada do usuário, calcula-se a similaridade do cosseno entre o vetor de entrada e os vetores dos sintomas do dataset.
- **Recomendação:** o sistema retorna os medicamentos cujos sintomas associados são mais semelhantes aos fornecidos, ordenando-os do mais ao menos relevante.

Resultados

Para testar a eficácia do algoritmo, foram inseridos sintomas reais (ex: “dor de cabeça e febre”). Logo depois, o sistema retornou uma lista com os medicamentos mais similares, considerando a base de dados fornecida.

O comportamento do sistema mostrou que sintomas comuns, como “tosse” e “febre”, retornaram medicamentos utilizados para gripes e resfriados. Já sintomas mais específicos retornaram medicamentos de uso mais direcionado, como antiinflamatórios ou analgésicos.

Esses resultados reforçam a eficácia da similaridade do cosseno na identificação de relações semânticas entre listas de sintomas, possibilitando sugestões coerentes.

Conclusão

O desenvolvimento do sistema de recomendação de medicamentos baseado na similaridade do cosseno demonstrou a eficácia desse método na identificação de relações entre sintomas e tratamentos.

Através do processamento adequado dos dados e da aplicação de técnicas de vetorização, foi possível construir um algoritmo funcional capaz de sugerir medicamentos com base em sintomas informados pelo usuário.

Os testes realizados indicaram que o sistema é capaz de fornecer recomendações coerentes, especialmente para sintomas comuns, reforçando o potencial da abordagem como uma ferramenta de apoio.

O projeto ainda oferece espaço para melhorias, como a ampliação da base de dados, integração com interfaces mais intuitivas e a aplicação de técnicas mais avançadas de Processamento de Linguagem.

Mesmo assim, os resultados obtidos já demonstram a aplicação prática dos conceitos estudados em sala, em especial a utilização da Álgebra Linear no desenvolvimento de soluções computacionais voltadas à saúde.