

**Evento:** XXXIII Seminário de Iniciação Científica ▾

CHATBOT ACADÊMICO BASEADO EM RETRIEVAL-AUGMENTED GENERATION PARA INDICAÇÃO DE ORIENTADORES EM PROGRAMAS DE PÓS GRADUAÇÃO¹

Isadora Beckmann², Fabricia Carneiro Roos Frantz³

¹ Projeto de pesquisa desenvolvido na Unijui financiado através do programa de fomento: Programa Institucional de Bolsas de Iniciação Científica do Conselho Nacional de Desenvolvimento Científico e Tecnológico - PIBIC/CNPq

² Bolsista; estudante do curso Engenharia de Software; Bolsista do programa de fomento: Programa Institucional de Bolsas de Iniciação Científica do Conselho Nacional de Desenvolvimento Científico e Tecnológico - PIBIC/CNPq

³ Professora do Programa de Pós-Graduação em Modelagem Matemática e Computacional (PPGMMC), Grupo de Pesquisa em Computação Aplicada (GCA) - UNIJUÍ

INTRODUÇÃO

Escolher um Programa de Pós-Graduação (PPG) é uma etapa fundamental na trajetória acadêmica e profissional de um estudante, mas o processo pode ser desafiador devido à ampla gama de opções disponíveis. Além da escolha do programa em si, a definição do orientador é um fator crucial, uma vez que a afinidade entre as linhas de pesquisa do estudante e do docente pode impactar diretamente no sucesso da pesquisa. Geralmente, essa seleção começa pela identificação da área de interesse do aluno, seguida por uma análise dos docentes que atuam nesse campo. Para isso, é necessário um exame detalhado dos currículos acadêmicos dos professores, que, no Brasil, estão disponíveis na Plataforma Lattes.

Com o intuito de tornar essa jornada mais ágil e assertiva, propõe-se o uso de um modelo computacional capaz de sugerir possíveis orientadores com base no perfil acadêmico e nos interesses do estudante. Esse sistema não apenas reduz o tempo de busca, mas também aumenta a precisão das recomendações, garantindo que o estudante encontre um docente cuja expertise esteja alinhada com seus interesses de pesquisa.

Este projeto tem como objetivo desenvolver um chatbot simples implementando um modelo de recomendação e incorporando técnicas de inteligência artificial e processamento de linguagem natural, utilizando os currículos acadêmicos dos docentes atuantes na UNIJUÍ. Com isso, busca-se otimizar a experiência do usuário, e elevar a qualidade das sugestões oferecidas, promovendo conexões mais eficientes entre estudantes e orientadores.



METODOLOGIA

A elucidação do sistema de recomendação de tutores foi realizada em etapas sucessivas, que englobam o estudo teórico, a análise de ferramentas, a implementação prática e a comprovação experimental. Ocorreu, inicialmente, um estudo detalhado sobre o Processamento de Linguagem Natural (PLN), tomando como base o livro Processamentos de Linguagem Natural (MARTINS et al., 2020), com o intento de embasar as decisões técnicas que envolvem o tratamento dos dados textuais.

A segunda etapa consistiu em revisar a literatura sobre Sistemas de Recomendação (SR), abrangendo desde os métodos mais tradicionais como filtragem colaborativa e baseada em conteúdo até abordagens mais modernas que incluem métodos de aprendizagem de máquina (MOTTA, 2012). Na mesma linha, as arquiteturas e as estruturas dos chatbots foram analisadas, visando entender como funcionam, para que são utilizados e quais as linguagens de programação que devem ser empregadas na proposta (BIRD et al., 2009).

Nesse sentido, também se pode destacar a dissertação de mestrado de Rehbein (2022), que propôs um modelo computacional para recomendação de orientadores em programas de pós-graduação stricto sensu. Esse trabalho foi relevante e inspirou diretamente o desenvolvimento inicial deste chatbot. Pela troca de ideias com o autor, surgiu a ideia de conceitualizar a solução em uma estrutura similar, valendo-se de estratégias de recomendação utilizadas no âmbito dos programas de pós-graduação.

A extração dos dados foi realizada através da coleta manual das informações presentes nos resumos dos currículos Lattes dos docentes da UNIJUÍ. A coleta englobou a síntese de dados como área de atuação, formação e linhas de pesquisa, utilizando-se do Google Gemini com o seguinte prompt: “Síntetize de maneira sucinta a área de estudo, atuação e linhas de pesquisa a seguir”.

Com estes dados organizados em arquivos de texto, foi iniciada a etapa de construção do sistema. A aplicação foi desenvolvida com base em um chatbot já existente, desenvolvido pelo professor Régis Rodolfo Schuch para a disciplina de Big Data, que foi adaptada às necessidades do projeto. Para o novo modelo, foram utilizadas tecnologias como Python, Streamlit, LangChain, ChromaDB, Ollama e GPT4All.

A arquitetura utilizada foi a Retrieval-Augmented Generation (RAG), que possibilita gerar respostas fundamentadas em dados reais. Os textos extraídos dos



currículos foram quebrados em segmentos (chunks) com o uso do LangChain, transformados em embeddings com o modelo GPT4All Embeddings e então arquivados localmente em um banco vetorial com ChromaDB. A partir da descrição fornecida pelo usuário, o sistema cria um vetor correspondente e faz a busca semântica para encontrar o(s) docente(s) que mais compatibiliza, retornando uma resposta embasada pelo conteúdo armazenado.

A interface foi criada com Streamlit, o que viabilizou ter uma aplicação web funcional e interativa que pode ser acessada diretamente pelo navegador. Adicionalmente, utilizou-se o Ollama para permitir a execução dos modelos de linguagem em modo local, sem depender de servidores externos.

RESULTADOS E DISCUSSÃO

O sistema proposto mostrou-se funcional durante os testes iniciais. Ao simular a introdução dos interesses acadêmicos por um estudante fictício, o chatbot conseguiu identificar o pedido e fornecer de volta uma recomendação compatível com as áreas mencionadas. Em um deles, os interesses em redes e sistemas de informação levaram o chatbot a propor o professor Tiago Mallmann Rohde, provando a eficácia do modelo para aproximar descrições semânticas aos pesquisadores de linha de pesquisa dos docentes, como é possível visualizar na imagem a seguir.

Assistente de Pós-Graduação UNIJUÍ

Descreva suas áreas de interesse para uma Pós-Graduação e lhe retornamos um Orientador (isso pode demorar um pouquinho)

Possuo interesse em áreas de tecnologia, ciência da computação com foco em redes de computadores e sistemas de informação, bem como computação aplicada

Mostrar Resultado

Orientador sugerido:

assistant: Com base nos interesses acadêmicos do aluno em tecnologia, ciência da computação com ênfase em redes de computadores e sistemas de informação, bem como computação aplicada, o professor Tiago Mallmann Rohde parece ser o mais compatível para sua orientação. Ele possui uma ênfase clara na área de redes de computadores e análise de sistemas, além de ter experiência em ciência da computação e computação aplicada. Essa combinação se encaixa melhor com os interesses do aluno, fornecendo um ambiente acadêmico e pesquisador que pode lhe proporcionar mais oportunidades de desenvolvimento e sucesso na área que deseja explorar.

Figura 1 - Chatbot funcionando



A interface desenvolvida com Streamlit mostrou-se adequada, contando com uma navegação intuitiva e um retorno conciso das recomendações. O uso do ChromaDB como armazenamento vetorial e o LangChain para tratar os documentos contribuiu com a efetivação da busca semântica, enquanto o modelo GPT4All, executado localmente, servido pelo Ollama, realizou as respostas sem a necessidade de conexão a servidores externos. Este ponto reforça a viabilidade das soluções locais e seguras para o uso em contextos acadêmicos.

Apesar do resultado positivo, saliente-se como limitação a restrição da base de dados da UNIJUÍ (devido à coleta manual e limitada dos currículos). Contudo, o sistema se mostra promissor e com possibilidade de expansão para outras instituições e bases de dados automatizadas.

CONSIDERAÇÕES FINAIS

O ChatBot Orientador é um bom exemplo do uso da Inteligência Artificial e do PLN como apoio à orientação acadêmica em Programas de Pós-Graduação de forma simplificada e intuitiva. A referida solução conseguiu integrar tecnologias como LangChain, ChromaDB, Streamlit, GPT4All Embeddings e Ollama de forma eficaz e estabelecer recomendações personalizadas a partir da correlação dos interesses do usuário e os perfis dos docentes.

A utilização da arquitetura RAG possibilitou a união entre busca semântica e geração de respostas corretas, enquanto a realização local do modelo colaborou com a privacidade e a independência da tecnologia utilizada, a interface intuitiva e o bom desempenho da recomendação evidenciam a viabilidade prática da solução em ambientes acadêmicos.

Este projeto trouxe à tona o potencial da utilização de modelos de linguagem locais e de bancos de vetores como ferramentas acessíveis e eficazes ao suporte das decisões pertinentes ao percurso formativo dos estudantes, constituindo uma contribuição relevante no uso de tecnologias emergentes na personalização dos serviços educacionais.

Palavras-chave: Chatbot. Pós-Graduação. Orientação. Arquitetura RAG. Inteligência Artificial.

AGRADECIMENTOS

À Professora Doutora Fabrícia Carneiro Roos Frantz, pela orientação capacitada, pela disponibilidade e pelas contribuições significativas ao desenvolvimento dessa pesquisa.



Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pelo fomento à pesquisa, por meio da concessão da bolsa de Iniciação Científica.

REFERÊNCIAS BIBLIOGRÁFICAS

BIRD, Steven; KLEIN, Ewan; LOPER, Edward. *Natural language processing with Python: analyzing text with the Natural Language Toolkit*. 2. ed. Sebastopol: O'Reilly Media, 2009.

MARTINS, Júlio S. et al. *Processamentos de linguagem natural*. Porto Alegre: SAGAH, 2020. E-book. 1. ed. 1 recurso online (p. 16). ISBN 9786556900575.

MOTTA, Claudia et al. Sistemas de recomendação. In: PIMENTEL, Mariano; FUKS, Hugo (org.). *Sistemas colaborativos*. 1. ed. Rio de Janeiro: Elsevier, 2012. cap. 15, p. 230–244.

REHBEIN, Matheus H. *Modelo computacional para recomendação de potenciais orientadores em programas de pós-graduação stricto sensu: um estudo de caso dirigido ao COMUNG*. 2022. 103 f. Dissertação (Mestrado em Computação Aplicada) – Universidade Regional do Noroeste do Estado do Rio Grande do Sul, Ijuí, 2022.

SCHUCH, Régis Rodolfo. Assistente IA. Disponível em: <https://github.com/regis-schuch/disciplinas/tree/main/BigData/AssistenteIA>. Acesso em: 24 jul. 2025.