

ChatBot sobre o vestibular da Unicamp de 2025

Introdução

Este relatório descreve a implementação de um chatbot utilizando o framework Streamlit e a biblioteca LangChain. O objetivo é criar um assistente virtual capaz de responder dúvidas sobre o edital do vestibular da Unicamp de 2025. O chatbot utiliza um modelo de linguagem natural que busca informações em documentos relevantes para gerar respostas contextuais.

Implementação

Modelo de Linguagem

generate_chain: Esta função configura a cadeia de recuperação e processamento de documentos. Ela envolve a criação de embeddings com o modelo HuggingFace, a divisão dos documentos em chunks (partes menores), e a utilização de um sistema de recuperação baseado em vetores para buscar informações relevantes.

Interface com o Usuário

A interface do chatbot é desenvolvida com Streamlit, apresentando duas abas:

Chat: Permite a interação direta com o chatbot.

Referências: Exibe o conteúdo do edital para consulta direta.

Fluxo de Funcionamento

Recuperação de Documentos

O sistema carrega o conteúdo do edital do vestibular a partir de um arquivo de texto gerado pela função **save_pdf_as_txt**. Este conteúdo é dividido em segmentos menores utilizando a classe de segmentação de texto **RecursiveCharacterTextSplitter**, permitindo uma recuperação mais eficiente das informações.

Geração de Respostas

As perguntas dos usuários são processadas pela cadeia de recuperação, que utiliza um modelo de linguagem da OpenAI para gerar respostas baseadas no contexto encontrado nos documentos. A função principal deste processo é o **conversational_retrieval_chain**, que combina as capacidades de recuperação de documentos com a geração de respostas.

Avaliação de Desempenho

O código inclui uma seção para geração e validação de dados, que serve para avaliar a precisão e a relevância das respostas fornecidas pelo chatbot. As perguntas são pré-definidas e as

respostas são geradas e validadas automaticamente sem uso de contexto ou referência, com os resultados armazenados em arquivos JSON.

ChatGPT

O uso de uma chat bot para o desenvolvimento do projeto teve várias aplicações. Dado o contexto de bibliotecas e abstrações desconhecidas, realizei a busca da sintaxe e funcionalidades ao lado da documentação oficial.

Conclusão

Este chatbot é uma ferramenta robusta para ajudar candidatos ao vestibular da Unicamp em 2025 a obterem informações detalhadas e precisas diretamente do edital. A implementação faz uso de técnicas de processamento de linguagem natural e recuperação de informações, garantindo que as respostas sejam contextualmente relevantes e baseadas em dados reais.

Próximos Passos

Deploy em Cloud: Realizar o deploy do chatbot em uma plataforma cloud, como Azure, para garantir escalabilidade e acesso global.

Melhoria da Avaliação: Implementar uma avaliação mais sofisticada que considere o contexto, permitindo uma análise mais precisa da relevância das respostas geradas.

Otimização de Performance: Testar e ajustar o desempenho do chatbot, especialmente em cenários com alto volume de consultas, incluindo como cachear o modelo e embeddings.