

DESENVOLVIMENTO DE CHATBOT PARA GERAÇÃO DE DOCUMENTOS JURÍDICOS USANDO O GPT-4

Misael Possidonio de Souza¹; Júnia Cristina Ortiz Matos²

¹ Bolsista; Centro de Competência em Inteligência Artificial - SENAI CIMATEC; misael.geof@gmail.com;

² Centro Universitário SENAI CIMATEC; Salvador - BA; junia.matos@fieb.org.br;

RESUMO

Este estudo apresenta uma metodologia para contornar as limitações da Inteligência Artificial Generativa no processo de elaboração de documentos. Um chatbot avançado foi desenvolvido, aproveitando a capacidade do GPT-4 e a biblioteca langchain para facilitar o acesso e a geração de informações a partir de documentos jurídicos e técnicos. Por meio da integração com a Azure Open AI, são utilizados serviços como Document Intelligence e AI Search. Além disso, o chatbot é estruturado em uma arquitetura dividida entre Front End e Back End, utilizando frameworks como React e Next.js para a interface do usuário, e Flask para a API do servidor. O chatbot não apenas consulta informações em bases de dados, mas também gerencia o armazenamento do histórico de conversas, permitindo a geração de documentos complexos em poucos minutos. A partir dos testes realizados, identificou-se a possibilidade de expansão da base de conhecimento acessível do modelo pré-treinado, incorporando detalhes específicos de casos judiciais, por exemplo. Otimiza-se, assim, a interação do usuário e a eficácia na criação de documentos corporativos relevantes.

PALAVRAS-CHAVE: GPT-4, Langchain, Azure Open AI, Chatbot Jurídico.

1. INTRODUÇÃO

A competição entre empresas no mercado atual tem sido cada vez mais impactante levando as indústrias a buscarem na tecnologia resultados melhores. A otimização de processos abrange várias etapas e metodologias tomadas para aumentar a eficiência em seus processos trazendo como consequência um bom desempenho geral do negócio, entre outros benefícios como a redução de custos, gerenciamento dos resultados e dos riscos, reduzindo a margem de erros¹.

Por outro lado, estamos presenciando o papel da Inteligência Artificial (IA) em transformar o sucesso empresarial, sendo que a área de Processamento de Linguagem Natural (PLN) permite que as máquinas consigam compreender e interagir com a linguagem humana, impactando nas relações com cliente, colaboradores e o mercado como um todo. Neste contexto, surgiu a Inteligência Artificial Generativa (IA Generativa)¹, a qual refere-se a sistemas de IA que são capazes de criar conteúdo novo e original, podendo variar desde texto, imagens e música até designs de produtos e códigos de software. Uma aplicação particularmente notável é o aprimoramento da comunicação entre empresas e clientes através do uso de chatbot e assistentes virtuais². São sistemas capazes de entender e responder às perguntas de forma rápida e eficaz, oferecendo um serviço de atendimento ao cliente mais eficiente e gratificante.

Paralelamente, tem se notado nos últimos anos no mercado a tendência de empresas consumirem serviços de Computação em Nuvem e Inteligência Artificial como Serviço (IA as a Service - AlaaS)³, os quais são dois conceitos poderosos que estão remodelando o panorama tecnológico atual. Eles permitem que empresas e indivíduos acessem recursos computacionais avançados e capacidades de inteligência artificial sem o custo e a complexidade de desenvolver e manter a infraestrutura de TI própria.

Neste estudo, mostramos como contornar a limitação da IA generativa na quantidade de tokens e demonstramos um chatbot avançado que emprega assistentes virtuais para a criação de documentos corporativos. Os recursos utilizados para o desenvolvimento foram provisionados na plataforma de Computação da Nuvem Microsoft Azure. Esse sistema segue um fluxograma detalhadamente planejado, estabelecendo uma interação dinâmica com o usuário final por meio de questionamentos estratégicos ao longo do processo. Adicionalmente, o chatbot é projetado para consultar informações em uma base de dados de forma autônoma sempre que necessário, garantindo assim a geração de documentos precisos e personalizados de acordo com as demandas específicas do usuário.

2. METODOLOGIA

Este estudo destaca a criação de uma funcionalidade, aproveitando a expertise do ChatGPT, especialmente a versão GPT-4, em processamento de linguagem natural para facilitar o acesso a informações de documentos jurídicos e técnicos em formato PDF. Para enfrentar o desafio de processar textos extensos, incorporamos o uso da biblioteca Python langchain, que é especialmente útil para manipular grandes volumes de texto. Utilizando a plataforma Azure Open AI, juntamente com módulos do Document Intelligence e o

reforço da langchain, foi viável empregar o GPT para extrair informações pertinentes desses documentos e gerar documentos relevantes⁴.

Langchain é uma biblioteca open source desenvolvida para facilitar a integração de Modelos de Linguagem de Grande Escala (LLMs), com fontes de dados externas, uma funcionalidade crucial para os objetivos deste projeto. Através do processo de segmentação de documentos em trechos de texto e sua subsequente indexação em um banco de dados vetorial, langchain proporciona uma solução eficaz para essa expansão de conhecimento, permitindo que o GPT-4 acesse detalhes precisos sobre documentos judiciais específicos¹. Contudo, foi necessário enfrentar a limitação imposta pela capacidade de processamento do GPT-4, que restringe o número de tokens por consulta. Para contornar essa barreira, os documentos foram divididos em múltiplos segmentos de tokens, com a equivalência de que cada palavra representa, em média, 8,192 tokens. A integração da langchain permitiu uma manipulação mais eficiente desses segmentos, otimizando o processamento e a análise de documentos extensos, permitindo enviar arquivos de até 30,000 tokens.

Nesse trabalho, não realizamos o treinamento do modelo utilizado pois o GPT-4 é um modelo pré-treinado disponibilizado pela OpenAI, que já passou por um extenso processo de aprendizado com um grande volume de dados de texto extraídos da internet. Usamos as APIs dos assistentes disponibilizadas no Azure Open AI para criar, configurar assistentes virtuais que são acionados de acordo com a necessidade durante a interação do usuário e chatbot. Para os assistentes foram criadas funções de busca na base de dados contendo os documentos, funções de reasoning contendo as instruções do assistente para verificar as informações mais importantes e necessárias para redigir determinado documento bem como acessar aos modelos do documento disponível e indagar sempre ao usuário quando for necessário.

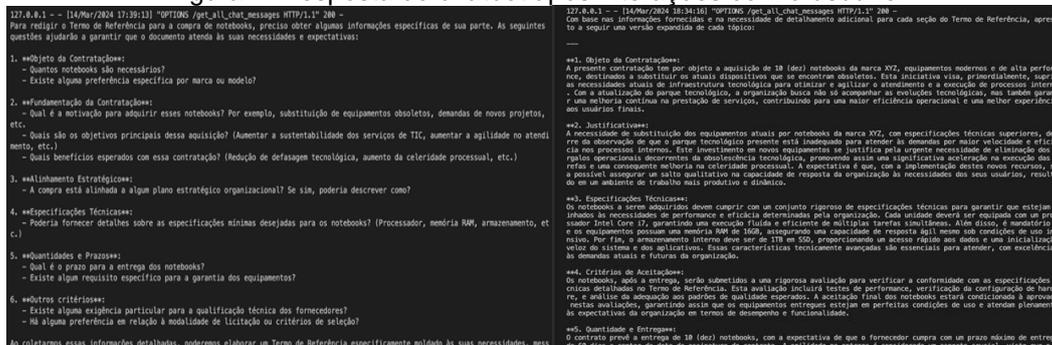
O desenvolvimento do chatbot foi estrategicamente segmentado em duas áreas principais: Front End e Back End. Na esfera do Front End, desenvolvemos módulos dedicados à interação com o usuário. Essa interface é encarregada de armazenar temporariamente (em cache) as informações fornecidas pelo usuário, encaminhá-las ao Back End através de uma API REST, gerenciar possíveis erros e apresentar as respostas processadas pelo servidor ao usuário. Para construir essa interface de usuário, optamos inicialmente pelo uso dos frameworks React e Next.js, conhecidos por sua eficácia no desenvolvimento de interfaces dinâmicas em JavaScript. Implementamos uma interface simplificada que inclui um campo de texto, permitindo que o usuário insira suas informações. Do lado do servidor, o Back End foi arquitetado para se comunicar com o Front End via API REST e integrar-se com serviços de Inteligência Artificial como Serviço (IA as a Service) oferecidos pela Azure, além de gerenciar o armazenamento do histórico de conversas em um banco de dados. Para isso, utilizamos o Flask, um framework Python leve e eficiente, para criar a API que intermedia a comunicação com os serviços Azure e o banco de dados.

3. RESULTADOS E DISCUSSÃO

Para demonstrar o chatbot com documentos reais, utilizamos alguns documentos público disponível através do link <https://www.tjba.jus.br/transparencia/wp-content/uploads/2022/05/Termo-referencia-3.pdf>, que consiste em um modelo de Termo de referência de compra de equipamentos eletrônicos.

Os resultados obtidos entre a interação do usuário com o chatbot inteligente são mostrados a seguir:
Usuário: "Gere um TR de compra de notebooks"

Figura 1: Resposta do chatbot após interações com o usuário.



Fonte: Próprio Autor.

Como podemos ver, o assistente acionado elabora várias perguntas de acordo com o conteúdo do documento e traz no output os pontos a serem acrescentados pelo usuário para assim prosseguir com a elaboração do documento. Na sequência temos: Usuário: "Serão necessários 10 novos notebooks, marca XYZ, processador

intel i7, 16gb de ram com 1Tb de SSD. substituição de equipamento com o objetivo de aumentar a agilidade do atendimento. Espera-se aumentar a celeridade processual principalmente. Gere um TR mais detalhado"

E no final o chatbot indaga se há um plano estratégico organizacional, ou se há suporte técnico ou serviço pós-venda como suporte técnico. Na prática, o usuário interagiu com o sistema apenas três vezes, podendo prosseguir com a interação até ficar satisfeito com a versão do documento gerado que é no caso o Termo de Referência. Todo o processo até esse ponto levou 10min incluindo o tempo levado para leitura por parte do usuário.

4. CONSIDERAÇÕES FINAIS

O desenvolvimento deste chatbot inteligente que integra a capacidade do GPT-4 Turbo com a biblioteca Langchain e os serviços da Azure Open AI, demonstra um avanço significativo na automatização e na eficiência da criação de documentos corporativos a partir de documentos jurídicos e técnicos disponíveis em PDF. O projeto não apenas otimiza a interação entre o usuário e o sistema, mas também eleva a precisão e a personalização dos documentos gerados, refletindo diretamente na agilidade e eficácia dos processos organizacionais.

A aplicação prática do chatbot, demonstrada através da elaboração de um Termo de Referência para compra de notebooks, evidencia o potencial deste sistema para acelerar a criação de documentos complexos. Este estudo abre caminhos para futuras pesquisas e desenvolvimentos na área de assistentes inteligentes, sugerindo a exploração de novas funcionalidades e a integração com outros sistemas e bases de dados.

Em conclusão, a implementação do chatbot jurídico com o apoio de tecnologias de ponta em IA e computação em nuvem apresenta uma inovação promissora para o campo do processamento de documentos e gestão do conhecimento, marcando um avanço importante na maneira como as organizações interagem com a informação e otimizam seus processos.

5. REFERÊNCIAS

¹SILVA, Cláudio Albués Alves da. Desenvolvimento de Assistente Jurídico Inteligente utilizando o modelo GPT-3. Salvador, BA: Universidade Federal da Bahia, 2023.

²ZHU, Q.; LUO, J. Generative Pre-Trained Transformer for Design Concept Generation: An Exploration. Dubrovnik: International Design Conference, 2022. Disponível em: <https://www.cambridge.org/core/services/aop-cambridge-core/content/view/41894D82DCBC0610B5B6E68967B7047F/S2732527X22001857a.pdf/generative-pre-trained-transformer-for-design-concept-generation-an-exploration.pdf>. Acesso em: 10 de janeiro de 2024.

³FERRARO, Gabriela et al. Automatic Extraction of Legal Norms: Evaluation of Natural Language Processing Tools. Springer International Publishing, 2020. Disponível em: <http://library.lol/main/2241D3ECDCC3FC327F9D3DBAA2CABBC6>. Acesso em 20 de janeiro de 2024.

⁴HERYANDI, A. Developing Application Programming Interface (API) for Student Academic Activity Monitoring using Firebase Cloud Messaging. Bristol: IOP Publishing, 2018. Disponível em: <https://iopscience.iop.org/article/10.1088/1757-899X/407/1/012149/pdf>. Acesso em: 08 de Fevereiro de 2024.