XXXI Seminário de Iniciação Científica da UFERSA

Núcleo de Avaliação: Núcleo II
Área temática: Ciências Exatas e da Terra
Área do Conhecimento: Ciência da Computação

Assistente Farmacêutico Virtual Baseado em RAG

Pablo Lucas de Souza Ernesto, Silvio Roberto Fernandes, Sérgio Roberto Fernandes de Araujo, Paulo Gabriel Gadelha

A baixa adesão ao tratamento farmacológico e o acesso limitado a informações confiáveis sobre medicamentos motivaram o desenvolvimento do FarmacIA, um assistente farmacêutico virtual baseado em Inteligência Artificial voltado ao apoio farmacoterapêutico. O sistema é composto por um front-end móvel, que possibilita a interação do paciente com a IA, e um back-end baseado na abordagem de Geração Aumentada por Recuperação (Retrieval-Augmented Generation – RAG), responsável por gerar respostas fundamentadas exclusivamente em bulas oficiais da Anvisa. Este trabalho teve como foco o desenvolvimento e avaliação do back-end do FarmacIA. A arquitetura foi implementada em Python (Flask), integrando módulos de pré-processamento e indexação de bulas, realizadas a partir da API pública da Anvisa, com armazenamento local e indexação vetorial via ChromaDB e embeddings Sentence-BERT. No primeiro acesso, a bula é obtida, segmentada em seções e convertida em chunks indexados; em consultas posteriores, duas buscas semânticas são executadas — uma para identificar a seção relevante e outra para recuperar o conteúdo específico utilizado na geração da resposta. A geração foi conduzida por modelos de linguagem de pequeno porte executados localmente (Llama 3.2:3b, Gemma 3:4b e Qwen 2.5:3b), preservando a privacidade e independência de serviços externos. Para avaliação, elaborou-se um conjunto de 45 perguntas em três níveis de dificuldade (fácil, médio e difícil), geradas pelo ChatGPT e validadas por um farmacêutico, contendo respostas esperadas e seções correspondentes da bula. As respostas dos modelos foram comparadas ao ground truth e avaliadas segundo a técnica LLM-as-a-judge, por meio do ChatGPT 4o, Gemini 2.5 Flash e de um avaliador humano (farmacêutico), com pontuações discretas (1, 3, 5, 7 e 10) em precisão factual, completude e clareza. Os resultados mostraram que, embora todos os modelos tenham obtido alta precisão em perguntas simples, o desempenho decaiu em questões de maior complexidade, revelando limitações na inferência e integração de informações clínicas. Conclui-se que a arquitetura RAG implementada é viável para consultas farmacêuticas de baixa complexidade, oferecendo rastreabilidade, segurança e execução local, além de estabelecer uma base sólida para os demais módulos do sistema FarmacIA desenvolvidos em trabalhos paralelos.

Palavras-chave: Inteligência Artificial; RAG; LLM; Assistente Farmacêutico; Bulas de Medicamentos.

Agência financiadora: PICI-UFERSA

XXXI Seminário de INICIAÇÃO CIENTÍFICA DA UFERSA

Campus: Mossoró

UFERSA 20 anos