KALASALINGAM
ACADEMY OF RESEARCH AND EDUCATION
(DEEMED TO BE UNIVERSITY)
Under sec. 3 of UGC Act 1956. Accredited by NAAC with "A++" Grade
Anand Nagar, Krishnankoil, Srivilliputtur (Via), Virudhunagar (Dt) - 626126, Tamil Nadu | info@kalasalingam.ac.in | www.kalasalingam.ac.in
CENTRE FOR DISTANCE AND ONLINE EDUCATION (CDOE)

Name of the Learner: Anandu O A **Register number: 91241460036**

Name of the Program: M.Sc. DS **Semester IV**

Name of the Course: Project work **Batch: Jan 2024**

Course Code: 225MDS4201-Project (External) **Regulation: 2022**

Project Title: Medical Knowledge Chatbot using RAG with MLOps & DevOps Integration

Abstract

This project aims to develop a medical knowledge chatbot that provides accurate, evidence-based responses by retrieving information from trusted medical sources rather than generating answers solely from a language model. The system uses a Retrieval-Augmented Generation (RAG) approach, where user queries are matched with relevant medical documents such as clinical guidelines and research articles, and the retrieved content is used to generate reliable responses with proper context. This helps reduce hallucinations and improves the trustworthiness of medical information delivery.

The chatbot is built using Python and leverages data science and natural language processing techniques such as text preprocessing, embeddings, semantic search, and large language models. A vector database is used to store and retrieve medical document embeddings efficiently. The system also integrates DevOps and MLOps practices including Docker-based containerization, automated data ingestion, model and index versioning, and continuous monitoring to ensure scalability, reliability, and ease of deployment.

This tool is intended for educational and medical decision-support purposes only and does not provide diagnosis or treatment recommendations. It can be useful for medical students, healthcare professionals, and researchers seeking quick access to reliable medical knowledge. The project highlights the importance of combining data science with DevOps practices to build maintainable and production-ready AI systems in the healthcare domain.

This system includes four main modules:

- User Interface Module
- Document Ingestion and Preprocessing Module
- Retrieval and Chatbot Engine Module
- Monitoring and Update Module

Tools & Technologies Used:

- **Frontend (Interface):** Streamlit or Web-based UI
- **Backend (Processing & Storage):** Python, Vector Database (FAISS / Chroma)
- **Machine Learning:** Sentence-BERT / BioBERT, Open-source LLM
- **DevOps & MLOps:** Docker, GitHub Actions, MLflow (optional)
- **Database:** PostgreSQL / SQLite for logs and metadata