SYNOPSIS

Date of Joining: 3rd July 2024

Program: B.Tech (CSE) AIML

Subject: Project from July 3rd, 2024 to August 31st, 2024

Student Information:

Name: DINU YADAV

Enrollment No.: 220160223006

Residential Address: [Your Address]

Contact No.: [Your Contact Number]

E-mail Id: [Your Email]

Organization Information:

Name: Yugasa Software Labs Pvt. Ltd.

Address: 3rd Floor, Yugasa Software Labs, Tower-B, UNITECH CYBER PARK, Sector 39, Gurugram, Haryana

Department Allotted: Data Science and Machine Learning

Project Head:

Name: Mr. Vivek Mittal

Contact No(s): 8800332227

E-mail Id: vivek@yugasa.com

Major Functions of the Organization:

- Custom Software Development
- Mobile Application Development
- Web Development Services

- AI and Machine Learning Solutions
- E-commerce Solutions
- Cloud Services
- Product Consulting
- UI/UX Design

Major Functions of the Department Allocated:

- Data Collection and Management
- Model Development and Deployment
- Statistical Analysis and Insights
- Model Optimization and Experimentation
- Automation and Decision Support
- Research and Development

Responsibilities during Project:

- Develop and implement a Retrieval-Augmented Generation (RAG) pipeline.
- Collect, clean, and manage internal policy documents.
- Develop vector embeddings for document chunks and queries.
- Integrate a Large Language Model (LLM) for answer generation.
- Design semantic search workflow for efficient retrieval.
- Build a user-friendly chatbot interface.
- Evaluate chatbot performance and accuracy.
- Optimize pipeline for speed and relevance.
- Document the implementation for knowledge transfer.

Nature of the Project/Work:

This project aims to design and implement a smart chatbot using Retrieval-Augmented Generation (RAG) to help employees access information from internal policy documents efficiently. By combining semantic search with an LLM, the system retrieves relevant context and generates accurate, human-like responses.

Key Objectives:

- Develop a robust end-to-end RAG pipeline using Python, PyTorch, and Hugging Face Transformers.
- Import and preprocess internal policy documents.
- Implement vector-based similarity search.
- Use an LLM for precise, context-aware responses.

- Build a user-friendly chat interface.
- Run the pipeline locally for privacy and speed.

Tools and Libraries Used:

- PyMuPDF
- SpaCy
- sentence-transformers
- FAISS or similar vector database
- Hugging Face Transformers
- PyTorch
- Gradio/Streamlit

Expected Outcome:

A functional RAG-based chatbot that answers employee queries using internal policy documents, returns context and sources, runs locally for privacy, and improves information access across the organization.

A Brief Summary of the Project/Work:

This project delivers an AI-powered chatbot for company employees by combining information retrieval with language generation. The system processes internal documents into embeddings and uses semantic search to retrieve relevant content, allowing the LLM to generate context-rich, accurate responses. This improves how employees access critical information.