

Mihir Kapile

USA | mihirkapile@gmail.com | +1 6693880479 | linkedin.com/in/mihirkapile | github.com/MihirKapile

Professional Summary

Software Engineer with a Master's in Computer Science, skilled in developing LLM applications, Java-based web apps, and ReactJS. Expertise in Prompt Engineering, RAG systems, and fine-tuning models like BERT. Proficient in building scalable APIs using Spring Boot, Flask, and Django, and deploying applications on AWS and Kubernetes.

Education

Northeastern University, Masters of Science in Computer Science Sept 2022 – May 2024

- GPA: 3.87/4.0

MIT ADT University, Bachelors of Technology in Computer Science Aug 2018 – July 2022

Experience

Software Engineer, Ipser Labs LLC , CA June 2023 – Current

- Developed a web application for universities and industries to sell research papers and patents, using ReactJS, NodeJS, and MongoDB, serving 100+ users. Integrated a web crawler that analyzed 500+ external websites for patent interest.
- Led the development of a website builder platform with Java Spring Boot, ReactJS, and MySQL, improving deployment efficiency by 30% and enhancing collaboration among 5+ cross-functional teams.
- Designed and implemented a customer assistance chatbot using LangChain and RAG, automating responses for 200+ customer queries per week and improving response time by 40%.

Projects

Agentic Voice Assistant with Task Management and Search Capabilities github.com/VoiceAssistant

- Built an agentic application for automating tasks like setting alarms, scheduling events, sending emails, and performing searches, leveraging real-time voice input to reduce manual execution time by 60%.
- Tools Used: Python, LangChain, GPT-4, Speech Recognition

Resume Changes Suggestion Chrome Extension github.com/ResumeExt

- Designed a Chrome extension that automated resume optimization, comparing job descriptions in real-time and suggesting targeted improvements.
- Tools Used: LangChain, Pinecone Vector DB, GPT-4 ,Javascript ,HTML ,CSS , Python

Hindi Chatbot for Law and Medical Assistance github.com/HindiBot

- Engineered a chatbot that provided critical legal and medical information to Hindi-speaking users, increasing accessibility for underserved communities. Optimized response accuracy by 85% using fine-tuned LLMs.
- Tools Used: Ollama LLMs, GPT-4 ,Javascript ,HTML ,CSS , Python, PyTorch, LoRA, QLoRA

Knowledge Graph Q&A and RAG with Tabular Data

- Architected a chatbot capable of extracting insights from large tabular datasets, reducing manual data lookup time by 60%. Implemented a flexible knowledge graph that improved response accuracy by 30%, enhancing decision-making efficiency.
- Tools Used: Neo4j Graph DB, LangChain, GPT-4, Gradio, Python

Technologies

Languages: C, C++ , C#, Java, Python, HTML, CSS, JavaScript, R, PHP, Go, Rust, PHP

Databases: MySQL, Oracle, MongoDB, GraphQL, ChromaDB, PineCone Vector DB

Frameworks: jQuery, Django, Flask, React, Angular, ExpressJS, NextJS, LangChain, LangGraph, LangFlow, LangSmith, LlamaIndex, RAG, RAGAs, PyTorch, TensorFlow, Spring Boot

Cloud and Deployment: AWS, Azure, Docker, Kubernetes