ZAINA PASHA

+91 9360860772 | zaina.bint.pasha@gmail.com | https://www.linkedin.com/in/zaina-pasha-34a875227/ | https://github.com/ZainaPasha

PROFILE SUMMARY

AI Engineer with hands-on experience delivering end-to-end Generative AI solutions, including RAG pipelines, LLM integration, backend APIs, and production-oriented testing. Proficient in Python, FastAPI, and vector databases, with exposure to databases, data preprocessing, model integration, deployment workflows, and documentation for handover. Comfortable taking ownership of use cases and collaborating across teams in fast-paced environments.

TECHNICAL SKILLS

- **Programming:** Python(Proficient), JavaScript, Java, C++
- **GenAI & LLMs:** Prompt Engineering, Embeddings, RAG Pipelines, HuggingFace embeddings, Model integration & inference workflows
- **Frameworks & Libraries:** FastAPI, LangChain, TensorFlow, PyTorch
- **Databases & Data Systems:** PostgreSQL, MongoDB, Vector Databases: Pinecone, SQL
- **Backend & APIs:** FastAPI, REST APIs Development, JSON-based integrations, Flask (basic)
- **Cloud & DevOps:** AWS (basic), Azure (basic – learning), Docker (basic), CI/CD fundamentals
- **Tools:** Git, GitHub, VS Code

PROJECTS

RAG-Based Resume Analyzer | Python · FastAPI · Pinecone · LangChain · Vector Search

- Designed and deployed an end-to-end RAG pipeline using LangChain, HuggingFace embeddings, and Pinecone.
- Built FastAPI backend services for ingestion, preprocessing, embedding generation, and query orchestration.
- Improved retrieval quality through data validation and preprocessing pipelines.
- Added structured JSON responses and error handling to support robust API consumption in production workflows.

AI-Powered Healthcare Referral System | Next.js · Node.js · TensorFlow · Python · Appwrite

- Developed a production-oriented AI system integrating a Python-trained DNN with a Node.js inference service.
- Implemented data preprocessing and feature validation pipelines to ensure reliable model inputs.
- Designed REST APIs to support real-time AI-driven decision workflows under constrained latency.
- Took ownership of model integration, API testing, and deployment readiness.

AI Interior Designer | Next.js · Node.js · ControlNet · Pytorch · PostgreSQL

- Integrated ControlNet-based generative models into backend services to support AI-powered image generation.
- Designed APIs to handle large payloads and inference workflows, improving stability and memory usage.
- Collaborated across frontend and backend layers to deliver a user-facing GenAI experience.

INTERNSHIPS

Programmer Analyst Intern - Cognizant Technology Solutions   Jan 2025 - July 2025
- Developed REST APIs using Node.js and improved backend performance for e-commerce workflows.
- Participated in SDLC, debugging, logging, testing, and code reviews in an Agile environment.
- Identified performance bottlenecks, resolved failures, and improved API throughput.

EDUCATION

B S Abdur Rahman Crescent Institute of Science and Technology, India   August 2021 - May 2025
B.Tech. Computer Science and Engineering   **CGPA: 9.81**

CERTIFICATIONS & PUBLICATION

- Oracle Cloud Infrastructure 2025 Certified AI Foundations Associate (2025)
- AWS S3 Basics, *Coursera Project Network (2024)*
- Publication: Prediction of Brain Diseases Using Machine Learning Models: A Survey, *Springer (2023)*