

Rutvi Rathod

📍 San Jose CA 95134 | 📞 669-272-4832 | ✉️ raturathod53@gmail.com | github:rutvi503

Technical Skills

Languages : JavaScript, Java, python
Framework : Springboot, Langchain, Flask, Django
Database : PostgreSQL, MongoDB, Pinecone
Cloud : Microsoft Azure, AWS, Docker, Kubernetes
Front end : React, HTML, CSS
Backend : NodeJs, Fast API, RestAPI, GQI
Developer Tools: Visual Studio Code, Jupyter Notebook, Tableau, PyCharm, Git, Agentic AI, MCP server

Education

Chhotubhai Gopalbhai Patel Institute of Technology
June 2020 - May 2024
Bachelor of Science in Software Engineering
Surat, Gujarat

Experience

Software Engineering Intern
Dec 2023 - June 2024
Ridaro Inc

- Developed and maintained scalable web applications using Java (Spring Boot) for backend services and React for frontend interfaces, enhancing application reliability, performance, and user experience.
- Designed and implemented RESTful APIs, integrating with React front-end applications for real-time user interactions.
- Built secure authentication systems featuring OTP verification and face recognition, enhancing user security and simplifying login processes.
- Managed data storage and retrieval with PostgreSQL and MongoDB, optimizing the system for large-scale data operations.
- Deployed cloud-based applications on AWS, ensuring high availability, scalability, and fault tolerance while optimizing API response times by 30%.
- **Impact:** Enabled Ridaro Inc. to achieve improved system performance and operational efficiency through scalable and secure backend systems. The features developed contributed to an enhanced user experience, robust data management, and streamlined workflows, driving productivity and customer satisfaction.

Projects

Attendance System Using Face Recognition

Dec 2023 - June 2024

Tech Stack: Springboot, Python, Django, FastAPI, PostgreSQL, AWS

- Developed an innovative face recognition-based attendance system leveraging Python libraries for facial data processing and analysis.
- Built real-time APIs using FastAPI to enable high-performance facial recognition and seamless data handling.
- Designed secure user authentication workflows and face enrollment processes to ensure accurate and reliable attendance tracking.
- Integrated the system with external cameras and devices, enabling real-time attendance monitoring across multiple locations.
- Optimized the application for scalability, supporting concurrent user operations and large-scale data storage using PostgreSQL and AWS services.
- **Impact:** Delivered a cutting-edge attendance solution that enhances operational efficiency and eliminates manual errors. The system provides organizations with a scalable, secure, and automated method for tracking attendance, reducing administrative workload, and improving accuracy.

Wedding App

June 2024 - Sep 2024

Tech Stack: Spring Boot, React, Node.js, HTML, CSS, AWS, Docker

- Developed a responsive web application for wedding planning, featuring guest list management, event scheduling, and dynamic itineraries.
- Designed and implemented a photo and video gallery module allowing users to upload, organize, and manage media by event/function.
- Enabled advanced media functionalities, including downloading, rotating, zooming, sharing, and marking items as favorites.
- Built secure server-side functionalities and RESTful APIs for efficient data management and user authentication.
- Deployed the application on AWS using Docker for scalability, reliability, and seamless performance.
- **Impact:** Delivered a versatile platform that simplifies wedding planning, offering users an intuitive and feature-rich solution to effectively manage their events and media.

AI-Powered Document Q&A System

June 2025 - Aug 2025

Tech Stack: Python, LangChain, OpenAI API, Pinecone, FastAPI, Asyncio, Docker, CI/CD, Pytest

- Developed and deployed an AI-powered document Q&A system enabling natural language queries across large-scale documents.
- Leveraged LangChain with OpenAI embeddings and Pinecone vector database to deliver accurate, context-aware responses.
- Optimized the pipeline with asynchronous batch processing, reducing query response time by 70% and ensuring high throughput.
- Built and deployed the system using FastAPI, Docker, and CI/CD, with automated testing for reliability.
- The project combined AI innovation with backend performance engineering, showcasing the ability to deliver scalable, production-ready solutions that transform unstructured data into actionable insights.
- Enhanced system reliability and fault tolerance with error handling, logging, and retry mechanisms.
- Collaborated in an end-to-end lifecycle from design to deployment, ensuring smooth handoff and maintainability.
- **Impact:** Implemented and deployed a scalable AI document Q&A platform using Python, LangChain, OpenAI API, and Pinecone; improved system performance by 70% via async pipeline optimization and ensured reliability with CI/CD and automated tests.