Mihir Deshpande

📞 +91 9623179829   📍 Pune   ✉ Email   66 Blog   in Linkedin   ⌂ GitHub

Software Engineer with 4 years of experience designing and developing scalable, high-performance backend systems. Skilled in integrating LLMs into production applications, with practical experience in prompt engineering and GenAI implementation. Adept at leading projects, improving system reliability, and collaborating cross-functionally to deliver solutions.

EDUCATION

Pune Institute of Computer Technology   **Pune, India**
B.E. Electronics and Telecommunication — CGPA: 9.15/10   *Aug '17 - May '21*
- Data structures and Algorithms, Object Oriented Programming, Operating Systems, Computer Networks, Machine Learning

WORK EXPERIENCE

Software Engineer   **Pune, India**
HelpShift   *Mar '23 - Present*
- Architected and optimized a scalable architecture for integrating a third-party LLM into a chatbot service handling 1.5K RPS, to achieve a **99.99% uptime**, reduce latency by **30%**, and ensure fault tolerance.
- Designed and developed a Distributed Asynchronous Task Queue, leveraging Kafka and Redis that **reduced system latency by 80%** and **load average by 30%** through offloading I/O and compute
- Contributed to a database migration from MongoDB to YugabyteDB, implementing a **dual-write strategy** to **maintain zero downtime**, resulting in a **60% improvement** in query performance and seamless data consistency
- Mentored **10+** new engineers, providing guidance on best coding practices, code reviews, and troubleshooting techniques, leading to faster onboarding, improved code quality, and increased team productivity
- Resolved **20+** on-call incidents and collaborated with SRE teams to implement auto-scaling policies, reducing overnight CPU spikes by **35%.**

Software Engineer   **Pune, India**
miniOrange Inc.   *Aug '21 - Feb '23*
- Enhanced system security by enforcing OAuth 2.0, JWT-based authentication, and multi-factor authentication (MFA), reducing unauthorized access incidents.
- Optimized Hibernate queries and indexed critical authorization tables, reducing database query execution time by **30%**, improving role lookup speed.

SKILLS

Programming Languages: Java, Clojure, Python
Databases and Streaming: MongoDB, PostgreSQL, Apache Kafka, Redis, Elasticsearch, YugabyteDB
Frameworks and Libraries: Spring Boot, Django
Core Competencies : Product Development, LLMs, Technical Leadership, Functional Programming, Distributed Systems

PROJECTS

Clj-deepwalk   **Clojure, Graphs, Machine Learning**
⌂ *View Project*
- A Graph Embedding Library that converts complex Networks into Vector Embeddings using Cosine Similarity, for Machine Learning analysis. Recognized in Clojure Deref

MCP Client for Ollama   **Python, Ollama, MCP**
⌂ *View project*
- Extensible MCP client for Ollama models
- Support for Git operations through MCP Git server, with interactive command-line assistant interface

AIde   **Python, OpenAI, OOP, Prompt Engineering**
⌂ *View project*
- An Intelligent code practice tool that harnesses OpenAI to dynamically generate coding challenges, evaluate solutions, and provide detailed feedback across multiple difficulty levels.

PUBLICATIONS

3D Modelling using Generative Adversarial Networks   **Python, Keras, Deep Learning, Tensorflow**
📖 *View Paper*
- Resolved limitations in existing 3D GANs and engineered a novel solution for generating varied 3D models of target objects, enhancing the efficiency of 3D modeling processes