

NIKHIL HEGDE

Software Engineer

✉ nikhilhegde989@gmail.com

☎ +91 7975438343

📍 Hyderabad

🌐 nikhilhegde.com

Backend engineer with 2+ years of production experience building event-driven systems, AI/LLM pipelines, and high-throughput APIs. Hands-on with distributed systems (Kafka, Redis, WebSockets), PostgreSQL internals, and LangGraph-based agentic workflows. Speaker at Hyderabad Postgres Days 2025. Published researcher in Scopus Q1.

★ TECHNICAL TALKS

08/2025

MVCC in PostgreSQL: The Hidden Engine Behind Your Transactions [🔗](#),
Speaker at Hyderabad Postgres Days 2025 (August)

- Delivered a deep technical talk on PostgreSQL internals, covering MVCC, tuple visibility, transaction isolation, indexing, storage layout, and VACUUM mechanics.

📁 WORK EXPERIENCE

08/2024 – Present
Hyderabad

Software Engineer, Stryv.ai

Technologies: Python, FastAPI, PostgreSQL, Redis, Kafka, WebSockets, Azure Functions, SQLAlchemy, LangGraph, Qdrant

- Recognized with the **Key Contributor Award** for impactful contributions to backend systems.
- Optimized the backend, achieving **78.6% reduction in request time under concurrent load**, **63% reduction for individual requests**, and **4.7× throughput increase**.
- **Architected and implemented semantic search engine** end-to-end, integrating Qdrant **vector search**, **multi-source retrieval** (Sanity CMS + AWS S3), and an automated content **ingestion pipeline**, cut inference costs significantly by replacing cloud **embeddings** with a local model without **accuracy regression**.
- **Designed and implemented a supply-chain scoring engine** to compute order-level business metrics from complex business rules.
- Built a **LangGraph-based agentic AI chatbot** for a hospitality platform, **automating customer support workflows** with multi-step reasoning and **tool-calling** capabilities.
- Built a real-time user-facing event **streaming** system using **Kafka and WebSockets** to surface live process updates in the UI.
- Reduced database load and improved API latency through **Redis caching**, **async design**, and **PostgreSQL query optimization**.

- Migrated legacy .NET modules into **FastAPI-based microservices**, improving maintainability and system performance.

01/2024 – 07/2024
Manipal

Software Engineer Intern, Ganglia Technologies Pvt Ltd

Technologies: Python, Node.js, Express, MongoDB, React, LLMs, RAG

- Built a speech-to-text system for a healthcare application to auto-populate clinical forms from voice input.
- Designed and implemented backend APIs for a livestock marketplace platform, supporting product listings, search, filtering, and transaction workflows.
- Contributed to LLM-based retrieval workflows using RAG for structured data access.

SKILLS

Programming Languages — Python, JavaScript, SQL, Java • **Backend Technologies** — FastAPI, Node.js, Express.js, Microservices, WebSockets, Async Programming (asyncio), Kafka, Redis, Event-Driven Architecture • **AI/LLM** — Lang Chain, Lang Graph, Large Language Models (LLMs), Vector Databases (Qdrant), RAG Pipelines (Retrieval, **Reranking**, Context Optimization), LLM Observability (Langfuse, LangSmith) **Embedding & Retrieval Pipelines (Chunking, Vectorization, Reranking, Optimization)** • **Databases** — PostgreSQL, MongoDB, Vector Databases (Pinecone, Qdrant) • **DevOps & Cloud** — Git, Docker, AWS, Google Cloud Platform, Azure • **Core Concepts** — System Design (HLD/LLD), Distributed Systems, Performance Optimization • **Familiar** — React

PROJECTS

BookARide - Ride Booking Platform (Uber-like system) [🔗](#),

Tech: React, Node.js, MongoDB, WebSockets, Razorpay, AWS, NGINX

- Built a full-stack ride-hailing platform with real-time driver tracking, WebSocket-based ride updates, and end-to-end ride lifecycle management (booking, cancellation, reassignment).
- Implemented secure authentication & payments, using JWT with role-based access and token blacklisting, and Razorpay integration with order and signature verification.
- Designed scalable backend infrastructure with MongoDB geospatial matching and heatmaps, deployed on AWS with NGINX reverse proxying.

Inspire - AI-Powered Social Platform [🔗](#),

Tech: Node.js, FastAPI, MongoDB, TensorFlow, Docker, GCP, Python

- Developed a production-ready social platform with ML-powered real-time toxicity filtering, FastAPI model inference, scalable backend services, image optimization via Sharp, and Docker-based deployment on GCP.

RESEARCH PUBLICATIONS

Analyzing AI Regulation Through Literature and Current Trend [🔗](#)

- Published in Journal of Open Innovation (Scopus Q1, Top 1%)

Hospital-Acquired Infection Management with Artificial Intelligence: A Comprehensive Review

- Presented at International Conference on IDEAAS 2024

Hybrid Approach—Diabetic Retinopathy Classification Through Activation Function Optimization

- Presented at Springer Nature SIGMAA 2023

EDUCATION

2022 – 2024
Manipal

**Master of Computer Applications (MCA),
Manipal Institute of Technology, Manipal**
CGPA - 9.06 (Top 10% of the batch)

2019 – 2022
Dharwad

Bachelor of Science (Physics, Maths, Comp. Science), JSS, Dharwad
Percentage - 84

CERTIFICATIONS

**NodeJS Internals and
Architecture**
By Hussein Nasser

**AWS Cloud Technical
Essentials**
Issued by Amazon Web
Services (AWS)

Machine Learning
Issued by Internshala
Trainings