

Sudhanshu Batra

✉ sudhanshubatra27@gmail.com [in LinkedIn](#) [GitHub](#)

Summary

Software Engineer specializing in full-stack development, backend systems, and applied AI. Experienced in building scalable web applications, LLM-powered workflows, data pipelines, and analytics platforms using FastAPI, Next.js, Python, and modern cloud technologies. Strong foundation in Machine Learning, system design, and production-grade software development.

Education

Vellore Institute of Technology (VIT) – 8.56 Cgpa 2023 – 2027
B.Tech in Computer Science (Business Systems)
Relevant Coursework: DSA, DBMS, OOP, Cloud Computing, Computer Networks, Machine Learning

Professional Experience

Full Stack Intern - Automentation Services May 2026 – Jun 2026

- Engineered and deployed new frontend and backend functionalities for a Medical Insurance Claim platform using modern full-stack technologies, integrating RESTful APIs and optimizing core application workflows.
- Collaborated in delivering a production-ready solution by resolving critical bugs, enhancing performance and user experience, and successfully integrating the application into the company's primary website.

Oracle Cloud Infrastructure – Associate Certified 2025

Projects

Forke – Developer Collaboration & Task Management Platform [Live](#) | [GitHub](#)

- Built a full-stack platform connecting developers with startups through structured engineering tasks, integrating GitHub-based submissions, secure escrow payments, and automated task lifecycle management.
- Developed contribution tracking, portfolio generation, and submission verification systems, with AI-assisted code review to streamline evaluation and quality assurance workflows.

Retail Data Hub – End-to-End Analytics Suite [Live](#) | [GitHub](#)

- Built a FastAPI + Next.js retail analytics platform with DuckDB and Apache Parquet, reducing query latency from 12s to under 2s on 10M+ rows through a Medallion Architecture pipeline.
- Developed 12+ interactive dashboards and integrated a PyTorch LSTM forecasting model (8% MAPE over a 90-day horizon) via FastAPI, enabling real-time analytics and AI-assisted business insights.

Retrieval-Augmented Generation (RAG) Pipeline

- Designed a RAG architecture using dense embeddings and ANN-based vector search for low-latency semantic retrieval over custom datasets.
- Improved contextual grounding and reduced hallucinations by 85%+ through optimized chunking, embedding selection, and prompt conditioning.

Technical Skills

Languages: Python, Java, C/C++, JavaScript, SQL

AI/ML: PyTorch, TensorFlow, Scikit-Learn, OpenCV, YOLO

Full Stack Development: React.js, Next.js, Node.js, Express.js, FastAPI, REST APIs, WebRTC, GSAP

Databases: PostgreSQL, MongoDB, DuckDB

Cloud & DevOps: AWS, Oracle Cloud, Vercel, Railway, Git/GitHub, CI/CD

Certifications

- [Oracle Cloud Infrastructure AI Foundations Associate](#) — Oct 2025
- [Oracle Cloud Infrastructure Foundations Associate](#) — Oct 2025