

Nikhil Kumar Singh

Software Engineer | Backend Developer | ML Engineer
+91 83285 76370 | nikhilkumarsingh2914@gmail.com | Portfolio | GitHub | LinkedIn | LeetCode

Summary

Software Development Engineer Intern and CS undergraduate (CGPA 9.16/10) building production backend and full-stack systems end to end. Strong in Python, C++, and JavaScript with hands-on FastAPI, RESTful API design, async services, Redis caching, and Docker. Experience across PostgreSQL data modeling, RAG/LLM pipelines, and system design — comfortable owning features from API design through containerized deployment.

Experience

Software Development Engineer Intern – Autometa AI

Mar 2026 – Present

- Shipped a real-time voice-interaction feature with token-based authentication in a production client web app — session management, protected routes, and low-latency audio I/O end to end.
- Designed and shipped **10+** async RESTful endpoints and a LinkedIn profile-data ingestion module powering core user flows for a React Native mobile app.

Research Intern – Adaptive Streaming Algorithms

May 2025 – Jul 2025

- Designed and benchmarked a Python adaptive-bitrate (ABR) algorithm for VR streaming that cut bandwidth consumption **30%** and reduced buffering **15%** across **6** simulated network conditions.
- Built a reproducible simulation framework and backend evaluation pipelines to compare streaming strategies under multi-scenario network tests.

Projects

SAHAY – AI Launcher for Elderly Safety (Funded Project)

GitHub

Kotlin, Jetpack Compose, AccessibilityService, NLP, SensorManager

- Engineered “Genie,” a voice assistant pairing Android SpeechRecognizer with an NLP intent parser that drives an AccessibilityService to automate third-party apps end to end — opening, searching, and adding to cart, halting at checkout for user confirmation.
- Built passive safety monitoring via a foreground service and SensorManager, applying Gaussian-based anomaly detection over vital-sign streams to trigger a real-time voice check-in and automated SOS (call, SMS, WhatsApp, live location).

ContinuityX-AI – Semantic Code Search & RAG Chat System

Demo | GitHub

Python, FastAPI, LangChain, ChromaDB, Docker

- Built a scalable async FastAPI backend that indexes and semantically searches **500+** files across code repositories using RAG over ChromaDB, with an ingestion pipeline for efficient retrieval.
- Cut repeated-query latency via response caching and batched embedding/retrieval, and containerized all services with Docker for reproducible deployments.

PolicyGraph – Document Retrieval & Semantic Query Engine

GitHub

Python, FastAPI, PostgreSQL, Redis, LangChain

- Built a FastAPI document-retrieval and semantic-query engine returning top-k ranked passages, with a modular query → retrieval → response-generation pipeline.
- Eliminated redundant database reads with a Redis caching layer for frequent policy lookups, improving response times under repeated queries.

Calmora – NLP-Based Mental Health Chatbot Platform

Live | GitHub

Python, Flask, NLP, SQLite

- Built an NLP chatbot backend using intent classification across **10+** intents for structured conversations, with keyword detection to flag high-risk interactions.
- Designed SQLite-backed user-data persistence and deployed the full stack via Vercel/Netlify with backend integration.

Technical Skills

Languages: Python, C++, JavaScript, SQL

Backend: FastAPI, Django, Flask, Node.js, Express.js, RESTful APIs, Async Services

Frontend: React.js, React Native, Bootstrap

Databases & Caching: PostgreSQL, MySQL, MongoDB, Redis, ChromaDB

AI/ML: PyTorch, TensorFlow, Scikit-learn, OpenCV, LangChain, RAG, NLP

DevOps & Tools: Docker, Git, GitHub, Linux, Postman, Vercel, Netlify

Core CS: Data Structures & Algorithms, OOP, Operating Systems, DBMS, Computer Networks, System Design, Distributed Systems, Concurrency

Education

SRM University, AP

B.Tech in Computer Science and Engineering

2023 – 2027

CGPA: 9.16 / 10 (Top 1%)

Achievements & Certifications

- Dean’s List – Top 1% of batch.
- Oracle Certified: Java SE 17 Developer.
- Winner – Code for Connection Hackathon | Selected Contributor – GSSoC 2026.