

SANKALP GIRI

☎ +91 8149148479 ✉ sankalpgiri110@gmail.com  <https://www.linkedin.com/in/sankalp-giri10/>
 <https://github.com/Sankalp1025>

Professional Summary

Computer Science undergraduate with strong skills in Data Structures and Algorithms and scalable backend systems. Built and deployed ArchCode, featuring asynchronous evaluation pipelines and LLM-driven feedback.

Technical Skills

Languages: C/C++, TypeScript/JavaScript, Python.
Frontend Technologies: React.js, Next.js, Tailwind CSS, HTML.
Backend Technologies: Node.js, Express.js, REST APIs, JWT Authentication, BullMQ, SSE.
Databases: PostgreSQL, Prisma ORM, Redis, Vector Databases.
AI/ML: RAG, Vector Search, Embeddings, LLM Integration.
Tools and Infrastructure: Git, GitHub, Postman, Vercel, Render.

Relevant Coursework

- | | | | |
|-------------------|--------------------|---------------------------|-----------------------|
| • Data Structures | • System Design | • Artificial Intelligence | • Distributed Systems |
| • Algorithms | • Machine Learning | • Operating System | • OOP |

Projects

AI-Powered System Design Platform | *TypeScript, Node.js, PostgreSQL, Prisma, Redis, BullMQ.* January 2026

- Engineered ArchCode, an AI-powered system design platform enabling users to design distributed systems and receive real-time automated architectural evaluations, processing **200+ evaluations/hour** in production.
- Architected a scalable and distributed evaluation pipeline leveraging BullMQ and Redis, decoupling request handling from AI inference workloads maintaining **45 ms API latency** and reliable background job processing.
- Designed an Architecture Pattern Detection and Linting Engine capable of identifying design patterns, architectural anti-patterns, and scalability bottlenecks in real time, achieving sub-millisecond analysis with 0.195 ms pattern detection and 0.425 ms linting engine.
- Implemented real-time AI evaluation streaming via Server-Sent Events (SSE) with multi-layered fallback mechanisms, delivering architecture scores, strengths, weaknesses, and actionable recommendations with 11 s end-to-end processing time and **4.8 s average LLM latency**.
- Live Demo: <https://archcode-ai.vercel.app/>
- Github: <https://github.com/Sankalp1025/Archcode>

Vector Database and RAG Engine built from Scratch | *RAG, REST APIs, Ollama, Vector Search.* April 2026

- Built a Vector Database from scratch in C++, implementing HNSW, KD-Tree, and Brute Force search algorithms to compare the trade-offs between **approximate and exact nearest-neighbor search** for semantic retrieval.
- Designed a production-style **HNSW indexing engine** using hierarchical graph traversal, enabling efficient similarity search and reproducing retrieval techniques used by modern vector databases such as Pinecone, Weaviate, and Milvus.
- Developed a full-stack semantic search platform with REST APIs for vector insertion, deletion, search, benchmarking, and index inspection; integrated Cosine, Euclidean, and Manhattan distance metrics along with PCA-based visualization of vector embeddings.
- Implemented an end-to-end RAG pipeline using Ollama, nomic-embed-text embeddings, HNSW-based context retrieval, and Llama 3.2, enabling document-grounded question answering over user-provided data through a fully local AI stack.
- Github: <https://github.com/Sankalp1025/VectorDB>

Education

Chhtrapati Shivaji Maharaj University.
B.Tech in Computer Science and Engineering (CSE)

August 2023 – June 2027
Panvel, Navi Mumbai

Achievements / Certificates

2023 - present:

- * **Solved 250+ Data structures and algorithms problems** across Leetcode and GeeksforGeeks focusing on time/space optimization and problem-solving patterns. [Leetcode: <https://bit.ly/3Sc6dgc>]
- * Secured **All-India Rank 3067** among 2 lakh+ students in the Internship Studio's Common Internship Test (Top 1.5 percent nationality). [Certificate shared on linkedin]
- * Completed Google Cloud's "Introduction to Generative AI", earned a badge via Google Cloud Skills Boost. [Link: <http://bit.ly/3PZQCj5> (Completion Badge)]