

Sriyansh Raj

www.linkedin.com/in/sriyanshraj/
github.com/sriyansh-dev
leetcode.com/u/sriyansh-dev
codolio.com/profile/sriyansh_dev

Email : sriyanshraj.dev@gmail.com

SUMMARY

Generative AI and Backend Engineer (final-year B.Tech CSE, CGPA 8.2) who has built several LLM-powered applications and completed 3 AI/ML internships in 2026. Architected S.A.G.E., an open-source agentic code-review engine with a custom Model Context Protocol (MCP) server, and Vault, an autonomous agent with two-stage RAG and dynamic model routing. Focused on clean API contracts, low-latency real-time systems, and production-grade RAG pipelines.

Education

- DR A.P.J ABDUL KALAM TECHNICAL UNIVERSITY** (Vishveshwarya Group of Institutions) GB NAGAR, UP
Bachelor of Technology in Computer Science & Engineering | 4th YEAR | CGPA: 8.2 Aug. 2023 – May 2027

Skills

- Languages:** Python, Java, SQL, JavaScript, HTML, CSS
- AI/ML Frameworks:** LangChain, Groq LLMs, Google Gemini API, FAISS, Chroma, TensorFlow, PyTorch, Scikit-learn, Hugging Face
- Competencies:** Generative AI, Agentic AI, Large Language Models (LLMs), RAG Pipelines, Machine Learning, NLP, Computer Vision, API Development
- Tools & Technologies:** Model Context Protocol (MCP), Spring Boot, Redis, WebSockets, REST APIs, Streamlit, Gradio, Git & GitHub, Google Colab

EXPERIENCE

Gen AI Intern | IIT Kanpur Prabandhan & EiSystems Technologies | June 2026 – August 2026

- Engineered a real-time AI Interview System on the MERN stack, using WebSockets for low-latency, two-way candidate-interviewer communication.
- Architected Google Gemini-powered question generation with LangChain and a RAG pipeline to deliver dynamic, context-aware interviews.
- Optimized backend throughput and response time by implementing a Redis caching layer.
- Tech Stack:** Python, MERN Stack, Google Gemini, LangChain, RAG, Redis, WebSockets

AI & ML Intern | Codomax Digital Solutions | July 2026

- Delivered project-based ML engineering work, building and evaluating models in Python and Scikit-learn from data preprocessing through validation.

AI & ML Intern | Cognifyz IT Solutions | Summer 2026

- Executed hands-on ML projects end to end, from data preparation and model training to performance evaluation, in a project-driven internship.

Projects

- S.A.G.E. (Secure Agentic Git Engine):** Architected an open-source agentic AI code-review and repository orchestration tool in Python on Groq LLMs, coordinating 3 agents (Planner, Patcher, Reviewer) through an async DAG scheduler.
- Built a custom Model Context Protocol (MCP) server exposing 5 tools (credential scanning, test verification, PR creation, report generation) to automate review-to-pull-request workflows locally.
- BitSwarm RAG:** Engineered a Retrieval-Augmented Generation (RAG) system that grounds LLM responses in retrieved context for accurate, context-aware answers.
- Vault:** Engineered an autonomous AI agent with dynamic model routing and live web search integration, grounded by a two-stage RAG pipeline and delivered through a Gradio web interface.
- Document Analysis System:** Built a LangChain RAG application with FAISS vector search, deployed on Streamlit, to extract insights from uploaded files.
- AI Code Reviewer Web App:** Developed a full-stack tool with React JS, Tailwind CSS, Monaco Editor, and the Google Gemini API for automated bug detection and explanation.
- AI Fact-Checking Browser Extension:** Built an extension that verifies claims on web pages using NLP, LangChain, and RAG-based retrieval.

Certifications and Achievements

- Oracle Cloud Infrastructure Agentic AI Certified Foundations Associate:** Oracle, Aug 2026
- AI Readiness Foundation Course:** IICT / Ministry of I&B / Google, Aug 2026
- Introduction to Generative AI & LLMs:** Google Cloud, 2025
- Data Analytics Job Simulation:** Deloitte (Forage), Feb 2026
- Bootcamp on Big Data & Data Science:** C-DAC Noida (FutureSkills Prime), Feb 2025
- Advanced C++ Training (Score: 85%):** IIT Bombay – Spoken Tutorial Project, Jan 2025

Open Source Contributions

- PyGitLoc:** Published a Python package that returns lines of code for any GitHub repository.
- ATS-Checker:** Built a resume ATS scoring tool.
- Actively maintain version-controlled projects using Git and GitHub.