

Pranav Suri

Email: pranavsuri195@gmail.com
New Delhi, India

LinkedIn: linkedin.com/in/pranavsuri1
GitHub: github.com/pranavsuri05

Education

- **Maharaja Agrasen Institute of Technology (G.G.S.I.P.U.)** Delhi, India
B.Tech'27 in Computer Science & Technology, CGPA: 9.27/10.0 Sep 2023 – Jun 2027
- **Apeejay School, Panchsheel Park** Delhi, India
Class 12th C.B.S.E., Percentage: 89.2% Apr 2022 – May 2023

Experience

- **Tech Mahindra** [techmahindra.com] Noida, UP, India
Software Developer Intern Jun 2026 – Aug 2026
 - Built a multi-agent software engineering platform automating **8 SDLC stages**, designing agent orchestration, workflow state management, and structured artifact handoffs across requirements, design, development, testing, and closure.
 - Developed FastAPI-based REST APIs and backend services to orchestrate **5+ specialized agents**, manage intermediate artifacts, validate structured inputs/outputs, and maintain versioned workflow state across SDLC stages.
 - Reduced **LLM token consumption by up to 25%** by optimizing inter-agent context propagation, eliminating redundant payloads, and selectively passing workflow state between backend services.
 - **Project: AI-SDLC Pipeline** — [GitHub](#)
- **Sparrow Risk Management Pvt. Ltd** [sparrowrms.in] Gurgaon, India
Software Development Intern (AI/ML) Feb 2025 – Mar 2025
 - Improved **RAG-based chatbot accuracy by 30%** by redesigning document chunking, optimizing embedding pipelines, and improving FAISS-based vector retrieval.
 - Reduced **LLM hallucination by 15%** through automated evaluation workflows and A/B testing across retrieval configurations and prompt templates.

Achievements & Research

- **AWS Hackathon — 1st Place** : Built an LLM-powered banking assistant using AWS PartyRock.
- **Flipkart GRiD 8.0 — National Semi-Finalist** : Advanced through 3 rigorous rounds of coding and software engineering challenges.
- **Adobe India Hackathon — Finalist** : Built an AI-powered PDF intelligence platform.
- **Research Publication — MA-LoRA** : Co-authored Multi-Attribute Debiasing of LLMs via Bias-Aware Low-Rank Adaptation, accepted at ICCNet 2026 at Manchester Metropolitan University, UK (Springer LNNS, Scopus-indexed).

Projects

- **IntelliDoc — AI-Powered PDF Intelligence Pipeline** (*Python, PyMuPDF, Docker*) [[GitHub](#)]
 - Engineered an offline PDF processing pipeline in Python that reconstructed document hierarchy using font metadata and rule-based H1–H3 detection, generating structured JSON document outlines.
 - Implemented semantic section ranking using all-MiniLM-L12-v2 embeddings to retrieve persona-relevant content without external API dependencies.
- **Healthcare Diagnostic Booking Backend** (*FastAPI, PostgreSQL, SQLAlchemy, JWT*) [[Github](#)]
 - Engineered a production-style REST API for diagnostic test bookings with JWT authentication, role-based access control, centre–test mappings, appointment scheduling, and payment workflows.
 - Implemented transactional booking and payment flows with SQLAlchemy, database constraints, webhook idempotency, and comprehensive pytest coverage to ensure reliable API behavior.

Skills

- **Programming Languages** : C++, SQL, Python
- **Backend & Databases** : FastAPI, REST APIs, PostgreSQL, SQLAlchemy, SQLite, Docker, Git, GitHub
- **Core CS Concepts & AI** : Data Structures & Algorithms, Object Oriented Programming, Operating Systems, Computer Networking, Database Management Systems, Multi-threading, Multi-Agent Systems, RAG