

ASHISH

(+91) 97175 01442 • New Delhi, India • soniashish37066@gmail.com • [LinkedIn](#) • [GitHub](#) • [LeetCode](#)

Education

Delhi Technological University
B.Tech in Computer Science Engineering

Nov 2022 – Jun 2026

Experience

Indian Space Research Organization (ISRO) | ML Intern

June 2025 – Aug 2025

- Built end-to-end **6-DoF spacecraft pose estimation** pipeline, achieving **ER 3.97°** and **ET 2.26m** on a **12K+ dataset**.
- Engineered **dual-head CNN architecture** for orientation and translation prediction, improving domain generalization.
- Designed evaluation pipeline with augmentation and validation across diverse orbital conditions to ensure robustness.
- **Tech Stack:** PyTorch, Python, OpenCV, NumPy, CNN, Deep Learning, Pose Estimation, CUDA

DTU USIP Program | Software Engineering Intern

May 2024 – July 2024

- Built production-grade **Hostel Management System** used by **2,000+ students** for complaints and hostel operations.
- Architected backend with Node.js, MongoDB, and **JWT-based authentication** with role-based access control.
- Developed full-stack system with Next.js, ensuring scalable deployment and smooth multi-user experience.
- **Tech Stack:** Next.js, TypeScript, Node.js, Express, MongoDB, JWT, REST APIs, Vercel

Projects

OmniForce — Autonomous Enterprise AI Workforce | [Live](#) | [GitHub](#)

- Built a production-grade **multi-agent AI workforce platform** automating KYC/AML compliance, invoice processing, and sales workflows using LangGraph-based orchestration pipelines.
- Developed **3 autonomous AI agents** integrated with Airtable CRM, Gmail, Slack, OFAC/UK Treasury sanctions datasets, Basel AML risk scoring, and Mindee OCR for end-to-end compliance automation.
- Engineered a **7-stage KYC/AML pipeline** combining sanctions screening, document verification, ChromaDB-powered compliance RAG, and Groq-accelerated **LLaMA 3.3 70B inference**.
- **Tech Stack:** Python, FastAPI, LangGraph, LLaMA 3.1, ChromaDB, Airtable API, Gmail API, Slack API, Streamlit, Docker

PatentPilot AI — Distributed Multi-Agent Patent Intelligence System | [Live](#) | [GitHub](#)

- Architected a **6-agent parallel reasoning system** for automated patent analysis, enabling novelty search, infringement detection, prior-art retrieval, and Freedom-to-Operate (FTO) assessment.
- Built a **semantic RAG retrieval pipeline** over 7,000+ patent chunks using FAISS vector indexing and sentence-transformer embeddings for citation-grounded legal reasoning and similarity search.
- Reduced end-to-end analysis latency from **30s to under 3s** via Groq-accelerated LLaMA 3.1 inference and concurrent execution using `asyncio.gather`, enabling near real-time patent analysis.
- **Tech Stack:** Python, FastAPI, FAISS, LLaMA 3.1, Groq, Streamlit, Sentence-Transformers, asyncio, Pydantic

Scientific Paper Knowledge Extractor | [Live](#) | [GitHub](#)

- Built a scalable **LLM-powered extraction pipeline** over 200M+ scientific papers, generating structured JSON outputs from unstructured research text using Groq-accelerated LLaMA 3.1 and prompt engineering.
- Reduced end-to-end processing latency from **45s to under 4s** by migrating from local Ollama inference to Groq, and developed a hybrid pipeline supporting PDF upload, ArXiv parsing, and raw text input with NLP-based fallback mechanisms.
- Engineered a **reliability layer** with retry logic, normalization, and schema validation ensuring robust structured outputs even under LLM failure across diverse input formats.
- **Tech Stack:** Python, LLaMA 3.1 (Groq), Streamlit, Sentence-Transformers, Semantic Scholar API

Open Source Contributions

DSPy — Stanford NLP | [Merged PR](#) | [Issue](#) | [Framework](#)

- Contributed a production-grade fix to **DSPy**, Stanford NLP's declarative LLM framework (**34K+ GitHub stars**) — resolving a critical **silent state-corruption bug** in `load_state()` causing inconsistent RAG and inference behavior.
- Implemented a **validate-first + deepcopy transactional loading strategy** ensuring all-or-nothing state restoration, analogous to database BEGIN/ROLLBACK/COMMIT semantics — shipped with **47 passing tests**.

Skills

Coursework: OOPs, Computer Networks, Machine Learning, DBMS, Operating System

Programming Languages: Python, TypeScript, JavaScript, C++, SQL

Databases: MongoDB, PostgreSQL, Pinecone, ChromaDB, FAISS, SQLite

AI & Agentic Engineering: LLMs, RAG Pipelines, AI Agents, Multi-Agent Workflow Orchestration, Prompt Engineering, Context Engineering, Function Calling, LangChain, LangGraph, MCP, tool calling, Semantic Search, Gemini API, Claude Code

Full-Stack web development: ReactJS, NodeJS, NextJS, REST APIs, Supabase, Django, Docker, Git, GitHub, Unit testing, SDLC, Django, Linux, Redis, AWS (EC2, S3), Vercel, Railway, microservices

Achievements

- Solved **1500+** DSA problems across LeetCode, Codeforces, and CodeChef with a maximum LeetCode rating of **1989**.
- Finished in the **top 2%** in LeetCode Weekly Contest 401, securing a global rank of **546**.
- Currently a **3-Star** rated coder at [CodeChef](#) with a peak rating of **1723**.
- Reached a peak rating of **1426** on [Codeforces](#), earning the title of **Specialist**.