

A. Sharan

hemsharan2050@gmail.com | linkedin.com/in/ahsharan | github.com/Ahsharan

SUMMARY

Undergraduate software engineer focused on AI/ML-driven systems and rapid prototyping. Hands-on experience building and deploying full-stack applications with backend APIs, databases, and applied machine-learning features, with emphasis on reliability and performance.

EDUCATION

Symbiosis Institute of Technology - OS, AI, Python

Bachelor of Technology, Robotics

Pune, India

2024 - Pursuing

IIT Ropar - NLP, CNN, RL, NN, CV, LLM

Minor in AI

India

2024 - Pursuing

TECHNICAL SKILLS

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

Machine Learning: TensorFlow, scikit-learn, NumPy, Pandas, OpenCV

Backend & APIs: Node.js, Flask, FastAPI, RESTful API Design

Frontend: React

Tools & Systems: Git, Linux, SQL, NVIDIA BUILD, Qualcomm AI Hub

EXPERIENCE

Intern - Tech Mahindra | *Research and Development*

Oct. 2025 – Dec. 2025

- Implemented **QLoRA** and advanced quantization techniques to optimize student models (Llama 3B/8B), achieving a **2% performance gain** on MMLU benchmarks compared to baseline reference models.
- Architected a PDF-to-JSON VLM pipeline using **Nvidia BUILD** and OpenCV, reducing processing latency by **70%** (from 30s to 7s per page) while correctly handling complex Devanagari scripts.
- Researched and profiled **on-device edge inference** strategies using Qualcomm AI Hub, optimizing transformer models for deployment on Snapdragon NPUs to enable low-latency, offline processing.

Intern – Symbiosis Centre for Applied Artificial Intelligence (SCAAI)

Apr. 2025 – Present

- Designed a multi-agent cognitive simulation system modeling psychological and behavioral dynamics using interacting probabilistic feature agents.
- Trained and evaluated 18 agents on 30 days of real-world user data, achieving **70% accuracy** in individual-level behavior prediction.
- Developing agentic intervention mechanisms focusing on explanation filtering, reliability, and decision consistency.

Head of Technical Department – SIT FOSS (Student Open Source Club)

Aug. 2024 – Present

- Directed a 30-member development team to architect and deploy the official SIT FOSS website.
- Conducted Git training workshop for 12 members, strengthening collaborative development workflows.

PROJECTS

Flow.AI - AI-Powered Workflow Builder

- Designed and deployed a production-grade full-stack system for generating and editing workflow diagrams from natural language.
- Built an asynchronous **FastAPI** backend with **REST APIs**, integrating **Google Gemini (gemini-2.0-flash)** for structured flow generation.
- Developed a **Next.js + TypeScript** frontend using **React Flow**, supporting **10+ interactive features** including undo/redo, autosave, export (PNG/JSON), and real-time graph editing.
- Implemented persistent storage and auth using **Supabase (PostgreSQL)** with session enforcement and row-level security.
- Designed a deterministic node-edge state model to ensure reliable autosave and low-latency UI updates; deployed on **Vercel** and **Render**.

Vision-Based Advanced Adaptive Traffic Control (VAATC)

- Selected as a **Top 10 project in a national hackathon**; built using **YOLOv8, OpenCV, Python**, achieving a **30–40% reduction** in congestion in simulated Indian traffic environments.

- Developed a realistic "India Mode" simulation to mimic complex native behaviors like **abrupt stops and chaotic lane merging**; introduced **dynamic lane control with adaptive signal timing**.
- Published the simulator on **PyPI** under Apache 2.0; created custom analytics tools using **Matplotlib** and **Pygame** for real-time traffic visualization and time series analysis.
- Built a tailored time-series analysis system to study **temporal variations in traffic flow**, average waiting times, and crossing efficiency—supporting smarter decision-making.

HackRx 6.0 – Smart Agentic RAG Model (Top 100 Finalist)

Jul. 2025

- Secured a position in the **Top 100 out of 46k+ participants** at HackRx 6.0, hosted by Bajaj Finserv.
- Engineered a **Retrieval-Augmented Generation (RAG) system** for scalable and fault-tolerant document understanding.
- Tech stack included: **FastAPI backend**, Pinecone vector DB, OpenAI + Sentence Transformers embeddings, **Gemma**, TableGPT for tabular parsing, and multi-PDF ingestion pipelines for robust accuracy.
- Integrated reranking for contextual precision, migrating storage to **ClickHouse** for optimized query performance.
- Designed agentic solutions to minimize latency, achieving **78% retrieval accuracy** with an average **4000 ms response time**.

PUBLICATIONS

The Python Wizardry - Book

Aug. 2023

The Python Wizardry is tailored for professionals and enthusiasts looking to deepen their Python expertise. It covers essential topics including Flask, DBMS, Automation techniques, Machine Learning, Data Visualization using Matplotlib.

Project Manager - Python Package

Jun. 2024

Published a Python package on PyPI for streamlined project directory and metadata management. It enhances productivity by organizing project structures and automating metadata handling.

Waterinator

Oct. 2021

This book teaches beginners how to build a smart plant watering system with clear instructions, diagrams, and a final quiz, covering Arduino, soil sensors, motor drivers, and essential programming concepts.