Shoaib Sadiq Salehmohamed

9502264416 | shoaibkulsums@gmail.com | linkedin.com/in/shoaib-ssm | github.com/shoaib2000857 | shoaibssm.me

Education

Indian Institute Of Technology–Madras — Chennai, Tamil Nadu
Bachelor of Science (BS) in Data Science and Applications — ***CGPA: 8.85/10*** — *May. 2024 – September 2028*

Keshav Memorial Institute of Technology — Hyderabad, Telangana
Bachelor of Technology in Computer Science and Engineering — ***CGPA: 8.75/10*** — *Aug 2024 – Aug 2028*

Technical Skills

Languages: Java, Python, C, SQL (PostgreSQL, MySQL), JavaScript, HTML/CSS, TypeScript
Frameworks: Next.js, Express.js, Vue.js, React, Node.js, Flask, FastAPI, ROS, ArduPilot
Developer Tools: Git, Docker, Google Cloud Platform, Azure, AWS, Linux (Ubuntu, Arch), Redis, Celery, WebSockets
Libraries: Pandas, NumPy, Transformers, Matplotlib, PyTorch, OpenCV

Projects

TattvaDrishti | *Python, PyTorch, FastAPI, Next.js* — **Winning Project – Smart India Hackathon 2025**

- Built an end-to-end AI-driven threat intelligence platform **to detect and analyze LLM-powered misinformation and coordinated influence operations** using a modular FastAPI-based analysis pipeline.
- Fine-tuned DeBERTa-v3-Large transformer classification models **to distinguish AI-generated content from human-written text** achieving **95% accuracy**, and **to perform model-family attribution** with **98% accuracy** on a **multi-million-sample corpus** (**2.7–3M rows**) with **long-context inputs (12K tokens per sample)**.
- Designed a graph intelligence engine **to identify coordinated campaigns, narrative clusters, and influence hubs** across actors, regions, and content.
- Engineered explainable risk scoring and decision rationale by fusing **10+ linguistic and stylometric features**, statistical signals, and model confidence to support high-stakes analyst decision-making.
- Implemented real-time analyst dashboards using **Server-Sent Events (SSE)** and a custom blockchain-based audit layer **to ensure tamper-proof logging, traceability, and accountable intelligence sharing**.

Rakshak — **Autonomous Surveillance Rover** | *Python, Jetson Nano, LiDAR, GPS, Pixhawk*

- Built a real-world autonomous surveillance rover integrating **LiDAR, GPS, IMU, and camera sensors** for obstacle detection, navigation, and perimeter monitoring.
- Served as **Team lead** for the project, which culminated in a **published Indian patent application**.
- Implemented a **hybrid control architecture** where a Pixhawk flight controller handled mission logic while a Jetson Nano performed perception, decision-making, and safety overrides.
- Developed a **LiDAR-based obstacle detection and avoidance pipeline** capable of dynamically overriding manual and autonomous commands to prevent collisions.
- Engineered **real-time video streaming and telemetry pipelines using WebSockets** for live remote monitoring and alert generation.
- Implemented **on-device face recognition using OpenCV-based computer vision pipelines** to detect and identify intruders, triggering real-time alerts and evidence capture.
- Designed the system with **multiple communication modes** (Wi-Fi, RC, telemetry) to ensure robust operation in constrained and offline environments.

LingualLink — **Real-Time Multilingual Chat Platform** | *FastAPI, Next.js, Redis, PostgreSQL, MongoDB*

- Built a real-time multilingual chat system enabling users to communicate in their native languages with automatic language detection and translation.
- Designed an event-driven backend using **Redis queues and Pub/Sub** to decouple message ingestion from translation workers and ensure low-latency delivery.
- Implemented a scalable data model using **PostgreSQL** for authentication and **MongoDB** for flexible, multi-language message storage.
- Delivered real-time message updates to clients using **WebSockets/SSE**, ensuring seamless user experience across languages and devices.

Experience

Research Intern – LLM Lens Feb 2025 – Present

AI Research Intern (LLM Reliability & Hallucination Detection)

- Conducting research on **hallucination detection and reliability analysis in Large Language Models (LLMs)** by studying internal transformer representations and uncertainty signals across multiple layers.
- Developed **PyTorch-based experimentation pipelines** to extract and analyze hidden-state embeddings from transformer models (e.g., LLaMA, Qwen), enabling token-level and sequence-level hallucination analysis.
- Designed and implemented **cross-layer and hierarchical transformer architectures** to model correlations across hidden states for improved hallucination classification and uncertainty estimation.
- Built scalable **dataset generation pipelines** using Hugging Face models to synthesize prompts, responses, and labeled hallucination datasets for training and benchmarking detection models.
- Implemented comprehensive **evaluation frameworks** including ROC-AUC, PR-AUC, F1 score, Expected Calibration Error (ECE), Brier score, and threshold optimization to assess model reliability.
- Conducted large-scale experiments on **GPU-accelerated environments (CUDA, PyTorch, Colab A100/H100)** while managing checkpoints, experiment logs, and reproducible training pipelines.
- Explored **representation-level uncertainty signals** such as covariance structures and layer-wise feature statistics to improve interpretability and robustness in LLM outputs.

Member — ChatGPT Lab for Students (Inaugural India Cohort) Aug 2025 – Sep 2025

OpenAI *Remote*

- Selected among **28 students nationwide** for the inaugural ChatGPT Lab India cohort focused on understanding and advancing how students use LLMs for learning.
- Explored advanced usage patterns of ChatGPT for academic problem-solving, research assistance, and productivity, contributing insights and feedback to the program.

Open Source Contributor July 2024 – Oct 2024

GirlScript Summer of Code (GSSoC) *Remote*

- Selected among **top 8% contributors** (**185 / 2613**) in GirlScript Summer of Code, contributing under a structured mentorship program.
- Implemented features, bug fixes, and documentation improvements across open-source projects using **GitHub pull requests, issue tracking, and code reviews**.

Open Source Contributor Oct 2024

Hacktoberfest *Remote*

- Recognized as a **Hacktoberfest Super Contributor** for sustained, high-quality contributions across multiple open-source repositories.

Core Team Member – Domain Expert (Web & ML) Sep 2025 – Present

Recurse, KMIT Technical Club *Hyderabad, India*

- Coordinated the design and refinement of **technical problem statements** for a 24-hour offline hackathon, ensuring clarity, feasibility, and appropriate difficulty levels.
- Evaluated participant submissions and solutions, providing technical feedback and assisting in shortlisting and judging based on correctness, approach, and implementation quality.
- Supported participants during the hackathon by resolving **Git, tooling, and development workflow issues**, enabling smooth progress under time constraints.
- Contributed to club web initiatives and collaborated with design and operations teams on event logistics, coordination, and sponsor outreach.
- Currently serve as a **domain expert**, mentoring students in web development and machine learning through guidance, doubt resolution, and technical reviews.

Web Development Intern October 2025 – Present

PR Committee, KMIT *Hyderabad, India*

- Contributed to frontend development for college web platforms as part of a collaborative development team.

Publications & Intellectual Property

RAKSHAK: An Autonomous Security and Surveillance Rover Patent Application Published

Indian Patent Application No. 202541112724 A *Jan 2026*

- Co-inventor on a published patent describing an autonomous ground surveillance rover integrating LiDAR, vision sensors, hybrid localization, sensor fusion, and dual-mode autonomous/manual control for perimeter security.