

SHASHVAT SINGHAL

+91 9582724000

Shashvat990@gmail.com

[Google Scholar](#) | [Github](#) | [LinkedIn](#)

EDUCATION			
B.TECH(Computer Engineering)	2022-2026	Delhi Technological University, New Delhi	8.1
CBSE (Class XII)	2022	DPS Rohini	96%
CBSE (Class X)	2020	DPS Rohini	92%
EXPERIENCE			

AI Research Engineer, Massachusetts Institute of Technology (MIT) July 2025 - Present

- Working under the mentorship of Prof. Manolis Kellis on free-form **knowledge graph** generation; improved efficiency and performance by 200x, developed a **custom entity aggregation** algorithm using pattern matching and structured LLM outputs, and contributed minor enhancements to the frontend.

Technical Intern, Makers Lab, Tech Mahindra May 2025 - July 2025

- Worked on the **Flagship Project Indus** (India's LLM), where I optimized file processing, cleaning, and translation using multithreading, async I/O, and semaphores, and performed **supervised fine-tuning** of Indus 2.0 using **PEFT** and **LoRA**, achieving a **30% increase** in model accuracy.
- Researched Agentic and MCP integration strategies for enhancing Indus's capabilities using **LangChain** framework.

Software Developer Intern, Grain Analytics April 2024 - June 2024

- Designed and developed the frontend application for Grain Analytics' official website using **Next.js** and **Tailwind CSS**, improving user experience and brand consistency.
- Also provided inputs for product development and integration of AI and analytics.

AI/ML Research Intern, Samsung Lab September 2023 - February 2024

- Developed **advanced ensemble models** using ML/DL techniques for Soil Organic Carbon (SOC) prediction with hyperspectral satellite imagery, applying noise and dimensionality reduction methods to optimize performance and achieved **state-of-the-art accuracy**.
- Research paper of the same has been published in International IEEE conference (ARIIA-2024), and received **best paper award**.

Data Science Intern, CEI India July 2023 - September 2023

- Worked on multiple data-driven projects, applying data processing, cleansing, and visualization techniques using Python and frameworks like Pandas, Matplotlib, and Seaborn. Also applied machine learning algorithms with Scikit-learn.

ACADEMIC PROJECTS

Advanced Citation-Grounded RAG for Large-Scale Official Document Collections ([Link](#))

- Developed a scientific **citation-based RAG system** using LangChain to retrieve **grounded** answers from a custom DTU knowledge base by **ingesting and indexing** all official PDFs (over 800+) and notices via web scraping and parallelization; built a Streamlit frontend to enable citation-backed queries with direct PDF source access.

Real-Time Hand Gesture-Based Drone Control System

- Designed a **real-time drone control** system by integrating OpenCV for video preprocessing and augmentation, and **MediaPipe** for hand landmark extraction with normalization; trained ML/DL models on the processed landmarks to recognize gestures and control a Tello drone.

ACADEMIC ACHIEVEMENTS AND AWARDS

- National Winner** of **Smart India Hackathon 2023** (World's biggest Open Innovation Model) ([Link](#))
- 2nd Runner Up**, IIT Roorkee Cognizance 2023 Techfest, CCD - Cloud Coverage Determination ([Link](#))
- Grand Finalist** of **Smart India Hackathon 2024**, selected in the **Top 5 teams** all over India
- Co-authored a research paper that won the **Best Paper Award** at the **International IEEE Conference** (ARIIA-2024) ([Link](#))
- Secured a **global rank of 19** in Codechef Starters 130 ([Link](#))
- Achieved **3 Star** rating on **CodeChef (1770+)** in competitive programming contests ([Link](#))
- Solved over **650+ questions** on **Leetcode**, with a contest rating of **1750+**. ([Link](#))
- Achieved an average typing speed of **100+ WPM**. ([Link](#))
- Achieved a 99+ percentile in JEE Mains 2022**, ranking among the top from over a million students.

PUBLICATIONS

- Audio-based Machine Fault Diagnosis using Hybrid Feature Extraction and Ensemble Learning** ([Link](#))
15th International Conference on Computing, Communication and Networking Technologies (ICCCNT 2024)
- DSeP-xNet: A Feature Optimized Ensemble Framework for Hyperspectral-Based Soil Organic Carbon Prediction** ([Link](#))
2024 IEEE International Conference on Augmented Reality, Intelligent Systems, and Industrial Automation (ARIIRA) (Best Paper Award)
- Boosting ML Techniques for Dementia Prediction Using Modifiable and Non-Modifiable Risk Factors** ([Link](#))
AAIC Neuroscience Next 2025 Conference
- Autonomous Cheating Detection for Online Examinations using Knowledge Distillation and Multi-task Learning**
16th International Conference on Computing, Communication and Networking Technologies (ICCCNT 2025)

SKILLS

- Courses:** Machine learning, Deep learning, Data structures and Algorithms, Object Oriented Programming, Discrete Maths, Operating system, Database management systems.
- Programming Languages/Frameworks:** Python, C++, C, SQL, Tensorflow, Keras, Pandas, Scikit-learn, NumPy, HTML, CSS, Tailwind, Javascript (basic), React, FastAPI, Express, LaTeX overleaf.