

EDUCATION

- **International Institute of Information Technology, Hyderabad** Hyderabad, India
B.Tech + M.S. (by research) in Computer Science and Engineering *September 2022 – May 2027*
CGPA: 8.17
→ Undergraduate Researcher at **Precog Research Lab**
→ Coordinator, **Programming Club**
→ Teaching Assistant for Data Structures and Algorithms Course Spring'25
- **Army Public School** Ahmednagar, India
Secondary and Higher Secondary Education *April 2019 - May 2022*
Percentage: Class X : 92.2, Class XII : 93.2

EXPERIENCE

- **Indian AI Olympiad (INAIO)** Remote
Technical Committee *January - March 2025*
 - Created and tested problems to test candidates core understanding of ML, and reviewed participant solutions.
- **Microsoft Research** Bangalore, India
Research Intern *July - September 2024*
 - Developed Labeling Copilot – a unified framework for rapid synthetic data creation and annotation, integrating data discovery, augmentation, and foundation models to overcome data scarcity and enable the development of robust computer vision models across industries.
- **SmartTerra** Bangalore, India
Machine Learning Intern *January - May 2024*
 - Developed a conversational chatbot to assist non-technical users in analyzing MongoDB datasets via natural language queries. Enabled in-depth data visualization and manipulation with large datasets, without requiring manual scripting or data analysis expertise.
- **Indian Computing Olympiad** Remote
Problem Setter and Author *2023*
 - Created problems for Zonal Informatics Olympiad'2023 and published articles for the official INOI (Indian National Olympiad in Informatics) guide including topics like basic graph theory
- **InterviewMocha** Remote
Algorithmic and Data Structural Problem Setter *June 2022 - June 2023*
 - Created engaging and original problems for companies to use during their hiring process
 - Involved problem setting, writing editorials, providing solutions in multiple languages
- **Codechef** Remote
Teaching Assistant *June 2022 - October 2022*
 - Assisted with and resolved 300+ user doubts and involved in debugging, clarifying problem solutions, and resolving general programming doubts.

PUBLICATIONS

- [4] **SEMMA: A Semantic Aware Knowledge Graph Foundation Model** [PDF]
Arvindh Arun*, **Sumit Kumar***, Mojtaba Nayyeri, Bo Xiong, P. Kumaraguru, Antonio Vergari, Steffen Staab
Currently under review.
- [3] **A Cognac shot to forget bad memories: Corrective Unlearning in GNNs** [PDF]
Varshita Kolipaka*, Akshit Sinha*, Debangan Mishra, **Sumit Kumar**, Shashwat Goel, Arvindh Arun, P. Kumaraguru
42nd International Conference on Machine Learning (ICML 2025)
- [2] **Labeling Copilot: Tooling for a Rapid Synthetic Data Creation & Annotation**
Debargha Ganguly*, **Sumit Kumar***, Ishwar Balallapanawar*, Shashank Kambhatla, Srinivasan Iyengar, Shivkumar Kalyanaraman, Vipin Chaudhary, P. Kumaraguru
Currently under review.
- [1] **Sanity Checks for Evaluating Graph Unlearning** [PDF]
Varshita Kolipaka*, Akshit Sinha*, Debangan Mishra, **Sumit Kumar**, Shashwat Goel, Arvindh Arun, P. Kumaraguru
3rd Conference on Lifelong Learning Agents Workshop Track (CoLLAS-W '24)

PROJECTS

- **Network File System:** Created a robust and efficient system that allows users to seamlessly interact with files and folders over a network. The NFS architecture consists of three major components: Clients, Naming Server, and Storage Servers.
- **StyleSynth.AI (Flipkart 5.0):** Developed an AI fashion tool that incorporates text and image Generative AI models to provide personalized outfit recommendations to users; Integrates LLaMa-based chat bot by GPT4ALL as well as the Stable Diffusion API from HuggingFace for a smooth user experience
- **Custom Shell:** Built a user-defined interactive shell program that can create and manage new processes in C using system calls in linux. Implemented several shell comands for functions like I/O redirection, process management, piping, etc.
- **Traffic Management System (IoT):** Built an end-to-end IoT system that models traffic management and congestion control for free passage of emergency vehicles; handles real-time monitoring and integration with emergency services
- **MatchBox - Tinder for Roommates:** Developed an end-to-end web application for peers to find their ideal roommate; features include a matching algorithm, platform to share roommate preferences, and a user-friendly interface
- **MOOOSIC - Internet Music database:** Designed a web portal for an Internet Music Database. This website has information about various artists, albums, and songs, with features like search, the ability to rate songs and view existing ratings, to comment on songs and view existing comments, etc.

SKILLS SUMMARY

- **Languages:** C, C++, Python, Kotlin, Assembly, HTML, CSS, Javascript, Flask, SQL, Bash
- **Skills:** Competitive Programming, Web Development, Scripting in Bash/Python, DBMS, Git, PyTorch

ACHIEVEMENTS

- **International Collegiate Programming Contest (ICPC):** Secured a national rank of 28 out of 2500 Indian college teams, qualified for Asia Amritapuri Regionals 2022, awarded with honorable mention
- **Indian National Olympiad in Informatics(INOI):** Ranked top 30 nationally, awarded a Silver medal at INOI 2022, prepared 4+ years for IOI
- **Football:** Awarded Best Player in the college for the year 2024 and I play as center attacking midfield for the college football team.
- **Google Code-In 2019:** Collaborated with various open source organisations and completed 18+ tasks
- **Meta HackerCup 2022:** Qualified for Round 2, secured rank 1274 out of 30,000 participants
- **Codechef Snackdown 2022:** Qualified for Round 2 of Global Programming Contest conducted by Codechef
- **Max rating of 1894(Expert) on Codeforces.:** Top 4% globally
- **Max rating of 2023(5*) on Codechef.:** Top 0.9% globally
- **CodeChamps - Codingal:** Secured rank 1 out of 187 students globally