

Jivitesh Jain

✉ jivitesj@cs.cmu.edu | 🌐 jivitesh.dev | 🌐 /jiviteshjain | 📄 Google Scholar | 🌐 jiviteshjain

EDUCATION

Carnegie Mellon University Aug 2024 – Jun 2026
MS in Language Technologies, School of Computer Science | Funded Research & Thesis Pittsburgh

International Institute of Information Technology, Hyderabad Aug 2018 – Jul 2022
Bachelor of Technology (Honors), Computer Science and Engineering | GPA 10.0/10.0 India

- **Institute Gold Medalist** and on the **Dean's List** in all semesters; the first student to graduate with a **perfect GPA**.
- Received the **Research Award**. TA for Operating Systems and Computational Social Science courses.
- Qualified for **ACM ICPC Asia-West** regionals. In top 500 ranks worldwide in **Google Kick Start & Hash Code**. Invited to participate in **Google Foo Bar** and attend **Google Research Week 2022**.
- Open-source contributor to the Julia Programming Language as part of **Major League Hacking (MLH) Fellowship**.

EXPERIENCE

Palantir Sep 2022 – Aug 2024
Software Engineer, Foundry & AI-Platform London

- Led the development of the **Data Exports** infrastructure which can generalize to **150+** destination systems, robustly handle **multi-TB scale data**, and is **1.5x faster** than the previous method, which could only connect to **<10** systems and was difficult to maintain and use. Collaborated with power users and teams across organizations (Product, Network Infrastructure, Security) and spearheaded the effort from design to delivery; **used Spark and Java**.

Google May 2021 – Jul 2021
Software Engineer Intern, Apps Search & Intelligence Bengaluru

- Designed and built a **model-testing framework** to test Gmail search auto-suggest models on user logs for recall and quality of suggestions, **shortening the testing cycle from weeks to hours**. Used **C++** and **MapReduce**.
- Ran ablation studies on several models. **Accomplished all stretch-goals** and productionized the pipelines.

RESEARCH

Language Technologies Institute, CMU – Prof. Mona Diab Aug 2024 – Present

- Classifying and **minimizing hallucinations in LLMs** through an interpretability-based analysis of model activations.
- Adapting and analyzing LLMs for privacy and compliance monitoring of software systems and processes.

3D Computer Vision & Learning Lab, Brown University – Prof. Srinath Sridhar Jul 2021 – May 2022

- Developed a tensor field deep-learning network that achieved SOTA results on pose-canonicalization of 3D objects.
- **Publication:** *ConDor: Self-Supervised Canonicalization of 3D Pose for Partial Shapes* at **IEEE/CVF CVPR 2022**.

Precog Social Computing Lab, IIIT Hyderabad – Prof. Ponnurangam Kumaraguru Jan 2021 – May 2022

- Used multilingual NLP and graphical models to analyze the characteristics and use of social media networks in India.
- **Publication:** *What's Kooking? Characterizing India's Emerging Social Network, Koo* won the **Best Student Paper Award** at **IEEE/ACM ASONAM 2021**.
- **Publication:** *Urbanization and Literacy as factors in Politicians' Social Media Use [...]* at **ACM COMPASS 2022**.

PROJECTS

Retrieval Augmented Generation: End-to-end implementation and analysis of **RAG for Q/A** from scratch, with knowledge corpus curation, **synthetic data generation and model finetuning**, SOTA embedding and indexing methods, query rewriting using **Hypothetical Document Embeddings (HyDE)**, document summarization, and model quantization.
Mini Llama: An implementation of **Llama 3.1** in **PyTorch**, with **Q-LoRA PEFT**, pretrained and finetuned from scratch.

SKILLS & COURSEWORK

Courses: On-Device ML, Advanced Natural Language Processing, Machine Learning, Optimization & Statistical Methods, Information Retrieval, Data Systems, Computer Vision, Data Structures & Algorithms, OS & Networks, Distributed Systems.
Skills: Python, C/C++, Java, JavaScript, ML, NLP, CV, Information Retrieval, PyTorch, Transformers, Quantization/Pruning, Efficient Training, Metal, Llama.cpp, ML Ops, Backend Architecture, OS & Networks, Databases, Spark, Linux, Git.