Aditya Bharat Soni

adityabs@cs.cmu.edu | 412-918-0307 | LinkedIn | GitHub

Education

Carnegie Mellon University - School of Computer Science **Dec 2025 (expected)**
Master of Science in AI and Natural Language Processing | **GPA: 4.17/4.00** Pittsburgh, PA
Courses: Introduction to ML, Advanced NLP, Introduction to DL, Generative AI, Multimodal ML

Indian Institute of Technology Kharagpur **Apr 2024**
Bachelor of Technology in Computer Science and Engineering | **GPA: 9.80/10** | **Department Rank: 2/89** Kharagpur, India

Experience

Google **May 2023 - Aug 2023**
Software Engineering Intern | Google Cloud Hyderabad, India
- Devised an innovative approach to efficiently compute storage consumption metrics for Google Cloud customers.
- Optimized the metric computation pipeline to significantly reduce the delay by 73% from 30 hours to 8 hours.
- Evaluated the approach on actual customer data and computed accurate metric values for >99% of storage buckets.

All Hands AI | OpenHands - Building Open-Source AI Agents **Sept 2024 - Present**
Graduate Research Assistant | Advisor: Prof. Graham Neubig CMU
- Developed an AI-powered browsing agent that can process multimodal data, including web-pages, images, and text.
- Benchmarked the browsing agent on VisualWebArena dataset and achieved a notable 26% success rate.
- Aim to improve agent's performance on benchmarks that evaluate document processing capabilities for diverse file types.

Tr^{2}AIL Lab | Trustworthy LLMs: Evaluating Robustness Against Malicious Inputs **Jul 2023 - Apr 2024**
Research Assistant | Nominated for Best Bachelor's Thesis Award (Top-5 CS Students) IIT Kharagpur
- Designed an algorithm to evaluate the adversarial robustness of various API-based LLMs including GPT-3.5 Turbo.
- Deployed gradient-based attacks against LLMs to learn universal input perturbations across multiple tasks and languages.
- Achieved >25% reduction in LLM performance for translation, question-answering, news-topic classification, and NLI tasks.

CNERG Lab | Social Biases in Knowledge Representations of Wikidata
Separate Global North from Global South | ACM Web Science 2025 (under review) **Aug 2022 - Sept 2023**
Undergraduate Research Assistant IIT Kharagpur
- Analyzed impact of social biases in knowledge graphs on occupation-based link prediction classifiers for 21 countries.
- Designed occupation classifiers with >80% accuracy from knowledge graph embeddings pre-trained using GNNs.
- Leveraged fairness metrics and clustering techniques to study patterns in occupation-related biases across countries.

Projects

Retrieval Augmented Generation for Question-Answering Systems | Advanced NLP **Oct 2024**
- Implemented a question-answering system using RAG and LLMs to answer factual queries about Pittsburgh and CMU.
- Utilized web-scraping to collect 35K documents and generated synthetic training dataset of 40K QA-pairs using LLMs.
- Deployed RAG pipeline using LangChain and fine-tuned Llama 3.1 on synthetic data to achieve 62% exact-match score.

Leveraging LLMs for Non-functional Code Requirements | Advanced NLP **Nov 2024**
- Designed a novel LLM-based method to classify which code satisfies non-functional requirements like security & latency.
- Achieved state-of-the-art accuracy on the NoFunEval benchmark using GPT-4o (71.3%) and DeepSeek-Coder (61.2%).

Section-Specific Ranking System for CS Research Papers | Information Retrieval **Feb 2023 - Mar 2023**
- Designed an approach to rank CS papers based on similarity with a query paper for a specific section like method/result.
- Leveraged citation data to fine-tune BERT-based document encoders via triplet loss to implement the ranking system.

Implementing the Linux Shell | Operating Systems Lab **Feb 2023**
- Implemented the Linux Shell using C++ system calls to support input/output redirection, piping and command-line history.
- Added support to handle keyboard interrupts and control foreground/background processes using custom signal handlers.

End-to-End CPU Development | Computer Architecture Lab **Oct 2022 - Nov 2022**
- Formulated the low-level design of a 32-bit single-cycle processor to support a custom assembly instruction set.
- Utilized Verilog to code the processor design and tested merge sort and binary search programs using a Nexys-4 FPGA.

Academic Achievements

Chanakya Fellowship: Awarded for my Bachelor's Thesis research on evaluating trustworthiness and robustness of LLMs.
IIT-JEE Entrance Exam: All India Rank 233 in JEE Advanced and **All India Rank 93** in JEE Mains among 1.2M students.

Skills and Relevant Coursework

Languages: Python, C++, C, Java, SQL, Bash, Verilog
Frameworks/Tools: PyTorch, HuggingFace, LangChain, NumPy, Pandas, Scikit-Learn, AWS, Docker, C++ STL, Linux, Git
Courses (undergraduate): NLP, Deep Learning, Machine Learning, Information Retrieval, OS, Networks, Algorithms