Darsh Agrawal

+1 (412) 918-0910 | darsha@andrew.cmu.edu | linkedin/darsh-agrawal

EDUCATION

Carnegie Mellon University - School of Computer Science **Aug 2024 - Dec 2025 (Exp.)**
Master of Science in Artificial Intelligence and Natural Language Processing Pittsburgh, PA

Relevant Courses: Machine Learning, Advanced Natural Language Processing, Multimodal Machine Learning.

Indian Institute of Technology (IIT), Jodhpur **July 2017 - May 2021**
Bachelor of Technology in Computer Science & Engineering Jodhpur, India

Relevant Courses: Linear Algebra and Calculus, Probability Statistics and Random Processes, Complex Analysis and Differential Equation, Algorithm Design and Analysis, Artificial Intelligence.

PROFESSIONAL EXPERIENCE

Goldman Sachs
Bengaluru, India

Software Engineer **June 2021 - June 2024**

- Core contributor to the firm's Natural Language Distributed Search Engine, providing both syntactic and semantic search capabilities.
- Built the search engine UI to enable search across seven distinct structured and unstructured datasets.
- Developed an infinite scroll model for loading and displaying results, which was universally adopted across the entire Compliance division.
- Developed document lookup and search dashboard services with 98% and 96% code coverage.
- Designed and integrated the query parser logic, retrieval qualifiers and ranking mechanisms used for driving information retrieval decisions.
- Developed distributed batch indexing jobs for large unstructured datasets (~ 1 TB data per day).
- Improved the ranking algorithm by capturing user interaction signals and training a deep learning model to identify query abandonment.
- Authored a versatile automated backfill script, reducing data backfill time by over 80%.
- Built services for ingestion and indexing of vector embeddings, supporting generative question answering.

Software Engineer Intern **May 2020 - July 2020**

- Fine-tuned Natural Language Processing (NLP) models to enable user query parsing.
- Developed Vision, an end-to-end user feedback tool, that more than halved the issue resolution time.

PROJECTS

Code Generation using LLMs.

- Implemented multi-GPU task-based fine-tuning for DeepSeek and Codestral models on Diachrony data.
- Created synthetic data for multi-law fine tuning and further fine-tuned the models.
- Evaluated MCTS for multi-law synthesis and code generation.

Min-LLaMa 2 Large Language Model.

- Implemented Attention, RMSNorm, RoPE, Llama layers, AdamW optimizer, and a sentiment classifier.
- Performed zero-shot sentiment analysis and task-based fine-tuning, with a 28% increase in accuracy.

MESH – Domain Specific Search Engine.

- Modified Google's page rank algorithm to improve score for domain-specific relevance.
- Customized the crawler algorithm to derive domain-specific signals, used for accurate ranking.
- Optimized the index algorithm for faster retrieval (2.2 times faster) using MD5 hashes.

TECHNICAL SKILLS

Programming Languages & OS: Java, Python, C, C++, Javascript, Typescript, Windows, Linux.

Frameworks & Libraries: Spring, Spring Boot, Hadoop, Map Reduce, Spark, Kafka, Angular, React, Django, Node.js, Docker, Kubernetes, Slurm, Elasticsearch, Vespa, Prometheus, Terraform, HuggingFace, Transformers, Deepspeed, PyTorch, NumPy, Pandas, SciPy, Scikit-learn.

Data Stores: Hbase, HDFS, Presto, MongoDB, Sybase, AWS S3.