

Aditya Srivastava

📞 718-200-1908 ✉ aditya.srivastava@colorado.edu [in linkedin.com/in/lamAdiSri](https://www.linkedin.com/in/lamAdiSri) github.com/lamAdiSri

Education

| | |
|---|-------------------------------|
| University of Colorado, Boulder <i>MS in Computer Science [GPA: 4.00 / 4.00]</i> International Institute of Information Technology <i>BTech in Computer Science and MS (Honors) in Computational Linguistics</i> <ul style="list-style-type: none">Dean's List | Expected May 2024 May 2022 |
|---|-------------------------------|

Technical Skills

Languages: Python, C/C++, Javascript, Java, PHP
Web: HTML, CSS, Django, ReactJS, Flask, FastAPI, jQuery, Bootstrap, Node.js
ML and NLP: Tensorflow, PyTorch, Keras, HuggingFace, LangChain, NumPy, Sklearn, spaCy
Data Warehousing: SQL, NoSQL (Redis, MongoDB, Neo4j), Elasticsearch, Spark
DevOps and MLOps: Amazon Web Services, Google Cloud Platform, Kubernetes, Terraform, Ansible, GitHub CI/CD
Miscellaneous: GNU-Linux, Windows, Git, LaTeX, Arduino, Docker, gRPC

Experience

| | |
|---|--|
| EthAirNet <i>Software Engineering Student Associate</i> <ul style="list-style-type: none">Developing a set of unicast and multicast and in-network-computation protocols at the data-link layer for datacenter HPC and AI applications in P4lang and C++.Employing stateless flow-zone switching using software defined addressing for highly efficient routing.Presented the work at the IEEE 802 Plenary Session 2024, at Denver. | Aug 2023 – Present Boulder, CO |
| Dropbox <i>DevOps Engineering Intern</i> <ul style="list-style-type: none">Built enterprise infrastructure and networks using Infrastructure-as-Code tools, such as Terraform and Ansible.Automated deployments using GitHub Actions CI/CD.Designed and implemented an end-to-end ChatOps pipeline in Python. | May 2023 – Aug 2023 San Francisco, CA |
| SentiSum <i>Machine Learning Engineer</i> <ul style="list-style-type: none">Built systems for translation, classification, anomaly detection and unsupervised information extraction.Made 3x improvement to translation throughput and improved performance by 15%, by switching away from online translation services and building custom models.Scaled and parallelized deep learning based natural language processing systems for efficient inference on production loads.Built APIs in Flask and FastAPI and employed cloud platforms such as AWS and GCP to serve live machine learning models. | Aug 2021 – Sep 2022 London, UK |

Projects

| | |
|--|----------|
| Retrieval Augmented Generation for File Search <ul style="list-style-type: none">Built an OpenAI GPT-4 based RAG pipeline using LangChain to perform natural language search over user data stored in a distributed file system.Won the AI First Award for the project at DropBox's Hack Week, 2023 hackathon. | Jul 2023 |
| HF-Trim github.com/lamAdiSri/hf-trim <ul style="list-style-type: none">Created an open-source Python library for trimming the vocabulary on pretrained language models, lowering memory requirements of models by up to 90% at minimal cost to performance.Employed in both research and production environments, with over 5k downloads.Actively maintained on GitHub and hosted on the Python Packaging Index. | Aug 2022 |
| Federated Learning on Private Chat Data github.com/lamAdiSri/federated_learning_bda <ul style="list-style-type: none">Privacy preserving ML on personal chat data, using federated learning algorithms.Implemented Map-Reduce routines in Apache Spark for highly efficient model aggregation and evaluation.Built with PyTorch and deployed to GCP, with extensive quantitative and qualitative testing on millions of samples. | Jan 2023 |
| TLDR for CODWOE, SemEval Workshop 2022 aclanthology.org/2022.semeval-1.6 <ul style="list-style-type: none">Achieved first place on the shared task of optimizing a large language model from scratch on dictionary data alone.Augmented the transformer architecture and explored techniques such as unsupervised pretraining, multitask learning and contrastive learning.Published at the CODWOE NLU shared task in the SemEval Workshop 2022, NAACL. | Jul 2022 |

Achievements

Ranked 8th and 11th globally in the Labor and SCRBL challenges respectively at Alcrowd, a competitive ML platform.
Winner of Howzhack 2019, India's biggest online hackathon, and **Megathon 2018**, a national university-level hackathon.