Aditi Tripathi

(+1) 412-708-7569 | aditit@andrew.cmu.edu | karasuno7.github.io | karasuno7 | linkedin.com/in/tripathi-aditi/

Education

Carnegie Mellon University (CMU)
Pittsburgh, PA

Masters in Electrical and Computer Engineering — *2019 - 2020*

- 11785 - Introduction to Deep Learning • 10601 - Introduction to Machine Learning • 15319 - Cloud Computing • 15640 - Distributed Systems • 18793 - Image and Video Processing • 18889 - Internet Services • 17681 - Data Structures • 18613 - Introduction to Computer Systems

National Institute of Technology (NIT)
Bhopal, India

Bachelors in Electrical Engineering — *2012 - 2016*

Skill Set

- **Languages** : Python, Java, C, C++, Scala,SQL, Matlab
- **Tools** **:** Scikit-Learn, Pytorch, Tensorflow, OpenCV, Apache Spark, AWS, GCP, Bash, Git, Vim etc.
- **Skills** **:** Multimodal Machine Learning, Natural Language Processing, Image Processing, Distributed Computing

Work Experience

Research Assistant - advised by Prof. Carolyn Rose, CMU SCS, Pittsburgh, PA
Feb 2021 - Present

- Developing intelligent conversational agents to support collaborative learning using machine learning and probabilistic models.
- Designed a performant JSON-RPC API to generate automatic hints for the students' code in real time using clustering based on modified tree-edit distance for abstract syntax trees.

Data Scientist Intern, Boeing, Bellevue, WA
June 2020 - Aug 2020

- Developed a more memory efficient and a 4x faster image processing framework (Using different Clustering algorithms for segmentation and use of OCR based image stitching instead of overlap detection).
- Achieved a cross validation accuracy of 95% with an XG-Boost clustering model for airplane part price prediction.

Graduate Research Assistant - advised by Prof. Richard Stern, CMU, Pittsburgh, PA
Oct 2019 - Dec 2019

- Developed a java user-interface for a GIS based air traffic noise prediction model for overlaying helicopter trajectory on top of raster maps created using LIDAR data points to display noise levels around Allegheny General Hospital.

Research Consultant, Worldquant LLC, Virtual Research Center, India
Dec 2018 - Jul 2019

- Developed new ML models to seek out sources of inefficiencies, and build predictive profitable strategies.
- Developed several low turnover high quality alphas for trading in the equity market used in daily re-balancing long-short algorithmic trading strategies on US, Europe, Japan and other markets.

C&I Engineer, L&T Mitsubishi Hitachi Power Systems, Chennai, India
Aug 2016 - Nov 2018

- Developed anomaly detection models for sensor based systems to enable predictive maintenance, early fault detection, time to failure prediction and resource optimization.
- Deployed a neural network based wall temperature prediction model with prediction accuracy of 92% on the experimental data.
- Coordinated with cross-functional Indo-Japanese team in creating SOPs, RFQs and design documents.

Academic Projects

Deep Learning
CMU, Pittsburgh

Tag: PyTorch, AWS, Tensorboard — *Fall 2020*

- **MyTorch**— Developed a custom library similar to PyTorch with computational graph functionality for automatic differentiation for MultiLayer Perceptron, 1D/2D CNN, and RNN networks with activation functions and ADAM/ SGD optimizers.
- **Attention-based End-to-End Speech-to-Text Deep Neural Network**– Implemented the Listen, Attend and Spell paper to design a system for speech-to-text transcription using Locked Dropout, Teacher Forcing and padded Cross Entropy Loss to achieve a Levenshtein Distance of 17.7 on Librivox dataset and top 11% on Kaggle Leaderboard. (kaggle)
- **Multimodal Visual Question Answering System**— A question answering system using BERT embeddings and multistage fusion for user's photo albums and photo metadata to enable multiple-choice question-answering and providing grounding photos based on the user's question. Achieved 5% improvement for "how many" and "where" type questions over the SOTA implementation.

Cloud Computing and Distributed Systems
CMU, Pittsburgh

Tag: MapReduce, Scala, Java, AWS — *Fall/Spring 2020*

- **Fault-tolerant 2PC Application** — Implemented a fault tolerant server with write-ahead-logging which publishes multi-client group photo-collage using 2PC protocol.
- **Java RPC File-Caching Proxy** — Developed the open-close semantics protocol between proxy and server where the multi-threaded proxy emulates C library calls and maintains an LRU file cache for concurrent clients to read and modify files without interference.
- **Big Data Analytics with MapReduce** — Processed 120GB dataset of Wikipedia traffic log to analyze current popular events using robust OS independent parallel algorithms running as MapReduce jobs on AWS EMR cluster.
- **Multithreading and Distributed Key-Value Store**— Built a globally distributed in-memory key-value store with strong consistency, using java multithreading API , to store sales records for a company.
- **Pagerank in Scala** — Analyzed the Twitter social graph by implementing the PageRank algorithm in Spark (Scala) using RDDs and DataFrame by performing iterative processing.

Dynamic Storage Allocator

CMU, Pittsburgh

Tag: C, Computer Systems

Fall 2019

- Realized a general-purpose dynamic storage allocator for C by implementing a self-defined “Malloc” library
- Applied segregated explicit lists (data structure) and first-fit block search (algorithm) for space management in heap to achieve better performance (30000 KOPS) and smaller memory utilizations (56.3%)
- Optimized the space utilization to 74.1% by eliminating footer in allocated blocks and handling small-size blocks.

NYC Cab Fare prediction App

CMU, Pittsburgh

Tag: Machine Learning, GCP

Fall 2020

- Deployed an end-to-end ML Flask App using a pipeline of cloud ML APIs and trained a XGBoost model on Google App Engine, with an interface to accept cab-price/location user queries as speech and respond with an audio result. Performed hyperparameter tuning with Bayesian optimization to improve the accuracy of the predictor.

Other Experience

Course Developer, CMU, Pittsburgh, PA

Jan 2021 - Present

- Teel Lab AI Practitioner Course - Developed course content for Information Retrieval and Question Answering Systems module.
- Teaching Assistant for 18474 - Developed Reinforcement Learning Car Racing Environment with DDPG and DQN agents in Matlab with Python bindings.

Undergraduate Research Assistant - advised by Prof. Tripta Thakur, NIT Bhopal

May 2014 - Jul 2014

- Developed optimal bidding strategy by comparing Genetic Algorithm and Monte Carlo approach in solving bi-level stochastic optimization problem. Paper presented in i-Fast Savishkar 2015 conference.