Anurag Katakkar

akatakka@alumni.cmu.edu | (412) 773-1297 | anuragkatakkar.github.io

EDUCATION

Carnegie Mellon University — Pittsburgh, PA
MS, Computational Data Science — GPA: 3.9/4.33 — May, 2021
Coursework: Intro. to Deep Learning (A+, Spring '21 TA), ML with Large Datasets (A), Neural Networks for NLP (A+)

Pune University — Pune, India
B.E., Computer Engineering — GPA: 9.0/10, First Class with Distinction — June, 2019

TECHNICAL SKILLS

Python (PyTorch, PySpark, SpaCy, NLTK, Scikit-Learn, Pandas, Matplotlib, TensorFlow 2.0), **C/C++** (CUDA, ISPC, OpenMP, MPI), **Cloud Services and DevOps** (AWS EC2, EMR, Azure HDInsight, Docker, Kubernetes, Kafka, Samza)

WORK EXPERIENCE

NVIDIA — Santa Clara, CA
Deep Learning Software Engineer, ***TensorRT*** — July, 2021 – Present
- Working on safety certification and integration of cuTensor math library into TensorRT backend

SWE Intern, ***TensorRT*** — May, 2020 – August, 2020
- Extended Fuzzed Neural Network Testing pipeline (generates random, verifiably correct neural networks) to support autonomous driving applications
- Built a tool to convert simple YAML specifications into fully functional ONNX models

IBM India — Pune, India
Project Intern — Fall 2018 - Spring 2019
- Re-trained a Text-to-Image GAN on the CelebA dataset to replace sketch artists in developing countries
- Deployed as an online incident report platform for expediting criminal investigations (among top-10 Senior Year Capstones)

NTT Data — Tokyo, Japan
Intern, ***Advanced AI Team*** — June, 2018 – July, 2018
- Implemented an end-to-end Information Retrieval system on Wikipedia data, using a hybrid TFIDF-BM25 algorithm for the NFCorpus medical dataset.

PUBLICATIONS

- **Towards Language Modelling in the Speech Domain Using Sub-word Linguistic Units**, Anruag Katakkar, Alan W Black, arXiv
- **Practical Benefits of Feature Feedback Under Distribution Shift**, Anruag Katakkar*, Clay H Yoo*, Weiqing Wang*, Zachary C. Lipton, Divyansh Kaushik, arXiv

PROJECT WORK

Explainable Machine Learning Machine Learning — Spring 2021
- Benchmarked current explanation methods including CXPlain and DiCE, against novel causal method (under submission at FaCCT 2022),
- Constructed an Amazon Mechanical Turk study for crowd-sourcing human preferences to various explanation methods

Probing Neural IR Models Deep Learning, IR — Spring 2021
- Constructed probing tests for contemporary neural IR models, such as Doc2Query, based on query and question categories

MyTorch Framework Deep Learning — Fall 2020
- Built a fully functional autograd-driven DL framework with forward/backpropagation implementation for Conv1D, Conv2D, LSTM, GRU and optimisation algorithms including Adam, SGD, RMSprop

Speech Processing Deep Learning, Speech, Textless NLP — Fall 2020
- Built a linguistic speech language model purely from speech melspectrograms (Capstone Project, **received A+ grade**)
- Speech-to-Text : Implemented Attention-based encoder-decoder model with teacher forcing and beam search decoding

Pedestrian Trajectory Prediction Deep Learning, Vision (top-5 best course projects, Intro. to DL) — Fall 2020
- Implemented Convolutional-LSTM spatiotemporal architectures for predicting future trajectories of agents (project video)

Hyperparameter Optimisation for Image Classification Machine Learning — Fall 2020
- Implemented the Hyperband algorithm for DenseNet-121, with 30% accuracy boost and 4-8x speedup over Bayesian Optimisation (github)

Accelerating Sparse DNN Inference Systems/DL *15-618 Course Project* — Spring 2020
- Developed a CUDA+MPI solution to achieve 10-400x speedup over ref CPU Implementation for inference on sparse MNIST

Scalable Twitter Web Service Systems *15-619 Course Project* — Spring 2020

- Developed a web service using an AWS EC2 web tier, and AWS EMR and RDS database tier to achieve upto 50k RPS on compute intensive queries, and upto 12k RPS on IO intensive queries, at a cost of only $1.24/hour
- Performed ETL on 1.4TB of raw twitter data using Azure HDInsight Spark Clusters to prepare the database tier

SERVICE
- Organizer of the Speech for Social Good Workshop, proposal under submission to INTERSPEECH 2022,
- Served as a reviewer for the Second Grand-Challenge and Workshop on Multimodal Language at ACL 2020

TEACHING & MENTORSHIP

11-785 Introduction to Deep Learning *Teaching Assistant* CMU, Spring 2021
- Conducted introduction lecture, recitations, and office hours. Class size $\approx$ 200 students
- Developed course homework modules, and mentored group projects to drive project direction and ensure course standards
- Developed a new autodiff library, "NewGrad" (github), that was used as the main bonus assignment for the course

MCDS Capstone - Egocentric Trajectory Prediction *Project Mentor* CMU, Spring 2020
- Novel methods for automatic annotation of video data, and deep learning methods for egocentric trajectory prediction using that data,
- Submitted as a long paper to CVPR 2022

11-631 Data Science Seminar *Teaching Assistant* CMU, Fall 2020
- Organised guest lectures, conducted office hours. Class size $\approx$ 80 students
- Designed key course modules, and redesigned grading system. Developed course website