ADITYA ANANTHARAMAN

adityanant@gmail.com · +44 7884900911 · aditya5558.github.io · linkedin.com/in/aditya1997

Education

Carnegie Mellon University (CMU), School of Computer Science — Pittsburgh, USA
Master of Computational Data Science (MCDS) | *GPA:* 4.11/4.33 — 2019 - 2020

National Institute of Technology Karnataka, Surathkal (NITK) — Surathkal, India
Bachelor of Technology Information Technology | *GPA:* 9.54/10, *Class Rank:* 5/103 — 2015 - 2019

Experience

Applied Scientist II, Amazon — Feb 2021 - Present | London, GB

- Leveraging Large Language Models (LLMs) to develop contextual advertising products at Amazon.
- Developed Mixture of Experts (MoE)-based Behavioral Foundation Models pre-trained on customer behavior tasks to boost click-through rate (CTR) prediction in e-commerce advertisement. Used cross-architecture knowledge distillation (KD) to distill them to 350x smaller cache-friendly factorized models to deploy in low-latency settings.
- Proposed a novel method called Reverse Distillation (RD) which makes use of knowledge from smaller models to train large encoder-based transformer models (upto 50 billion parameters) for CTR prediction.
- Leveraged LoRA-based parameter-efficient fine-tuning to adapt LLMs for product retrieval and ranking applications.

Applied Scientist Intern, Amazon — May 2020 - Aug 2020 | Seattle, USA

- Developed co-teaching-based and BERT-based models to link context-of-use entities with products to improve search experience. Proposed approaches improved both precision & coverage of links compared to lexical matching.

Research Intern, IIT-Hyderabad — Aug 2018 - Dec 2018 | Hyderabad, India

- Developed a novel multi-space model for Zero-Shot Object Detection (ZSD) which outperformed the state-of-the-art on Pascal VOC by 14% in mean average precision (mAP).

Software Engineering Intern, Microsoft — May 2018 - July 2018 | Hyderabad, India

- Developed a plug and play service for effective management, monitoring and usage of test clusters for the Azure Networking team.

Selected Publications

- J Zhang, A Muhamed, **Aditya Anantharaman**, G Wang, C Chen , K Zhong, Q Cui, Y Xu, B Zeng, T Chilimbi, Y Chen *"ReAugKD: Retrieval-Augmented Knowledge Distillation For Pre-trained Language Models"*, ACL 2023 (Oral) [Paper] [Blog]
- M Yang*, **Aditya Anantharaman***, D C Robert*, Z Kitowski* *"Graph Relation Transformer: Incorporating pairwise object features into the Transformer architecture"*, Visual QA Workshop, CVPR 2021 [Paper]
- D Gupta, **Aditya Anantharaman**, N Mamgain, S Kamath, V Balasubramanian, C V Jawahar *"A Multi-Space Approach to Zero-Shot Object Detection"*, WACV 2020 [Paper]
- M Vikram, **Aditya Anantharaman**, Suhas BS, S Kamath, *"An Approach for Multimodal Medical Image Retrieval using Latent Dirichlet Allocation"*, India KDD CoDS-COMAD 2019 (Oral). Short paper at AI for Social Good Workshop, NeurIPS 2018 [Paper] [Poster]

Academic Projects

Graph Relation Transformer for Text Visual Question Answering (Text-VQA) — Paper | Fall 2020

- Proposed a multimodal Graph Relation Transformer which leverages transformer layers for graph attention computation with rich edge and node information for the Text-VQA task.

End-to-End 2D to 3D Video Conversion — Github | Spring 2020

- Extended Deep-3D model using monocular depth estimation and segmentation masks from Mask-RCNN.

Fact Extraction and Verification (FEVER shared task) — Github | Spring 2020

- Implemented a BERT-based model and strengthened claim verification module using Multi-Task Deep Neural Networks (MT-DNN) and Stochastic Answer Networks (SAN) in addition to multi-hop evidence reasoning.

Skills

Programming Languages:	Python, C++, C, Java, MySQL
Deep Learning:	PyTorch, TensorFlow, Deepspeed
Cloud Platforms and Tools:	AWS, Azure, Google Cloud Platform, Hadoop MapReduce, PySpark