

# ADITYA ANANTHARAMAN

adityanant@gmail.com • +44 7884900911 • [aditya5558.github.io](https://github.com/aditya5558) • [linkedin.com/in/aditya1997](https://linkedin.com/in/aditya1997)

## EDUCATION

<b>Carnegie Mellon University (CMU), School of Computer Science</b> Master of Computational Data Science (MCDS)   <i>GPA: 4.11/4.33</i>	Pittsburgh, USA 2019 - 2020
<b>National Institute of Technology Karnataka, Surathkal (NITK)</b> Bachelor of Technology Information Technology   <i>GPA: 9.54/10, Class Rank: 5/103</i>	Surathkal, India 2015 - 2019

## EXPERIENCE

<b>Applied Scientist II, Amazon</b> • 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.	Feb 2021 - Present   London, GB
<b>Applied Scientist Intern, Amazon</b> • 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.	May 2020 - Aug 2020   Seattle, USA
<b>Research Intern, IIT-Hyderabad</b> • 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).	Aug 2018 - Dec 2018   Hyderabad, India
<b>Software Engineering Intern, Microsoft</b> • Developed a plug and play service for effective management, monitoring and usage of test clusters for the Azure Networking team.	May 2018 - July 2018   Hyderabad, India

## 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

<b>Graph Relation Transformer for Text Visual Question Answering (Text-VQA)</b> • 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.	<a href="#">Paper</a>   Fall 2020
<b>Fact Extraction and Verification (FEVER shared task)</b> • 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.	<a href="#">Github</a>   Spring 2020

## 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

## ACADEMIC SERVICE

Reviewer: ACL 2024, EMNLP 2024, ES-FoMo-II Workshop ICML 2024