ADITYA ANNAVAJJALA

Atlanta, GA

adityaas@gatech.edu \leq linkedin.com/asaditya/ \leq www.adityaas.com

SUMMARY

Graduate student with unique strengths in Systems and Artificial Intelligence. Solid 5-year experience in building Systems for Natural Language Processing (Language Models) and Computer Vision (Object detection, Semantic Segmentation, 3D reconstruction) pipelines in award-winning cross-functional teams. Brings a unique mix of technical, critical thinking, leadership, and creative skills that help him thrive in ambiguity.

EDUCATION

Master of Science in Computer Science, Georgia Institute of Technology Aug 2021 - Dec 2023 (expected) Relevant Coursework: Statistical Machine Learning, First-order Optimization for ML, ML with Limited Supervision, Algorithms, Global Entrepreneurship & Innovation, Computer Vision

Bachelor of Engineering (Hons.) Computer Science, BITS Pilani Aug 2013 - May 2017 Relevant Coursework: Data Mining, Distributed Computing, Artificial Intelligence, Algorithms

PUBLICATIONS

- A. Annavajjala*, A. Khare*, H. Latapie, M. Lee, A. Tumanov DES: Delayed E-shrinking for Faster Once-for-all training, 2023 (Under review at NeurIPS '23)
- A. Khare, A. Agrawal, A. Annavajjala, H. Latapie, M. Lee, A. Tumanov FedOdin: <u>Fed</u>erated Learning For Efficient <u>On-Device Inference</u>, 2023 (Under review at AAAI '24)
- Y. Khandelwal, M. Arvind, A. Gupta, S. Kumar, P. Bagad, A. Madan, M. Lunayach, A. Annavajjala, A. Maiti, S. Jain, A. Dalmia, N. Deka, J. White, J. Doshi, A. Kanazawa, R. Panicker, A. Raval, S. Rana, M. Tapaswi, *NurtureNet: A Multi-task Video-based Approach for Newborn Anthropometry*, 2023 (Manuscript)
- A. Annavajjala, P. Goyal, S. Kumari, A. Wani, J.S. Challa, S. Islam, N. Goyal μDBSCAN: An exact scalable DB-SCAN algorithm for big data exploiting spatial locality In 2019 IEEE International Conference on Cluster Computing (IEEE Cluster 2019), Sep 22-26, Albuquerque, USA (27% acceptance rate)

ACHIEVEMENTS AND EXTRA CURRICULARS

- Selected as a student volunteer to attend ICCV 2023 in Paris, France
- Selected as 1/200 Young researchers around the world, **Heidelberg Laureate Forum**, 2023
- Recipient of a \$8K USD grant from Partnership for Advanced Computing (PACE), Georgia Tech
- Recipient of the Anne Robinson Clough Scholarship for International Students, Georgia Tech
- Member of the team that won the FastCompany Best World Changing Idea APAC 2021
- Recipient of \$735 USD Compute Grant and a registration for NeuRIPS 2020 as part of the MineRL Competition
- Member of a team that won a \$5M grant from the Bill & Melinda Gates Foundation
- Competed at the the \$5M IBM AI XPrize 2019 and was one of the top 32 teams qualified worldwide (and the only team from a developing world)
- Invited Talk at BITS Pilani on how to use AI for the next 4 billion Youngest speaker at the annual Alumni Research Talks (ART 2020) that invites outstanding alumni conducting research in computer science
- Invited Talk at Perceiving Systems, Max Planck Institute for Intelligent Systems titled Neonate weight estimation using monocular RGB video
- Google Summer of Code mentee, Apache Software Foundation, 2017

Graduate Researcher Aug 2021 - Present

Systems for Artificial Intelligence Lab, Georgia Institute of Technology

Atlanta, GA

Advisor: Prof. Alexey Tumanov

Efficient Convolutional Networks

• Developed mechanisms for efficient training of large convolutional models using weight sharing, reducing deployment cost to O(1) from O(N)

- Discovered theoretical insights to enable faster weight-shared training in large convolutional models that are generalizable and cost efficient
- Achieved SOTA performance at SOTA training cost on ImageNet training. Compared our approach with existing methods such as Once-for-all, and BigNAS

Efficient Transformers

- Identified latency vs. accuracy tradeoffs in BERT based transformer models
- Developed training strategies to train transformers cost efficiently to enable zero additional deployment cost for real time applications

Applied Scientist II (Level 5) (Internship) Amazon May 2022 - Aug 2022

Seattle, WA

- Built a latency-aware image retrieval algorithm with an AUC-ROC of 93% with latency ≤ 10 ms, as part of the Last Mile Technology org impacting all (≈10M/day) deliveries in North America (NA) region
- Implemented a deep metric-learning algorithm on a curated data set of over 20M images

Research Fellow & Machine Learning Scientist (Full-time)

Aug 2018 - Aug 2021

Collaborator: Prof. Angjoo Kanazawa

Mumbai, India

Wadhwani Institute for Artificial Intelligence

Advisors: Dr. Rahul Panicker, Dr. Makarand Tapaswi

- Founding member of a team that built an AI-based visual weighing machine. The vision is to identify the 10 million low birth weight babies missed each year without using any specialized equipment with an explicit goal to reduce neonatal mortality rate
- Developed 3D reconstruction algorithms obtaining a MAE of ≤ 100 grams, beating the least count of the conventional spring balances
- **Technical lead** building data collection and processing pipeline with about 5000 videos (400k images) collected to date. The abstraction schosen enabled the development of a data visualization tool
- Technical lead for building a configurable machine learning framework to increase flexibility and adaptability. Adopted by a research team of 20 researchers decreasing development time by 2-3x
- Negotiated and closed an annotation agency contract worth \$30k USD/year.
- Winner of a \$5M grant from the Bill and Melinda Gates Foundation, \$300K grant from the Botnar Foundation, Best World Changing Idea (APAC) by FastCompany, and qualified for the \$5M AI XPrize
- Invited talks at Max Planck Institute of Intelligent Systems, Tuebingen, and my alma mater BITS Pilani

Project Assistant (Full-time)

Aug 2017 - Aug 2018

Machine & Language Learning Lab, Indian Institute of Science

Bengaluru, India

Advisor: Prof. Partha Pratim Talukdar

- Built an auto-scalable system to construct a knowledge graph given a small set of seed documents. The process involved a distributed crawler, a one-class classifier and infrastructure support for auto scaling
- Devised an algorithm for document classification with an accuracy of 94% beating state of the art i.e. OCSVM, PUSVM, SVM on the NYT dataset

PROJECTS

Domain Adaptation for Semantic Segmentation using Sparse Neural Networks | PyTorch, Python Dec 2021, Advisor: Prof. Judy Hoffman

- Proposed an efficient way for source-free domain adaptation by enforcing sparsity constraints obtaining a 4% performance boost on SOTA
- Video overview of our method and results can be found here

SKILLS

Languages Python, C/C++, Scala

Frameworks Ray, PyTorch, JAX, Blender, Tensorflow, Amazon Web Services (EC2, RDS, S3, Sagemaker, Lambda)

Tools Docker, Kubernetes, VS Code, PyCharm, Git & GitHub, CircleCI

Skills Software Design, Computer Vision, Deep Learning, Product Management, Leadership

TEACHING EXPERIENCE

Graduate Teaching Assistant - Computer Vision (400 students)

Graduate Teaching Assistant - Web Search and Text Mining (110 students)

Head Teaching Assistant - Systems for Machine Learning (40 students)

Undergraduate Teaching Assistant - Data Mining (200 students)

Fall 2016, Spring 2017

Undergraduate Teaching Assistant - Principles of Programming Languages (300 students)

Fall 2016

FEATURED ARTICLES AND AUTHORED BLOGS

Featured: Best World Changing Idea (APAC), Fast Company, 2021

Featured: MIT Emerging Technologies Conference, 2019

Featured: \$5M AI XPrize for solving grand challenges facing humanity

Blog: Technical blog on ML System Design

Blog: Technical blog on data and annotation pipelines

Blog: Technical blog on identifying low birth weight babies using AI

LEADERSHIP

- Member, Leadership Education and Development program at Georgia Tech
- Co-founder, Gyanbodh Built and led a team of 25 volunteers responsible for enrollment and tutoring of 7 child labourers into schools via the Right To Education act. The children have been going to school regularly for 5 years now