Aditya Kusupati

Bill & Melinda Gates Center for Computer Science & Engineering
Paul G. Allen School of Computer Science & Engineering
University of Washington, Seattle, WA, USA 98195

kusupati@cs.washington.edu
adityakusupati.com
Google Scholar

Research Interests

I focus on designing fundamental *Machine Learning* algorithms with strong empirical performance & real-world deployability geared towards enabling **efficient, elastic and contextual intelligence**.

Education

University of Washington, Seattle *2019 - present*
PhD Candidate in Computer Science and Engineering
Advisors: Prof. Ali Farhadi & Prof. Sham Kakade
Committee: Prof. Luke Zettlemoyer, Prof. Zaid Harchaoui & Dr. Rahul Sukthankar

Indian Institute of Technology Bombay *2013 - 2017*
B.Tech (Honours) in Computer Science and Engineering with Minor in Electrical Engineering
Advisor: Prof. Soumen Chakrabarti

Work Experience

Google Research *January 2022 - Present*
Student Researcher in Perception
Advisors: Dr. Prateek Jain, Tom Duerig & Dr. Rahul Sukthankar
Work on making fundamental Machine Learning algorithms elastic, flexible, and end-to-end differentiable for efficient and accurate deployment in web-scale systems like Google Search & Ads along with top-tier publications. Leading multiple research and product adoption efforts with a team of interns, research, and software engineers.

Berkeley Artificial Intelligence Research Lab *Summer 2022*
Visiting Researcher
Advisor: Prof. Alexei A. Efros
Worked on understanding the underlying discretization of the natural visual world through variable length representations.

NVIDIA Toronto AI Lab *June - September, 2020*
Research Scientist Intern
Advisor: Prof. Sanja Fidler
Explored missing modality (audio/visual/control) generation with cross-modal supervision for Atari video games using GANs.

Microsoft Research India *2017 - 2019*
Research Fellow in Machine Learning and Optimization
Advisors: Dr. Manik Varma & Dr. Prateek Jain
Worked on resource-efficient and large-scale machine learning resulting in top-tier publications & deployment in Bing.

Publications

Preprints

* - equal contribution

2. **MatFormer: Nested Transformer for Elastic Inference**
Devvrit*, Sneha Kudugunta*, **Aditya Kusupati***, Tim Dettmers, Kaifeng Chen,
Inderjit Dhillon, Yulia Tsvetkov, Hannaneh Hajishirzi, Sham Kakade, Ali Farhadi and Prateek Jain.
Under Review, ***ICML*** *2024.*
Efficient Natural Language and Speech Processing workshop @ NeurIPS 2023 **(Oral, 🏆 Best Paper Award)**.
Workshop on Advancing Neural Network Training @ NeurIPS 2023 **(Oral)**.

1. **EHI: End-to-end learning of Hierarchical Index for Efficient Dense Retrieval**
Ramnath Kumar*, Anshul Mittal*, Nilesh Gupta, **Aditya Kusupati**, Inderjit Dhillon and Prateek Jain.
Under Review, ***ICML*** *2024.*

Conference Publications

17. **Gecko: Versatile Text Embeddings Distilled from Large Language Models**
Jinhyuk Lee*, Zhuyun Dai*, Xiaoqi Ren*, Blair Chen, Daniel Cer, Jeremy R. Cole, Kai Hui, Michael Boratko, Rajvi Kapadia, Wen Ding, Yi Luan, Sai Meher Karthik Duddu, Gustavo Hernandez Abrego, Weiqiang Shi, Nithi Gupta, **Aditya Kusupati**, Prateek Jain, Siddhartha Reddy Jonnalagadda, Ming-Wei Chang and Iftekhar Naim.
Google Technical Report, 2024.

16. **SHARCS: Efficient Transformers through Routing with Dynamic Width Sub-networks**
Mohammadreza Salehi, Sachin Mehta, **Aditya Kusupati**, Ali Farhadi and Hanna Hajishirzi.
*Empirical Methods in Natural Language Processing (**EMNLP**) Findings, 2023.*

15. **Objaverse-XL: A Universe of 10M+ 3D Objects**
Matt Deitke, Ruoshi Liu, Matthew Wallingford, Huong Ngo, Oscar Michel, **Aditya Kusupati**, Alan Fan, Christian Laforte, Vikram Voleti, Samir Yitzhak Gadre, Aniruddha Kembhavi, Carl Vondrick, Georgia Gkioxari, Kiana Ehsani, Ludwig Schmidt and Ali Farhadi.
*Neural Information Processing Systems (**NeurIPS**) Dataset and Benchmarks Track, 2023.*

14. **MADLAD-400: Monolingual And Document-Level Large Audited Dataset**
Sneha Kudugunta, Isaac Caswell, Biao Zhang, Xavier Garcia, Christopher A. Choquette-Choo, Katherine Lee, Derrick Xin, **Aditya Kusupati**, Romi Stella, Ankur Bapna and Orhan Firat.
*Neural Information Processing Systems (**NeurIPS**) Dataset and Benchmarks Track, 2023.*

13. **Neural Priming for Sample-Efficient Adaptation**
Matthew Wallingford*, Vivek Ramanujan*, Alex Fang, **Aditya Kusupati**,
Roozbeh Mottaghi, Aniruddha Kembhavi, Ludwig Schmidt and Ali Farhadi.
*Neural Information Processing Systems (**NeurIPS**), 2023.*

12. **AdANNS: A Framework for Adaptive Semantic Search**
Aniket Rege*, **Aditya Kusupati***, Sharan Ranjit, Alan Fan, Qingqing Cao,
Sham Kakade, Prateek Jain and Ali Farhadi.
*Neural Information Processing Systems (**NeurIPS**), 2023.*
Practical ML for Developing Countries workshop @ ICLR 2023 **(Oral)**.

11. **FLUID: A Unified Evaluation Framework for Flexible Sequential Data**
Matthew Wallingford, **Aditya Kusupati***, Keivan Alizadeh-Vahid*, Aaron Walsman,
Aniruddha Kembhavi and Ali Farhadi.
*Transactions on Machine Learning Research (**TMLR**), 2023.*

10. **Neural Radiance Field Codebooks**
Matthew Wallingford, **Aditya Kusupati**, Alex Fang, Vivek Ramanujan,
Aniruddha Kembhavi, Roozbeh Mottaghi and Ali Farhadi
*International Conference on Learning Representations (**ICLR**), 2023.*

9. **Matryoshka Representation Learning.**
Aditya Kusupati*, Gantavya Bhatt*, Aniket Rege*, Matthew Wallingford, Aditya Sinha, Vivek Ramanujan, William Howard-Snyder, Kaifeng Chen, Sham Kakade, Prateek Jain, and Ali Farhadi.
*Neural Information Processing Systems (**NeurIPS**), 2022.*
Vision Transformers: Theory and Applications workshop @ NeurIPS, 2022 **(Oral)**.
Self-Supervised Learning - Theory and Practice workshop @ NeurIPS, 2022.
Computer Vision in the Wild workshop @ ECCV, 2022.

8. **MERLOT RESERVE: Neural Script Knowledge through Vision and Language and Sound**
Rowan Zellers, Jiasen Lu, Ximing Lu, Youngjae Yu, Yanpeng Zhao, Mohammadreza Salehi, **Aditya Kusupati**,
Jack Hessel, Ali Farhadi and Yejin Choi.
*IEEE Conference on Computer Vision and Pattern Recognition (**CVPR**), 2022* **(Oral)**.

7. **ProtoSound: Personalized, Scalable Sound Recognition for d/Deaf and Hard of Hearing Users through In-the-Wild Few-Shot Interactions.**
Dhruv Jain, Khoa Nguyen, Steven Goodman, Rachel Grossman-Kahn, Hung Ngo, **Aditya Kusupati**, Ruofei Du,
Alex Olwal, Leah Findlater and Jon Froehlich.
*ACM CHI Conference on Human Factors in Computing Systems (**CHI**), 2022* **(Talk)**.

6. **LLC: Accurate, Multi-purpose Learnt Low-dimensional Binary Codes**
Aditya Kusupati, Matthew Wallingford, Vivek Ramanujan, Raghav Somani, Jae Sung Park, Krishna Pillutla,
Prateek Jain, Sham Kakade and Ali Farhadi.
*Neural Information Processing Systems (**NeurIPS**), 2021* **(Virtual Talk)**.

5. **RNNPool: Efficient Non-linear Pooling for RAM Constrained Inference**
Oindrila Saha, **Aditya Kusupati**, Harsha Vardhan Simhadri, Manik Varma and Prateek Jain.
*Neural Information Processing Systems (**NeurIPS**), 2020* **(Virtual Spotlight)**.
WiCV workshop @ CVPR, 2020.

4. **Soft Threshold Weight Reparameterization for Learnable Sparsity**
Aditya Kusupati, Vivek Ramanujan*, Raghav Somani*, Mitchell Wortsman*,

Prateek Jain, Sham Kakade and Ali Farhadi.
*International Conference on Machine Learning (**ICML**), 2020* **(Virtual Talk)**.

3. **Extreme Regression for Dynamic Search Advertising**
Yashoteja Prabhu, **Aditya Kusupati**, Nilesh Gupta and Manik Varma.
*International Conference on Web Search and Data Mining (**WSDM**), 2020* **(Long Oral)**.
eXtreme Classification: Theory and Applications workshop @ ICML, 2020.

2. **One Size Does Not Fit All: Multi-Scale, Cascaded RNNs for Radar Classification**
Dhrubojyoti Roy*, Sangeeta Srivatsava*, **Aditya Kusupati**, Pranshu Jain, Manik Varma and Anish Arora.
*International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation (**BuildSys**), 2019.*
🏆 **Best Paper Runner-Up Award**.

1. **FastGRNN: A Fast, Accurate, Stable and Tiny Kilobyte Sized Gated Recurrent Neural Network**
Aditya Kusupati, Manish Singh, Kush Bhatia, Ashish Kumar, Prateek Jain and Manik Varma.
*Neural Information Processing Systems (**NeurIPS**), 2018.*

Workshop Publications

2. **Are "Hierarchical" Visual Representations Hierarchical?**
Ethan Shen, Ali Farhadi and **Aditya Kusupati**.
Workshop on Symmetry and Geometry in Neural Representations @ NeurIPS 2023.

1. **Disrupting Model Training with Adversarial Shortcuts**
Ivan Evtimov, Ian Covert, **Aditya Kusupati** and Tadayoshi Kohno.
Workshop on Adversarial Machine Learning @ ICML 2021.

Journal Publications

1. **One Size Does Not Fit All: Multi-Scale, Cascaded RNNs for Radar Classification**
Dhrubojyoti Roy*, Sangeeta Srivatsava*, **Aditya Kusupati**, Pranshu Jain, Manik Varma and Anish Arora.
*ACM Transactions on Sensor Networks (**TOSN**), 17(2), January 2021.* **(Best Paper Nomination)**.

Demos

1. **Lightweight, Deep RNNs for Radar Classification**
Dhrubojyoti Roy*, Sangeeta Srivatsava*, Pranshu Jain, **Aditya Kusupati**, Manik Varma and Anish Arora.
*International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation (**BuildSys**), 2019.*

Theses

1. **Efficient Spatial Representation for Entity-Typing**
Anand Dhoot*, **Aditya Kusupati*** and Soumen Chakrabarti.
Undergraduate Thesis, Computer Science and Engineering, IIT Bombay, 2016 - 17.

Software

1. **EdgeML: Machine Learning for Resource-constrained Edge Devices**.
Dennis et al., including **Aditya Kusupati**.
Microsoft Research India, 2017.
Stats: ★ >1,200, >320, >300,000, >4,500.

Select Awards and Honors

- **Best Paper Award** at ENLSP workshop @ NeurIPS '23 *2023*
- **JUWELS Booster Compute Grant** worth 100K A100 GPU hours *2023*
- **Best Poster Award** and a Research Grant worth $25,000 at Citadel Securities PhD Summit *2023*
- Google **Level 3: Accelerate** Research Grant worth $300,000 extendable up to 1M dollars *2022*
- Academic Research **GCP Credit Award** worth $100,000 *2021 - 2023*
- **Expert Reviewer** for **ICML '21** *2021*
- **Best Paper Runner-Up Award** at **BuildSys '19** *2019*
- Young Researcher at Heidelberg Laureate Forum (**HLF '19**) with Romberg Grant & MSR Travel Grant *2019*
- **Facebook AI Research International Scholarship** for DPhil at VGG, Oxford (2019 - 22, declined) *2019*
- IIT Bombay CSE **Teaching Assistant of the month** (Feb '16 and Feb '17) award *2016 - 2017*

- **All India Rank 44** in JEE Advanced (IIT-JEE) 2013 among 150,000 candidates qualified from 1.5 million *2013*
- Gold Medal and rank **6 out of top 40** in India at OCSC for International Chemistry Olympiad '13 *2013*
- KVPY Fellowship from Government of India - All India Rank 22. *2011*
- NTSE Scholarship from Government of India. *2008*
- **Best/Top/Outstanding Reviewer** award for **NeurIPS '19, '20, '22**; **ICML '20, '21**; **CVPR '21** & **ICLR '22**

Talks

- **Towards Adaptive Intelligence**
  - UT Austin Computer Science Colloquium *April 2024*
  - Microsoft Research India *March 2024*
  - IIT Bombay Computer Science & C-MInDS *March 2024*
  - Harvard University Computer Science & Kempner Institute Lecture *February 2024*
  - Columbia University Computer Science Colloquium *February 2024*
- **Indexing the World**
  - Hazy Research Lab @ Stanford *November 2023*
  - Scaled Foundations *October 2023*
  - MIT Vision and Graphics Seminar *March 2023*
  - Harvard Machine Learning Foundations Seminar *March 2023*
  - Google Research India *February 2023*
  - H2Lab Seminar @ UW CSE *January 2023*
- **Matryoshka Representation Learning**
  - Jina AI *March 2024*
  - ThursdAI *February 2024*
  - Weaviate Podcast *February 2024*
  - Mosaic ML *June 2023*
  - Neural Information Processing Systems (NeurIPS) *December 2022*
  - Pinterest Labs *September 2022*
  - Perception Spotlight @ Google Research *August 2022*
  - DeepPhenomena @ Google Research *August 2022*
  - Image Understanding @ Google Research *June 2022*
- **LLC: Accurate, Multi-purpose Learnt Low-dimensional Binary Codes**
  - Image Understanding @ Google Research *February 2022*
  - Neural Information Processing Systems (NeurIPS) *December 2021*
  - Microsoft Research India *November 2021*
  - UC Berkeley Computer Vision Seminar *November 2021*
  - University of Washington CSE Colloquium *October 2021*
- **Soft Threshold Weight Reparameterization for Learnable Sparsity**
  - International Conference on Machine Learning (ICML) *July 2020*
  - NVIDIA Research *July 2020*
  - Deep Learning: Classics and Trends *June 2020*
- **The Edge of Machine Learning**
  - University of Washington CSE Colloquium & Sensor Systems Seminar *October 2019*
  - VGG @ Oxford University, UK *April 2019*
  - Microsoft Research Redmond *March 2019*
  - Microsoft Research India *August 2018*
- **The Extremes of Machine Learning**
  - Microsoft Bing *March 2019*

Teaching Experience

- *Co-instructor* – Computer Science and Engineering, University of Washington
  - CSE 493G1/599G1: Deep Learning w/ Prof. Ali Farhadi *Fall 2023*
  - CSE 493G1/599G1: Deep Learning w/ Prof. Ranjay Krishna *Spring 2023*
- *Undergraduate Teaching Assistantship* – Computer Science and Engineering, IIT Bombay
  - Digital Logic Design - *Prof. Supratik Chakraborty* - **TA of the month, Feb '17** *Spring 2017*
  - Software Systems Lab - *Prof. Sharat Chandran* *Autumn 2016*
  - Digital Logic Design - *Prof. Supratik Chakraborty* - **TA of the month, Feb '16** *Spring 2016*
  - Computer Programming and Utilisation - *Prof. Varsha Apte* *Autumn 2015*
  - Computer Programming and Utilisation - *Prof. Kavi Arya* *Spring 2015*

Professional Service

- *Reviewing:* IEEE TPAMI, TMLR, NeurIPS (2019 - present), ICML (2020 - present), ICLR (2021 - present), CVPR (2021 - present), ICCV/ECCV (2021 - present).
- *Workshop Organization*
  - ML in India Social *NeurIPS 2021*
  - Rethinking ML Papers *ICLR 2021*
- *Mentorship*
  - Students (Position → Next Placement)
    - Ethan Shen [W.2]
      BS Student, UW CSE *2023 -*
    - Devvrit [P.2]
      PhD Student, UT Austin CS *2023 -*
    - Alan Fan [C.12, C.15]
      BS Student, UW CSE → Software Engineer @ LinkedIn *2023 - 24*
    - Pruthvi Raju
      Software Engineer, Google *2022 - 23*
    - Sharan Ranjit [C.12]
      MS student, UW ECE → Machine Learning Engineer @ Autodesk *2022 - 23*
    - Venkata Sailesh Sanampudi
      Software Engineer, Google *2022 -*
    - Umangi Jain
      Pre-doc Researcher, Google Research India → PhD Student @ UofT CS *2022 - 23*
    - Avishree Khare
      Software Engineer, Google → Research Fellow @ MSR India → PhD Student @ UPenn CS *2022*
    - Gantavya Bhatt [C.9]
      PhD Student, UW ECE *2022 - 23*
    - Aniket Rege [C.9, C.12]
      MS Thesis, UW ECE → PhD Student @ UW–Madison CS *2022 - 23*
    - William Howard-Snyder [C.9]
      BS/MS Student, UW CSE *Fall 2021*
    - Sahil Verma
      PhD Student UW CSE *2021 - 22*
    - Oindrila Saha [C.5]
      Research Fellow, MSR India → PhD Student @ UMass CS *2019 - 21*
    - Sachin Goyal
      Research Fellow, MSR India → PhD Student @ CMU MLD *2019 - 21*
    - Nilesh Gupta [C.3, P.1]
      Research Fellow, MSR India → PhD Student @ UT Austin CS *2019 - 20*
    - Sahil Bhatia
      Research Fellow, MSR India → PhD Student @ UC Berkeley EECS *2018 - 20*
    - Sheshansh Agrawal
      Bachelor's Thesis, IIT Bombay → RSDE @ MSR India *2018 - 19*

    * Manish Singh [C.1]
      Bachelor's Thesis, IIT Delhi → PhD Student @ MIT EECS *2017 - 18*
      🏆 **Best Undergraduate Thesis Award (2018), IIT Delhi**
  - New In ML session @ NeurIPS '19 *2019*
  - MSR India Summer Workshop 2018: Machine Learning on Constrained Devices *Summer 2018*
- *Faculty Recruiting Liaison* - Paul G. Allen School of CSE, University of Washington *2020 - 2022*
- *Student Area Chair (ML/AI): PhD Admissions* - Paul G. Allen School of CSE, University of Washington *2020 - 2022*
- *Co-Founder & Organizing Committee Member* - Allen School PhD Pre-Application Mentorship Service (PAMS) *2021*
- *Co-Founder & Co-Lead* - Allen School PhD Pre-Application Review Service (PARS) *2020*
- *Department General Secretary* - Computer Science and Engineering, IIT Bombay *2016 - 2017*