Krishnateja Killamsetty

CONTACT INFORMATION

Department of Computer Science
The University of Texas at Dallas
Richardson, TX 75080 USA

Voice: (469) 456-6190
E-mail: krishnateja.killamsetty@utdallas.edu
WWW: `https://krishnatejakillamsetty.me/`

RESEARCH INTERESTS

My research focuses primarily on the development of robust, data-efficient, and compute-efficient machine learning systems. Currently, I am involved in the creation of data subset selection techniques that promote efficient and generalizable machine learning. By leveraging the properties of submodularity, I have been able to achieve promising results in both supervised and semi-supervised learning contexts. Additionally, my interests also extend to label-efficient machine learning, which I explore through the development of innovative active, semi-supervised, and unsupervised learning algorithms.

EDUCATION

The University of Texas at Dallas, Richardson, Texas USA

Ph.D., Computer Science, GPA: 4.00/4.00 — August, 2023

- Dissertation Topic: "Data Subset Selection for Compute Efficient Deep Learning"
- Advisor: Rishabh Iyer

Indian Institute of Technology, Kharagpur, West Bengal, India

B.Tech., Electronics and Electrical Communications Engineering, GPA: 8.43/10 — May, 2015

PROFESSIONAL EXPERIENCE

IBM Almaden Research Center, San Jose, California USA

AI Research Intern — **May, 2022 - August, 2022**

- Spearheading the development of efficient AutoML algorithms for text datasets, utilizing subset selection and performance predictors.
- Devised an innovative subset selection technique, "MILO", delivering state-of-the-art efficiency for training and tuning machine learning models, with speedups ranging from 5x to 80x with minimal performance loss.

Microsoft Research, Redmond, Washington USA

Research Intern — **May, 2021 - August, 2021**

- Pioneered the development of a GUI screen understanding framework and established a data synthesis pipeline for GUI screens for training object detection and visual relationship detection models.
- Advanced visual relationship detection models for identifying relations between UI elements.
- Engaging in the creation of a performance-driven data synthesis pipeline for Visual relationship model training.
- Enhanced the efficiency of object detection model training using subset selection approaches.

Mercedes Benz Research and Development Pvt Ltd, Bengaluru, India

Data Scientist — **February, 2018 - December, 2019**

- Analyzed data from diverse sensors (Lidar, Radar, Camera) across the Mercedes Fleet, delivering insights to enhance existing ADAS (Advanced Driver Assistance Systems) Algorithms.
- Generated innovative business concepts to monetize data produced by the Mercedes fleet.

- Conducted spatial analysis of Mercedes Cars data to identify worldwide Hotspots and areas where Driver Assist Functions are failing.
- Utilized Cloud-Based Technologies, Big Data Processing, Computer Vision, and Machine Learning Algorithms.

Robert Bosch Engineering and Business Solutions Pvt Ltd, Bengaluru, India

Senior Software Engineer (Computer Vision) **July, 2015 - February, 2018**

- Contributed to Signal Processing, Image Processing, and Computer Vision algorithms for video-based Driver Assistance projects.
- Managed algorithm prototyping, development, and SIL testing.
- Developed Image Stitching and Harmonization Algorithms for Surround View Camera for BMW and Daimler.
- Created lane detection algorithms to identify lanes from video input, useful for functions like Lane Departure Warning and Lane Keep Assist.
- Designed a comprehensive web-based evaluation framework for assessing Lane Detection Algorithms used in Bosch.

Honors and Awards

- **Outstanding Ph.D. Applicant Award, 2020:** Awarded by the Department of Computer Science, University of Texas at Dallas, in recognition of exceptional doctoral application.
- **ABCD (Above and Beyond the Call of Duty) Award, 2018:** Presented by the Mercedes Benz R&D, India, acknowledging significant contributions within the first six months of service.
- **Digital Life Day Product Innovation Excellence Award, 2018:** Bestowed by Manu Saale, Vice President of Daimler AG, in recognition of significant advancements in product innovation.
- **Innovation Excellence Driver Award, 2018:** Honored by the Mercedes Benz R&D, India, in acknowledgment of pioneering approaches and groundbreaking ideas.
- Graduated with Honors in Electronics and Electrical Communications Engineering from the Indian Institute of Technology, Kharagpur, 2015.
- Attained an **All India Rank of 679** in the highly competitive IITJEE examination in 2011, out of 4,68,240 students.
- Secured a **State Rank of 25** in the EAMCET examination in 2011, among 2,78,974 students.

Academic Experience

University of Texas at Dallas, Richardson, Texas USA

Graduate Student **January, 2020 - present**

Includes current Ph.D. research, Ph.D. and Masters level coursework and research/consulting projects.

Research Assistant **May, 2020 - present**

- Engaged in research on Data Efficient Learning, focusing on Active Learning, Data Selection, and Data Partitioning.
- Currently investigating Robust Learning methodologies, particularly in scenarios involving Noise, Outliers, and Class Imbalance.
- Conducted comprehensive research on Labeling Functions and Data Programming paradigms.

Teaching Assistant **January, 2020 - May, 2020**

Course: CS4375 - Introduction to Machine Learning

- Duties have included office hours, grading, and leading weekly assignments.

Peer Reviewed Publications

Google Scholar: `https://scholar.google.com/citations?user=cHDE-2YAAAAJ`
DBLP: `https://dblp.org/pid/273/3972.html`

Krishnateja Killamsetty, Guttu Sai Abhishek, Aakriti, Alexandre V. Evfimievski, Lucian Popa,

Ganesh Ramakrishnan, Rishabh Iyer. "AUTOMATA: Gradient Based Data Subset Selection for Compute-Efficient Hyper-parameter Tuning". In Neural Information Processing Systems, NeurIPS 2022. (25.6% Acceptance Rate)

Athresh Karanam*, **Krishnateja Killamsetty***, Harsha Kokel*, Rishabh K Iyer. "Orient: Submodular Mutual Information Measures for Data Subset Selection under Distribution Shift". In Neural Information Processing Systems, NeurIPS 2022. (25.6% Acceptance Rate)

Xujiang Zhao*, **Killamsetty Krishnateja***, Rishabh Iyer, Feng Chen. "How Out of Distribution Data Hurts Semi-Supervised Learning". In IEEE International Conference on Data Mining, ICDM 2022. (9% Acceptance Rate)

Rishabh Tiwari, **Krishnateja Killamsetty**, Rishabh Iyer, Pradeep Shenoy, "GCR: Gradient Coreset based Replay Buffer Selection for Continual Learning". In Conference on Computer Vision and Pattern Recognition, CVPR 2022.

Ayush Maheshwari*, **Krishnateja Killamsetty***, Ganesh Ramakrishnan, Rishabh Iyer, Marina Danilevsky, Lucian Popa. "Learning to Robustly Aggregate Labeling Functions for Semi-supervised Data Programming". In Findings of the Association for Computational Linguistics: ACL 2022. (Long paper)

Krishnateja Killamsetty*, Changbin Li*, Chen Zhao, Rishabh Iyer, Feng Chen. "A Nested Bi-level Optimization Framework for Robust Few Shot Learning". In Thirty-Sixth AAAI Conference on Artificial Intelligence, AAAI 2022. (15% Acceptance Rate)

Krishnateja Killamsetty, Xujiang Zhou, Feng Chen, and Rishabh Iyer, "RETRIEVE: Coreset Selection for Efficient and Robust Semi-Supervised Learning". In Neural Information Processing Systems, NeurIPS 2021. (26% Acceptance Rate)

Suraj Kothawade, Nathan Beck, **Krishnateja Killamsetty**, Rishabh Iyer, "SIMILAR: Submodular Information Measures Based Active Learning In Realistic Scenarios". In Neural Information Processing Systems, NeurIPS 2021. (26% Acceptance Rate)

Ayush Maheshwari, Oishik Chatterjee, **Krishnateja Killamsetty**, Ganesh Ramakrishnan, Rishabh Iyer."Semi-Supervised Data Programming with Subset Selection". In Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021, 4640–4651. doi:10.18653/v1/2021.findings-acl.408. (Long paper)

Krishnateja Killamsetty, Durga Sivasubramanian, Ganesh Ramakrishnan, Abir De, Rishabh Iyer. "GRAD-MATCH: Gradient Matching based Data Subset Selection for Efficient Deep Model Training". In Proceedings of the 38th International Conference on Machine Learning, ICML 2021, 18-24 July 2021, Virtual Event, 139:5464–5474. Proceedings of Machine Learning Research. PMLR, 2021. 21% acceptance rate)

Krishnateja Killamsetty, Durga Sivasubramanian, Ganesh Ramakrishnan, Rishabh Iyer. "GLISTER: Generalization based Data Subset Selection for Efficient and Robust Learning". In Thirty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2021, Virtual Event, February 2-9, 2021, 8110–8118. AAAI Press, 2021. 21% acceptance rate)

Workshop Papers

Krishnateja Killamsetty, Alexandre Evfimievski, Tejaswani Pedapati, Kiran Kate, Lucian Popa, Rishabh K Iyer. "MILO: Model-Agnostic Subset Selection Framework for Efficient Model Training and Tuning". The 4th workshop of Practical ML for Developing Countries Workshop(PML4DC), In Conjunction with ICLR 2023. (**ORAL**)

H S V N S Kowndinya Renduchintala, **Krishnateja Killamsetty**, Sumit Bhatia, Milan Aggarwal, Ganesh Ramakrishnan, Rishabh K Iyer, Balaji Krishnamurthy. Using Informative Data Subsets for Efficient Training of Large Language Models: An Initial Study. The 2nd workshop of Efficient Natural Language and Speech Processing (ENLSP), In Conjunction with NeurIPS 2022.

Krishnateja Killamsetty*, Changbin Li*, Chen Zhao, Rishabh Iyer, Feng Chen. "A Nested Bi-level Optimization Framework for Robust Few Shot Learning". Fifth Workshop on Meta-Learning at the Conference on Neural Information Processing Systems, In Conjunction with NeurIPS 2021

Savan Amitbhai Visalpara, **Krishnateja Killamsetty**, Rishabh Iyer. "A Data Subset Selection Framework for Efficient Hyper-Parameter Tuning and Automatic Machine Learning". Workshop on Subset Selection in Machine Learning, SubSetML 2021, In Conjunction with ICML 2021

Krishnateja Killamsetty, Durga Sivasubramanian, Baharan Mirzasoleiman, Ganesh Ramakrishnan, Abir De, Rishabh Iyer. "A Gradient Matching Framework for Efficient Learning". Workshop on Hardware Aware Efficient Training, In Conjunction with ICLR 2021

Preprints and Working Papers

Krishnateja Killamsetty, Alexandre V. Evfimievski, Tejaswini Pedapati, Kiran Kate, Lucian Popa, Rishabh Iyer. "MILO: Model-Agnostic Subset Selection Framework for Efficient Model Training and Tuning". arXiv:2301.13287.

HSVNS Kowndinya Renduchintala, **Krishnateja Killamsetty**, Sumit Bhatia, Milan Aggarwal, Ganesh Ramakrishnan, Rishabh Iyer, Balaji Krishnamurthy. "INGENIOUS: Using Informative Data Subsets for Efficient Pre-Training of Large Language Models". arXiv:2305.06677.

Nathan Beck, **Krishnateja Killamsetty**, Suraj Kothawade, Rishabh Iyer. Beyond Active Learning: Leveraging the Full Potential of Human Interaction via Auto-Labeling, Human Correction, and Human Verification

Open Source Repositories

Krishnateja Killamsetty, Dheeraj N Bhat, Rishabh Iyer. "CORDS: COResets and Data Subset selection". GitHub repository. GitHub, 2021.

Citations

Please see my Google Scholar page for more details: `https://scholar.google.com/citations?hl=en&user=cHDE-2YAAAAJ`.

Professional Activities

I have served as a Program Committee member and Reviewer for several conferences and Journals.

- Reviewer for Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2023
- Programme Committee Member for Conference on Neural Information Processing Systems (NeurIPS), 2023
- Reviewer for the Journal of Machine Learning and Research (JMLR), 2023
- Reviewer for the International Journal of Computer Vision (IJCV), 2023
- Programme Committee member for International Conference on Computer Vision (ICCV), 2023
- Programme Committee member for Conference on Computer Vision and Pattern Recognition (CVPR), 2023
- Reviewer for the Data Mining and Knowledge Discovery Journal (DAMI), 2023
- Programme Committee member for ACM SIGKDD CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING (KDD), 2022
- Programme Committee member for AAAI Conference on Artificial Intelligence (AAAI), 2022
- Programme Committee Member for Conference on Neural Information Processing Systems (NeurIPS), 2021
- Programme Committee Member for International Conference on Artificial Intelligence and Statistics (AISTATS), 2021
- Programme Committee member for AAAI Conference on Artificial Intelligence (AAAI), 2021

Technical Skills

- Proficient in a variety of programming languages and tools, including Python, C++, C, JAVA, MATLAB, along with libraries and frameworks such as PyTorch, TensorFlow, Pandas, and NumPy.

- Solid grasp of Machine Learning, Computer Vision, and Deep Learning algorithms, coupled with the ability to construct advanced and innovative algorithms for a range of machine learning applications.
- Experienced in addressing complex problems in sectors such as autonomous vehicles, utilizing state-of-the-art machine learning technologies.
- Comprehensive understanding of current challenges and emerging technologies in the autonomous vehicles domain, keeping up-to-date with advancements and industry trends.

Presentations and Talks

"Practical ML for Developing Countries Workshop," ICLR - 2023.

"AAAI Conference on Artificial Intelligence (AAAI) Tutorial," 2023.

"Neural Information Processing Systems (NIPS)," 2022.

"International Conference on Machine Learning (ICML) Tutorial," 2021.

"AAAI Conference on Artificial Intelligence (AAAI)," 2022.

"Neural Information Processing Systems (NIPS)," 2021.

"International Conference on Machine Learning (ICML)," 2021.

"Subset Selection in Machine Learning: From Theory to Applications," ICML - 2021.

"AAAI Conference on Artificial Intelligence (AAAI)," 2022.

"Scenario Simulation Engine," Presented at Digital Life Day Daimler at Bengaluru & Germany to Mr. Dieter Zetsche, CEO of Daimler, Mr. Ola Kallenius, R&D Head at Daimler AG, Mr. Manu Saale, Vice president at Daimler AG, and other high-level management in Daimler.

Extra Curricular Activities

- Active participant in the National Sports Organization (NSO), IIT Kharagpur.
- Volunteered for a rural development organization at IIT Kharagpur, contributing to community upliftment.
- Enthusiastic reader and cinema enthusiast, with a wide range of genres and styles.
- Passionate about playing video games and cricket, enjoying both the competitive and team-building aspects.
- Regularly enjoys outdoor activities such as long walks and hiking, cherishing the peace and rejuvenation of nature.

References

- Prof. Rishabh Iyer, Assistant Professor, Department of CS, University of Texas at Dallas (rishabh.iyer@utdallas.edu)
- Prof. Ganesh Ramakrishnan, Institute Chair Professor, CSE Department, IIT Bombay (ganesh@cse.iitb.ac.in)
- Prof. Jeff Bilmes, Professor, Department of EE, University of Washington (bilmes@uw.edu)
- Prof. Feng Chen, Associate Professor, Department of CS, University of Texas at Dallas (feng.chen@utdallas.edu)
- Dr. Lucian Popa, Principal Research Scientist, IBM Research, Almaden (lpopa@us.ibm.com)
- Dr. Alexandre V. Evfimievski, Research Staff Member, IBM Research, Almaden (evfimi@us.ibm.com)