

Krishnateja Killamsetty

+1 469-456-6190 | krishnateja.killamsetty@utdallas.edu | <http://www.linkedin.com/in/krishnateja-killamsetty/>

RESEARCH INTERESTS

My research centers on developing techniques and algorithms that enable data-efficient, compute-efficient, and robust machine learning systems. My current work on data subset selection for efficient and generalizable machine learning leverages submodularity to achieve this goal in the context of supervised and semi-supervised learning scenarios. I am also interested in label-efficient machine learning through active and semi-supervised learning algorithms.

EDUCATION

The University of Texas at Dallas
Ph.D., Computer Science

4.0/4.0
December 2023

Indian Institute of Technology, Kharagpur
BE, Electronics and Communication Engineering

8.43/10
May 2015

WORK EXPERIENCE

IBM Almaden Research Center, San Jose (USA)
AI Research Intern

May 2022 - August 2022

- Working on developing efficient AutoML algorithms for text datasets using subset selection and performance predictors.

Microsoft Research, Redmond (USA)
Research Intern

May 2021- August 2021

- Worked on developing a GUI screen understanding framework. Developed a data synthesis pipeline for GUI screens that can be used as training data for object detection and visual relationship detection models.
- Worked on visual relationship detection models for detecting the relations between UI elements in the screens.
- Currently working on a new performance-driven data synthesis pipeline for Visual relationship model training.
- Improved the efficiency of object detection model training using subset selection approaches.

Mercedes Benz Research and Development Private Limited, Bengaluru (India)

Data Scientist

February 2018 - December 2019

- Worked on data from different sensors like Lidar, Radar, Camera, etc. from the whole Mercedes Fleet
- Supplying insights by analyzing the vast data from Mercedes Cars to improve existing ADAS (Advanced Driver Assistance Systems) Algorithms of Daimler.
- Producing new business ideas to monetize the data being generated by the Mercedes fleet.
- Spatial Analysis of Mercedes Cars data to generate Hotspots throughout the world which helps find areas throughout-out the world where Driver Assist Functions are failing.
- I worked on 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

- Worked on Signal Processing, Image Processing, and Computer Vision algorithms for Video-based Driver Assistance projects.
- My responsibilities include algorithm prototyping, algorithm development, and SIL testing.
- Worked on Image Stitching and Harmonization Algorithms of Surround View Camera for BMW and Daimler.
- Worked on lane detection algorithms to detect lanes using video input, beneficial for functions like Lane Departure Warning and Lane Keep Assist.
- Developed a complete web-based Evaluation framework to evaluate Lane Detection Algorithms used in Bosch.

RESEARCH EXPERIENCE

The University of Texas at Dallas

Ph.D. Researcher

Richardson, Texas

January 2020 – present

- I am currently working on Data efficient Learning: Data Selection, Active Learning, and Data partitioning research Problems.
- I am currently working on Robust learning in the presence of Noise, Outliers, Class Imbalance et,c.
- Worked on the Labeling Functions and Data Programming research problems.

IIT (Indian Institute of Technology) Kharagpur

Undergraduate Researcher

Kharagpur, India

June 2014 – May 2015

- Worked on Blind Image forgery detection research problems.
- Worked on High-frequency crystal oscillator design.

PUBLICATIONS

PEER-REVIEWED PUBLICATIONS

- Rishabh Tiwari, Krishnateja Killamsetty, Rishabh Iyer, Pradeep Shenoy. **GCR: Gradient Coreset based Replay Buffer Selection for Continual Learning.** <https://arxiv.org/abs/2111.11210>. To Appear 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.** <http://arxiv.org/abs/2109.11410>. To Appear 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.** <https://arxiv.org/abs/2011.06782>. To Appear in Thirty-Sixth AAAI Conference on Artificial Intelligence, AAAI 2022. (15% Acceptance Rate)
- Killamsetty, K., Zhao, X., Chen, F., & Iyer, R. (2021). **RETRIEVE: Coreset Selection for Efficient and Robust Semi-Supervised Learning.** arXiv [cs. LG]. <http://arxiv.org/abs/2106.07760> (Accepted to NeurIPS 21)
- Kothawade, S., Beck, N., Killamsetty, K., & Iyer, R. (2021). **SIMILAR: Submodular Information Measure Based Active Learning in Realistic Scenarios.** arXiv [cs. LG]. <http://arxiv.org/abs/2107.00717> (Accepted to NeurIPS 21)

- Krishnateja Killamsetty, Durga Sivasubramanian, Ganesh Ramakrishnan, Abir De, Rishabh K. Iyer: **GRAD-MATCH: Gradient Matching based Data Subset Selection for Efficient Deep Model Training. ICML 2021: 5464-5474**
- Ayush Maheshwari, Oishik Chatterjee, Krishnateja Killamsetty, Rishabh Iyer, & Ganesh Ramakrishnan. **Semi-Supervised Data Programming with Subset Selection. ACL/IJCNLP (Findings) 2021: 4640-4651**
- Krishnateja Killamsetty, Durga Sivasubramanian, Ganesh Ramakrishnan, Rishabh K. Iyer: **GLISTER: Generalization based Data Subset Selection for Efficient and Robust Learning. AAAI 2021: 8110-8118**

WORKSHOPS

- 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

PRE-PRINTS

- Athresh Karanam*, Krishnateja Killamsetty*, Harsha Kokel*, Rishabh Iyer. **Orient: Submodular Mutual Information Measures for Data Subset Selection under Distribution Shift.**
- 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**
- 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.** <https://arxiv.org/abs/2203.08212>. arXiv [cs.LG], 2022. arXiv:2203:08212.
- Xujiang Zhao, Killamsetty Krishnateja, Rishabh Iyer, & Feng Chen. (2020). **Robust Semi-Supervised Learning with Out of Distribution Data**, arXiv: 2010.03658 (In review at KDD)

PATENTS

- Boddeti Mohanvarakrishna, Gautam Raju, Krishnateja Killamsetty, Swadeep Kumar, **SYSTEM AND METHOD FOR ASSIGNING RESPONSIBILITY SCORES TO USERS OF VEHICLE**, Indian Patent (Application No: 201911022751)
- Boddeti Mohanvarakrishna, Gautam Raju, Krishnateja Killamsetty, Swadeep Kumar, **SYSTEM AND METHOD FOR PREDICTING BEHAVIOUR OF A USER OF A VEHICLE**, Indian Patent (Application No: 201911022492)
- Boddeti Mohanvarakrishna, Gautam Raju, Krishnateja Killamsetty, Kishore Subramanian, **METHOD AND SYSTEM FOR EXTRACTING AND GENERATING CRITICAL TEST SCENARIOS FOR AUTONOMOUS VEHICLES**, Indian Patent (Application No: 201841048116)

OPEN-SOURCE REPOSITORIES

- Krishnateja Killamsetty, Dheeraj N Bhat, Rishabh Iyer. [“CORDS: COResets and Data Subset selection”](#). GitHub repository. GitHub, 2021.

PRESENTATIONS

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

TECHNICAL SKILLS

- Adept in programming skills (Python, C++, C, MATLAB, PyTorch, TensorFlow, Pandas, NumPy)
- Good Knowledge of Machine Learning, Computer Vision, and Deep Learning Algorithms.
- Able to develop advanced and sophisticated algorithms for various machine learning applications to tackle problems in areas such as Driverless Cars.
- I have a good knowledge of challenges in the domain of driverless cars currently and am thorough with the newer technologies in the same domain.

CITATIONS

- The total number of citations (both self and independent) from international conferences and journals is 92. Please see my [Google Scholar page](#) for more details.

PROFESSIONAL ACTIVITIES

- I served as a Programme Committee member for KDD 2022.
- I served as a Programme Committee member for AAAI 2022.
- I served as a Programme Committee member for NeurIPS 2021.
- I Served as a Programme Committee member for AISTATS 2021.
- I served as a Programme Committee member for AAAI 2021

AWARDS, LEADERSHIP & ORGANIZATIONS

- July 2018: ABCD (**Above and Beyond the call of Duty**) Award from my Department in Mercedes Benz within six months of my joining
- June 2018: **Digital Life Day Product Innovation Excellence Award** from Manu Saale, Vice President of Daimler AG.
- November 2018: **Innovation Excellence Driver Award** from my Department in Mercedes Benz.