Satyapriya Krishna

Webpage: satyapriyakrishna.com
Email: skrishna@g.harvard.edu
Google Scholar : satyapriya-krishna
GitHub: github.com/y12uc231

Education

Harvard University — Cambridge, MA
Doctor of Philosophy in Computer Science — 2026 (Expected)

- Research Area : Trustworthy Machine Learning - Explainable, Fair, Robust and Reliable Machine Learning
- Advisor: Prof. Himabindu Lakkaraju and Prof. Finale Doshi-Velez

Carnegie Mellon University — Pittsburgh, PA
Master of Science, School of Computer Science, Advisor: Prof. Anatole Gershman — 2018

- Thesis: "Conversational Agents based on Deep Reinforcement Learning"
- GPA : 3.91/4.00

LNM Institute of Information Technology — Jaipur, India
Bachelors in Computer Science & Engineering — 2016

- Thesis: "Fake Content Verification"
- GPA: 9.29/10.00
- Rank: 3/296

Publications

[1] S. Krishna, T. Han, A. Gu, J. Pombra, S. Jabbari, S. Wu, and H. Lakkaraju, "The disagreement problem in explainability: A practitioner's perspective", *(Under Review) The ACM Conference on Fairness, Accountability, and Transparency (ACM FAccT)*, 2022.

[2] S. Krishna, R. Gupta, A. Verma, J. Dhamala, Y. Pruksachatkun, and K.-W. Chang, "Measuring fairness of text classifiers via prediction sensitivity", *The Joint Conference of the 60th Annual Meeting of the Association for Computational Linguistics (ACL 2022)*, 2022.

[3] U. Gupta, J. Dhamala, V. Kumar, A. Verma, Y. Pruksachatkun, S. Krishna, R. Gupta, K.-W. Chang, G. V. Steeg, and A. Galstyan, "Equitable text generation with distilled language models via counterfactual role reversal", *The Joint Conference of the 60th Annual Meeting of the Association for Computational Linguistics (ACL 2022 - Findings)*, 2022.

[4] U. Gupta, J. Dhamala, V. Kumar, A. Verma, Y. Pruksachatkun, S. Krishna, R. Gupta, K.-W. Chang, G. V. Steeg, and A. Galstyan, "Mitigating gender bias in distilled language models via counterfactual role reversal", *Under Review at The Joint Conference of the 60th Annual Meeting of the Association for Computational Linguistics (ACL 2022)*, 2022.

[5] S. Krishna, R. Gupta, and C. Dupuy, "Adept: Auto-encoder based differentially private text transformation", *The 16th Conference of the European Chapter of the Association for Computational Linguistics*, 2021.

[6] S. Krishna, Y. Pruksachatkun, J. Dhamala, R. Gupta, and K.-W. Chang, "Does robustness improve fairness? approaching fairness with word substitution robustness methods for text classification", *The Joint Conference of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (ACL-IJCNLP 2021)*, 2021.

[7] J. Dhamala, T. Sun, V. Kumar, S. Krishna, Y. Pruksachatkun, R. Gupta, and K.-W. Chang, "Bold: Dataset and metrics for measuring biases in open-ended language generation", *The ACM Conference on Fairness, Accountability, and Transparency (ACM FAccT)*, 2021.

[8] Y. Pruksachatkun, S. Krishna, A. Ramakrishna, J. Dhamala, and R. Gupta, "Workshop on trustworthy natural language processing", *Workshop proposal accepted to The North American Chapter of the Association for Computational Linguistics*, 2021.

[9] J. Payan, Y. Merhav, H. Xie, S. Krishna, A. Ramakrishna, M. Sridhar, and R. Gupta, "Towards realistic single-task continuous learning research for ner", in *The 2021 Conference on Empirical Methods in Natural Language Processing*, 2021.

[10] A. Patel, R. Gupta, M. Harakere, S. Krishna, A. Alok, and P. Liu, "Towards classification parity across cohorts", *ML-IRL Workshop, International Conference on Learning Representations (ICLR)*, 2020.

[11] Y. Tao, S. Gupta, S. Krishna, X. Zhou, O. Majumder, and V. Khare, "Finetext: Text classification via attention-based language model fine-tuning", *arXiv preprint arXiv:1910.11959*, 2019.

[12] S. Krishna, Q. Hu, Y. Zhang, B. Yin, and H. Rangwala, "Document expansion with keyword extraction from massive product catalog", *Submitted to The 33rd Conference on Neural Information Processing Systems, Vancouver, Canada*, 2019.

[13] S. Krishna, M. de Jong, and A. Agarwal, "Grounding complex navigational instructions using scene graphs", *arXiv preprint arXiv:2106.01607*, 2018.

Work Experience

Amazon Alexa AI — Boston, MA
Sr. Researcher — Dec. 2019 –Aug 2021

- Led Trustworthy Natural Language Understanding (NLU) team to build robust, fair and interpretable language processing models for Amazon Alexa
- Launched the first fairness/bias measurement and mitigation pipeline for Alexa

Amazon Search (A9.com) — Palo Alto, CA
Applied Scientist - 1 — Oct. 2018 –Nov. 2019

- Launched the first set of document summarization models on amazon retail worldwide (15 countries)
- Launched the first version of Amazon Product Graph which improved search recall by 27%

Amazon Web Services — Seattle, WA
Machine Learning Engineer — Jul 2018 –Sep 2018

- Launched universal state-of-the-art text-classification models on AWS Sagemaker

Carnegie Mellon University — Pittsburgh, PA
Teaching Assistant — Nov 2016 –May 2018

- Teaching assistant in Math for Machine Learning(10-600) to Prof. Geoffrey J. Gordon
- Teaching assistant in Practical Data Science(15-688) to Prof. Zico Kolter
- Teaching assistant in PhD-Intro to ML(10-701) to Prof. Manuela M. Veloso and Prof. Pradeep Ravikumar

Amazon Research — Seattle, WA
Research Intern — May 2017 –Aug 2017

- Developed software to generate alerts for potential physical-virtual geolocation mismatches, when the physical location of a particular logistic (trailer/package) is not consistent with the system
- Applied clustering algorithms(EM, K-Means) to investigate factors behind mismatches and developed a real-time geo-location tracker which tracked logistic movement for Amazon fulfillment centers worldwide

Indian Institute of Technology (IIT) Chennai, India
Research Intern May 2015 –Sep 2015

– Developed state-of-the-art face recognition algorithms under the supervision of Prof. Sukendu Das from Visualization and Perception Lab

National Institute of Technology (NIT) Rourkela, India
Research Intern May 2014 –Sep 2014

– Developed statistical models to detect and remove shadow regions from low-resolution videos from security video sensors to maximize feature extraction in Computer Vision Lab, supervised by Prof. Banshidhar Majhi

Scholarships and Awards

- Awarded Nicole A. Chen and Karina A. Chen Graduate Student Research Fellowship 2021–2022
- Director's Scholarship, LNM Institute of Information Technology (India) - ₹80000 2012–2016

Coursework

- **Undergraduate:** Design & Analysis of Algorithms, Data Structures, Genetic Programming, Software Engineering, Graph Theory, Probability & Statistics, Database Management , Machine Learning
- **Graduate:** Bayesian Statistics, Multivariate Analysis, Causal Inference, Deep Reinforcement Learning, Deep Learning, Conversational agents, Machine Learning, Big Data Analytics, Robotics & Machine Learning, Natural Language Processing, Applied Machine Learning, Math for Machine Learning, Independent Research : Machine Learning for Social Good
- **Independent:** Deep Learning - Google, Parallel Computing - NVIDIA, Search Engine- Google

Notable Projects

- Grounding Complex Navigational Instructions Using Scene Graphs — Ranked 2nd out of 40 projects
  Tools: Python, Tensorflow
  – Trained agents to navigate in gaming environment(ViZDoom) using natural language instructions
  – Achieved state-of-the-art accuracy in reaching target objects with 20% faster travel time

- **SynSem** : Semi- Goal Oriented Conversational Agent — Masters Research Project
  Tools: Python, CuDNN, C++, PyTorch
  – Developed chat agents with better context-transitions and initiations, improving conversation time by 5 times compared against other end-to-end chatbots, along with 70% improvement in end-to-end training speed
  – Experimented with Asynchronous Actor-Critic (A3C) and Proximal Policy Optimization (PPO) learning algorithms, and improved inference time with C++/cuDNN programmed modules

- Universal Semantic Parser — Capstone Project
  Tools: Pytorch
  – Developed seq-2-seq based semantic parser trained for multiple formalisms (ATIS, GeoQuery, Overnight) to achieve faster parsing across different formalisms by maximizing parameter-sharing
  – Reduced parameters by 30% against state-of-the-art baseline with comparable performance

Skills

- **Programming:** Python • Java • C++ • C • R

- **Frameworks:** PyTorch • Keras • OpenGym • CUDA, cuDNN • MXNET • Tensorflow • Pandas • Hadoop • Spark • MongoDB • Apache Hive • AJAX • jQuery • Kubernetes • Docker
- **Tools & Utilities:** Bash • AWS Services • Spring • Apache Pig • Node.js • SQL • Apache Hive • Vim • Conda • Jupyter • Rstudio • tmux • Git • Slurm Workload Manager • Microsoft Office.
- **Operating Systems :** Unix (MacOS) • Linux (CentOS and Ubuntu) • Windows 10

Service to the Community, Organization, and Profession

Natural Language Processing Research — Boston, MA
Reviewer — Jul 2020 –Oct 2020

- Reviewed research papers on speech and language processing problems submitted to the Association for the Advancement of Artificial Intelligence Conference 2021
- Reviewed research papers on benchmarks and datasets submitted to NeurIPS - 2021

Machine Fairness - Amazon Research — Seattle, WA
Panelist & Tutor — Jul 2018 –current

- Participated in panel discussions on social bias in machine learning at Amazon Research. Also organized hackathons, poster presentations, and tutorials to motivate research in the area of Fairness, Accountability, Transparency and Ethics (FATE)

Speech & Language Research - Amazon Research — Seattle, WA
Mentor — Dec 2018 –current

- Mentoring research scientists on conducting high-quality & effective research at Amazon through weekly discussions

LNM Institute Of Information Technology — Jaipur, India
Mentor — May 2020 –current

- Mentoring students from my undergrad university on various topics related to career path, graduate applications, and academic research.

ConvAI - Carnegie Mellon University — Pittsburgh, PA
Organizer & Speaker — 2016 - 2018

- Organized reading group to invite discussions on research problems related to conversational AI. This group was joined by students and researchers from various departments in the university, which also resulted in several research publications

COVID-19 Volunteer — Boston, MA
Miscellaneous — March 2020 –current

- Helped local hospitals and stores with emergency medical supplies.