Amrith Setlur

(+1) 412-636-6235 | asetlur@cs.cmu.edu | ars22.github.io
ars22 | amrith-setlur | @setlur_amrith | Amrith Setlur

Research Interests

My research interests lie in bridging the gap between machine learning (ML) theory and its applications in NLP, speech, and vision. I am particularly interested in multi-task learning, meta-learning algorithms for fast adaptation, and learning fair, invariant representations capturing the true factors in the data distribution that may further influence causal mechanisms.

- **ML Theory:** Distributional robustness, Bayesian methods, meta-learning, federated learning, large-scale optimization, learning theory, differential privacy, fairness, causal inference, generative models, interpretability.
- **ML Applications:** Controlled text/speech synthesis (style transfer), multilingual representations, zero shot NLP, computer vision, GANs, speech processing, healthcare (Survival Analysis).

Education

Carnegie Mellon University *Pittsburgh, PA*
M.S. in Language Technologies *Aug 2019 - May 2021*
CGPA: 4.12/4.0 Advisors: Prof. Alan W Black, Prof. Barnabás Póczos and Prof. Virginia Smith

National Institute of Technology (President of India Gold Medallist) *Trichy, India*
B.S. (Honors) in Computer Science and Engineering *Aug 2013 - May 2017*
CGPA: 9.91/10.0; Institute Rank: 1

Publications (including workshops)

Explaining The Efficacy of Counterfactually Augmented Data
Divyansh Kaushik, Amrith Setlur, Eduard Hovy, Zachary C. Lipton; *To be presented at* ***ICLR 2021*** [pdf] [reviews]
Tags: Causal Inference, Theory, NLP
TL;DR: Analyzes linear Gaussian models to reveal relationships between causal models and measurement noise. Analysis suggests that adding noise to causal features degrades out-of-domain performance, while adding it to non-causal features enhances model robustness.

Nonlinear Independent Subspace Analysis with Auxiliary Variables for Learning Speech Representations (Finalist for Best Student Paper Award)
Amrith Setlur, Barnábas Poczós, Alan W Black; ***Interspeech 2020*** [pdf] [video]
Tags: Theory, Independent Subspace Analysis, ICA, Speech
TL;DR: Separation theorem for independent subspaces in the presence of auxiliary variables. An extension of non-linear ICA applicable to problems involving observed mixtures of high-dimensional non-stationary latent sources.

Politeness Transfer: A Tag and Generate Approach
Amrith Setlur, Aman Madaan*, Tanmay Parekh*, Barnabás Póczos, Graham Neubig, Yiming Yang, Ruslan Salakhutdinov, Alan W Black, Shrimai Prabhumoye*; ***ACL 2020*** [pdf] [code] [news]
Tags: NLP, Controlled Text Synthesis
TL;DR: Introduces a new task and dataset for politeness transfer. Proposes a simple pipeline of paired seq2seq models that are trained on weak rationales to first identify style attributes and then replace them in a way that preserves most of the content.

Is Support Set Diversity Necessary for Meta-Learning?
Amrith Setlur, Oscar Li*, Virginia Smith*; ***Meta-learning Workshop NeurIPS 2020*** [pdf] [video] [code]
Tags: Meta-learning, Optimization, Computer Vision
TL;DR: Questions the role played by the inherent redundancy in a task distribution on a meta-learner's generalizability. By reducing the support set diversity and optimizing a biased meta-learning objective we improve performance for several base learners. We are now theoretically analysing the impact of support variance reduction in the convex setting of meta-linear regression **(in preparation for ICML).**

Covariate Distribution Aware Meta-learning
Amrith Setlur, Saket Dingliwal*, Barnabás Póczos*; ***Lifelong Learning Workshop ICML 2020 (Under review at AAAI)*** [pdf] [video]
Tags: Meta-learning, Hierarchical Bayes, Computer Vision
TL;DR: Bayesian method to model uncertainty in the post adaptation parameters on a novel task. Models the latent shared structure in the covariate distribution across tasks to better inform the task-specific initialization for gradient-based meta-learners like MAML.

An efficient fault tolerant workflow scheduling approach using replication heuristics and checkpointing in the cloud
Amrith Setlur, S. Jaya Nirmala, Har Simrat Singh, Sudhanshu Khoriya; ***Journal of Parallel and Distributed Computing*** [pdf]
Tags: Scheduling Algorithms, Fault Tolerance, Scientific Workflows, Cloud Computing
TL;DR: Unsupervised prediction of replication heuristics to improve resource wastage with roughly the same makespan as HEFT. Additionally, we provide a high probability bound on the optimal checkpoint interval.

ReStGAN: A step towards visually guided shopper experience via text-to-image synthesis (US Patent)
Shiv Surya, Amrith Setlur, Arijit Biswas, Sumit Negi; ***IEEE & CVF WACV 2020*** [PDF] [PATENT]: **US 10,713,821 B1**
TAGS: GENERATIVE ADVERSARIAL NETS, COMPUTER VISION
TL;DR: A conditional recurrent GAN that learns an implicit distribution over sequences of evolving and coherent images.

Robust Handwriting Recognition with Limited and Noisy Data
Hai Pham, Amrith Setlur, Saket Dingliwal, Tzu-Hsiang Lin, Barnabás Póczos, Kang Huang, Zhuo Li, Jae Lim; ***ICFHR 2020*** [PDF]
TAGS: COMPUTER VISION, OBJECT DETECTION, APPLICATION
TL;DR: Leverages architectures from scene-text detection with a modified CTC loss for text recognition in noisy documents.

Better Approximate Inference for Partial Likelihood Models with a Latent Structure
Amrith Setlur, Barnabás Póczos; ***Workshop on Temporal Point Processes NeurIPS 2019 (Oral)*** [PDF] [VIDEO]
TAGS: TEMPORAL POINT PROCESSES, SURVIVAL ANALYSIS, NON-PARAMETRIC STATISTICS
TL;DR: Computationally efficient (linear in size of discrete latent space) approximate inference for partial likelihood models. Provides an upper bound on the approximation and using convex conjugates proposes an objective to minimize a smooth relaxation of the bound.

Multi-dimensional Zero-shot Transfer Learning for Named Entity Recognition
Amrith Setlur, Tanmay Parekh*, Aman Madaan*, A. Chaudhary, B. Poczos, Y. Yang, G. Neubig, A. Black*; ***Under Review at EACL***
TAGS: ZERO-SHOT, MULTILINGUAL NLP
TL;DR: Analysis on Indic and European language families to understand which layers of a multilingual feature extractor are most suited for fusing phonetic/orthographic features that improve zero-shot NER performance by resonating complimentary views of the same text.

Work Experience

Graduate Research Assistant (Auton Lab) *Pittsburgh, PA*
CARNEGIE MELLON UNIVERSITY, ADVISOR: PROF. BARNABÁS PÓCZOS *Sep 2019 - Present*
Working in close collaboration with The Boeing Company to leverage multi-task learning methods for object detection in the noisy and low data regime. Developed detection and recognition pipelines for a robust solution to handwriting recognition.

Machine Learning Engineer *Bengaluru, IND*
AMAZON RESEARCH *Aug 2017 - May 2019*
As part of the Ad Placement Optimization team, worked on the advancement of NLP solutions for problems like advertiser claim detection and diversified product recommendations conditioned on user queries. Developed deep learning pipelines to match products for the "Frequently Bought Together" widget leading to a 0.4% improvement in Click-Through Rate (CTR), and subsequent revenue gains from Sponsored Ads on Amazon.com. The research on semi-supervised click-bait detection was presented at ICON 2018 conference (oral).

Software Engineering Intern *Chennai, IND*
AMAZON KINDLE *May 2016 - Aug 2016*
Contributed to the development of an efficient, distributed product/vendor attribute storage and low latency (reducing ≈ 50ms) retrieval engine for the billion plus e-books on Kindle.

Notable Accolades

- **Academic:** Recipient of three consecutive Institute Medals for the years 2014–15, 2015–16, 2016–17; President of India Gold Medal for securing the highest CGPA (9.91) across all departments; and the RECAL Alumni Award for "Best Academic Performance" and highest GPA in Computer Science Department.
- **Scholarships:** Graduate Research Fellowship covering Tuition + Stipend from 2019-2021. Secured All India Rank 196 at JEE Mains Examination (AIEEE) 2013 and as a result received the AIEEE Merit Scholarship from HRD Ministry, Government of India (INR 140,000). CEDI TATA Industrial Grant of INR 40,000 for *Most Innovative Undergraduate Thesis* on extractive text-summarization. Best poster award at the LTI Student Research Symposium 2019.
- **ACM ICPC:** Represented the college at the ACM ICPC Regionals 2014 (Team Leader) and 2015 (Coach). Secured 1st positions at the Delta Algothon and the Algorithmic Coding Triathlon (Vortex) with 500+ participants across the nation.
- **Extra-curricular:** Part of the school debate team (multiple Best Debater awards), conducted TEDx symposiums at college level and contributed to the National Service Scheme by teaching high school math to students preparing for AIEEE in Trichy.
- **IASc Research Fellowship (2015)**: Proposed and implemented a polynomial time algorithm for K-median with outliers using two dimensional local search heuristics, specifically for large clusters (report). Advisor: Prof. Naveen Garg, IIT Delhi.

Relevant Coursework

- **Graduate:** 10725 Convex Optimization (A+), 10708 Probabilistic Graphical Models (A+), 10701 Machine Learning (PhD) (A+), 11731 Machine Translation (A+), 11747 Neural Networks for NLP (A+), 11777 Multimodal Machine Learning (A).
- **Undergraduate:** MA101 Mathematics-I Calculus (A+), MA102 Mathematics-II Linear Algebra (A+), MA204 Probability Theory (A+), CS064 Artificial Intelligence & Expert Systems (A+), CS065 Natural Language Processing (A+), CS201 Data Structures & Algorithms (A+), CS212 Combinatorics & Graph Theory (A+), CS203 Discrete Structures (A+), MA304 Operations Research (A+).