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 and few-shot learning approaches for fast adaptation as well as learning principled representations imbibing generative/causal mechanisms.

- **ML Applications:** Controlled text/speech synthesis (style transfer), multilingual representations, zero shot NLP, visual language understanding, computer vision, GANs, speech processing.
- **ML Theory:** Non-convex optimization, Bayesian methods, graphical models, learning theory, privacy and fairness.

Education

Carnegie Mellon University *Pittsburgh, PA*
M.Sc. 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 Gold Medallist) *Trichy, India*
B.Sc. (Honors) in Computer Science and Engineering *Aug 2013 - May 2017*
CGPA: 9.91/10.0; Institute Rank: 1

Publications

Refereed Conference or Journal

Nonlinear ISA 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, 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 content.

ReStGAN: A step towards visually guided shopper experience via text-to-image synthesis (US Patent)
Shiv Surya, Amrith Setlur, Arijit Biswas, Sumit Negi; ***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.

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: Fault Tolerance, Scientific Workflows, Cloud Computing
TL;DR: Unsupervised prediction of replication heuristics to improve resource wastage with roughly the same makespan as HEFT.

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 and a modified CTC loss for text recognition in noisy documents.

Workshops

Covariate Distribution Aware Meta-learning
Amrith Setlur, Saket Dingliwal*, Barnabás Póczos*; ***Lifelong Learning Workshop ICML 2020*** [pdf] [video]
Tags: Hierarchical Bayes, Meta-learning
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.

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

TL;DR: Questions the role played by the redundancy in existing task distributions on a meta-learner's generalizability by reducing the support set diversity for tasks. The surprising observations harbinger further discussion of the validity of existing benchmarks.

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. Proposes an auxiliary objective that minimizes a smooth relaxation of the approximation bound using convex conjugates.

Preprints

Explaining The Efficacy of Counterfactually-Augmented Data

Divyansh Kaushik, Amrith Setlur, Eduard Hovy, Zachary C. Lipton; *Preprint* ***Under Review at ICLR 2020*** [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.

Work Experience

Graduate Research Assistant — *Pittsburgh, PA*

CARNEGIE MELLON UNIVERSITY, ADVISOR: PROF. BARNABÁS PÓCZOS — *Sep 2019 -*

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 — *Pittsburgh, PA*

AMAZON RESEARCH — *Sep 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.

Software Engineering Internship — *Chennai, India*

AMAZON KINDLE — *May 2016 - Aug 2016*

Member of the team that developed 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:** Awarded the President of India Gold Medal for securing the highest CGPA (9.91) across all departments for the 2013–17 batch. Recipient of three consecutive Institute Medals for the years 2014–15, 2015–16, 2016–17 and RECAL Alumni Award medal for "Best Academic Performance" and highest GPA in Computer Science Department. Best poster award at the LTI Student Research Symposium 2019.
- **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). TATA Industrial Grant of INR 40,000 for *Most Innovative Undergraduate Thesis*.
- **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+), 10701 Machine Learning (A+), Seq2Seq Models and Machine Translation (A+), Neural Networks for NLP (A+), Multimodal Machine Learning 11777 (Fall 2020), Probabilistic Graphical Models 10708 (Fall 2020).
- **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+), CS203 Discrete Structures (A+), MA304 Operations Research (A+).