ARUSHI JAIN

Mila, McGill University, Montreal, Canada

arushi.jain@mail.mcgill.ca ⋄ https://arushijain94.github.io/

RESEARCH INTERESTS

Reinforcement Learning (RL), Safe RL, Off-Policy RL, Constrained MDPs.

EDUCATION

McGill University, Montreal, Canada

Ph.D., Computer Science	GPA: 4.00/4.00	*Sept 2019 - Present*
MSc, Computer Science	GPA: 4.00/4.00	*Sept 2017 - Sept 2019*

Mila

Supervisor: *Doina Precup*, *Pierre-Luc Bacon*

Indraprastha Institute of Information Technology Delhi (IIIT-D), India

Bachelor of Technology, Computer Science and Engineering   GPA: 9.42/10.00   *Aug 2012 - May 2016*

Image Analysis and Biometrics (IAB) lab

Supervisor:*Mayank Vatsa*, *Richa Singh*

AWARDS

- **Bourse du Centre de Recherche Informatique de Montréal (CRIM) pour Études Supérieures (2018-2019)** Graduate Scholarship provides financial support to only one master and a doctoral student.

PUBLICATIONS

1. **Towards Painless Policy Optimization for Constrained MDPs**[Paper]
   Arushi Jain, Sharan Vaswani, Reza Babanezhad, Csaba Szepesvari, Doina Precup

   Accepted in Conference on Uncertainty in Artificial Intelligence (**UAI**), 2022.
   Accepted at **RLDM** 2022 [Paper]

2. **Variance Penalized On-Policy and Off-Policy Actor-Critic**[Paper][Poster][Slide]
   Arushi Jain, Gandharv Patil, Ayush Jain, Khimya Khetarpal, Doina Precup

   Accepted to the Thirty-Fifth AAAI Conference on Artificial Intelligence (**AAAI**), 2021.

3. **Safe Option-Critic: Learning Safety in the Option-Critic Architecture**[Paper][Poster][Slide]
   Arushi Jain, Khimya Khetarpal, and Doina Precup

   Accepted in Adaptive Learning Agents (ALA), **ICML** Workshop, 2018.
   Accepted in **The Knowledge Engineering Review (KER) Journal** 2021. [Impact Factor 1.07]

4. **Safe Actor-Critic**[Paper][Poster][Slide]
   Arushi Jain, Ayush Jain, and Doina Precup

   Accepted in Safety, Risk and Uncertainty in RL, Uncertainty in AI (**UAI**) Workshop, 2018.
   Accepted in Women in ML (WiML), (**NeurIPS**) Workshop, 2018.

5. **Learning Options using Constrained Return Variance**[Paper][Poster]
   Arushi Jain and Doina Precup

   Accepted in Safety and Robustness in Decision Making, **NeurIPS** Workshop 2019.

6. **Safe Policy Learning with Constrained Return Variance**[Paper][Poster]
   Arushi Jain
   Accepted in Graduate Student AI Symposium held at **Canadian AI Conference** 2019. The proceeding is also published in the **LNAI** Series by Springer Verlag.

7. **Safe Hierarchical Policy Optimization using Constrained Return Variance in Options**[Paper]
   Arushi Jain, Doina Precup
   Accepted at **RLDM** 2019.

8. **Safety using constraint variance in policy-gradient methods**[Thesis]
   Arushi Jain
   Master Thesis, McGill University, Montreal, March 2020.

RESEARCH & WORK EXPERIENCE

Mila, RL Lab, McGill University   September 2017 - Present
Graduate Student *advised by* *Prof. Doina Precup, Pierre-Luc Bacon*

· Real-world tasks with different competing objectives or constraints, it is not always feasible to evaluate the performance of a given policy/behavior. For example, one wants to assess the performance of a given policy in recommendation systems. However, allowing a potentially bad policy that does not abide system's constraints to interact with customers can lead to unpredictable losses. To overcome this problem, we propose to predict by learning from a different **off-policy** that follows constraints and also leads to smaller variability in the performance than the standard on-policy approach.

· We designed a novel and generic **on-policy and off-policy safe actor-critic** framework to learn safe policies such that objective function trades off between minimizing the effects of model uncertainty and maximizing the mean return.

· We worked on introducing novel safety solutions in Options (**hierarchical RL**) by deriving a policy-gradient algorithm where variance in the temporal difference (TD) error is regularized.

Mila, Amii   August 2021 - March 2022
Mentored by *Sharan Vaswani, Reza Babanezhad, Csaba Szepesvri, Doina Precup*

We study planning in infinite horizon Constrained MDPs, where we use parameter-free algorithm from online linear optimization to learn both the policy and follow the constraints. Unlike gradient descent-ascent methods, it is robust to hyperparameter tuning and environment misspecification.

SPORTLOGiQ, Montreal, Canada   June 2019 - Sept 2019
RL Research Intern *mentored by* *Norm Ferns*

Worked on formally comparing and contrasting the agents in a given Markov Decision Process (MDP) by coming up with behavioral pseudo-metrics following the work on *lax-bisimulation*.

Borealis AI, Edmonton, Canada   May 2018 - Aug 2018
AI Research Intern *mentored by* *Nidhi Hedge*

Worked on **safe recommendations** in RL setting which would help provide recommendations with certain guarantees. This would provide life-long learning as well as long-term goal maximization which would be particularly useful in financial applications.

Microsoft Research (MSR), India   June 2016 - July 2017
Research Fellow *advised by* *Sundararajan Sellamanickam, Arun Iyer*

· Developed a **service monitoring and diagnostic tool** for unsupervised hierarchical monitoring of services which replaced a heuristic-based system failing to jointly model multivariate time-series.

· Worked on **email solutions using DNN** to build intelligent approaches to solve problems like multi-label classification, learning a useful representation of email content, etc.

TECHNICAL SKILLS & COURSEWORK

Programming Languages: Python, Tensorflow, Pytorch, C

Coursework: Reinforcement Learning, Probabilistic Analysis of Algorithms, Applied ML, Theoretical ML, Probabilistic Graphical Modeling, Matrix Computation, Mathematical Foundation of ML, Reinforcement Learning and Optimal Control.