

## Research Interests

Reinforcement Learning (RL), Large Language Models (LLMs), Reasoning, Code Generation, Reinforcement Learning from Human Feedback (RLHF), Inverse Reinforcement Learning

## Education

2022–present **Mila, Université de Montréal.**

PhD student supervised by *Prof. Irina Rish*

2013–2018 **Indian Institute of Technology Kharagpur.**

Integrated B.Sc. and M.Sc. in Mathematics and Computing

## Research & Work Experience

Feb'25 – **Research Intern, Cohere.**

July'25 Working on multi-turn LLM agents for code generation

Aug'24– **Visiting Researcher, PoRTaL lab, Cornell University.**

May'25 Visited Sanjiban Choudhury's lab for collaboration on IRL via Successor Features Matching (SFM), model-based IRL with diffusion policies (SAILOR) and using RL to train LLM agents for Code Generation ( $\mu$ Code).

Aug'20 – **Data & Applied Scientist 2, Microsoft IDC.**

Dec'20 Collaborated with Dr. Manik Varma at MSR India to develop scalable and accurate eXtreme Classification algorithms for web-scale recommendation system (published at TheWebConf'21)

Jun'18 – **Data & Applied Scientist, Microsoft IDC.**

Jul'20 Algorithms to improve cross-lingual retrieval of relevant keywords for a query.

May'17 – **Research Intern, HyperVerge Inc..**

Jul'17 Object detection algorithms for video surveillance systems.

May'16 – **Research Intern, ParallelDots.**

Jul'16 Lung nodule detection in 3D CT scans and mitosis detection in histology images.

## Patents

2021 Extreme Classification Processing using Graphs and Neural Networks,  
*Kushal Dave, Deepak Saini, **Arnav Kumar Jain**, Amit Singh, Jian Jiao, Ruofei Zhang and Manik Varma.*  
Pending Approval

## Preprints

2025 A Smooth Sea Never Made a Skilled SAILOR: Robust Imitation via Learning to Search | [Paper](#),  
**Arnav Kumar Jain\***, Vibhakar Mohta\*, Subin Kim, Atiksh Bhardwaj, Yunhai Feng, Sanjiban Choudhury, Gokul Swamy.  
Preprint.

## Publications

2025 Multi-Turn Code Generation Through Single-Step Rewards | [Paper](#),  
**Arnav Kumar Jain\***, Gonzalo Gonzalez\*, Wayne Chen, Alexander M Rush, Wenting Zhao, Sanjiban Choudhury.  
International Conference on Machine Learning (ICML), 2025 ([Spotlight](#), 2.6%)  
VerifAI, SSI-FM and RAP Workshops, ICLR 2025.

- 2025 Non-Adversarial Inverse Reinforcement Learning via Successor Feature Matching | [Paper](#),  
**Arnav Kumar Jain**, Jesse Farebrother, Harley Wiltzer, Irina Rish, Glen Berseth, Sanjiban Choudhury.  
International Conference on Learning Representations (ICLR) 2025  
MHFAI Workshop, ICML 2024
- 2023  $\eta\psi$ -Learning: Maximum State Entropy Exploration using Predecessor and Successor Representations | [Paper](#),  
**Arnav Kumar Jain**, Lucas Lehnert, Irina Rish, Glen Berseth.  
Neural Information Processing Systems (NeurIPS), 2023  
Frontiers4LCD Workshop, International Conference on Machine Learning (ICML), 2023
- 2022 Learning Robust Dynamics through Variational Sparse Gating | [Paper](#),  
**Arnav Kumar Jain**, Shivakanth Sujit, Shruti Joshi, Vincent Michalski, Danijar Hafner, Samira Ebrahimi-Kahou.  
Neural Information Processing Systems (NeurIPS), 2022  
Deep RL Workshop, Neural Information Processing Systems (NeurIPS), 2021
- 2021 GalaXC: Graph neurAL networks with Labelwise Attention for eXtreme Classification | [Paper](#),  
Deepak Saini\*, **Arnav Kumar Jain\***, Kushal Dave\*, Amit Singh, Jian Jiao, Ruofei Zhang, Manik Varma.  
The Web Conference (TheWebConf), 2021
- 2020 Graph Regularization for Multi-lingual Topic Models | [Paper](#),  
**Arnav Kumar Jain\***, Gundeep Arora\*, Rahul Agrawal.  
SIGIR Conference on Research and Development in Information Retrieval, 2020
- 2020 Prior guided GAN based Semantic Inpainting| [Paper](#),  
Avishek Lahiri\*, **Arnav Kumar Jain\***, Sanskar Agrawal, Prabir Kumar Biswas, Pabitra Mitra.  
Computer Vision and Pattern Recognition (CVPR), 2020
- 2019 Faster unsupervised semantic inpainting: A GAN based approach | [Paper](#),  
Avishek Lahiri\*, **Arnav Kumar Jain\***, Divyashree Nadendla, Prabir Kumar Biswas.  
International Conference on Image Processing (ICIP), 2019
- 2018 Optimal Spline Trajectories by Modelling Kinematic Constraints in Robot Soccer | [Paper](#),  
Abhinav Agarwalla\*, **Arnav Kumar Jain\***, KV Manohar, Arpit Saxena, Jayanta Mukhopadhyay.  
Conference on Data Science and Management of Data (CoDS-COMAD), 2018
- 2016 KgpKubs Team Description Paper,  
Abhinav Agarwalla, Kumar Abhinav, **Arnav Jain**, Kaustubh Mundhadha, Dhananjay Yadav, ....  
RoboCup, 2016

## Workshop Papers

- 2020 Predicting Regional Locust Swarm Distribution with Recurrent Neural Networks | [Paper](#),  
Hadia MO Samil\*, Annabelle Martin\*, **Arnav Kumar Jain\***, Susan Amin, and Samira Ebrahimi-Kahou.  
AI+HADR Workshop, Neural Information Processing Systems (NeurIPS), 2020
- 2017 Recurrent Memory Addressing for describing videos | [Paper](#),  
**Arnav Kumar Jain\***, Abhinav Agarwalla\*, Kumar Krishna Agrawal\* and Pabitra Mitra.  
DeepVision Workshop, Computer Vision and Pattern Recognition (CVPRW), 2017

## Reviewing

ICLR'25, ICML'25, NeurIPS'25, ICML'24, RLC'24, NeurIPS'24, ICML'23, NeurIPS'23 (Top Reviewer),  
ICML'22, NeurIPS'22.

## Mentoring

- 2025- **Atiksh Bhardwaj**, (BTech student at Cornell University).  
Scaling up Model-based Inverse Reinforcement Learning (SAILOR).

- 2025 **Gonzalo Gonzalez-Pumariega**, (*PhD student at Cornell University*).  
Multi-turn Code Generation ( $\mu$ Code)
- 2024-2025 **Vibhakar Mohta**, (*MS CMU, Now Nuro*).  
Scaling up Model-based Inverse Reinforcement Learning (SAILOR).
- 2024-2025 **Subin Kim**, (*MS student at KAIST*).  
Scaling up Model-based Inverse Reinforcement Learning (SAILOR).
- 2024 **Mahdi Kleit**, (*MS student at Mila*).  
Finetuning Diffusion Models.

## Talks

- 2024 **PoRTaL group, Cornell University**, Switching RL WorkFlow to Jax
- 2023 **ServiceNow Research**, Efficient Exploration with Successor and Predecessor Representations
- 2023 **Mila RL Workshop**, Efficient Exploration with Successor and Predecessor Representations

## Awards & Achievements

- 2025 **ICLR Travel Assistance**.  
Received the financial assistance from ICLR to present SFM.
- 2024 **FRQNT Fellowship**.  
Received the Fonds de recherche du Québec PhD fellowship.
- 2023 **Top Reviewer, NeurIPS 2023..**
- 2019 **Excellence in Innovation, Microsoft**.  
Awarded for creating models resulting in business impact and reducing defect rate on Bing Ads platform
- 2018 **ACM India Student Travel Grant**.  
Received travel grant to present accepted paper in ACM IKDD CoDS COMAD 2018
- 2017 **Data Science Bowl 2017**.  
Received 5000\$ for the 3<sup>rd</sup> highest voted kernel on *Candidate generation and LUNA16 preprocessing*
- 2015 **FIRA RobotSoccer WorldCup, South Korea**.  
Participated in FIRA, 2015 in SIMUROSOT league and won Bronze in MIROSOT league, and were the first Indian team to have a podium finish
- 2015 **SudoCode, Kshitij, IIT Kharagpur**.  
Secured 1<sup>st</sup> Position (2015) and Best Freshers' (2014) in a national event to develop AI algorithms.
- 2013 **Innovation in Science Pursuit for Inspired Research (INSPIRE) Scholarship**.  
Scholarship awarded by the Department of Science and Technology, Government of India

## Other Activities

- 2014 **Texas Instruments certified Winter Workshop**.  
Mentored 60 freshmen to develop a bot that can follow lanes and detect shapes.
- 2013 – 2015 **National Service Scheme**.  
Organized two medical camps with free checkups and medicines, and volunteered to teach school children.