Videh Raj Nema

da25d003@smail.iitm.ac.in • G-Scholar • vrn25 • vrn25 • vrn25.github.io

My research focuses on building an Ātmanirbhara Bhārata in STEM. To this end, I am interested in AI grounded in the Indian values to meet her interests, needs, and goals. Presently, I work on multimodal machine learning for information synthesis. Further, my research methodology focuses on building intelligent autonomous agents that are compatible (and not necessarily indistinguishable) with humans and other intelligent agencies in a wide range of dynamic environments. Previously, I have worked on reinforcement learning with multiple learning agents and humans integrated in a learning loop. Concurrently, I am also interested in Indian Knowledge Systems (Bhāratīya Jñāna Paramparā) towards the effort of bringing in the Bhāratīya way of thought, speech, and action in STEM along with restoring the lost Indian knowledge and applying it in the present context.

Education

Indian Institute of Technology Madras • Chennai, India — July '25 – Present
Doctor of Philosophy (PhD) • *Data Science and Artificial Intelligence* • CGPA: **9.64**/10.00
Advisor: Prof. Balaraman Ravindran

University of Alberta • Edmonton, Alberta, Canada — Sept '22 – Jan '25
Master of Science (Thesis-based) • *Computing Science* • CGPA: **4.0**/4.0
Advisor: Prof. Matthew Taylor, IRLL, RLAI, AMII
Thesis: Continual Preference-based Reinforcement Learning with Hypernetworks [PDF]
Nominated for the UAlberta Computing Science MS Outstanding Thesis Award

National Institute of Technology Karnataka • Surathkal, India — July '18 – May '22
Bachelor of Technology • *Computer Science and Engineering* • CGPA: **9.02**/10.00

Publications

Interactive Robust Policy Optimization for Multi-Agent Reinforcement Learning [PDF]
Videh Raj Nema *and Balaraman Ravindran*
Deep Reinforcement Learning Workshop, **NeurIPS 2021**
Learning and Decision-Making with Strategic Feedback Workshop, **NeurIPS 2021**

Understanding the effect of varying amounts of replay per step [PDF]
Animesh Kumar Paul and ***Videh Raj Nema*** *(equal contribution)*
arXiv:2302.10311, 2023

Research Experience

Intelligent Robot Learning Laboratory & AMII — Edmonton, CA
Graduate Researcher | Supervisor: Prof. Matthew Taylor — Jan '23 – Jan '25

- Formalized a continual preference-based reinforcement learning setting involving continual learning of reward functions (and subsequently policies) from the non-stationary preferences of a teacher (e.g., a human) over different behaviors.
- Proposed using regularized hypernetworks that learn to simultaneously acquire new and preserve old knowledge and learn reward functions better aligned with the teacher's preferences than several continual learning baselines.
- Performed a study analyzing the mechanics of continually learned reward functions by varying the amount of teacher feedback. Demonstrated reward overfitting in the continual setting, the associated risks, and preventive measures.
- Worked on human-in-the-loop reinforcement learning algorithms that incorporate nuanced objectives aligned well with user preferences into agents to avoid meticulous reward engineering and hacking via a mixture of modalities.

California Institute of Technology — Pasadena, USA
Research Intern | Supervisors: Prof. Animashree Anandkumar & Sahin Lale — June '21 – Aug '21

- Worked at the intersection of model-based reinforcement learning (MBRL) and game theory to develop algorithms treating MBRL as an interactive game between policy and model to solve the problem of distribution shift (source).
- Extended interactive optimization framework of CGD to effectively solve the MBRL game [PDF]. Prepared a literature survey on various MBRL algorithms such as MBPO, PETS, MB-MPO, ME-TRPO, MBMF, & STEVE.

Purdue University — West Lafayette, USA
Research Intern | Supervisors: Prof. Kamyar Azizzadenesheli & Manish Prajapat — Mar '21 – May '21

- Developed a framework for turn-based strategic optimization (TB_Opt) inspired by cognitive hierarchy. The derived family of algorithms together encompass and generalize various existing game-theoretic models of optimization.
- Realized the update rules for various levels of reasoning via a regularized local linear or multilinear approximation of agent's objective and substituting the estimated future learning steps of the other agents.
- Extended TB_Opt to a turn-based multi-agent reinforcement learning setting where exact optimization objectives and derivatives are not available and must be estimated via policy gradients.
- Experimented in polynomial optimization settings demonstrating that naive models like gradient descent/ascent are either prone to oscillatory behavior or converge very slowly to the Nash equilibrium, while going to higher levels of reasoning with TB_Opt enables faster convergence. [PDF] [GitHub].

Robert Bosch Center for Data Science and AI IIT Madras, India
Research Intern | Supervisor: Prof. Balaraman Ravindran April '20 – Aug '20

- Worked at the conjunction of multi-agent reinforcement learning and robust control for developing interactive policies robust to destabilizing adversarial disturbances/perturbations in simulated robotic and continuous control domains.
- Developed interactive policy optimization (IPO) algorithms that address non-stationarity in multi-agent systems by recovering the classical game-theoretic infinite-order recursive reasoning between the agents and converge to better approximations of the Nash equilibria in multi-agent games.
- Extended the interactive optimization framework of CGD to derive practical policy gradient (stochastic and deterministic), natural policy gradient, and trust-region algorithms for multi-agent reinforcement learning using the paradigm of centralized training and decentralized execution.
- Combined IPO with opponent modeling via maximum likelihood for application in adversarial settings and integrated it with RARL to provide robustification to destabilizing adversarial disturbances from the environment or opponents at test time. Analyzed the applicability of IPO-RARL for multi-agent Sim2Real transfer. [PDF].

Selected Projects

Increasing Replay *Prof. Adam White*

- Worked on improving the sample efficiency of off-policy deep reinforcement learning methods that use experience replay by increasing the amount of replay per step (replay frequency or replay ratio).
- Investigated the effect of changing replay frequency on performance and hyperparameter robustness using sensitivity curves. Empirically showed that higher replay frequency results in less sensitivity across a range of hyperparameters.
- The experiments were based on good empirical practices in reinforcement learning (source). [PDF] [Slides].

CrUiSE: A Comprehensive & Unified Metric for Summarization Evaluation *Prof. Shashidhar Koolagudi*

- Worked on a robust research setup for automatic text summarization via the introduction of new informative evaluation protocols overcoming the limitations of existing gram-based metrics like ROUGE and its variants (source).
- Developed a novel learning-based metric (CrUiSE) that comprehensively evaluates summaries across multiple dimensions – *fluency*, *consistency*, *coherence*, *relevance*, and unifies them according to the user's requirement.
- Compared CrUiSE with 34 existing metrics based on correlation with human judgment (source), showing that it outperforms several baselines in 3 of 4 dimensions. [PDF] [Slides]

Interactive Multi-Agent Reinforcement Learning *IEEE NITK Student Branch*

- Worked on multi-agent reinforcement learning (MARL) algorithms that enable interactive, collaborative, and negotiating behaviors in mixed cooperative-competitive multi-agent games.
- Implemented MARL algorithms with CGD in two-player zero-sum games like Markov soccer, LQR, bilinear game, rock-paper-scissors, and matching pennies, observing better convergence to Nash equilibrium and interactive behavior.
- Implemented CGD and LOLA in general-sum social dilemmas like one-step and iterated prisoner's dilemma, observing better and faster convergence to the socially optimal equilibrium (Tit-for-Tat) than independent reinforcement learning.

Adversarial Reinforcement Learning *IEEE NITK Student Branch*

- Worked on analyzing adversarial attack and defense algorithms in reinforcement learning in pixel-based & physically realistic multi-agent simulated robotic settings. Wrote a survey illustrating the internal mechanism of such algorithms.
- Implemented pixel-based attacks on policies trained using PPO and DQN algorithms with the L_1, L_2, and L_∞ variations of the FGSM in Atari, which reduces the performance in both white-box and black-box settings.
- Implemented black-box attacks in two-player zero-sum MuJoCo environments by training adversarial policies that induce natural observations that are adversarial to the victim. The attacks reduce the performance of agents explicitly trained to be robust to opponents via self-play methods. [Project page]

Coursework

Graduate courses at the Indian Institute of Technology Madras (ongoing)

- DA 7016: Recent Advances in Generative AI (S)[1]
- EE 5179: Deep Learning for Imaging (S)
- CS 6852: Theory and Applications of Ontologies (A)

Graduate courses at the University of Alberta

- CMPUT 566: Intro to Machine Learning (A+)
- CMPUT 609: Reinforcement Learning II (A+)
- CMPUT 603: Teaching & Research Methods (CR)[2]
- CMPUT 655: Reinforcement Learning I (A+)
- CMPUT 605: Human in the Loop RL (A+)
- INT D 710: Ethics & Academic Citizenship (CR)

[1]The grade S is equivalent to A+ (grade point 10) at IIT Madras.
[2]CR implies that the course is taken for credits only without grading.

Teaching

- Teaching Assistant for DA 5401: Data Analytics Lab (Fall 2025) by Dr. Sudarsun Santhiappan at IITM.
- Mentoring for CMPUT 656: Human in the Loop Reinforcement Learning (Fall 2023) by Prof. Matthew Taylor at UofA.
- Teaching Assistant for CMPUT 267: Basics of Machine Learning (Fall 2022) by Prof. Nidhi Hegde at UofA.

Achievements

- Nominated for the University of Alberta, Computing Science MSc Outstanding Thesis Award, 2025.
- Secured first rank in CMPUT 609: Reinforcement Learning II (Winter, 2023) offered by Prof. Rich Sutton at UAlberta.
- Awarded the University of Alberta Graduate Recruitment Scholarship, 2022.
- Selected for Caltech's **Summer Undergraduate Research Fellowship** (SURF), 2021 (award declined due to COVID).
- Selected for Research Week with Google 2022, organized by Google Research India.
- Selected for **CIFAR DLRL** (2023) **EEML** (2022) and **M2L** (2021, 2022) summer schools.
- Secured **322/340** in GRE and **113/120** in TOEFL iBT organized by the ETS.
- Secured **AIR 1939** out of 1 million candidates in JEE Mains 2018 (**99.80** percentile).
- Secured **94.8%** in AISSCE 2018 conducted by the Central Board of Secondary Education.
- Secured **AIR 1127** out of 50,000 candidates in the KVPY SX stream examination 2017 (**97.75** percentile).
- Awarded the NTSE (2016) scholarship, being among the **750** selected out of 1 million students appearing in the country.
- Awarded DELF (Diplôme d'Études en Langue Française) A2 by French ministry of education (**92,00/100**). [Certificates].

Extra Curricular

- **Intelligence SIG Head, Web Enthusiasts' Club, NITK (2020-22; 100 members)**
  - Organized the Expert talk series at NITK and invited *Prof. Balaraman Ravindran* for a talk and interactive session on *Why do computers learn to play games* (YouTube).
  - Gave a talk on *Learning and Optimization in Multi-agent systems* for WEC NITK (YouTube, slides).
  - Organized the Research Talk series to enhance the research culture at NITK; Conducted Kaggle mentorship sessions.
  - Organized workshops on version control systems & Open Source programs like GSoC as a part of Hacktoberfest NITK.
- **Executive Member, Computer Society, IEEE NITK Student Branch (2020-22; 60 members)**
  - Mentored 25+ students in reinforcement learning and path-finding AI algorithms in maze-games (Website, Slack).
  - Co-organized the *Around Her World in 10 Ways* technological event as a part of Women in Engineering, NITK.
  - Gave an *INSIGHT* talk guiding undergraduate students with my research journey (YouTube, slides).
  - Gave a talk on *Learning to learn via meta-learning* for the computer society at IEEE NITK (YouTube, slides).
  - Published technical blogs on Generative Adversarial Networks (GANs) and adversarial reinforcement learning.
- **Technical reviewer** for the 4$^{\text{th}}$ *Conference on Lifelong Learning Agents* (**CoLLAs, 2025**).
- **Volunteer** for the 35$^{\text{th}}$ conference on *Neural Information Processing Systems* (**NeurIPS, 2021**).