Mudit Verma

Arizona, USA • +1-480-494-7411 • mudit.verma2014@gmail.com • linkedin.com/in/iammudit • famishedrover.github.io

Research : Human Aware AI, Preference based Reinforcement Learning, Large Language Models & HRI

EDUCATION

Arizona State University, Tempe, AZ — 2019 – Expected Fall 2024
Ph.D in Computer Science, Advisor : Dr. Subbarao Kambhampati
GPA : 4.0/4.0

Delhi Technological University (Delhi College of Engineering), New Delhi, India — 2015 – 2019
B.Tech in Information Technology, Thesis Advisor : Dr. Seba Susan
GPA : 9.6/10 (Gold Medalist)

WORK EXPERIENCE

Apple Inc — May 2023 – Sept 2023
Machine Learning Research Intern, Cupertino, CA

Research work with Machine Learning Research (MLR) Group. Advised by Rin Metcalf and Barry Theobald.

- Hindsight PRIORs for Reward Learning from Human Preferences. *(ICLR 2024)*

Apple Inc — May 2022 – Aug 2022
Machine Learning Research Intern, Cupertino, CA

Preference based Reinforcement Learning research with Machine Learning Research (MLR) group
Advised by Rin Metcalf Barry Theobald.

- Symbol Guided Hindsight Priors for Reward Learning from Human Preferences at IROS RLCONFORM, NeurIPS HILL 2022.

Intel Corporation — May 2021 – Aug 2021
Deep Learning Software Engineering Intern, Santa Clara (Remote), CA

Deep Learning Acceleration and implications to explainability. Advised by Wei Wang

- First analysis of float32 ResNet50 architecture on Intel IceLake (ICX) with several BFloat16 optimizations.
- Showed discrepancy in Saliency Based explanations between original ResNet50 and Quantized ResNet50.

Samsung Semiconductor India Research — May 2018 – Aug 2018
Machine Learning Intern, Bangalore, India

Novel Monte Carlo Tree Search (MCTS) with Residual Network approximation and state space reduction schemes based approach to Redundancy Analysis (RA) Algorithms for DRAM. Advised by : Atishay Kumar

- Implemented C++ based DRAM Bank Simulator, (400x faster) with enhanced Fault Classes.
- Awarded Best Intern Project at SSIR.

Samsung Semiconductor India Research — May 2017 – Aug 2017
Machine Learning Intern, Bangalore, India

LSTM based Stream Selection for Smart Data Categorization (STRASDAC) algorithm to reduce write-wearing in SSDs and improve Garbage Collection (GC). Advised by : Sandeep Sammatshetti

- Identified fault classes in SSDs. C++/Python based SSD to simulate reads/write/garbage collection
- Runners up Best Intern Project at SSIR.

Indraprastha Institute of Information Technology (IIIT Delhi) — Aug 2017 – May 2019
Visiting Researcher, New Delhi, India

Worked on Indian regional spoken language identification using Capsule Networks. Worked on optimizing smart homes for humans in the loop. Advised by : Dr. Arun Balaji Buduru

CONFERENCE PROCEEDINGS : ICLR / NeurIPS / HRI / AAAI / AAMAS

Preference Based / Advisable Reinforcement Learning :

- **Verma, M.**, Metcalf, K. (2024). Hindsight PRIORs for Reward Learning from Human Preferences. *(ICLR 2024)*
- **Verma M.***, Bhambri S.* Kambhampati S. (2024) Theory of Mind abilities of Large Language Models in Human-Robot Interaction : An Illusion *(In ACM/IEEE International Conference on Human-Robot Interaction (HRI))*
- Guan, L., **Verma, M.,** Kambhampati, S. (2021). Widening the Pipeline in Human-Guided Reinforcement Learning with Explanation and Context-Aware Data Augmentation. *Advances in neural information processing systems (NeurIPS). (Spotlight)*

Symbolic Lingua Franca / Explainable AI

- Sreedharan, S., Soni, U., **Verma, M.**, Srivastava, S., Kambhampati, S. (2022). Bridging the Gap: Providing Post-Hoc Symbolic Explanations for Sequential Decision-Making Problems with Inscrutable Representations. In 2020 ICML Workshop on Human in the Loop Learning (ICML HILL 2020). In International conference on learning representations (ICLR) 2022.
- Kambhampati, S., Sreedharan, S., **Verma, M.**, Zha, Y., Guan, L. (2021). Symbols as a Lingua Franca for Bridging Human-AI Chasm for Explainable and Advisable AI Systems. *In Proceedings of the AAAI Conference on Artificial Intelligence (Blue Sky Track).*

Trust Aware AI

- Zahedi Z.,**Verma M.**, Sreedharan S. Subbarao Kambhampati. Trust-Aware Planning: Modeling Trust Evolution in Iterated Human-Robot Interaction. *(2023, HRI)*
- Zadehi, Z., Sreedharan, S., **Verma, M.**, Kambhamapati, S. (2022) Modeling the Interplay between Human Trust and Monitoring. *In Companion of the 2022 ACM/IEEE International Conference on Human-Robot Interaction (HRI).*

Human Aware AI

- Thai T., Soni U., **Verma M.** Gopalakrishnan S., Shen M., Garg M., Kalani A.,, Vaidya N., Kambhampati S., Varshney N., Baral C., Sinapov J., Scheutz M. (2023). Methods and Mechanisms for Interactive Novelty Handling in Adversarial Environments. *(AAMAS 2023, Extended Abstract)*
- **Verma, M.**, Buduru, A. B. (2020). Fine-grained Language Identification with Multilingual CapsNet Model. *In 2020 IEEE Sixth International Conference on Multimedia Big Data (BigMM) IEEE.*

Others :

- **Verma, M., Sinha, P**., Goyal, K., Verma, A., & Susan, S. *(2019, June).* A novel framework for neural architecture search in the hill climbing domain. *In 2019 IEEE Second International Conference on Artificial Intelligence and Knowledge Engineering (AIKE) (pp. 1-8). IEEE.*

WORKSHOP PROCEEDINGS : ICML / AAAI / IROS / NeurIPS / ICAPS

Preference based RL / Advisable RL :

- **Verma M.***, Bhambri S.*, Murthy A., Kambhampati S. Benchmarking Multi-Agent Preference based Reinforcement Learning for Human-AI Teaming. *(arxiv Preprint 2024)*
- **Verma M.***, Bhambri S.* Kambhampati S. Preference Proxies: Evaluating Large Language Models in capturing Human Preferences in Human-AI Tasks. (*In Theory of Mind Workshop, Many Facets of Preference Learning (Oral) Workshop at ICML 2023.)*
- **Verma M.**, Bhambri S., Kambhampati S. Exploiting Action Distances for Reward Learning from Human Preferences. In Many Facets of Preference Learning Workshop at ICML 2023.
- **Verma, M.**, Bhambri, S. Kambhampati, S. (2023).Exploiting Unlabeled Data for Feedback Efficient Human Preference based Reinforcement Learning. (*In AAAI R2HCAI 2023.)*
- **Verma, M.**, Kambhampati, S. (2023). Data Driven Reward Initialization for Preference based Reinforcement Learning. *(In AAAI R2HCAI 2023.)*

- **Verma, M.**, Kambhampati, S. (2023). A State Augmentation based approach to Reinforcement Learning from Human Preferences. *(In AAAI R2HCAI 2023.)*
- **Verma, M.,** Kharkwal, A., Kambhampati, S. (2022). Advice Conformance Verification by Reinforcement Learning agents for Human-in-the-Loop. *(In IROS RLCONFORM 2022.)*
- **Verma, M.**, Metcalf, K. (2022). Symbol Guided Hindsight Priors for Reward Learning from Human Preferences. *(In IROS RLCONFORM NerurIPS HILL 2022.)*
- S., Soni, U., Sreedharan, S., **Verma, M.**, Guan, L., Marquez, M., Kambhampati, S. (2022). Towards customizable reinforcement learning agents: Enabling preference specification through online vocabulary expansion. *(In NeurIPS HILL 2022.)*
- **Verma, M.***,Guan, L.*, Kambhampati, S. (2020). Explanation Augmented Feedback in Human-in-the-Loop Reinforcement Learning. *In 2020 ICML Workshop on Human in the Loop Learning (HILL 2020), NeurIPS 2020 HAMLETS, NeurIPS 2020 DRL*

Trust :

- Zahedi, Z., **Verma, M.**, Sreedharan, S., Kambhampati, S. (2021). Trust-Aware Planning: Modeling Trust Evolution in Longitudinal Human-Robot Interaction. *In ICAPS 2021 Workshop on Explainable AI Planning.*

Human Aware AI :

- Gopalakrishnan, S., **Verma, M.**, Kambhampati, S. (2021, June). Synthesizing Policies That Account For Human Execution Errors Caused By State Aliasing In Markov Decision Processes. *In ICAPS 2021 Workshop on Explainable AI Planning.*
- **Verma, M.**, Bhambri, S., Buduru, A. B. (2019). Making Smart Homes Smarter: Optimizing Energy Consumption with Human in the Loop. *arXiv preprint arXiv:1912.03298.*

HONORS & AWARDS

Nomination ASU PhD Block Grant, 2024
ASU SCAI Doctoral Fellowship, 2023
ASU Engineering Graduate Fellowship, 2022
ASU CIDSE Doctoral Fellowship, 2019
DTU/DCE Gold Medalist 2019
DTU/DCE Department Merit Rank Scholarship 2019, 2018, 2017
Pramod Jain Scholarship, Best Student at DTU 2017

First, Smart India Hackathon, 2019. Varanasi, India. (37000+ submissions)
Fourth, Hack in the North (IIIT Allahabad), 2018, Allahabad, India (200+ participants)
Education Innovation Mentorship Programme, ReadAlliance USAID, 2018 (highly selective)
First, READing Hackathon, 2017. USAID, NASSCOM
15th, World Food India Hackathon, 2017. 150+ teams pan India

PROJECTS

- **Perfect Observability is a Myth** : Proposed a method to deal with partial observability of humans for Reinforcement Learning domains, when providing advice to agents.
- **Term Paper, Dr. Joshua Daymude :** Randomly Wired Networks are on the rise, have we been creating wrong Networks all along?
- **Term Paper, Dr. Dimitri Bertsekas :** Diverging Emerging Field of Multi-Task Reinforcement Learning
- **Colors of Desert, Dr. Sharon Hsiao :** Used D3 to highlight deserts are indeed colorful.
- **TAC :** App that adapts and teaches children (dyslexic) to read/write/recognize using ML techniques.
- **CookHub :** Open source community for Recipes where one can chat, push, pull, fork, collaborate and view trending recipes and contributors.
- **Shut The Fake Up :** App/Website Wisdom of Majority based AI for Fake News detection.

TEACHING & SERVICE

Teaching Assistant, CSE 471 with Dr. Subbarao Kambhampati. Fall 2019
Reviewer/PC Member, at several Computer Science research conferences.
*ICML (2023, 2022); NeurIPS, (*2023, 2022); ICLR(2023, 2022), IJCAI(2024), *ICAPS, (*2023, 2022, 2021); *AAAI* (2023, 2022), *HRI* (2022)
and several IEEE venues.

SKILLS

Research Areas:
Reinforcement Learning, Interactive Machine Learning, RLHF, Large Language Models (LLM), Multi-modal Learning, Preference Learning, Safe AI, Recommender Systems, Natural Language Processing, Human Robot Interaction

Programming : C++ / Python / Java / PDDL
Misc : Android App Development, Bootstrap, Flask, D3, HTML/CSS, SQL
Frameworks : PyTorch, Jax, Tensorflow, Mujoco-Gym, HuggingFace-Transformers, OpenCV, Scikit-Learn