Daman Arora

Indian Institute of Technology, Delhi

WORK EXPERIENCE

Research Fellow, Microsoft Research India **Jul '23 – Present**
Working on AI Alignment in the context of Retrieval Systems *Bengaluru, India*

EDUCATION

Indian Institute of Technology, Delhi **Jul '18 – Jul '23**
*B.Tech and M.Tech Computer Science and Engineering - **9.4*** *Delhi, India*

Delhi Public School, R.K. Puram **Jul '16 – Jul '18**
*Central Board of Secondary Education **95.2%*** *Delhi, India*

PUBLICATIONS

Evaluating and Providing a Problem Solving Benchmark for LLMs [*Accepted at EMNLP 2023*] **Oct '23**
- Generated a novel dataset for evaluating LLMs on JEE Advanced problems
- Evaluated and analysized performance of multiple open source and closed source models on this dataset

Verifiers to Improve Planning Capabilities of LLMs [*Accepted at ICML KLR workshop 2023*] **May '23**
- Discovered that fine-tuned language models are unable to follow pre-conditions while generating trajectores
- Trained verifiers using random sampling to distinguish between valid and invalid actions

Exploiting Long-Range Influences for Generalized Neural Policies [*Accepted at UAI 2023*] **Nov '22**
- Devised a novel metric called **Influence Distance** which captures dependency length for RDDL RMDPs
- Provided a novel GNN architecture capable of discovering and exploiting long range dependencies
- Demonstrated improved explainability of the policies learnt using Imitation Learning

Generalized Neural Policies for Relational MDPs [*Accepted at UAI 2022*] **Jan '22**
- Augmented state representation for RMDPs in RDDL incorporating non-fluents to improve generalization.
- Improved performance using imitation learning(behavioral cloning) from MCTS-based planners.
- Received student scholarship from Uncertainty in Artificial Intelligence (UAI) to attend the conference

PROJECTS

Diffusion Models for Downscaling for Climate Data (Prof. Aditya Grover, UCLA) **Ongoing**
- Trained Stacked Diffusion Models to generate SOTA performance on downscaling for precipitation data.
- Proposed Bayesian inference from Diffusion models by marginalizing over noise variables via MC sampling.
- Proposed and implemented uncertainty bounds for diffusion models via iterative Jacobian computation.

Visual Sudoku using Graph Neural Networks (Prof. Parag Singla) **May '21**
- Trained a Recurrent Relational Network to solve sudoku from unlabelled images, achieving 93.4% accuracy.
- Created a dataset for LeNet using GAN trained on augmentations of grounded images to provide a warm start
- Jointly trained RRN & LeNet with contrastive loss to handle noisy labels by preventing representation collapse

chaii - Hindi and Tamil Question Answering (Prof. Mausam) **Nov '21**
- Competed and topped among classmates on Google's chaii competition as part of the course on NLP.
- Fine-tuned SOTA Multi-lingual model on data collated from multiple sources on span prediction task.
- Experimented with Ensembling and Label smoothing in order to counter class imbalance b/w Hindi & Tamil.

TECHNICAL SKILLS

Languages: Python, Java, C, C++ **Tools:** Linux, HPC, PyTorch, Tensorflow

SCHOLASTIC ACHIEVEMENTS

* Achieved **AIR 158** among 2 lakh students in JEE Advanced and **AIR 118** in JEE Mains 2018 examination
* Awarded the Merit Award for scoring in top 7 percentile, achieving 9.535 and 10 GPA in 1st & 2nd semester
* Achieved an All India Rank of 15 in the Kishore Vaigyanik Protsahan Yojana(KVPY) Examination, 2018.