(+91)-8100199177
agarwalsiddhant10@gmail.com
agarwalsiddhant10

Siddhant Agarwal

siddhant-agarwal
agarwalsiddhant10.github.io
Siddhant Agarwal

EDUCATION

B. Tech. + M. Tech. (Dual Degree) in Computer Science *(Ongoing)* 2017-2022
Indian Institute of Technology Kharagpur
GPA: **9.89/10.00** (**Dept. Rank 1**) / *Thesis:* Reinforcement Explanation Learning *Advisor:* Prof. Abir Das

FIELDS OF INTEREST

Reinforcement learning, Robotics, Computer Vision, Adversarial Attacks

PUBLICATIONS

[1.] Agarwal S., Courville A., and Agarwal R. **Behavior Predictive Representations for Generalization in Reinforcement Learning**. *NeurIPS 2021 DeepRL and EcoRL workshops* [PDF]

[2.] Agarwal S. et al. **Reinforcement Explanation Learning**. *NeurIPS 2021 workshop on eXplainable AI approaches for debugging and diagnosis* [*Oral*] [PDF]

[3.] Sun M., Agarwal S., and Kolter Z. **Poisoned Classifiers are not only backdoored, they are fundamentally broken**. *ICLR 2021 workshop on Security and Safety in Machine Learning Systems, Under Review at ICLR 2022* [PDF]

[4.] Raman M., Agarwal S., et al. **Learning to Deceive Knowledge Graph Augmented Models via Targeted Perturbations**. *ICLR 2021, KR2ML workshop at NeurIPS 2020* [Best paper nomination] [PDF]

[5.] Agarwal S. et al. **Traffic Sign Classification using Hybrid HOG-SURF Features and Convolutional Neural Networks**. *International Conference on Pattern Recognition Applications and Methods, 2019* [Best poster nomination] [PDF]

EXPERIENCE

- **Generalization in Model-Based RL** - *Advisor:* Rishabh Agarwal, Prof. Aaron Courville *March 2021 - present*
  Research Areas: **Model-based Reinforcement Learning, Predictive Representations, Generalization**
  - Trained the agent to learn predictive representations capturing long term behavior to improve generalization in RL.
  - Achieved state of the art performance on the jumping task, both with and without data augmentations and currently improving on complex environments like OpenAI procgen benchmark and distracting dm-control.

- **Goldman Sachs, Data Engineering Team** *June 2021 - July 2021*
  Developed a python client to use the services of an ETL platform to quickly transfer data from pandas dataframe to a MemSQL store. Created RESTful server endpoints for the ETL platform and accessed them from the client.

- **LocusLab, Carnegie Mellon University** - *Advisor:* Prof. Zico Kolter *May 2020 - October 2020*
  Research Areas: **Adversarial Attacks, Random Smoothing, Data Poisoning, Feature Visualization**
  - Showed that *backdoored classifiers* are fundamentally broken and can be attacked by anyone rather than only the adversary.
  - Used the perceptual property of robust classifiers and smoothing techniques to construct alternate triggers.
  - Tested for BadNet, CLBD and HTBD poisoning on ImageNet. Achieved high performance on Trojai-AI dataset.

- **INKLab, University of Southern California** - *Advisor:* Prof. Xiang Ren *March 2020 - May 2020*
  Research Areas: **Visual Commonsense Reasoning, Knowledge Graphs, Graph Neural Networks, Visual Question Answering, Visual Linguistic Encoding, Deceiving Knowledge graph models, Reinforcement Learning**
  - Incorporated common sense knowledge from KGs (ConceptNet) into visual linguistic models like VL-BERT using MHGRN and RN to deduce common sense inferences from images and answer questions from the VCR dataset.
  - Trained DQN policies to heavily perturb KGs without changing the downstream performance of the different KG models. Was able to deceive MHGRN, GRN and RN (on CSQA and OBQA) and KGCN and RippleNet (on MovieLens and Music).

- **Reinforcement Explanation Learning** - *Advisor:* Prof. Abir Das *August 2019 - May 2021*
  Research Areas: **Explainable AI, Reinforcement Learning, Computer Vision, Deep Learning**
  - Trained RL agents to intelligently search for the optimal masks for generating saliency maps to explain classification.
  - Showed tremendous speedup (10x) over the existing black box techniques without hurting performance.

- **Autonomous Ground Vehicles Research Group** *March 2018- current*
  Research Areas: **Computer Vision, Deep Learning, Reinforcement Learning, Planning, Controls**
  - Developed a traffic sign detection system for Indian Traffic Signs combining deep learning and object tracking.
  - Developed a Urban Lane Segmentation Pipeline using Data Augmentation for Domain Adaption and EL-GAN.

- Adapted the Frenet Optimal Trajectory Planner for a full scale Ackermann drive vehicle both in real world and simulator.

TEACHING EXPERIENCE AND SERVICES

- **Teaching Assistant, CS60077: Reinforcement Learning, IIT Kharagpur** *August 2021 - present*
  Organized the course on Reinforcement Learning, which included contributing to the lecture content, setting up homework, designing tests and quizzes and evaluating students' performance.
- **The International Summer School on Situational Awareness in Cognitive Technologies, 2021** *Speaker*
  Conducted a hands on session with on the contemporary techniques in explainable AI like CAM, GradCAM and RISE.
- **Dynamic Neural Networks meet Computer Vision Workshop at CVPR 2021** *Volunteer and Reviewer*
  Volunteered to help the organizers with the logistics of the workshop, created the workshop poster and reviewed papers.

COURSE PROJECTS

- **Cluster Management System** - Distributed Systems Project *March - April 2021*
  Research areas and tools: **Operating Systems, Computer Networks, Consistency Models**
  Developed a fault tolerant cluster management system. It had two master nodes (active-passive with sequential consistency), several compute nodes and a file server and could handle failure of one master and any number of compute nodes.
- **Accelerating Graph Algorithms on GPU** - CUDA Programming Project *March - April 2021*
  Research areas and tools: **CUDA Programming, GPU, Compilers**
  Accelerated BFS, Dijkstra and Floyd Warshall by extensive parallelization of these algorithms. Showed significant speedup on a dataset consisting of large graphs taken from SNAP or generated using Mersenne Twister pseudorandom number generator.
- **Just Another Rather Very Intelligent Chatbot** - Software Engineering Project *March 2019*
  Research areas and tools: **Deep Learning, Language models, Android Development, DBMS**
  Developed an automated conversational application that detects and analyses the emotions of the users. Used Bayes' classifier for classifying emotions and seq2seq models for generating replies conditioned on the emotion.

COMPETITIONS

- **Bosch Traffic Sign Detection,** *Inter IIT Tech Meet - **Bronze*** *2021*
  Developed a web based tool to add real life data augmentations on a traffic sign dataset, train classifiers and debug them. The tool helps the users to visualize the mistakes of the model and debug the model.
- **Intelligent Ground Vehicles Competition,** *Oakland University, Michigan - **Runners Up*** *2018 and 2019*
  Developed an autonomous ground vehicle for a constraint environment using simple Computer Vision and Image Processing techniques for Lane navigation and Obstacle avoidance, TEB local planner for planning. First team to qualify.

ACHIEVEMENTS

- **IUSSTF - Viterbi Scholarship** - *Only around ten students selected from India* *2020*
- **Class of 1970 Alumni (US) Association Prize** for **best student in order of merit** among third year B.Tech.(Hons.)/B.Arch.(Hons.) courses securing highest CGPA *2019*
- **IIT Kharagpur Alumni (California Chapter) Award** for being the **best student in order of merit** among third year B.Tech.(Hons.)/B.Arch.(Hons.) courses securing highest CGPA *2019*
- **Student Par Excellence Award by Computer Science & Engineering Dept** *2018*
- **KVPY Fellowship Award** - *Among the top 0.6% of the applicants* *2017*

RELEVANT COURSEWORK AND SKILLS

- **Courses** - Linear Algebra, Calculus, Probability and Statistics, Convex Optimization, Machine Learning, Deep Learning, Reinforcement Learning, Operating Systems, CUDA Programming, Distributed Systems, Networks
- **Relevant libraries and Frameworks** - pytorch, Tensorflow, ScikitLearn, opencv, ROS, OpenAI-gym, Mujoco, Carla, gazebo, git, UI/UX, POSIX