

# SOUMYA RANI SAMINENI

✉ [ssamine4@asu.edu](mailto:ssamine4@asu.edu) | ☎ +1 623-265-4001 | 🌐 [soumyarani.github.io](https://soumyarani.github.io) | 🔗 [soumya-rani](https://soumya-rani.github.io)

## SUMMARY

Second year PhD student at Arizona State University with 3+ years of professional experience, including over 2 years of research experience in Artificial Intelligence (AI), Reinforcement Learning (RL), Large Language Models (LLMs), Machine Learning (ML) and Robotics.

## EDUCATION

**Arizona State University, Tempe** 2024-2029  
Ph.D., *Computer Science and Engineering* CGPA: **4.0/4.0**  
Advisors: [Prof. Subbarao Kambhampati](#)

**Indian Institute of Science, Bangalore** 2019 – 2021  
Master of Technology, *Computer Science and Engineering*. CGPA: **8.5/10**  
Advisors: [Prof. Shalabh Bhatnagar](#) & [Prof. Shishir Kolathaya](#)  
🏆 Awarded A+ grade for M. Tech thesis.

**National Institute of Technology, Warangal** 2012 – 2016  
Bachelor of Technology, *Civil Engineering* CGPA: **8.55/10**

## RESEARCH & WORK EXPERIENCE

**Graduate Researcher**, School of Computing and Augmented Intelligence, ASU Aug'24 - Present

- **Associated Lab:** [Yochan](#) under [Prof. Subbarao Kambhampati](#).
- Topic: Reinforcement Learning, Reasoning & Planning abilities of LLMs.
- Projects: Post training LLMs with RL, Inference Time Scaling, Agentic LLMs & Latent RL for LLMs.

**Machine Learning Research Engineer**, Quantiphi Analytics, India July'23 - Aug'24

- Developed an **RL solution** for flexible job shop scheduling using graph attention networks.
- Co-authored a patent on [System and method for intelligent scheduling of manufacturing jobs](#)
- Tools Used: [TensorFlow](#), [Python](#), [Plotly](#), [Dash](#), [Google Cloud Platform](#)

**Research Fellow**, Microsoft Research, India Feb'22 – July'22

- **Mentors:** [Tanuja Ganu](#) & [Akshay Nambi](#), **Project:** [Vasudha](#)
- Contributed to **RL Framework** in Decision Management Platform for renewable energy monetization.
- Optimized for use cases like carbon arbitrage, demand matching, bidding and profit maximisation.
- Tools Used: [Python](#), [Scikit](#), [Pandas](#), [Data Science](#), [PyTorch](#), [MLOps \(Microsoft Azure\)](#).

**Solution Leader**, Brane Enterprises, India July'21 – Feb'22

- Developed a controller for quadrupedal locomotion by designing lightweight C++ libraries optimized for microcontroller deployment inspired from MIT Cheetah's [pattern modulation & impedance control](#).
- Implemented communication protocols such as CAN and UART for efficient data exchange.
- Improved controller latency to microseconds compared to milliseconds reported in the original paper.
- Tools Used: [C++ Design Patterns](#), [YONO](#), [ROS](#), [OpenCV](#), [Docker](#) and [TensorFlow Serving](#).

**Graduate Researcher**, Computer Science & Automation, IISc Bangalore March'20 – July'21

- **Associated Labs:** [Stochastic Systems Lab](#), Prof. Shalabh & [Stochastic Robotics Lab](#), Prof. Shishir.
- Developed **DeMo RL Algorithm**, achieving a 30% improvement in sample efficiency in simulations.
- Demonstrated a twofold performance gain in real-world experiments on 2R Link and quadruped.
- Accepted at ICRA 2022, NeurIPS Deep RL & Offline RL Workshops 2021.
- Tools Used: [Python](#), [TensorFlow](#), [OpenAI Gym](#), [RL Libraries: RLlib & Stable Baselines](#).

**Assistant Executive Engineer**, Government of Telangana, India June'18 – July'19

- Civil Engineer at Inspection & Quality Control, Roads & Buildings Department.
- Directed progress of major infrastructure projects, with high quality standards across half of the State.

## RESEARCH PROJECTS

---

### Analytical study on LLM MDP, Post Training of LLMs with GRPO & SFT *Under Review at ICLR'26*

- Analysed key assumptions in RLVR based post training of LLMs and its similarity to Iterative SFT.
- Post-trained **Qwen 2.5**, **LLaMA 3.2**, and **DeepSeek R1** using **GRPO** and **Iterative SFT** on **GSM8K**.
- Identified length increase as a post training effect of structural assumptions in LLM MDPs.
- Tools Used: [PyTorch](#), [Verl](#), [vLLM](#), [Amazon AWS](#), [Hugging Face Transformers](#).

### Efficiency of VLM generated rewards in Inverse Reinforcement Learning *Aug'24 - Jan'25*

- Developed an Inverse RL Framework to compare VLM generated rewards with ground truth.
- Demonstrated 70% efficiency of VLMs in discriminating the RL policy and expert demonstration.
- Tools Used: [TensorFlow](#), [Python](#), [GPT-4V](#), [Clip](#), [Meta world](#).

### Inference Time Scaling of LLMs using Rollouts *In progress*

- **Mentors:** [Prof. Bertsekas](#)
- In contrast to Tree of Thoughts, used MCTS Rollouts at every time step in inference.
- Implemented Rollouts with LLMs as base policy and a verifier at inference stage of LLMs.
- Tools Used: [Python](#), [PyTorch](#), [Dataset: Game of 24](#).

### Agentic LLMs approach for solving Grid world problems *In Progress*

- Developed Multi Agent LLM Framework with agents iteratively processing clues & refining solutions.
- Tools Used: [Python](#), [PyTorch](#), [Dataset: Logic based Grid Puzzles](#).

## PUBLICATIONS & PATENTS

---

1. RL in Name Only? Analysing the Structural Assumptions in RL post-training for LLMs  
**Samineni S.R**, Kalwar D, Valmeekam K, Stechly K, Kambhampati S. [arXiv](#)  
NeurIPS 2025, LAW: Bridging Language, Agent, and World Models for Reasoning and Planning workshop.
2. Local Coherence or Global Validity? Investigating RLVR Traces in Math Domains.  
**Samineni S.R**, Kalwar D, Gangal V, Siddhant Bhambri S, Kambhampati S. [arXiv](#)  
NeurIPS 2025, 5th Workshop on Mathematical Reasoning and AI.
3. Stop Anthropomorphizing Intermediate Tokens as Reasoning/Thinking Traces!  
Kambhampati S, Stechly K, Valmeekam.K, Saldyt L, Bhambri S, Palod V, Gundawar A, **Samineni S.R**, Kalwar D, Biswas U. [arXiv](#)  
NeurIPS 2025, Workshop on CogInterp: Interpreting Cognition in Deep Learning Models
4. **US Patent 2024:** System and Method for Intelligent Scheduling of Manufacturing Jobs  
Dagnachew Birru, Anirudh Deodhar, Achint Chaudhary, **Soumya Rani Samineni**. [US20240319718A1](#)
5. Dynamic Mirror Descent based Model Predictive Control for Accelerating Robot Learning.  
**Samineni, S.R.**, Mishra U, Goel P, Kunjeti C, Lodha H, Singh A, Sagi A, Bhatnagar S, Kolathaya S.  
International Conference on Robotics and Automation (**ICRA**) 2022. [DOI](#) [website](#)

## CERTIFICATIONS & SHORT-TERM ACADEMICS

---

- |      |  |
|------|--|
| 2023 | <b>Applied Data Science</b> with Python specialisation certificate from University of Michigan |
| 2021 | <b>TensorFlow Developer</b> certificate from DeepLearning.AI                                   |
| 2020 | Winter School on <b>Hybrid Cloud</b> by IBM Research India & IISc Bangalore                    |
| 2019 | Summer School on <b>Machine Learning and Computer Vision</b> , IIIT Hyderabad                  |

## ACHIEVEMENTS

---

- |      |   |
|------|---|
| 2021 | 🏆 Secured a <b>Student Research Grant</b> from Robert Bosch Centre for Cyber Physical Systems, at IISc in recognition of the excellence demonstrated in my M. Tech project. |
| 2018 | 🏆 Secured a rank of <b>136</b> among <b>107,893</b> aspirants in GATE, Computer Science & Engineering.  |
| 2012 | 🏆 Secured an All-India Rank of <b>6062</b> in <b>IIT JEE'12</b> with a <b>98.67</b> percentile.   |