Harsh Patel

(+1) 858-247-9926 | h1patel@ucsd.edu | harshp1802.github.io | harshp1802 | harshp1802 | Harsh Patel

Education

University of California San Diego — *California, USA*

M.S. Computer Science and Engineering (GPA: 4/4) [***Transcript***] — *2022 - Mar 2024*

Indian Institute of Technology Gandhinagar — *Gandhinagar, India*

B.Tech. with Honours in Computer Science and Engineering (CPI: 9.01/10) [***Transcript***] — *2018 - 2022*

Technical Knowledge

Programming Languages: Python, C, C++, Go, MATLAB, SQL, Verilog, JavaScript, HTML, CSS, SLURM Scripting

Tools: Git, Airflow, PyTorch, Tensorflow, Spark, Docker, Kubernetes, ROS, OpenCV, GCP, LaTeX, Huggingface, Databricks

Relevant Experiences

TikTok U.S. Data Security Inc. — *Mountain View, California*

Machine Learning Engineer | E-Commerce Risk Control & Security — *July 2024 - Present*

- Building advanced buyer segmentation models with value and risk scoring to enable early and accurate fraud detection, significantly reducing platform losses. Balancing platform integrity with a seamless experience for high-value users to drive trust and growth.

Lucid Motors — *Newark, California*

Sr. Data Scientist | Manager: Dr. Anuradha Kodali — *April 2024 - June 2024*

- Joined the team and quickly made remarkable contributions by *leading the adoption of Generative AI* for automating customer care data analysis. This initiative reduced manual workforce effort by **90%**, optimized metrics like MTTR (Mean Time to Resolution), and provided valuable insights from customer feedback, resulting in potential significant process improvements.
- Enabled the transition from rule-based to *ML-driven anomaly detection* for vehicle fleet security, incorporating Responsible AI principles and guardrails to prevent catastrophic failures. This enhancement reduced false positives by **50%**, simplifying the validation of cybersecurity threats. Proposed and implemented feature importance techniques to enhance the explainability of these operations.

Nokia Bell Labs — *Murray Hill, New Jersey*

Autonomous Systems Research Intern | Mentor: Mrs. Buvaneswari Ramanan — *June 2023 - August 2023*

- Leveraged large language models (LLMs) to enhance Nokia's patent-pending, proprietary MLOps platform for the end-to-end operations of ML-based use cases. [*Manuscript under review at IEEE Transactions on Artificial Intelligence*]
- Developed innovative task-specific knowledge enrichment strategies, involving automatic retrieval using Langchain and vectorstores, to improve the performance of LLMs in complicated code generation tasks.

Nanyang Technological University — *Singapore*

Research Intern | Mentor: Prof. Erik Cambria — *May. 2021 - Jul. 2021*

- Developed a *deep multitask learning* framework that enhances the performance of Negation Scope Detection using POS tagging as an auxiliary task. Used *transformers* and *neural tensor fusions* to leverage the inter-task correlations. Achieved **5%** improvement over the baseline models.

Publications

Automating Code Adaptation for MLOps – A Benchmarking Study on LLMs [***arxiv***]

Accurate and Scalable Gaussian Processes for Fine-grained Air Quality Inference, **AAAI (2022)**[***doi***]

Enhancing Negation Scope Detection using Multitask Learning, **ICDMW (2021)**[***doi***]

Program Synthesis: Does Feedback Help?, **ACM CoDS-COMAD (2022)**[***doi***][***poster***]

Assessing the interplay between travel patterns and SARS-CoV-2 outbreak in realistic urban setting, **SpringerOpen (2021)** [***doi***]

Projects

Advancing Model-Agnostic Text Dataset Distillation | **GitHub Repo** | **Arxiv** — *UC San Diego*

Mentor: Prof. Jingbo Shang ▷ NLP | Deep Learning | Efficient Learning — *Apr. 2023 - June 2023*

- Developed novel text-dataset distillation techniques that demonstrate strong cross-architecture generalization capability, enhancing efficiency and performance in natural language processing tasks. Achieved a remarkable 95% distillation ratio with just 30 samples.

Robust, Scalable, & Fault-Tolerant Networked File Storage Service — *UC San Diego*

Mentor: Prof. George Porter ▷ Networked Systems Design | Distributed Cloud Computing — *Jan. 2023 - Mar. 2023*

- Developed a cloud-based file storage system, leveraging **gRPC** for streamlined communication, **Consistent Hashing** for efficient load balancing, and the **RAFT** consensus algorithm to ensure fault-tolerance and consistency.

Accurate and Scalable Gaussian Processes for Fine-grained Air Quality Inference | **Repo** | **GP-Viz** — *IIT Gandhinagar*

Mentor: Prof. Nipun Batra ▷ Applied & Data-Driven Machine Learning | Bayesian Modeling — *Aug. 2021 - May 2022*

- Implemented stationary & non-stationary probabilistic **Gaussian Process** models for urban air quality estimation - as spatio-temporal regression. Our **uncertainty-aware** approach outperformed conventional baselines on standard air quality datasets.

Exploring Constrained Reinforcement Learning for Autonomous Driving | **GitHub Repo** — *IIT Gandhinagar*

Mentor: Prof. Nipun Batra ▷ Reinforcement Learning | Policy Optimization & Evaluation — *Jan. 2021 - May. 2021*

Honors and Community Engagement

2024 **Teaching Assistant - Unsupervised Learning**, UC San Diego

2021 **Teaching Assistant - Machine Learning and Natural Language Processing courses**, IITGN

2021 **Pull Request (PR) accepted for PyMC**, Open source GitHub package for Bayesian statistical modeling

2019 **Leadership - Core Committee Member**, Amalthea - Tech Summit, Jashn - Cultural Fest at IITGN

2018 **Joint Entrance Exam (JEE)**, All-India-Rank 143 out of 1.13 million candidates