Naman Jain

Website: naman-ntc.github.io ◇ Email: naman_jain@berkeley.edu

EDUCATION

University of California, Berkeley *2022 - present*
Ph.D. in Computer Science

Indian Institute of Technology, Bombay *2016 - 2020*
B.Tech (Honors) in Computer Science and Engineering

WORK AND INTERNSHIPS

Microsoft Research (MSR), India *Aug '20 - Aug '22*
Research Fellow

University of Maryland, College Park *June '19 - Aug '19*
Research Intern

PUBLICATIONS

5. StaticFixer: From Static Analysis to Static Repair
Naman Jain, Shubham Gandhi, Atharv Sonwane, Aditya Kanade, Nagarajan Natarajan, Suresh Parthasarathy, Sriram Rajamani and Rahul Sharma
Submitted to PLDI 2023, Orlando, Florida [preprint]
4. Jigsaw: Large Language Models meet Program Synthesis
Naman Jain, Skanda Vaidyanath, Arun Iyer, Nagarajan Natarajan, Suresh Parthasarathy, Sriram Rajamani and Rahul Sharma
Proceedings of ICSE 2022, Pittsburgh, Pennsylvania [paper]
3. Learning Accurate Decision Trees with Bandit Feedback via Quantized Gradient Descent
Ajaykrishna Karthikeyan*, **Naman Jain***, Nagarajan Natarajan, and Prateek Jain
Proceedings of TMLR 2022 [paper]
2. What's in a Name? Are BERT Named Entity Representations just as Good for any other Name?
Sriram Balasubramanian*, **Naman Jain***, Gaurav Jindal, Abhijeet Awasthi and Sunita Sarawagi
Workshop Proceedings of ACL 2020, Virtual Conference [paper] [supplement]
1. On the Robustness of Human Pose Estimation
Naman Jain*, Sahil Shah*, Abhishek Sharma and Arjun Jain
Workshop Proceedings of CVPR 2019, Long Beach, California [paper] [supplement]

* joint first authors

RESEARCH PROJECTS

Interactive MultiModal Data Wrangler Sep'22 - Present
University of California, Berkeley *Student Researcher*

· Building an data-wrangling ingesting intent via both natural language and demonstrations interactively
· Using query-guided back-translation to perform unsupervised ranking of generated program snippets
· Performing query-guided summarization as interaction to ground system's prediction and alternatives

Static Repair of Dataflow Vulnerabilities Sep'21 - Nov'22
Microsoft Research, India *Research Fellow*

· Proposed static-analysis-witnessing for mining sanitizers and guards in public source code (using static-analysis) and collected examples of safe programs (and witnesses) along with their *unsafe* versions
· Built a DSL that defines repair strategies operating on unsafe programs and dataflow information
· Developed a programming-by-examples approach that learns repair strategies in this DSL deductively

- Repaired cross-site-scripting and unvalidated-call vulnerabilities in over 1000 repositories (90% recall)

Jigsaw – Combining Language Models with Program Synthsis Dec'20 - Sep'21
Microsoft Research, India *Research Fellow*

- Proposed architecture for augmenting black-box models with program analysis and synthesis-based post-processing block that provides correctness guarantees and also allows learning from user feedback
- Instantiated the architecture for Pandas with GPT-3/Codex and developed multi-modal synthesis tool
- Designed transformations for different error classes using enumerative search and Prose AST rule edits
- Released two datasets for benchmarking performance of system allowing offline and temporal evaluation

DGT – Versatile Decision Tree Learning Sep'20 - Oct'21
Microsoft Research, India *Research Fellow*

- Developed alternative conceptualization of decision tree problem that allows end-to-end gradient based learning while still performing competitive to SOTA tree methods and outperforming in bandit setting
- Combined straight through estimators (quantization), additive AND gates formulation, and linear over-parameterization to better condition the gradients and reach desirable performance with learnt trees

Robustness in Natural Language Processing August 2019 - June 2020
IIT Bombay (with Prof. Sunita Sarawagi) *Bachelor's Thesis*

- Studied and designed training algorithms robust to synonym and named entity replacement attacks on tasks including sentiment analysis, grammar correction, coreference resolution and question answering
- Demonstrated non-robustness of BERT based models on various tasks and analysed potential causes
- Developed simple yet effective replacement-ensembling algorithm to defend against named entity attacks

Adversarial Examples in Human Pose Estimation Aug 2018 - Dec 2018
IIT Bombay (with Prof. Arjun Jain) *Research Assistant*

- Released implementation of six SOTA pose estimation systems along with pretrained models
- Performed image agnostic & dependent adversarial attacks evaluating crucial design choices such as direct regression vs heatmap, imagenet pretraining, using compositional human body structure, etc.

Interacting Humans Video Prediction June 2019 - Dec 2020
University of Maryland, College Park (with Prof. Abhinav Shrivastava) *Research Intern*

- Worked on multi-human video prediction via pose forecasting and subsequent frame generation
- Introduced novel framework for incorporating multi-person context via additive conditional batchnorm

Virtual Try On May 2020 - August 2020
Uplara AI, Palo Alto *Remote Internship*

- Implemented image based virtual try on system using 2D warping and 3D mesh alignment methods
- Trained networks to warp cloth images to match human shape and pose using FLOW and TPS mappings

SELECTED SOFTWARE AND OPEN SOURCE

- Jigsaw server and Jupyter extension - Flask, SocketIO, React
- Human Pose Estimation - PyTorch (stats - ★>350, ᛘ>70)
- `torch.randint` in PyTorch - open source contribution

SCHOLASTIC ACHIEVEMENTS

- Awarded Undergraduate Research Award (**URA**) for Autumn 2018 (2 out of 121 students)
- Secured **All India Rank 36** in JEE Advanced 2016 among **0.15 million** candidates
- Awarded Gold Medal for being among top **35** students in Indian National Physics Olympiad (**INPhO**)
- Among the **top 300** in Indian National Chemistry and Astronomy Olympiads (**INChO** & **INAO**)
- Awarded KVPY Fellowship and NTSE Scholarship by Govt. of India

RELEVANT COURSES

- **Computer Science -** Automata Theory, Compilers, Parallelizing Compilers, Advanced Machine Learning, Learning Agents (RL), Computer Vision, Data Structures, Algorithms, Operating System
- **Mathematics -** Calculus, Linear Algebra, Differential Equations, Numerical Analysis, Discrete Maths

TECHNICAL STRENGTHS

Strong	Python, C, C++, C#, Racket, Prolog, Bash
Web	Flask, ASP.NET, Guincorn, Nginx, IIS, WebSockets, React, JSP
Tools	Git, LaTeX, OpenGL, OpenMP, Gnuplot, Doxygen

TEACHING & MENTORSHIP

- Teaching Assistant for the course Automatic Speech Recognition under Prof. Preethi Jyothi
- Teaching Assistant for the course Quantum Mechanics under Prof. Aftab Alam
- Mentored 6 students for project on statistical face recognition organized by WnCC, IIT Bombay

REFERENCES

Prof. Sunita Sarawagi
Indian Institute of Technology, Bombay
webpage ◇ email

Dr. Nagarajan Natarajan
Microsoft Research, India
webpage ◇ email

Dr. Sriram Rajamani
Microsoft Research, India
webpage ◇ email

Dr. Rahul Sharma
Microsoft Research, India
webpage ◇ email

Prof. Koushik Sen
University of California, Berkeley
webpage ◇ email

Prof. Aditya Parameswaran
University of California, Berkeley
webpage ◇ email