JIAN VORA

jianv@cs.stanford.edu | +1 650-334-9262 | linkedin.com/in/jianvora | github.com/jianvora

Education

Stanford University — Stanford, CA
M.S. in Computer Science (AI Specialization), GPA: 4.04/4.0 — *Sept 2021-Present*
- Distinction in Research (Advisor: Jure Leskovec)

Indian Institute of Technology, Bombay — Mumbai, India
B.Tech. in Electrical Engineering, Major GPA : 9.77/10.0, AI Minor GPA: 10.0/10.0 — *Jul 2017 - May 2021*
- Dual Minor in Computer Science and Machine Intelligence & Data Science (C-MInDS)

Academic Research Experience

Language Model Agents as Research Assistants (PAPER) — Stanford, CA
Guides: Prof. Jure Leskovec, Prof. Percy Liang — *Jan 2023 – Present*
- **Problem:** Can LLM agents do science: design and test experiments given some tools, prove or disprove a given hypothesis, reflect upon the observations, update its beliefs to run the next set of experiments, and how do we evaluate its performance?
- **AI Research Agent:** The proposed model (with reflection and fact check modules) based on GPT-4 can improve performance on CIFAR-10, code up Kaggle challenges from scratch, perform literature review + integrate its findings with the research pipeline. Released a benchmark to evaluate these agents on performance and efficiency.
- **BIO Research Agent:** Focusing on genome-wide CRISPR screens and using LLMs to optimally design batch-wise experiments. This simple method outperforms a lot of machine learning baselines, while at the same time providing interpretable plans and robustness to noisy readouts from experiments. Collaborating with biologists.

Policy Surgeon: Fixing LLMs as Policies for Robot Learning — Stanford, CA
Guides: Prof. Jiajun Wu, Prof. Fei Fei Li — *Sept 2023 – Present*
- **Problem:** Asking LLMs to output robot plans in a DSL typically fails due to hallucination, and the lack of semantic and logical constraints of the planning domain. Can we inject this information in the LLM to get better policies?
- Exploring adding control vectors to LLMs to use knowledge from the physical world (through some external knowledge base) to fix errors after characterizing them. Experimenting on VirtualHome and Behaviour environments.

Inference Time Robustness for GNNs (PAPER) — Stanford, CA
Guide: None — *June 2023 – Oct 2023*
- **Problem:** For images, simple test-time transformations help in increasing robustness to adversarial attacks such as quantization, and randomization. Does something similar hold for graphs?
- Showed that GNN predictions on a k-hop egonet around the node to classified significantly boost adversarial accuracies across many base models and attack types, indicating such a simple thing can filter a lot of adversarial edges

Scoring Black-Box Models for Adversarial Robustness (PAPER) — Stanford, CA
Guide: None — *June 2022 – Sept 2022*
- **Problem:** Adversarial robustness of a model is measured in a roundabout fashion – by first constructing adversarial inputs (most methods need white-box models) and then measuring the model performance on the constructed adversarial inputs
- Showed that adversarially robust models have sparser LIME weights and consequentially sharper explanations
- Use the sparsity of LIME weights as a scoring function, showed experiments on a variety of models from `robustbench`

Non-Myopic Recommendations using Reinforcement Learning — Stanford, CA
Stanford Design School, Guide: Prof. Soh Kim, Dr. Chandrayee Basu — *Jan 2022 – June 2022*
- **Problem:** Build a recommender system to suggest people food recipes that takes into account the individual's long-term health and are not just optimized to maximize the reward modeled as the amount of user interaction with the platform
- Posed the objective as a multi-task objective by trying to improve the health of the worst-case users (CVaR)
- Worked with `food.com` data, performed user behavior modeling, and deployed the model in the Stanford dining halls

PAC Mode Estimation and Best Arm Identification (PAPER) — IIT Bombay
Guide: Prof. Shivaram Kalyankrishnan — *Jan 2021 – Sept 2021*
- Adapted the prior-posterior ratio (PPR) as an α confidence sequence to determine the stopping rule for mode estimation and best arm identification (with an appropriate sampling rule) for Bernoulli bandits
- Proved the asymptotic optimality of the method for 2 arms, proposed two extensions for K arms, and showed that the one-vs-one approach provably outperforms the one-vs-rest approach (which many BAI methods use)
- Showed applications to election polling and verification of smart contracts in blockchains

Tractable Inference for Multimodal Generative Modeling (PAPER) — University of California Los Angeles, CA
Research Intern — Guides: Prof. Guy Van den Broeck, Prof. Antonio Vergari — *May 2020 – Oct 2020*
- **Problem:** Learn a joint distribution over multimodal data while allowing for queries such as marginalization, conditional sampling, exact likelihood evaluation, and a composition of any of these queries

- Trained probabilistic circuits to allow tractable inference on a joint distribution over multimodal data in the latent space; Implemented regularized encoders for each modality for a smooth latent space
- Outperformed multimodal VAEs (MVAE, MMVAE) on both qualitative and quantitative measures

Low-Rank Probability Density Tensor Recovery from Marginals (PAPER) IIT Bombay
Guide: Prof. Ajit Rajwade *Sept 2020 – Mar 2021*

- Showed that random projections of data are distributed as the radon transform of the original distribution and the original distribution is recoverable given that the projection is taken along many directions
- Modeled the joint PMF to be a low-rank tensor allowing for a canonical polyadic decomposition (CPD)
- Proposed an algorithm to recover the tensor modes from 1-D densities learned on random projections of data

Efficient Learning of Log-Concave Mixtures (REPORT) IIT Bombay
Guide: Prof. Vivek Borkar *Sept 2020 – Jan 2021*

- Proved that random projections of data drawn from a mixture of log-concave densities are provably distributed as a gaussian mixture in the subspace with the mixture weights being the same with a very high probability.
- Higher dimensional moments can be recovered by invoking the Johnson–Lindenstrauss Lemma

Blind Calibration of Perturbations in Compressed Sensing (PAPER) IIT Bombay
Guide: Prof. Ajit Rajwade *Jan 2020 – Jun 2020*

- Calibrated frequency offsets (gradient delays) and sensor gains (person in motion) in a linear compressive measurement framework which occur in MRI signal acquisition for faithful signal reconstruction using a single snapshot
- Proposed an alternating minimization algorithm and proved recovery guarantees and identifiability conditions

Industry Research Experience

Leveraging Sequential Data for Ranking in Recommendations (PAPER) San Francisco, CA
Applied Scientist II Intern, Twitch Interaction Inc. — Guide: Nikita Mishra, Saad Ali *June 2023 – Sept 2023*

- Incorporated a novel way to incorporate pageview history and minutes watched in the ranking pipeline.
- Outperformed the deep ranker model in production by 9% increase in NDCG values in offline experiments. Wrote a paper to report the finding which was published in the AMLC workshop on Personalization and Ranking.

Unified Pre-training for Speech and Text Modalities Santa Clara, CA
Applied Scientist Intern, Amazon Web Services (AWS) AI Labs — Guide: Prof. Katrin Kirchhoff *June 2022 – Sept 2022*

- Designed and implemented a multimodal self-supervised pre-training strategy for speech and text modalities. Experimented with various *coherence* losses for a small amount of paired data, fusion stages for both the modalities
- Incorporated adapters in the shared model to allow for text-only adaptation, dialog context incorporation for ASR
- Outperformed the multimodal models (SLAM, MAESTRO) on unimodal SUPERB (speech) and GLUE (text) benchmarks trained on comparable data and a similar number of parameters

Spatio-Temporal Action Detection and Classification (REPORT) Tokyo, Japan
Research Intern, Hitachi Central Research Lab — Guide: Dr. Martin Klinkigt *May 2019 – July 2019*

- Participated in the TRECVid'19 challenge which involved performing action detection and classification in videos
- Proposed a stage-wise architecture of object detection followed by tracking and activity classification, incorporated a state machine to get better temporal alignments of actions predicted by the deep learning model
- Ranked 3rd in the final leaderboard beating the previous edition's winners; 1st in temporal alignment sub-task

Publications

1. **AI Guided CRISPR Perturbation Experiments**
   Yusuf Roohani, Jian Vora, Qian Huang, Percy Liang, Jure Leskovec
   Under Preparation

2. **Benchmarking Large Language Models as AI Research Agents [Paper]**
   Qian Huang, Jian Vora, Percy Liang, Jure Leskovec
   accepted at *NeurIPS Foundation Model For Decision-Making Workshop, 2023*, under review at *ICLR 2024*

3. **GNN Predictions on k-Hop Egonets Boosts Adversarial Robustness [Paper]**
   Jian Vora
   accepted at the *2nd New Frontiers in Graph Learning (GLFrontiers) workshop, NeurIPS, 2023.*

4. **Scoring Black-Box Models for Adversarial Robustness [Paper]**
   Jian Vora, Pranay Reddy Samala
   accepted at the *2nd Workshop on New Frontiers in Adversarial Machine Learning, ICML, 2022.*

5. **Sequential Consumption-Aware Ranking Model for Recommendations at Twitch [Paper]**
   Jian Vora, Edgar Chen, Nikita Mishra, Saad Ali
   Accepted at *AMLC Workshop on Personalization and Ranking, 2023.*

6. **Plug&Play Multimodal Generative Model Allowing Tractable Inference [Paper]**
   Jian Vora, Isabel Valera, Guy Van den Broeck, Antonio Vergari
   Preprint, 2021.

7. **PAC Mode Estimation using PPR Martingale Confidence Sequences [Paper]**
   S. Jain*, R. Shah*, Jian Vora†, S. Gupta†, D. Mehta†, I. Nair†, S. Khyalia, S. Das, V. Riberio, S. Kalyankrishnan
   Accepted at *The 25th International Conference on Artificial Intelligence and Statistics (AISTATS) 2022.*

8. **Recovery of Joint Probability Distribution from One-Way Marginals: Low-Rank Tensors and Random Projections [Paper]**
   Jian Vora, Karthik S. Gurumoorthy, Ajit Rajwade
   Accepted at *2021 IEEE Statistical Signal Processing (SSP) (SSP 2021).*

9. **Compressive Signal Recovery Under Sensing Matrix Errors Combined With Unknown Measurement Gains [Paper]**
   Jian Vora and Ajit Rajwade
   Accepted at *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2021.*

Scholastic Achievements & Awards

- Stood among the nationwise top 1% in National Standard Examination in Physics and Astronomy (NSEP, NSEA)
- Achieved All India Rank 3 in ICSE 2015 out of 0.2 Million applicants
- Selected for undergraduate research internship programs at Caltech, USC, and NTU
- Bagged the silver medal in Homi Bhabha Young Scientist (top 0.04%) search for scientific and research aptitude
- Selected as a Clear Ventures Fellow aiming to serve as a bridge between investors, entrepreneurs, and researchers.

Skills

Languages: Python, C++, Java, VHDL, MATLAB, Bash, HTML/CSS, Embedded C, LaTeX
Tools/Frameworks: Pytorch, TensorFlow, JAX, Keras, Git, Docker, Quartus, Scilab, GDB

Selected Course Projects

Robust Experimental Design under Reward Misspecification (REPORT) CS332, Prof. Emma Brunskill

- Proved certain misspecified reward functions where a linear design can be provably bad. Proposed a soft-max version of experimental design which provides robustness to misspecifications in reward for the case of contextual bandits.

Tractable Cooperative Multi-Agent Reinforcement Learning (REPORT) CS333, Prof. Dorsa Sadigh

- Modelled the joint Q-function in the cooperative MARL setting as a factor graph over action potentials (low-rank tensor)
- This allowed us to learn a joint distribution over actions of all agents while allowing for tractable inference
- The policy was an EBM over the Q-function with probabilistic circuits as the variational family

Improving Inference in Latent Variable Models (REPORT) CS236, Prof. Stefano Ermon

- Improved inference in VAEs by reducing two gaps – approximation gap by using hierarchical VAEs (NVAE) and amortization gap by performing unamortized inference. Showed improvements for image inpainting and denoising as inference applications

A Study on Continuous Optimization Schemes for Machine Learning (REPORT) EE736, Prof. Vivek Borkar

- Literature survey of three papers – convergence in deep learning via over-parameterization, is Q-learning provably efficient compared to model-based methods and coupling of gradient and mirror descent

Conditional StyleGAN for Audio Generative Modeling (REPORT) EE782, Prof. Amit Sethi

- Modified StyleGAN to allow for conditioning and trained on audio spectrograms

Continual Learning for Keyword Spotting and Speaker Identification (REPORT) CS753, Prof. Preethi Jyothi

- Proposed a joint model to perform simultaneous KWS and SID based on an Interspeech 2021 challenge

Stochastic Approximation Algorithms for PCA Self-Project

- Implemented Oja's algorithm, Incremental PCA and Matrix Stochastic Gradient for empirical comparision

Relevant Coursework

Theoretical Machine Learning and Optimization

- CS229M (Theoretical Machine Learning), CS767 (Theoretical Machine Learning), CS769 (Optimization in Machine Learning), EE736 (Introduction to Stochastic Optimization)

Reinforcement Learning and Robotics

- CS332 (Advanced Survey in RL), CS333 (Safe and Interative Robotics), CS223A (Introduction to Robotics), CS748 (Advanced RL), CS747 (Introduction to RL)

Machine Learning and Deep Learning

- CS231N (Computer Vision), CS329T (Trustworthy ML), CS224W (Graph ML), CS236 (Deep Generative Models), EE782 (Advanced Topics in ML), CS753 (Automatic Speech Recognition), CS419 (Introducing to Machine Learning)

Teaching and Mentoring Experience

CS221: Artificial Intelligence (WEBPAGE) Jehangir Amjad

CS234: Reinforcement Learning (WEBPAGE) Prof. Emma Brunskill

CS224V: Conversational Virtual Assistants with Deep Learning (WEBPAGE) Prof. Monica Lam

CS236G: Generative Adversarial Networks (WEBPAGE) Sharon Zhou

MA111: Vector Calculus (Head TA) Prof. Rekha Santhanam

Institute Student Mentor & Department Academic Mentor SMP, IIT Bombay

- Responsible for co-ordinating with the department as a senior DAMP mentor, organized a lecture series on mental health awareness and stress management, mentoring students to clear off their academic backlogs
- Mentored 8 students under Summer of Science by providing them necessary guidance in machine learning; spoke at a 15 day long Machine Learning bootcamp with 250+ participants and developed assignments and projects

Leadership Positions

Manager, Electronics and Robotics Club IIT Bombay

- Led a team of 17 members and managed a budget of $0.3M for organizing student events, maintenance and procurement of equipment for Tinkerers Lab (a student-run 24 x 7 technical lab equipped with electronics, 3D printers, laser cutters)
- Initiated and maintained a reading group for discussing and implementing recent research in the field on computer vision

Editor, Times NIE Student Edition Mumbai

- Named budding journalist of the state by Times NIE, got a good fortune to interview a bunch of luminaries including leading economist Dr. Raghuram Rajan, educator and author Sudha Murthy among many others.
- Represented the state in national level debates, mock parliaments, and MUNs