

Vihari Piratla

University of Cambridge, Cambridge

@ vp421@cam.ac.uk

🌐 <https://vihari.github.io>

🎓 [Google Scholar](#)

Research Focus

I am interested in the reliable deployment of frontier AI models. My current research focus is on methods for model debugging and revision.

Research Areas: Generalization, Reliable and Robust ML systems, Trustworthy ML.

Education

- **Indian Institute of Technology, Bombay** June 2017 - July 2022
M.Tech. + Ph.D. Dual Degree, Department of Computer Science GPA: 9.75/10
Advisors: [Prof. Sunita Sarawagi](#) and [Prof. Soumen Chakrabarti](#)
Thesis: *Robustness, Evaluation and Adaptation of Machine Learning Models in the Wild* [PDF]
- **Indian Institute of Technology, Mandi** 2010 - 2014
B.Tech. Department of Computer Science

Professional Experience

- Research Associate Aug'22 - Present
Mentor: [Dr Adrian Weller](#) ML Group, University of Cambridge
Trustworthy ML.
○ By-fellow, [Churchill College](#) Feb'23-Present
○ Postdoctoral Research Associate, [Trinity Hall College](#) Sept'22-Jan'23
- Research Intern Aug'19 - Nov '19
Mentor: [Dr Praneeth Netrapalli](#) Microsoft Research, India
Worked on algorithms that enable efficient transfer and generalization to new domains.
☆ Work done during the internship was published at ICML 2020.
- Project Staff Oct'16 - July'17
Mentor: [Prof. Sunita Sarawagi](#) IIT Bombay
Developed a smart complaint management system.
- Full-stack Software Developer June'14 - Feb'16
Mentor: [Dr Sudheendra Hangal](#) Amuse Labs, Stanford University Libraries
Contributed the following language processing features to an open-source digital archival project called [ePADD](#).
(1) A fine-grained entity recognizer robust to domain shifts using binomial mixture models. (2) Cross document co-referencing and entity linking using context cues.
- Research Intern June'13 - Aug'13
Mentor: [Naoneeth S](#) GE Global Research
Implemented software for distortion correction and 3D registration of Ultrasound images.

Scholastic Achievements

- Best PhD thesis awards from the department (Vashee Award), institution ([Naik & Rastogi Excellence in Ph. D. Thesis](#)), and a national competition ([ACM SIGKDD Doctoral Dissertation In Data Science Award](#)) in 2023.
- One of the sixteen global [Google PhD fellowship](#) recipients in Machine Learning in 2020.
- Selected to receive Prime Minister's Fellowship for Doctoral Research, 2021 (declined for clerical reasons).
- Google and Microsoft travel grant awards for ICLR 2018 and ACL 2019.
- Highest GPA in the Computer Science MTech batch of hundred students in 2019 at IIT Bombay.
- Represented IIT Mandi, in a team of three, at ACM International Collegiate Programming Contest 2012-13 Kharagpur and 2013-14 Kanpur regional.

- Qualified for the Indian National Mathematical Olympiad after clearing Regional Mathematical Olympiad'09 with a state rank of 26.

Publications

i In the field of Machine Learning, conference publications are peer-reviewed and are reputed equally as journal publications. Acceptance rate at a conference is usually between 20-30%.

Preprints / Under Review

- [Estimation of Concept Explanations Should be Uncertainty Aware](#)
V Piratla, J Heo, S Singh, A Weller
 A shorter version of paper was accepted at the *NeurIPS 2023 workshop XAI in Action: Past, Present, and Future Applications*..

Conference Publications

11. [Use Perturbations when Learning from Explanations](#)
 J Heo*, **V Piratla***, M Wicker, A Weller [code][slides]
Neural Information Processing Systems (NeurIPS) 2023.
10. [Certification of Distributional Individual Fairness](#)
 M Wicker, **V Piratia**, A Weller
Neural Information Processing Systems (NeurIPS) 2023.
9. [Human-in-the-loop mixup](#)
 KM Collins, U Bhatt, W Liu, **V Piratla**, I Sucholutsky, B Love, A Weller [code]
Uncertainty in Artificial Intelligence (UAI) 2023.
8. [Focus on the Common Good: Group Distributional Robustness Follows](#)
V Piratla, P Netrapalli, S Sarawagi [code][talk][slides]
International Conference on Learning Representations (ICLR) 2022.
7. [Active Assessment of Prediction Services as Accuracy Surface Over Attribute Combinations](#)
V Piratla, S Chakrabarti, S Sarawagi [code][talk][slides]
Neural Information Processing Systems (NeurIPS) 2021.
 ☆ Recognised as an [AI Game Changer](#) by [NASSCOM](#) in ML Fundamentals category.
6. [Training for the future: A simple gradient interpolation loss to generalize along time](#)
 A Nasery, S Thakur, **V Piratla**, A De, S Sarawagi [code]
Neural Information Processing Systems (NeurIPS) 2021.
5. [Nlp service apis and models for efficient registration of new clients](#)
 S Shah, **V Piratla**, S Chakrabarti, S Sarawagi [code]
Findings of Empirical Methods in Natural Language Processing (EMNLP), 2020.
4. [Efficient Domain Generalization via Common-Specific Low-Rank Decomposition](#)
V Piratla, P Netrapalli, S Sarawagi [code][talk][slides]
International Conference on Machine Learning (ICML) 2020.
3. [Parallel iterative edit models for local sequence transduction](#)
 A Awasthi, S Sarawagi, R Goyal, S Ghosh, **V Piratla**. [code]
Conference on Empirical Methods in Natural Language Processing (EMNLP) 2019.
2. [Topic Sensitive Attention on Generic Corpora Corrects Sense Bias in Pretrained Embeddings](#)
V Piratla, S Sarawagi, S Chakrabarti. [code][talk][slides]
Annual Meeting of the Association for Computational Linguistics (ACL) 2019 (Oral).
1. [Generalizing Across Domains via Cross-Gradient Training](#)
 S Shankar*, **V Piratla***, S Chakrabarti, S Chaudhuri, P Jyothi, S Sarawagi [code][talk][slides]
International Conference on Learning Representations (ICLR) 2018.

Workshop

- [Untapped Potential of Data Augmentation: A Domain Generalization Viewpoint](#)
V Piratla, S Shankar
ICML 2020 Workshop on Uncertainty and Robustness in Deep Learning

Talks

- Are we teaching machines right? Role of supervision in training ML models. [\[slides\]](#)
Presented at IIT Bombay, Google Research India, University of Cambridge in 2023
- Research Challenges when scaling to millions of users through Prediction Service APIs [\[talk\]](#)[\[slides\]](#)
Presented at [Trust ML Rising Star Spotlights Series](#)

Development Experience

[ePADD: Digital Archival Project](#)
Stanford University Libraries

Aug'14 – Feb'16

- ePADD is an open-source project for collecting and processing digital archives.
 - Contributed tens of thousands of lines of code for seamless deployment across various operating systems, browsers, compute hardware and huge archives.
- ☆ The features I contributed continue to be a big part of the project: [browse here](#).

Academic Service/Experience

- **Teaching Assistant or Supervisions:** (2017-)
 - Advanced Machine Learning
 - Digital Image Processing
 - Organization of Web Information
 - Data Interpretation and Analysis
 - Learning with Graphs
 - Parallel Programming Paradigms
 - Data Science
- **Student Volunteer:** NeurIPS 2021, ICML 2020, ACL 2019.
- **Reviewer:** ICLR, NeurIPS, ICML, AAAI, IEEE Transactions on Multimedia.
☆ Top reviewer at NeurIPS 2023.

Technical Skills

- **Programming & Scripting:** Java, Python, C, Shell Scripting, PERL
- **Technologies :** Spring, HTML, CSS, JavaScript, Lucene
- **Programming Libraries:** PyTorch, GPyTorch, Tensorflow, NumPy, Pyro
- **Languages:** English, Telugu, Hindi

Other Activities

- *Athletics.* At IIT Bombay's PG sports events conducted in 2018, I won (or was part of a team that won) **Gold** in the 400m relay, **Silver** in 100m relay, **Bronze** in Volleyball, and finished fourth in the Long Jump. I finished 6th in the 2017 and 2018 PG sports 5km running event.
- *Mountains.* I am an enthusiastic trekker and camper with medium level expertise.
- *Activism/Volunteering:* Basic sanitation, Isha yoga, Rally for Rivers.

References

Dr Adrian Weller	Postdoc Supervisor	Director of Research, University of Cambridge
Prof. Sunita Sarawagi	PhD Supervisor	Professor, IIT Bombay
Prof. Soumen Chakrabarti	PhD Supervisor	Professor, IIT Bombay
Dr Praneeth Netrapalli	Collaborator	Research Scientist, Google Research