




# Vihari Piratla

Research Associate, University of Cambridge  
vp421@cam.ac.uk (  )

## Research Focus

I am interested in studying the performance of Machine Learning systems beyond well-represented training distributions. Towards this objective, I work on ML algorithms' generalization, evaluation, and adaptation aspects on unseen distributions.

**Research Areas:** Reliable, robust and secure ML systems, Trustworthy ML or Responsible AI

## 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](#)  
☆ Google PhD Fellow ☆ Department rank one in the M.Tech. class of >100 students.
- **Indian Institute of Technology, Mandi** 2010 - 2014  
B.Tech. Department of Computer Science

## Professional Experience

- Research Associate Aug'22 - Present  
Mentor: [Dr Adrian Weller](#) [CBL, University of Cambridge](#)  
Trustworthy Machine Learning.  
☆ Bye-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](#)  
Contributed features to a complaint management system, such as text or image based automatic complaint categorization, and deduplication.
- Research Member Staff June'14 - Feb'16  
Mentor: [Dr Sudheendra Hangal](#) [Amuse Labs](#)  
Worked on a digital archival project called [ePADD](#) and developed the following features. (1) A fine-grained entity recognizer that is robust to domain shifts, which is built using distantly supervised binomial mixture models. (2) Cross document co-referencing and entity linking using context cues.
- Research Intern June'13 - Aug'13  
Mentor: [Navneeth S](#) [GE Global Research](#)  
Implemented software for distortion correction and 3D registration of Ultrasound images.

## Scholastic Achievements

- One of the sixteen global recipients of [Google PhD fellowship](#) in Machine Learning in 2020.
- Selected to receive Prime Minister's Fellowship for Doctoral Research, 2021 (declined).
- Google, Microsoft travel grants to present at ACL 2019; ICLR 2018 Travel Award.
- Department Rank one among 100 students in the M.Tech. computer science batch of 2019.
- Represented IIT Mandi, as a member of a team, in 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.

## Selected 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%. Workshop publications are also peer-reviewed, but have higher acceptance rates.

Please see Google Scholar [page](#)  for a full list of publications.

## Conference

1. [Focus on the Common Good: Group Distributional Robustness Follows](#)  
V Piratla, P Netrapalli, S Sarawagi  
*International Conference on Learning Representations (ICLR) 2022.*
2. [Active Assessment of Prediction Services as Accuracy Surface Over Attribute Combinations](#)  
V Piratla, S Chakrabarty, S Sarawagi  
*Neural Information Processing Systems (NeurIPS) 2021.*  
☆ Recognised as an [AI Game Changer](#) by [NASSCOM](#) in ML Fundamentals category.
3. [NLP Service APIs and Models for Efficient Registration of New Clients](#)  
S Shah, V Piratla, S Sarawagi, S Chakrabarti  
*Findings at Empirical Methods in Natural Language Processing (EMNLP), 2020.*
4. [Efficient Domain Generalization via Common-Specific Low-Rank Decomposition](#)  
V Piratla, P Netrapalli, S Sarawagi  
*International Conference on Machine Learning (ICML) 2020.*
5. [Topic Sensitive Attention on Generic Corpora Corrects Sense Bias in Pretrained Embeddings](#)  
V Piratla, S Sarawagi, S Chakrabarti  
*Annual Meeting of the Association for Computational Linguistics (ACL) 2019 (Oral).*
6. [Generalizing Across Domains via Cross-Gradient Training](#)  
S Shankar\*, V Piratla\*, S Chakrabarti, S Chaudhuri, P Jyothi, S Sarawagi [\* Shared first author]  
*International Conference on Learning Representations (ICLR) 2018.*

## Workshop

7. [Untapped Potential of Data Augmentation: A Domain Generalization Viewpoint](#)  
V Piratla, S Shankar  
*ICML 2020 Workshop on Uncertainty & Robustness in Deep Learning.*
8. [Historical Research Using Email Archives](#)  
S Hangal, V Piratla, C Manovit, P Chan, M Lam, G Edwards  
*Conference on Human Factors in Computing Systems 2015 Case Studies.*

## Preprints

9. [Robust Learning from Explanations](#)  
J Heo\*, V Piratla\*, Matthew Wicker, Adrian Weller

## Talks

- Research Challenges when scaling to millions of users through Prediction Service APIs [\[talk\]](#)[\[slides\]](#)  
*Presented at [Trust ML Rising Star Spotlights Series](#)*
- Efficient Domain Generalization via Common-Specific Low-Rank Decomposition [\[talk\]](#)[\[slides\]](#)  
*Presented at ICML 2020 Conference*

- Topic Sensitive Attention on Generic Corpora Corrects Sense Bias in Pretrained Embeddings. [\[talk\]](#)[\[slides\]](#)  
*Presented at ACL, Florence*

## Development Experience

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

Aug'14 – Feb'16

- ePADD is an open-source project to develop tools for collecting and processing of digital archives.
- Contributed tens of thousands of lines of code for smooth functioning of the application across various operating systems, browsers, compute hardware, and for processing archives that are several gigabytes large.

☆ The features I contributed continue to be a big part of the project to this day: [browse here](#).

## Academic Service/Experience

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

## Technical Skills

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

## Other Activities

- 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.
- Trekker with medium level expertise.

## References

|  |              |  |                                     |
|--|--------------|--|-------------------------------------|
| <a href="#">Prof. Sunita Sarawagi</a>    | Advisor      | <a href="mailto:sunita@cse.iitb.ac.in">sunita@cse.iitb.ac.in</a>   | Professor, IIT Bombay               |
| <a href="#">Prof. Soumen Chakrabarti</a> | Co-Advisor   | <a href="mailto:soumen@cse.iitb.ac.in">soumen@cse.iitb.ac.in</a>   | Professor, IIT Bombay               |
| <a href="#">Dr Praneeth Netrapalli</a>   | Collaborator | <a href="mailto:pnetrapalli@google.com">pnetrapalli@google.com</a> | Research Scientist, Google Research |