Krishna Pillutla

Contact

Website: `https://krishnap25.github.io`
Email: `pillutla@cs.washington.edu`

Position

Visiting Researcher, Google Research *Sept 2022 - Date*

Education

University of Washington *2016-2022*
Ph.D. in Computer Science & Engineering
Thesis: From Enormous Structured Models to On-device Federated Learning: Robustness, Heterogeneity and Optimization
Advisors: Zaid Harchaoui and Sham Kakade

Carnegie Mellon University *2014-15*
M.S. in Computer Science (QPA: 3.95/4.00)
Thesis: Data Driven Resource Allocation for Distributed Learning
Advisor: Maria-Florina Balcan

Indian Institute of Technology, Bombay *2010-14*
B.Tech (Hons) in Computer Science & Engineering (QPA: 9.54/10.0)
Thesis: Distributed Machine Learning: Iterative Convex Optimization Methods
Advisor: J. Saketha Nath

Awards

Outstanding Paper at NeurIPS *2021*
Top 6 of 9000 submissions

J.P. Morgan PhD Fellowship *2019-20*
1 of 14 awardees worldwide

Anne Dinning - Michael Wolf Endowed Regental Fellowship *2016-17*
First-year PhD Fellowship awarded on merit

CBSE Merit Scholarship by the Central Board of Secondary Education in India *2010-14*
Awarded by the Govt. of India for the duration of undergraduate studies

Previous Positions

Research Intern, Facebook AI Research *2021*
Host: Lin Xiao
Project: Federated learning with partial model personalization

Research Intern, Facebook AI Research *2019*
Hosts: Awni Hannun, Gabriel Synnaeve
Project: Sequence level end-to-end training of speech recognition models.

Machine Learning Engineer Intern, Rocket Fuel Inc. *2016*
Project: Large-scale recommender systems to provide personalized advertisement recommendations, on both model-building and systems.

Research Intern, Microsoft Research *2013*
Hosts: Dhruv Mahajan, Sundararajan Sellamanickam
Project: Iterative parameter averaging for distributed machine learning

Publications

Working papers and manuscipts[1]:

- **Pillutla, K.***, Liu, L.*, Thickstun, J., Welleck, S., Swayamdipta, S., Zellers, R., Oh, S., Choi, Y., Zaid Harchaoui, Z.. (2023)
  MAUVE Scores for Generative Models: Theory and Practice.
- Fisher, J., Jiang, L., **Pillutla, K.**, Swayamdipta, S., Hessel, J., Harchaoui, Z., Choi, Y. (2022)
  On Model Revision: Content Re-mapping and Avoidance.

Peer-reviewed journal and conference papers:

- **Pillutla, K.***, Laguel, Y.*, Malick, J., & Harchaoui, Z. (2023)
  Federated Learning with Superquantile Aggregation for Heterogeneous Data.
  Machine Learning (to appear).
- **Pillutla, K.**, Roulet, V., Kakade, S. M., & Harchaoui, Z. (2023)
  Modified Gauss-Newton Algorithms under Noise.
  IEEE Statistical Signal Processing Workshop (to appear).
- Mehta, R., Roulet, V., **Pillutla, K.**, Liu, L. & Harchaoui, Z. (2023)
  Stochastic Algorithms for Ordered Empirical Risk Minimization.
  Artificial Intelligence and Statistics Conference (AISTATS).
- Fisher, J., Liu, L., **Pillutla, K.**, Choi, Y., Harchaoui, Z. (2023)
  Statistical and Computational Guarantees for Influence Diagnostics.
  Artificial Intelligence and Statistics Conference (AISTATS).
- **Pillutla, K.**, Malik, K., Mohamed, A., Rabbat, M., Sanjabi, M., & Xiao, L. (2022).
  Federated Learning with Partial Model Personalization.
  International Conference on Machine Learning (ICML).
- **Pillutla, K.**, Kakade, S. M., & Harchaoui, Z. (2022).
  Robust Aggregation for Federated Learning.
  IEEE Transactions on Signal Processing.
- **Pillutla, K.**, Swayamdipta, S., Zellers, R., Thickstun, J., Welleck, S., Choi, Y. & Harchaoui, Z. (2021).
  MAUVE: Measuring the Gap Between Machine Text and Human Text using Divergence Frontiers.
  Neural Information Processing Systems (NeurIPS). **Outstanding Paper Award (Top 6 of 9000)**.
- Liu, L., **Pillutla, K.**, Welleck, S., Oh, S., Choi, Y. & Harchaoui, Z. (2021).
  Divergence Frontiers for Generative Models: Sample Complexity, Quantization Effects, and Frontier Integrals.
  Neural Information Processing Systems (NeurIPS).
- Kusupati, A., Wallingford, M., Ramanujan, V., Somani, R., Park, J. S., **Pillutla, K.**, Jain, P., Kakade, S., & Farhadi, A. (2021).
  LLC: Accurate, Multi-purpose Learnt Low-dimensional Binary Codes.
  Neural Information Processing Systems (NeurIPS).
- Laguel, Y., **Pillutla, K.**, Malick, J., & Harchaoui, Z. (2021).
  Superquantiles in Action: Subdifferential Calculus in Practice and Applications in Machine Learning.
  Set Valued and Variational Analysis.
- Laguel, Y.*, **Pillutla, K.***, Malick, J., & Harchaoui, Z. (2021).
  A Superquantile Approach to Federated Learning with Heterogeneous Devices.
  IEEE Conference on Information Sciences and Systems (CISS).
- **Pillutla, K.**, Roulet, V., Kakade, S. M., Harchaoui, Z. (2018).
  A Smoother Way to Train Structured Prediction Models.
  Neural Information Processing Systems (NeurIPS).
- Jain, P., Kakade, S. M., Kidambi, R., Netrapalli, P., **Pillutla, V. K.**, & Sidford, A. (2017).
  A Markov Chain Theory Approach to Characterizing the Minimax Optimality of Stochastic Gradi-

[1] equal contribution denoted by * and alphabetical order by $^{\alpha}$

ent Descent (for Least Squares).
Foundations of Software Technology and Theoretical Computer Science (FSTTCS).

- Ruffalo, M., Stojanov, P., **Pillutla, V. K.**, Varma, R., & Bar-Joseph, Z. (2017).
Reconstructing cancer drug response networks using multitask learning.
BMC Systems Biology.
- Dick, T.$^{\alpha}$, Li, M.$^{\alpha}$, **Pillutla, V. K.**$^{\alpha}$, White, C.$^{\alpha}$, Balcan, M-F., & Smola, A. (2017).
Data Driven Resource Allocation for Distributed Learning.
Artificial Intelligence and Statistics Conference (AISTATS).
- **Pillutla, V. K.***, Fang, Z.*, Devineni, P., Faloutsos, C., Koutra, D., & Tang, J. (2016).
On Skewed Multi-dimensional Distributions: the FusionRP Model, Algorithms, and Discoveries.
SIAM International Conference on Data Mining.

Workshop papers:

- **Pillutla, K.***, Laguel, Y.*, Malick, J., & Harchaoui, Z. (2022)
Tackling Distribution Shifts in Federated Learning with Superquantile Aggregation.
NeurIPS 2022 Workshop on Distribution Shifts. **Spotlight Presentation**.
- **Pillutla, K.***, Laguel, Y.*, Malick, J., & Harchaoui, Z. (2022)
Differentially Private Federated Quantiles with the Distributed Discrete Gaussian Mechanism.
International Workshop on Federated Learning: Recent Advances and New Challenges (FL-NeurIPS).
- **Pillutla, K.**, Roulet, V., Kakade, S. M., & Harchaoui, Z. (2022).
The Trade-offs of Incremental Linearization Algorithms for Nonsmooth Composite Problems.
Order up! The Benefits of Higher-Order Optimization in Machine Learning (HOO-NeurIPS).
- **Pillutla, K.**, Kakade, S. M., & Harchaoui, Z. (2020).
Robust Aggregation for Federated Learning.
International Workshop on Federated Learning for User Privacy and Data Confidentiality (FL-ICML).
Long Presentation.
- Dick, T.$^{\alpha}$, Li, M.$^{\alpha}$, **Pillutla, V. K.**$^{\alpha}$, White, C.$^{\alpha}$, Balcan, N., & Smola, A. (2017).
Data Driven Resource Allocation for Distributed Learning.
AAAI Workshop on Distributed Machine Learning.

Software Released

Mauve: Measuring the Gap Between Neural Text and Human Text

- A package to compute the Mauve score for neural text generation
- Install as `pip install mauve-text`: **5000 monthly downloads**
- Software is available at `https://github.com/krishnap25/mauve` and documentation at `https://krishnap25.github.io/mauve/`
- Also available via the HuggingFace Evaluate package

Geom-Median: Fast and Differentiable Geometric Median in PyTorch and NumPy

- Implementation of the smoothed Weiszfeld algorithm to compute the geometric median, a high-dimensional median analogue
- Install as `pip install geom-median`: **265 monthly downloads**
- Software is available at `https://github.com/krishnap25/geom_median`

SQwash: Distributionally robust learning in PyTorch with a 1 additional line of code

- Implementation of the superquantile to add distributional robustness to PyTorch training loops
- Install as `pip install sqwash`: **65 monthly downloads**
- Software is available at `https://github.com/krishnap25/sqwash` and documentation at `https://krishnap25.github.io/sqwash/`

Casimir: Catalyst, smoothing, and inference.
A toolbox of selected optimization algorithms including Casimir-SVRG (as well as SVRG and Catalyst-SVRG as special cases) for unstructured tasks such as binary classification, and structured prediction

tasks such as named entity recognition and visual object localization. Available at `https://github.com/krishnap25/casimir`.

RFA: Robust Federated Aggregation
Implementation of robust aggregation for federated learning in *TensorFlow Federated.* Available at this URL.

Workshop/ Conference Organization

IFDS Workshop on Distributional Robustness in Data Science (website) *2022*
Local Organizer

Minisymposium on Federated Learning at ICCOPT *2022*
Main Organizer

Invited Talks

Federated Learning with Partial Model Personalization (2022).
Federated Learning One World Seminar.

Federated Learning with Superquantile Aggregation for Heterogeneous Data (2021-22).
IFDS Ethics and Algorithms, International Conference on Continuous Optimization.

From Enormous Structured Models to On-device Federated Learning: Robustness, Heterogeneity, and Optimization (2022).
Microsoft Research, Meta AI Research, Google Research.

MAUVE: Measuring the Gap Between Neural Text and Human Text (2022).
Stanford NLP Seminar, Microsoft Research Asia, IFML NSF Site Visit.

Mentoring

- Mentoring two junior graduate students in research *2021-date*
  - → Jillian Fisher
  - → Ronak Mehta
- Undergraduate Student Mentor: Mentored 12 freshmen in their overall development as they deal with academic pressures and transition into college life *2013-14*
- Department Academic Mentor: Served as a liaison between students struggling with academics and professors *2012-14*

Teaching

Statistical Learning with Differentiable Programming *Winter 2021 and 2022*
- Teaching assistant for a class of 25-30 students in the professional Masters in Data Science program
- Co-designed the syllabus; prepared and delivered the "lab" component of each lecture; designed homework and exams; co-managed the logistics; held office hours and graded

Machine Learning for Big Data *Spring 2018*
- Teaching assistant for 50 early-stage PhD students
- Prepared data and designed homework; co-designed, prepared, advised on course projects; designed exams; held office hours and graded

Reinforcement Learning and Bandits, Teaching Assistant *2019*

Algorithms and Foundation of Computing, Volunteer Tutor *2016-17*

Programming 101, Chemistry 101, Numerical Analysis, Teaching Assistant at IITB *2012-14*

Academic Honors

- Perfect 100 percentile (top 8 out of 174,000) in Common Admission Test (CAT) *2013*
- Gold medal at the Indian National Chemistry Olympiad (INChO). Part of initial shortlist for the International Chemistry Olympiad (Top 35 from 28,000) *2010*
- Secured All India Rank 22 in IITJEE, an exam taken by half million students *2010*
- Awarded the Certificates of Merit by the CBSE [2] for being in the top 0.1% in India in Mathematics and Chemistry in Grade 12 examinations, AISSCE *2010*

Service

- **GPU cluster co-maintainer** for Zaid Harchaoui's group (2018-2022)
- **Working Group Organizer**:
  - → Extending the Reach of Differentiable Programming (2020)
  - → Statistical Inference via Convex Optimization (2020)
  - → Computational Optimal Transport (2020)
- **Reviewer** for JMLR, Math. Prog., NeurIPS, AISTATS, JOTA, AISTATS, ICLR
- **Student Area Chair** for Machine Learning, UW CSE Graduate Admissions (2020-21) and application reader (2018 -20)
- **Organizer** for New Graduate Student Orientation at UW (2017) and Panelist (2018-20)

Relevant Coursework

Interactive Machine Learning, Convex Analysis and Nonsmooth Optimization, Deep Learning Systems, Reinforcement Learning and Bandits, Advanced Probability, Randomized Algorithms, Intermediate Statistics, Convex Optimization, Natural Language Processing, Computer Vision, Foundations of Machine Learning and Data Science, Probabilistic Graphical Models, Statistical Machine Learning, Advanced Introduction to Machine Learning, Data Mining

[2] CBSE is the Central Board of Secondary Education in India