Education

École Polytechnique Fédérale de Lausanne

École Polytechnique Fédérale de Lausanne, Switzerland
Ph.D. in Computer Science

- Presently working on Statistical Physics of Computation and Universality under Prof. Lenka Zdeborová and Prof. Florent Krzakala.
- Accepted to the ELLIS program. Main supervisor: Lenka Zdeborová, Secondary (host) supervisor: Bernhard Schölkopf.

Indian Institute of Technology Kanpur

Indian Institute of Technology Kanpur, Kanpur, India
Dual degree (B.Tech-M.Tech), Computer Science and Engineering

- Cumulative Performance Index: 9.6/10.0, Master’s CPI: 10.0/10.0
- Exchange and visiting student at École polytechnique fédérale de Lausanne (EPFL)

Interests

Theory of Deep Learning, Statistical Physics of Computation, Sampling, Optimization, Probability Theory, Causality, Dynamical Systems, Random Matrix Theory, Combinatorics.

Publications and Preprints

Universality laws for Gaussian mixtures in generalized linear models

Yatin Dandi*, Ludovic Stephan*, Florent Krzakala, Bruno Loureiro, Lenka Zdeborová (* denotes equal contribution)
Submitted to COLT, 2023

Single-spin-flip-stability in Random Graphs and Spin Glasses

Yatin Dandi, David Gamarnik, Lenka Zdeborová
In preparation. Presented at the meeting on “Random Graphs: Combinatorics, Complex Networks and Disordered Systems” at the Oberwolfach Research Institute for Mathematics in March.

Data-heterogeneity-aware Mixing for Decentralized Learning

Yatin Dandi, Anastasia Koloskova, Martin Jaggi, Sebastian Stich
NeurIPS 2022 Workshop: Optimization for Machine Learning

Implicit Gradient Alignment in Distributed and Federated Learning

Yatin Dandi*, Luis Barba*, Martin Jaggi (* denotes equal contribution)
Proceedings of the AAAI Conference on Artificial Intelligence (Preliminary version also accepted at the FL-IcML Workshop, 2021)

Generalized Adversarially Learned Inference

Yatin Dandi, Homanga Bhuradhwaj, Abhishek Kumar, Piyush Rai
Proceedings of the AAAI Conference on Artificial Intelligence (Preliminary version also accepted at the Neurips 2020 Workshop: Self-Supervised Learning – Theory and Practice)

Jointly Trained Image and Video Generation using Residual Vectors

Yatin Dandi, Aniket Das, Soumye Singhal, Vinay P. Namboodiri, Piyush Rai
2020 Winter Conference on Applications of Computer Vision (WACV ’20)

Model-Agnostic Learning to Meta-Learn

Arnout Devos*, Yatin Dandi* (* denotes equal contribution)
Pre-registration workshop, NeurIPS (2020). Full paper published in Proceedings of Machine Learning Research (PMLR).

NeurInt-Learning Interpolation by Neural ODEs (Spotlight)

Avinandan Bose*, Aniket Das*, Yatin Dandi, Piyush Rai
The Symbiosis of Deep Learning and Differential Equations: DLDE Workshop, NeurIPS 2021.

Understanding Layer-wise Contributions in Deep Neural Networks through Spectral Analysis

Yatin Dandi, Arthur Jacot
Arxiv Preprint: https://arxiv.org/abs/2111.03972
Honors & Awards

2017 Aditya Birla Scholarship, Awarded to 15 students from all IITs
2018, 2019 Academic Excellence Award, Awarded (twice) for exceptional performance
2016 KVY Scholarship, 2016, Indian Institute of Science
2015 NTSE Scholarship, Government of India
2017 All India Rank 135, JEE Advanced 2017
2016 Selected for Indian National Physics and Chemistry Olympiads, HBCSE

Projects and Internships

Single-spin-flip-stability in Random Graphs and Spin Glasses

PROJECT UNDER PROF. ZDEBOROVA AT EPFL AND PROF. GAMARNIK AT MIT
June 2022 - Present
• Rigorously proved a phase transition characterizing partitions with a large number of vertices having sufficient single-spin-flip-stability or friendliness/unfriendliness in Random Graphs and Spin Glasses.
• Proved universality of the threshold across different distributions of sparse weighted random graphs in the large degree limit.
• Proved that with appropriate re-scaling, the universality result further extends to dense fully-connected spin glass models such as the Sherrington-Kirkpatrick model.
• Utilized a range of theoretical techniques from Combinatorics, Theory of Large Deviations, and Probability Theory.

Optimization, Sampling, and Trajectory Inference algorithms

INTERNSHIP UNDER PROF. SIMON LACOSTE-JULIEN AND PROF. YOSHUA BENGIO
June 2022 - August 2022
• Worked on the analysis of Saddle Point Frank Wolfe algorithm.
• Worked on Sampling and Trajectory Inference Algorithms motivated by applications in single-cell Biology.

Data dependent mixing for Faster Heterogeneous Decentralized Learning

PROJECT UNDER PROF. JAGGI AT THE MLO LAB, EPFL
August 2021 - Present
• Extended the analysis for decentralized SGD to characterize the interaction between data heterogeneity and the mixing matrix.
• Formulated a novel criterion for periodically updating the mixing matrix based on an appropriate constrained optimization problem.
• Implemented the approach in toy settings as well as on large scale datasets using the torch.distributed and MPI framework.

Implicit Gradient Alignment in Distributed and Federated Learning

PROJECT UNDER PROF. JAGGI AT THE MLO LAB, EPFL
April 2021 - August 2021
• Studied several theoretical works aiming to explain the generalization benefits of SGD.
• Designed a novel algorithm that replicates the implicit regularization effect of SGD in distributed and federated learning settings.
• Proved theoretical results that show that the proposed algorithm approximates SGD and leads to descent on the regularized objective.
• Demonstrated an analogous implicit regularization effect for SCAFFOLD.

Generalized Adversarially Learned Inference

UNDERGRADUATE PROJECT UNDER PROF. PIYUSH RAI
October 2019 - May 2020
• Proposed a novel approach to introduce reconstructive, self-supervised, and knowledge-based feedback in the ALIGAN/BIGAN framework by modifying the discriminator’s objective to correctly identify more than just two joint distributions of image-latent vector pairs.
• Performed quantitative and qualitative evaluations over baselines on SVHN and CelebA datasets.
• Proved that the optimum for the proposed objective corresponds to simultaneously matching all the distributions.

Understanding Layer-wise Contributions in Deep Neural Networks through Spectral Analysis

PROJECT UNDER ARTHUR JACOT AND PROF. CLEMENT HONGLER AT EPFL
March 2020 - May 2021
• Utilized the properties of Hermite Polynomials and Spherical Harmonics to prove that the eigenvalues corresponding to high frequency functions of the contribution to the NTK from the first layer are larger than those corresponding to the second layers for a two-layer network with arbitrary activation functions.
• Validated the proposed theory using experiments on datasets of Spherical Harmonics as well as MNIST.

Causal Inference, Out of Distribution Generalization and Meta Learning: A Unified Perspective

SEMESTER PROJECT UNDER PROF. GROSSGLAUSER AT THE INDY LAB, EPFL
August 2020 - December 2020
• Studied the theory of causal inference, surveyed recent works on causal inference and out of distribution generalization for several classes of machine learning models.
• Analyzed the relationship between out of distribution generalization, meta-learning and causal inference.
• Formulated and empirically analyzed several algorithms under the novel paradigm of “out of task distribution generalization.”
On the Effect of Noise induced by Gradient Stochasticity on Optimizing 2-player
Differentiable Games

**PROJECT UNDER TATJANA CHAVDAROVA AT EPFL**

- Analyzed the continuous time limit of stochastic variants of first order algorithms for differentiable games using the theory of Stochastic Differential Equations (SDEs).
- Derived and studied first order approximations for second order algorithms such as SGA (Symplectic Gradient Adjustment) for n-player games.

**Towards a Proof of the Fourier Entropy Conjecture**

**COURSE PROJECT UNDER PROF. RAJAT MITTAL**

- Surveyed several papers related to the Fourier Entropy conjecture.
- Prepared a summary report and presentation on the recent work by Kelman et al.

**Jointly Trained Image and Video Generation using Residual Vectors**

**UNDERGRADUATE PROJECT UNDER PROF. PIYUSH RAI AND PROF. VINAY NAMBOODIRI**

- Designed a novel modelling technique for joint image and video generation that simultaneously learns to map latent variables with a fixed prior onto real images and interpolate over images to generate videos.
- Implemented Variational Autoencoder and Generative Adversarial Networks based video generation models in Pytorch utilizing the proposed technique.
- Performed quantitative and qualitative evaluations over previous VAE and GAN based video generation approaches.

**New York Office, IIT Kanpur**

**MACHINE LEARNING FOR LARGE SCALE LOGISTICS PLATFORM, UNDER PROF. MANINDRA AGARWAL**

- Implemented a state-of-the-art algorithm for online collaborative filtering based on Fast Matrix Factorization for Online Recommendation with Implicit Feedback (He et al.) using NumPy and improved the model with sentiment and frequency dependent weighting schemes.
- Used Kafka for real-time data processing and simulated interactions using locust.
- Implemented a recommender system based on deep autoencoders and compared the results with other models using metrics such as hit ratio.
- Implemented a Bidirectional LSTM model using Keras for sentiment analysis of user comments.
- Trained the Latent Dirichlet allocation model on Wikipedia articles for automatic extraction of topics.

**Social Situation Inference in Simple Animated Shapes**

**UNDERGRADUATE PROJECT UNDER PROF. NISHEETH SRIVASTAVA**

- Conducted experiments to characterise the nature and granularity of information about social interactions accessible to human observers from simple visual displays.
- Developed an animation engine capable of producing a vast range of social situations involving autonomous agents having long term and short term goals.

**Deep Reinforcement Learning for Atari Games**

**ASSOCIATION FOR COMPUTING ACTIVITIES, IIT KANPUR**

- Used NumPy to implement various reinforcement learning algorithms such as Dynamic Programming (Policy and Value iteration), Monte Carlo (Epsilon-greedy and off-policy), TD Learning (Q-Learning and SARS) and Q-Learning with Function Approximation.
- Implemented Deep Q-Learning and Policy Gradient methods for Atari Games using PyTorch and OpenAI Gym.

**Image Captioning with Visual Attention**

**PROGRAMMING CLUB, IIT KANPUR**

- Compared various CNN based models for image classification and implemented them using eager execution in Tensorflow.
- Implemented the model described in Show, Attend and Tell (Xu et al.2015) using Tensorflow’s estimator API and evaluated the model on MS COCO dataset.

**Microsoft code.fun.do hackathon**

**NATIONAL-LEVEL HACKATHON WINNER**

- Selected out of 120 students to represent IIT Kanpur at Microsoft Hyderabad center to showcase our project in their academia-industry collaboration event AXLE.
- Built an interactive interface using D3.js to display changing geopolitical relations and popularity of world leaders.
- Used Scrapy, a scraping framework to extract world news.
- Performed sentiment analysis and named entity-recognition on the extracted news to infer the effect of the concerned news on world politics.

**Positions of Responsibility**

**ICML, 2023**

**REVIEWER**

- Reviewed a paper on recovery from pooled data.

**ICLR, 2023**

**REVIEWER**

- Reviewed a paper on Noise Contrastive Estimation.
AISTATS, 2022

Reviewer
• Reviewed papers on Random Feedback and Adversarial Risk.

Introduction to Machine Learning
Teaching Assistant
• Graded assignments and exams for the introductory machine learning course.

Mathematics for Computer Science - 2, 3
Teaching Assistant
• Graded assignments and exams for courses on “Logic for Computer Science” and “Probability for Computer Science”

SIGML - Special Interest Group on Machine Learning
Coordinator
• Helped conduct and organize talks and lectures on various topics in machine learning.

Programming Club
Secretary
• Helped conduct and organize Linux fest, introductory workshops and various hackathons.

Association of Computing Activities
Project Mentor
• Mentored first-year students on a semester long project on Generative Adversarial Networks.

Programming Club
Project Mentor
• Mentored first-year students on a semester long project on Probabilistic Machine Learning.

Relevant Courses

| Course                                                                 | Grade |
|-----------------------------------------------------------------------|-------|
| Introduction to Programming - A                                      |       |
| Discrete Mathematics - A                                              |       |
| Computational Cognitive Science (audited)                             |       |
| Bayesian Analysis - A                                                 |       |
| Optimization For Machine Learning - 6.0/6.0                          |       |
| Mathematical Methods For Physicists                                  |       |
| Quantum Computing - A                                                |       |
| Foundations of Data Science - 6.0/6.0                                 |       |
| Probability for Computer Science - A*                                 |       |
| Logic for Computer Science - A*                                       |       |
| Linear Algebra and ODE - A                                            |       |
| Theory of Computation - A                                             |       |
| Learning Theory - 6.0/6.0                                             |       |
| Computational Complexity - A                                          |       |
| Modern Cryptology - A                                                |       |
| Introduction to Stochastic PDEs (ongoing)                             |       |
| Calculus and Real Analysis - A*                                       |       |
| Topics in Probabilistic Modeling and Inference - A                   |       |
| Data Structures and Algorithms - A                                   |       |
| Introduction to Machine Learning - A                                 |       |
| Statistical Physics of Computation - 6.0/6.0                         |       |
| Complexity Measures for Boolean Functions - A                        |       |
| Lattice Models                                                        |       |
| Mathematical Methods For Physicains                                  |       |
| Computational Complexity - A                                         |       |
| Modern Cryptology - A                                                |       |
| Introduction to Stochastic PDEs (ongoing)                             |       |

A*: grade for exceptional performance

Programming Skills

- Programming Languages: C, C++, Python, Javascript, MATLAB, Haskell
- Libraries and frameworks: Tensorflow, Pytorch, Jax, Scikit-Learn, Pillow, Keras, Numpy, Flask, HTML, CSS, JQuery, MySQL
- Web: LaTeX, Git, Docker
- Utilities: Linux shell utilities

Miscellaneous
• Delivered a talk on causal inference and discovery in a joint meeting of the SPOC, Idephics labs at EPFL.
• Delivered a talk on Zero Knowledge Proofs for the course Complexity Theory, IIT Kanpur and on its applications to machine learning at the MLO lab, EPFL.
• Delivered a talk on Brain-Computer Interface in an event organized by Science Coffee House, IIT Kanpur.
• Winner of Blackbox - a three hour high speed hackathon based on an Esoteric language organized by Programming Club IIT Kanpur.