Yunzhen Feng

✉ yf2231@nyu.edu · 📞 917-8917732 · Homepage · Google Scholar

EDUCATION

New York University (NYU), New York, United States   Sep'21 - present
- Ph.D. Student at the Center for Data Science.
- Advisor: Julia Kempe.
- Topics: Synthetic Data, Reinforcement Learning for scientific reasoning and inference time optimization, LLM for math and code generation.

École Normale Supérieure (ENS), Paris, France   Jan'24 - Aug'24
- Visiting student at the Centre Sciences des Données.

Peking University (PKU), Beijing, China   Aug'17 - June'21
- *B.S.* (honor track) in Applied Mathematics.
- Advisor: Bin Dong, Di He.

PUBLICATIONS

1. Elvis Dohmatob, **Yunzhen Feng**, Arjun Subramonian, Julia Kempe. Strong Model Collapse. *In submission*.
2. **Yunzhen Feng***, Elvis Dohmatob*, Pu Yang*, Francois Charton, Julia Kempe. Beyond Model Collapse: Scaling Up with Synthesized Data Requires Reinforcement. *ICML 2024 Workshop on Theoretical Foundations of Foundation Models*. Link.
3. Elvis Dohmatob, **Yunzhen Feng**, Julia Kempe. Model Collapse Demystified: The Case of Regression. *The Thirty-Eighth Annual Conference on Neural Information Processing Systems (NeurIPS), 2024*. Link.
4. **Yunzhen Feng***, Elvis Dohmatob*, Pu Yang, Francois Charton, Julia Kempe. A Tale of Tails: Model Collapse as a Change of Scaling Laws. *International Conference on Machine Learning (ICML), 2024*. Link.
5. Kai Yang, Jan Ackermann, Zhenyu He, Guhao Feng, Bohang Zhang, **Yunzhen Feng**, Qiwei Ye, Di He, Liwei Wang. Do Efficient Transformers Really Save Computation? *International Conference on Machine Learning (ICML), 2024*. Link.
6. **Yunzhen Feng**, Ramakrishna Vedantam, Julia Kempe. Embarrassingly Simple Dataset Distillation. *the Twelfth International Conference on Learning Representations (ICLR)*, 2024. Link.
7. **Yunzhen Feng**, Nikos Tsilivis, Tim G. J. Rudner, Julia Kempe. Attacking Bayes: Are Bayesian Neural Networks Inherently Robust? *Transactions on Machine Learning Research (TMLR), 2024*. Link.
8. Chizhou Liu, **Yunzhen Feng**, Ranran Wang, Bin Dong. Enhancing Certified Robustness of Smoothed Classifiers via Weighted Model Ensembling. *ICML 2021 workshop on Adversarial Machine Learning, 2021*. Link.
9. **Yunzhen Feng***, Runtian Zhai*, Di He, Liwei Wang, Bin Dong. Transferred Discrepancy: Quantifying the Difference Between Representations . *arXiv:2007.07446*. Link.
10. **Yunzhen Feng**, Yue M. Lu. A Precise High-dimensional Analysis of Laplacian Semi-Supervised Learning, 2020.

RESEARCH EXPERIENCE

Foundational AI Research (FAIR), Meta   May'24 - Sep'24
Research Scientist Intern   *Paris, France*
- Advisor: Yann Ollivier, François Charton, Julia Kempe.
- Develop novel methods for scientific reasoning using reinforcement learning and large language models, with a focus on inference time optimization.

Marketplace Investment Team, Uber Technologies   May'22 - Aug'22

Applied Scientist PhD Intern   *California, U.S.*

- Mentor: Jason Dowlatabadi, Chen Xu, Allen Zhang.
- Team Advisor: Stefan Wager, Peter Frazier.
- Design doubly robust CausalML model for learning heterogeneous incentive response across cities with both observational data and experimental data.

School of Engineering and Applied Science, Harvard University   Aug'20 - Dec'20

Research Assistant   *Remote*

- Advisor: Yue M. Lu.
- Investigated how unlabeled data help the classification for Gaussian mixtures and Laplacian regularization using high dimensional statistics.

Microsoft Research Asia   Jan'20 - June'20

Visiting   *Beijing, China*

- Advisor: Di He.
- Proposed to measure the difference between two networks' representation via their performances on a set of downstream tasks and investigate the effect of random initialization on the learned features.

TEACHING

Head TA for **Modern Topics in Statistical Learning Theory**	2023 Spring
Grader TA for **Modern Topics in Statistical Learning Theory**	2024 Spring

INVITED TALKS

Synthetic Data: Benefits and Challenges

ENS PhD seiminar	Mar'2024
FAIR, Meta	Sep'2024
NYU machine learning for language	Sep'2024

SELECTED HONORS AND AWARDS

National Science Foundation Research Traineeship (NRT) Future Program	2021
Data Science Fellowship ($ 190,000), New York Univeristy	2021
Outstanding Graduate Student, Peking University	2021
Bronze Medal, S.-T.Yau College Student Mathematics Contests (top 20 in China)	2020
Meritorious Award, Mathematical Contest in Modeling (top 9%)	2019
The Elite Undergraduate Training Program of Applied Mathematics (top 10%)	2018
Gold Medal, China Mathematics Olympiad (top 100 in China)	2016
Ranked First in First Prize, Chinese High School Mathematics League, Beijing (top 1)	2016

SPECIALIZED SKILLS

- **Programming**: Python (PyTorch), MATLAB, LaTeX.
- **Language**: English (Fluent), Mandarin (Fluent), French (Beginner).

ACTIVITIES AND INTERESTS

- Successfully climbed the Luodui Mount (6010m) at Tibet, China.
- Rock-climber and outdoor hiker trained by committees from Chinese Mountaineering Association.
- One-year volunteer at Arthur M. Sackler Museum of Art and Archaeology, Peking University.
- Teaching Assistant for the *Outdoor Exploration* course at PKU.