Jane H. Lee

Office: 219 Prospect St, KT1233
Email: jane.h.lee@yale.edu
Website: janehjlee.github.io

Education

Yale University — New Haven, CT
Ph.D. in Computer Science — In progress
- Research areas: statistical learning theory, machine learning theory

University of Pennsylvania — Philadelphia, PA
M.S.E. in Computer Science, GPA: 4.00/4.00 — May 2019
- Submatriculant: Completed Bachelor's and Master's degrees concurrently
- Thesis: "Noisy Labels in Multiclass Classification", advised by Dr. Shivani Agarwal

University of Pennsylvania — Philadelphia, PA
B.A. in Mathematics and in Computer Science, minor in Statistics, GPA: 3.84/4.00 — May 2019
- Summa Cum Laude

Selected Scholarships and Awards

- **Greenberg Fellowship**, *Yale Graduate School of Arts and Science (GSAS)* 2024
  Named Yale fellowships recognize graduate students' academic excellence. They are also a reflection of GSAS alumni and friends' commitment to support the individual student.
- **Distinguished Teaching Award**, *Yale Computer Science* 2023
  One (1) graduate student recognized for outstanding teaching during the 2022-2023 academic year.
- **Quad Fellowship by Schmidt Futures**, *Finalist* 2022
  Finalist (<5% of applicants from U.S.A., India, Australia, and Japan) for the Quad Fellowship 2022-2023 cycle.
- **Graduate Fellowship for STEM Diversity** 2022
  Formerly the National Physical Science Consortium Fellowship (NPSC), the GFSD is a national fellowship which provides 6 years of support and $20,000 stipend. Supported by the National Security Agency (NSA).
- **Ford Foundation Predoctoral Fellowship Competition**, *Honorable Mention* 2022
  Accorded honorable mention status in the 2022 Ford Foundation Fellowship Programs competition administered by the National Academies of Sciences, Engineering, and Medicine.
- **Department of Defense SMART Scholarship**, *Declined* 2020
  Offered up to 5 years of support in addition to a $38,000 living stipend. Awarded to the top 20% of Ph.D. applicants.
- **Twitter GHC Fellow** 2018
  A merit-based fellowship awarded to attend the 2018 Grace Hopper Celebration and a full-time offer of employment at Twitter.

Working Papers

1. **J. H. Lee**, A. Mehrotra, M. Zampetakis, "Learning with Positive and Imperfect Unlabeled Data" (under review)
2. Z. Chen, C. Zhao, K. Mo, Y. Jiang, **J. H. Lee**, K. C. Mahajan, N. Jiang, K. Ren, C. Li, W. Yang, "Massive Memorization with Hundreds of Trillions of Parameters for Sequential Transducer Generative Recommenders" (under review)

Publications

[1] **J. Lee**, B. Saglam, S. Pougkakiotis, A. Karbasi, and D. Kalogerias, "Risk-averse constrained reinforcement learning with optimized certainty equivalents", in *Advances in Neural Information Processing Systems (NeurIPS)*, forthcoming, Dec. 2025.

[2] **J. Lee**, A. Mehrotra, and M. Zampetakis, "Efficient statistics with unknown truncation: Polynomial time algorithms, beyond gaussians", in *65th IEEE Symposium on Foundations of Computer Science (FOCS)*, Oct. 2024.

[3] **J. Lee**, A. Wibisono, and M. Zampetakis, "Learning exponential families from truncated samples", in *Advances in Neural Information Processing Systems (NeurIPS)*, Dec. 2023.

[4] **J. Lee**, S. Haghighatshoar, and A. Karbasi, "Exact gradient computation for spiking neural networks through forward propagation", in *International Conference on Artificial Intelligence and Statistics, (AISTATS)*, Apr. 2023.

[5] M. Zhang, **J. Lee**, and S. Agarwal, "Learning from noisy labels with no change to the training process", in *International Conference on Machine Learning (ICML)*, Jul. 2021.

[6] K. Jaidka, S. C. Guntuku, **J. H. Lee**, Z. Luo, A. Buffone, and L. H. Ungar, "The rural–urban stress divide: Obtaining geographical insights through twitter", *Computers in Human Behavior*, vol. 114, p. 106 544, Jan. 2021, ISSN: 0747-5632.

[7] S. Chen, E. Dobriban, and **J. H. Lee**, "A group-theoretic framework for data augmentation", *Journal of Machine Learning Research (JMLR)*, vol. 21, no. 245, pp. 1–71, 2020.

[8] S. Chen, E. Dobriban, and **J. Lee**, "A group-theoretic framework for data augmentation", in *Advances in Neural Information Processing Systems (NeurIPS)*, **Oral Presentation**, 2020.

Technical Reports

[9] S. Mayhew, T. Tsygankova, F. Marini, Z. Wang, **J. Lee**, X. Yu, X. Fu, W. Shi, Z. Zhao, W. Yin, K. K, J. Hay, M. Shur, J. Sheffield, and D. Roth, "University of Pennsylvania LoReHLT 2019 Submission", Tech. Rep., 2019.

Professional Experience

Meta — Menlo Park, CA
Research Scientist, Video Recommendations (Internship) — Summer/Fall 2025

- Study algorithms for improving video recommendation model using user interest history.
- Developed methods to process noisy signal (short watch time) videos using sub-sequence-aware unbiased loss (SUL) as an auxiliary loss. Also contributed to work on VIrtual Sequential Target Attention (VISTA) model.

Morgan Stanley — New York, NY
Machine Learning Research Associate (Internship) — Summer 2024

- Studied algorithms for solving variational forms of conditional moment problems with applications to regression with heteroskedasticity.
- Implemented practical solutions on muni bond trading data while analyzing guarantees of algorithms for solving the variational method of moments problem.

Twitter — San Francisco, CA
Machine Learning Engineer II, Ads Targeting and Modeling (Full-Time) — Jul 2019-Aug 2021

- Built ML models and pipelines to help advertisers find their audience on Twitter.

- Worked fully end-to-end on ML pipelines: from data processing, data pipeline, model architecture and design, training, and serving, on a variety of targeting products including demographic targeting, mobile app recommendation systems, and early filtering models.

Goldman Sachs New York, NY
Investment Management Summer Analyst (Private Wealth Management Strats) (Internship) Summer 2018

- Collaborated with portfolio managers/traders to analyze portfolios, create investment algorithms, and build pricing and other models.
- Built a web tool for managing PWM (Private Wealth Management) clients' preferred stock portfolios. Designed machine learning models to predict proportion of private wealth clients likely to need margin/bank loans to grow the business.

Morgan Stanley New York, NY
Quantitative Finance Summer Analyst (Securitized Products Group Strats) (Internship) Summer 2017

- Worked with traders and data to advise new trading strategies in securitized products.
- Found discrepancies in real loan data to find trading opportunity in RMBS (residential mortgage-backed securities) and used machine learning models to predict loan term modification rate of Freddie Mac loans to find opportunity in CRT (credit risk transfer) products.

Teaching

I received the Yale Computer Science Department Distinguished Teaching Award for 2022-23.

- **Teaching Fellow** at Yale University — Spring 2023
  Probabilistic Machine Learning (CPSC 586)
- **Teaching Fellow** at Yale University — Fall 2022
  Introduction to Database Systems (CPSC 537)
- **Teaching Assistant** at University of Pennsylvania — Spring 2018, Spring 2019
  Machine Learning (CIS 520)
- **Teaching Assistant** at University of Pennsylvania — Spring 2019
  Algorithms (CIS 320)
- **Teaching Assistant** at University of Pennsylvania — Fall 2018
  Agent-Based Modeling and Simulation (ESE 520)
- **Teaching Assistant** at University of Pennsylvania — Spring 2018
  Internet and Web Systems (CIS 555)
- **Head Teaching Assistant** at University of Pennsylvania — Fall 2017
  Software Engineering (CIS 573)
- **Teaching Assistant** at University of Pennsylvania — Spring 2017
  Software Engineering (CIS 350)
- **Head Teaching Assistant** at University of Pennsylvania — Fall 2016, Spring 2017, Fall 2017
  Data Structures and Algorithms (CIS 121)

Skills

- **Technical Skills:** Proficient: Python (Tensorflow, PyTorch), Java; Basic: Scala, MATLAB, SQL, C/C++
- **Language Skills:** Native: English; Conversational: Korean

Service

- **Reviewer**
  I have served as a reviewer for the following:
  - **Conferences:** Innovations in Theoretical Computer Science (ITCS) 2023, IEEE Symposium on Foundations of Computer Science (FOCS) 2024, Conference on Learning Theory (COLT) 2025, Conference on Neural Information Processing Systems (NeurIPS) 2025, International Conference on Learning Representations (ICLR) 2026, ACM Conference on Economics and Computation (EC) 2026 Program Committee
  - **Journals:** Journal of Machine Learning Research (JMLR)
  - **Workshops:** Reliable ML from Unreliable Data Workshop (NeurIPS 2025)
- **Graduate and Professional Student Senate (GPSS)** 2024 –2025
  I am an elected member of the Graduate and Professional Student Senate (GPSS) for the 2024-2025 school year.
- **Graduate Student Assembly (GSA)** 2022–2025
  I have been elected to serve on the Graduate Student Assembly (GSA) as a representative of the Computer Science department for the 2022-2023, 2023-2024, and 2024-2025 school years.
  - **Service Committee Chair**, 2022-2023, 2023-2024
  - I served as the elected chair of the Service Committee for the 2022-2023 and 2023-2024 school years, focused on improving service to graduate students and New Haven residents at large.
- **Departmental Graduate Student Advisory Committee (GSAC)** 2022–2024
  I served as one of the first members of the newly formed Yale Computer Science Department's Graduate Student Advisory Committee for the 2022-2023 school year. I continued to serve as an elected member for 2023-2024.
- **Ackerman Teaching Award Student Review Committee** 2022–2023, 2024–2025
  I was nominated by my department to serve on Yale's School of Engineering and Applied Science (SEAS) graduate student committee to review nominations for the Ackerman Teaching Award for 2022-2023 and 2024-2025.

Volunteering and Outreach

- Graduate Student Mentor at **Yale Graduate Society of Women Engineers (SWE)** 2023–2025
  Volunteer to mentor an undergraduate student who identifies as a woman in engineering. I have committed to make time for advice, support, and meeting throughout the school year.
- Graduate Student Mentor at **Women in Science at Yale (WISAY)** 2021–2025
  Volunteer to mentor an undergraduate student who identifies as a woman in science. I have committed to make time for advice, support, and meeting regularly throughout the school year.
- Volunteer Teaching Assistant at **Microsoft TEALS Program** 2021–2024
  Volunteer 2-3 days a week to assist AP Computer Science A class at local high school. The TEALS program provides high school students with equitable access to computer science (CS) education and create a pathway to economic opportunity.
- Graduate Student Mentor at **STEM Mentors at Yale** 2021–2024
  Volunteer through annual events (2-part college essay writing workshops and career panel) to motivate and support local students grades 6-12 in STEM, especially underrepresented groups like women and minorities.
- Volunteer Tutor at **The SMART Program** 2019–2021
  Volunteer weekly to help tutor a high school student in mathematics, ranging from Algebra II + Trigonometry to Precalculus. The SMART program supports low-income middle and high school students in the San Francisco area to break the cycle of poverty.
- Alumna Mentor at **WiCS Alumni Mentorship Program (Penn)** Spring 2020
  Volunteered to be an alumni mentor for WiCS (Women in Computer Science) pilot program that aims to connect undergraduate upperclassmen with UPenn Alumni members who also studied CIS/Engineering. Made time for questions, advice, chatting for undergraduate students.