Jane H. Lee

17 Hillhouse Office 333
New Haven, CT 06511
jane.h.lee@yale.edu

Education

Yale University — New Haven, CT
Ph.D. in Computer Science — Aug 2021–Current
- 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

- **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.
- **PennApps XIX Third Grand Prize, Best AI/ML Hack** — 2019
  I conceptualized and wrote the ML models for an automated recruiting tool that incorporates fair machine learning techniques from recent research. Won third grand prize at PennApps XIX, as well as the overall best AI/ML prize.
- **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.

Publications

[1] 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.

[2] 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.

[3] 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.

[4] 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

[5] 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.

Presentations and Posters

1. **Lee, J.**, March 28, 2022. Exact gradient computation for spiking neural networks via forward propagation. Presented at The Institute for Learning-enabled Optimization at Scale (TILOS) Student and Postdoc Workshop. Online.

2. **Lee, J.**, September 4, 2020. A group-theoretic framework for data augmentation. Presented at Workshop for Equivariance and Data Augmentation, supported by NSF TRIPODS. Online.

Research Experience

Yale University, Yale Institute for Network Science (YINS) — New Haven, CT
Graduate Student Researcher, advised by Dr. Amin Karbasi — Aug 2021–Current

- Exact Gradient Computation for Spiking Neural Networks via Forward Propagation
- While a majority of prior literature believes that exact gradients cannot be computed for SNN, we prove that exact gradients can be computed by invoking the implicit function theorem and propose a novel forward propagation algorithm for training.

University of Pennsylvania, Wharton Statistics — Philadelphia, PA
Student Researcher, advised by Dr. Edgar Dobriban — Aug 2018–Dec 2020

- A Group-Theoretic Framework for Data Augmentation
- We proposed a novel framework for modeling data augmentation as a group of invariant transformations to mathematically prove the benefits of data augmentation, where previous research on data augmentation has been mostly empirical.

University of Pennsylvania, Computer and Information Science — Philadelphia, PA
Student Researcher, advised by Dr. Sanjeev Khanna — Summer 2019

- Influence Maximization in Graphs using Local Information
- Studied algorithms to find (near-)optimal seeding in graphs to maximize influence constrained to local information, focusing on expanding on current results which apply under limited conditions on the infection rate as well as the inter- and intracommunity edge probabilities.

Cognitive Computation Group (Penn) — Philadelphia, PA
Student Researcher, advised by Dr. Dan Roth — Summer 2019

- Language Processing for Low Resource Languages
- Participated in the Low Resource Languages for Emergent Incidents (LORELEI) 2019 evaluation, which is a DARPA-sponsored program. Worked specifically on named entity recognition for low-resource languages, which got the highest scores for both surprise languages.

World Well-Being Project (Penn) — Philadelphia, PA
Student Researcher, advised by Dr. Lyle Ungar — Aug 2017–Aug 2018

- The rural–urban stress divide: Obtaining geographical insights through Twitter
- Utilized Twitter data to predict regional stress levels and find insights on language surrounding stress. Used LDA for topic modeling, and compare predictive performance using different features (topics, 1grams, LIWC categories) to show differences in rural and urban regions' language surrounding stress.

Professional Experience

Twitter San Francisco, CA

Machine Learning Engineer II, Ads Targeting and Modeling 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) 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) 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

- **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 Fall 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; Fluent: Korean

Service

- **Graduate Student Assembly (GSA)** 2022–Current
  I have been elected to serve on the Graduate Student Assembly (GSA) as a representative of the Computer Science department for the 2022-2023 school year.
- **Departmental Graduate Student Advisory Committee (GSAC)** 2022–Current
  I serve 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.

Extracurricular Activities

- Graduate Student Mentor at **Women in Science at Yale (WISAY)** 2021–Current
  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–Current
  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.
- 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.
- 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.
- Member at **Smart Woman Securities** 2015–2019
  Attended weekly seminars that focused on investment philosophies, hedging strategies, and security analysis. Presented a stock pitch on Southwest Airlines after researching how its expansion into new market demographics and competitive advantage as a low-cost provider were undervalued by the market.
- Violinist at **Penn Chamber** 2015-2017
  Rehearsed in groups for 6 hours per week in preparation for ensemble concerts occurring once every semester. Led as a soloist in the performance of the Clarinet Quintet by Kryzysztof Penderecki.