JINGYANG LYU

(+1) 312-961-1294 | jlyu55@wisc.edu

EDUCATION

University of Wisconsin-Madison, WI, USA *09/2022 - Present*

- *Ph.D. in Statistics* (advised by: Yiqiao Zhong). GPA: 4.00/4.00

University of Chicago, IL, USA *09/2020 - 06/2022*

- *M.S. in Statistics* (advised by: Wei-Biao Wu). GPA: 4.00/4.00
- **Honors & Awards:** Tuition Scholarship Increased (2021)

Tsinghua University, Beijing, China *08/2016 - 06/2020*

- *B.Eng. in Industrial Engineering*. GPA: 3.71/4.00 (Rank: 7/63)
- *Minor in Statistics*. GPA: 3.94/4.00
- **Honors & Awards:** Outstanding Academic Performance & Progress Scholarship (2019), Excellent Student Leadership Award (2018), Excellent Social Worker Scholarship (2018)

RESEARCH INTERESTS

◊ Statistical foundation of Deep learning ◊ Large language models ◊ Applied Statistical modeling

RESEARCH EXPERIENCE

A Statistical Theory of Overfitting for Imbalanced Classification. *02/2023 - Present*
Supervised by *Prof. Yiqiao Zhong*

- Characterized overfitting of minority classes in deep learning by deriving asymptotic margin distribution from the last layer, in both separable and non-separable regimes.
- Revealed the impact of overfitting on generalization and model calibration, verified by extensive data analysis including convolutional neural networks and large language models with PyTorch.
- Propose a margin-adjustment approach to improve the performance of imbalanced classification with theoretical guarantee, especially for highly imbalanced datasets.

A Debiased Version for Online Covariance Matrix Estimation in SGD. *09/2021 - 12/2022*
Supervised by *Prof. Wei-Biao Wu*

- Improved the estimator of online SGD covariance matrix with a efficient single-pass algorithm.
- Derived the error bound for the proposed debiased estimator in serveral different settings.

THESIS

Online Bootstrap Confidence Intervals for Stochastic Gradient Descent (SGD). *06/2022*
Supervised by *Prof. Wei-Biao Wu*, Department of Statistics, University of Chicago

- Conducted statistical inference for SGD by using Bootstrap perturbed estimates, with mathematical details added and improved for the proofs of a previous study.
- Performed simulation experiments for regressions models and real data analysis by using R.

Fault Diagnosis of High-Speed Train Based on Imbalanced Learning. *06/2020*
Supervised by *Prof. Yan-Fu Li*, Department of Industrial Engineering, Tsinghua University

- Proposed an algorithm, i.e., Recursive Different-Error-Costs Support Vector Machine (R-DEC-SVM), which can handle highly imbalanced classification tasks, implemented in R.
- Outperformed previous methods on a train operation dataset in terms of fault detection.

PUBLICATIONS

1. Chen, F., **Lyu, J.**, Wang, T., Sze, N. N. (2023). Exploring the association between quantified road safety target attributes and their success: An empirical analysis from OECD countries using panel data. *Journal of safety research, 85*, 296-307.

2. Chen, F., Zhu, Y., Zu, J., **Lyu, J.**, Yang, J. (2022). Appraising road safety attainment by CRITIC-ELECTRE-FCM: a policymaking support for Southeast Asia. *Transport policy, 122*, 104-118.
3. Chen, F., **Lyu, J.**, Wang, T. (2020). Benchmarking road safety development across OECD countries: An empirical analysis for a decade. *Accident Analysis & Prevention, 147*, 105752.

TECHNICAL STRENGTHS & OTHERS

Programming Languages	Python, R, Matlab, Java, C, SQL, SAS
Software	Gurobi, CPLEX, AnyLogic, Plant Simulation
Tools	LaTeX