Curriculum Vitae
Luyang Fang

310 Herty Drive, Department of Statistics, Athens, GA, 30602
Email: Luyang.Fang@uga.edu
Personal Website: https://luyangfang.github.io/
Google Scholar: https://scholar.google.com/citations?user=jriFo4wAAAAJ&hl=en

Education

- University of Georgia, GA, USA 08/2021 - Present
  Ph.D. *Statistics*
  Advisors: Professor Ping Ma and Professor Wenxuan Zhong
- University of Wisconsin-Madison, WI, USA 09/2019 - 05/2021
  M.S. *Statistics*
- Nankai University, Tianjin, China 09/2016 - 07/2020
  B.S. *Mathematics and Applied Mathematics*

Major Research Areas

- Deep Learning; Large Language Model (LLM); Non-parametric Methods; Subsampling Methods in Big Data; Bioinformatics; AI-Enhanced Educational Assessment

Academic Experience

- **Research Assistant** Big Data Analytics Lab Sep. 2021 – Present
  University of Georgia, Athens, GA
- **Teaching Assistant** Department of Statistics Sep. 2021 – Dec. 2024
  University of Georgia, Athens, GA

Publications

- Peer-reviewed Journal & Conference

[1] **Fang, L.**, Meng C., Zhao, L., Wang, T., Liu, T., Zhong, W., Ma, P. (2025). SPOT: An Active Learning Algorithm for Efficient Deep Neural Network Training. *Big Data Mining and Analytics.*

[2] **Fang, L.**, Latif, E., Lu, H., Zhou, Y., Ma, P., Zhai, X. (2025). Efficient Multi-Task Inferencing: Model Merging with Gromov-Wasserstein Feature Alignment. *International Conference on Artificial Intelligence in Education.*
(Nominee for the best poster award, AIED 2025)

[3] Wang, B., Cai, J., **Fang, L.**, Ma, P., Leung, Y.F. (2025). Tensor analysis of animal behavior by matricization and feature selection. *Computers in Biology and Medicine.*

[4] Ruan, W., Lyu, Y., Zhang, J., Cai, J., Shu, P., Ge, Y., Lu, Y., Gao, S., Wang, Y., Wang, P., Zhao, L., Wang, T., Liu, Y., **Fang, L.**, Liu, Z., Liu, Z., Li, Y., Wu, Z., Chen, J., Jiang, H., Pan, Y., Yang, Z., Chen, J., Liang, S., Zhang, W., Ma, T., Dou, Y., Zhang, J., Gong, X., Gan, Q., Zou, Y., Chen, Z., Qian, Y., Yu, S., Lu, J., Song, K., Wang, X., Sikora, A., Li, G., Li, X., Li, Q., Wang, Y., Zhang, L., Abate, Y., He, L., Zhong, W., Liu, R., Huang, C., Liu, W., Shen, Y., Ma, P., Zhu, H., Yan, Y., Zhu, D., Liu, T. (2025). Large language models for bioinformatics. *Quantitative Biology.*

[5] **Fang, L.**, Chen, Y., Zhong, W., Ma, P. (2024). Bayesian Knowledge Distillation: A Bayesian Perspective of Distillation with Uncertainty Quantification. *Forty-first International Conference on Machine Learning (ICML).* (**Top** machine learning conference).

[6] Wang, Z.*, **Fang, L.***, Cai, J., Ma, P., Zhong, W. (2024). MultiCOP: An Association Analysis of Microbiome-Metabolome Relationships. *Statistics in Bioscience.* (* Joint-first Author)

[7] Latif, E., **Fang, L.**, Ma, P., Zhai, X. (2024). Knowledge Distillation of LLMs for Automatic Scoring of Science Assessments. *International Conference on Artificial Intelligence in Education.*

[8] Wang, S., Wu, S., Chen, Y., **Fang, L.**, Xiao, L., Li, F. (2024). Exploring Latent Constructs through Multimodal Data Analysis. *Journal of Educational Measurement.*

[9] Wu, S., **Fang, L.**, Zhang, J., Sriram, T. N., Coshatt, S. J., Zahiri, F., Mantooth, A., Ye, J., Zhong, W., Ma, P., Song, W. (2023). Unsupervised Anomaly Detection and Diagnosis in Power Electronic Networks: Informative Leverage and Multivariate Functional Clustering Approaches. *IEEE Transactions on Smart Grid.* (IF **9.8. Top 1** journal in Power Engineering).

- Preprints

[10] Lu, H., **Fang, L.**, Zhang, R., Li, X., Cai, J., Cheng, H., Tang, L., Liu, Z., Sun, Z., Wang, T., Zhang, Y., Zidan, A. H., Xu, J., Yu, J., Yu, M., Jiang, H., Gong, X., Luo, W., Sun, B., Chen, Y., Ma, T., Wu, S., Zhou, Y., Chen, J., Xiang, H., Zhang, J., Jahin, A., Ruan, W., Deng, K., Pan, Y., Wang, P., Li, J., Liu, Z., Zhang, L., Li, X., Zhao, L., Liu, W., Zhu, D., Xing, X., Dou, F., Zhang, W., Huang, C., Liu, R., Zhang, M., Liu, Y., Sun, X., Lu, Q., Xiang, Z., Zhong, W., Liu, T., Ma, P. (2025). Alignment and Safety in Large Language Models: Safety Mechanisms, Training Paradigms, and Emerging Challenges. arXiv:2507.19672.

[11] **Fang, L.**, Yu, X., Cai, J., Chen, Y., Wu, S., Liu, Z., Yang, Z., Lu, H., Gong, X., Liu, Y., Ma, T., Ruan, W., Abbasi, A., Zhang, J., Wang, T., Latif, E., Liu, W., Zhang, W., Kolouri, S., Zhai, X., Zhu, D., Zhong, W., Liu, T., Ma, P. (2025). Knowledge Distillation and Dataset Distillation of Large Language Models: Emerging Trends, Challenges, and Future Directions. arXiv:2504.14772

[12] Latif, E., Zhou, Y., Guo, S., Gao, Y., Shi, L., Nayaaba, M., Lee, G., Zhang, L., Bewersdorff, A., **Fang, L.**, Yang, X., Zhao, H., Jiang, H., Lu, H., Li, J., Yu, J., You, W., Liu, Z., Liu, V. S., Wang, H., Wu, Z., Lu, J., Dou, F., Ma, P., Liu, N., Liu, T., Zhai, X. (2024). A Systematic Assessment of OpenAI o1-Preview for Higher Order Thinking in Education. arXiv:2410.21287

[13] **Fang, L.**, Lee, G., Zhai, X. (2023). Using GPT-4 to Augment Unbalanced Data for Automatic Scoring. arXiv:2310.18365

- Under Review

[14] **Fang, L.**, Chen, Y., Cai, J., Ma, P., Zhong, W. (2025). Multi-Teacher Bayesian Knowledge Distillation for Protein Localization Prediction with Uncertainty Quantification. (Under review, *Journal of the American Statistical Association (JASA).*)

[15] Lu, H., **Fang, L.**, Zeng, O., Ma, P., Zhong, W., Yuan, G. (2025). Optimal Gene Panel Selection for Targeted Spatial Transcriptomics Experiments. (Under review, *Nature Methods.*)

[16] Cai, J., Chen, Y., **Fang, L.**, Zhong, W., Ma, P., Yuan, G. (2025). SpaDiff: Denoising for Sequence-based Spatial Transcriptomics via Diffusion Process. (Under review, *Nature Communications.*)

[17] **Fang, L.**, Latif, E., Wang, T., Ma, P., Zhai, X. (2025). Generalizable and Efficient Automated Scoring with a Knowledge-Distilled Multi-Task Mixture-of-Experts. (Under review, *EAAI 2026: Educational Advances in Artificial Intelligence, co-located with AAAI 2026.*)

[18] Wang, T., **Fang, L.**, Zhong, W., Ma, P. (2025). GASDU: Gauss–Southwell Dynamic Update for Efficient LLM Fine-Tuning. (Under review, *International Conference on Learning Representations (ICLR 2026, Top machine learning conference).*)

[19] Cheng, H., Yu, X., Wu, S., **Fang, L.**, Cao, C., Zhang, J., Liu, T., Zhu, D., Zhong, W., Ma, P. (2025). DCMM-Transformer: Degree-Corrected Mixed-Membership Attention for Medical Imaging. (Under review, phase 2, *Association for the Advancement of Artificial Intelligence (AAAI 2026, Top machine learning conference).*)

[20] **Fang, L.**, Lu, H., Chen, Y., Zhong, W., Ma, P. (2024). Knowledge Cascade: Reverse Knowledge Distillation on Nonparametric Multivariate Functional Estimation. (Under revision, *Journal of Machine Learning Research (JMLR).*)

- In Preparation

[21] **Fang, L.**, Ma, P., Zhong, W. A Hierarchical Bayesian Framework for Reliable and Interpretable Knowledge Distillation.

[22] **Fang, L.**, Lu, H., Zhong, W., Ma, P. Reliable Reasoning Transfer via Confidence-Filtered CoT Distillation.

[23] **Fang, L.**, Liu, Y., Lu, H., Zhang. Y., Ma, P., Li, J., Zhong, W. Efficient and Reliable Medical Dialogue Generation with Limited Supervision.

[24] **Fang, L.**, Zhong, W., Ma, P. Domain Generalization via Double Knowledge Distillation.

[25] Lu, H., Cheng, H., **Fang, L.**, Ma, P., Zhong, W. Double Generative Learning for Causal Inference in High Dimensions.

Research Collaborations

- AI for STEM Education Center, UGA 2023 - Present
  AI in Education
- NewGen Psychometrics and Data Science Analytics Lab, UGA 2023 - Present
  Data Science in Educational Measurement
- Guo-Cheng Yuan Lab, Icahn School of Medicine at Mount Sinai 2022 - Present
  Spatial Transcriptomics and Multimodal Data Analysis
- Torrance Center for Creativity and Talent Development, UGA 2024 - 2025
  Nurturing and Development of Creative Potential

- Sensorweb Research Lab, UGA 2022 - 2024
  Cyber-physical System, Computing and Security

Awards & Grants

- Early Career Travel Award, American Statistical Association. 2025
- James L. Carmon Scholarship Award ($4000), University of Georgia. 2025
  Only one recipient each year.
- Joint ATD/AMPS Workshop Travel Grant, National Science Foundation. 2024
- Graduate School Conference Travel Grant, University of Georgia. 2024, 2025
- Best Beginning PhD Student, University of Georgia. 2024
  Two recipients each year.
- Honorable Mention, Student Employee of the Year team, University of Georgia. 2024
- Best Student Poster Presentation Award, Georgia Statistics Day. 2023
  "Georgia Statistics Day" is a statistics conference jointly organized by the Georgia Institute of Technology, Emory University, and the University of Georgia.
- ICSA Travel Grant, International Chinese Statistical Association. 2023
- Student Poster Presentation Award, Georgia Statistics Day. 2022
- Scholarship for Academic Excellence, Nankai University. 2019

Presentations

- Invited Conference/Workshop Presentations
  - Joint Statistical Meetings, Nashville, TN 2025
  - International Conference on Machine Learning, Vienna, Austria 2024
  - Joint Statistical Meetings, Portland, OR 2024
  - International Chinese Statistical Association China Conference, Wuhan, China 2024
  - National Council on Measurement in Education Annual Meeting, Philadelphia, PA 2024
  - Georgia Statistics Day, Georgia Institute of Technology, Atlanta, GA 2023
  - International Chinese Statistical Association China Conference, Chengdu, China 2023
  - Georgia Statistics Day, University of Georgia, Athens, GA 2022
  - Georgia Statistics Day, Emory University, Atlanta, GA 2021

Outreach/Supervision Activities

- Supervising high school student Orlando Zeng from Johns Creek High School 2024 - Present
  Deep Learning Methods for Spatial Transcriptomics
  Improving LLM Data Augmentation Quality for Automatic Scoring.
- Supervising undergraduate research of Yingchuan Zhang from Nankai University. 2024 - 2025
  Dynamic target updates for reinforcement learning.

Service Activities

- Reviewer for: IEEE International Conference on Computer Communications, International Conference on Artificial Intelligence in Education, Journal of Educational and Behavioral Statistics, and Journal of Educational Measurement.
- Session Chair for: Joint Statistical Meetings (JSM) 2025.

Teaching

- **Teaching Assistant**
  - STAT 4365/6365 Modern Statistical Programming Fall 2024
  - STAT 8530 Advanced Statistical Inference I Spring 2023
  - STAT 8900 Special Topics in Statistics Spring 2023
  - STAT 4220 Applied Experimental Designs Fall 2022
  - STAT 2000 Introductory Statistics (Lab) Summer 2022, Spring 2022
    STAT 2000 is an introductory course in UGA with 2 hours of lab a week. My primary duty was teaching lab sessions.
  - STAT 6315 Statistical Methods Summer 2022
  - STAT 8070 Statistical Computing II Spring 2021, 2022
  - STAT 4230/6230 Applied Regression Analysis Fall 2021
  - STAT 610 Introductory to Statistical Inference (Grader) Spring 2021
    University of Wisconsin-Madison, Madison, WI