YE TIAN

ye.t@columbia.edu
https://yet123.com
Google Scholar
Room 901, Social Work Building, 1255 Amsterdam Avenue, New York, NY 10027
Last update: April 11, 2025

Education

2019 – 2025 (expected) **Ph.D. in Statistics, Columbia University**
M.Phil. in Statistics awarded on 2022/10
M.A. in Statistics awarded on 2022/04
Advisor: Yang Feng (NYU Biostatistics) and Zhiliang Ying (Columbia Statistics)

2015 – 2019 **B.S. in Statistics, University of Science and Technology of China (USTC)**
Advisor: Weiping Zhang
Thesis title: *Random Sparse Projection Ensemble Classification*.

Industry Experience

2023/06-2023/08 **Applied Scientist Intern, PayStation intelligence (Pi) team, Amazon Inc.**
Conducting time series forecasting and anomaly detection on payment data

Papers

Published/Accepted

1. **Tian**, **Y.**, & Feng, Y. (2024). Neyman-Pearson Multi-class Classification via Cost-sensitive Learning. *Journal of the American Statistical Association*, (just-accepted), 1–15. [PDF].
2. **Tian**, **Y.**, Rusinek, H., Masurkar, A. V., & Feng, Y. (2024). ℓ_1-penalized Multinomial Regression: Estimation, Inference, and Prediction, with an Application to Risk Factor Identification for Different Dementia Subtypes. *Statistics in Medicine*, *43*(30), 5711–5747. [PDF].
3. **Tian**, **Y.**, Weng, H., & Feng, Y. (2024). Towards the Theory of Unsupervised Federated Learning: Non-asymptotic Analysis of Federated EM Algorithms. *The 41st International Conference on Machine Learning (ICML)*. PMLR. [PDF].
4. **Tian**, **Y.**, & Feng, Y. (2023a). Comments on: Statistical Inference and Large-scale Multiple Testing for High-dimensional Regression Models. *Test*, *32*(4), 1172–1176. [Link].
5. **Tian**, **Y.**, & Feng, Y. (2023b). RaSE: A Variable Screening Framework via Random Subspace Ensembles. *Journal of the American Statistical Association*, *118*(541), 457–468. [PDF].
6. **Tian**, **Y.**, & Feng, Y. (2023c). Transfer Learning under High-dimensional Generalized Linear Models. *Journal of the American Statistical Association*, *118*(544), 2684–2697. (*The 3rd most cited article on JASA in the last 3 years*). [PDF].
7. **Tian**, **Y.**, Rusinek, H., Feng, Y., & Vedvyas, A. (2022). Prediction of Dementia Subtypes by Machine Learning: Feasibility of Error Control. *Alzheimer's & Dementia*, *18*, e069240. [Poster].
8. **Tian**, **Y.**, & Feng, Y. (2021). RaSE: Random Subspace Ensemble Classification. *Journal of Machine Learning Research*, *22*(45), 1–93. [PDF].
9. **Tian**, **Y.**, & Zhang, W. (2019). THORS: An Efficient Approach for Making Classifiers Cost-sensitive. *IEEE Access*, *7*, 97704–97718. [PDF].

Preprint

1. Li, M., **Tian**, **Y.**, Feng, Y., & Yu, Y. (2024). Federated Transfer Learning with Differential Privacy. *arXiv preprint arXiv:2403.11343*. [PDF].
2. **Tian**, **Y.**, Gu, Y., & Feng, Y. (2023). Learning from Similar Linear Representations: Adaptivity, Minimaxity, and Robustness. *arXiv preprint arXiv:2303.17765 (minor revision at Journal of Machine Learning Research)*. [PDF].
3. **Tian**, **Y.**, Weng, H., Xia, L., & Feng, Y. (2022). Unsupervised Multi-task and Transfer Learning on Gaussian Mixture Models. *arXiv preprint arXiv:2209.15224*. [PDF].

Softwares

R Packages

- **pemultinom**: Author, maintainer
  Latest version: 0.1.1 [CRAN link] [Paper link]
- **mtlgmm**: Author, maintainer
  Latest version: 0.1.0 [CRAN link] [Paper link]
- **npcs**: Author, maintainer
  Latest version: 0.1.1 [CRAN link] [Paper link]
- **glmtrans**: Author, maintainer
  Latest version: 2.1.0 [CRAN link] [Paper link]
- **RaSEn**: Author, maintainer
  Latest version: 3.0.0 [CRAN link] [Paper link 1] [Paper link 2] [Paper link 3]

Other Languages

- Python
- C

Honors

2025/01 SLDS Student Paper Award
Statistical Learning and Data Science (SLDS) Section of the ASA

2024/08 SIAM Student Travel Award
Society for Industrial and Applied Mathematics (SIAM)

2023/10 Howard Levene Outstanding Teaching Award
Department of Statistics, Columbia University

2023/06 NESS Student Research Award
The 36th New England Statistics Symposium (NESS 2023), Boston

2023/04 IMS Hannan Graduate Student Travel Award

2022/12 Student Paper Award
2022 IMS International Conference on Statistics and Data Science (ICSDS), Florence, Italy

2022/11 Columbia ASGC Graduate Student Travel Award
The Arts and Sciences Graduate Council, Columbia University

2022/10 Columbia GSAS Student Conference Travel Award
Graduate School of Arts and Sciences, Columbia University

Teaching

Instructor

2024/05 **Short Course: An Introduction to the Statistical Foundations of Transfer Learning**
Co-instructor (with Yang Feng), at the 37th New England Statistics Symposium
Course website

2022 Summer **Calculus-based Introduction to Statistics (STAT1201)**. Undergrad-level.
Instructor, Columbia University
Course website

Teaching Assistant

2022 Fall **Time Series Analysis (GU4221/GR5221)**. Undergrad/master-level.
TA, Columbia University

2022 Spring **Introduction to Statistics (without calculus) (STAT1101)**. Undergrad-level.
TA, Columbia University

2024/2023/2021/2020 Spring **Statistical Machine Learning (GR5241)**. Master-level.
TA, Columbia University

2025 Spring
2024/2021/2020 Fall **Linear Regression Models (GU4205/GU5205)**. Undergrad/master-level.
TA, Columbia University

2023/2019 Fall **Introduction to Probability & Statistics (GU4001)**. Undergrad/master-level.
TA, Columbia University

2019 Spring **Probability Theory**. Undergrad-level.
TA, USTC

2018 Spring **Linear Algebra**. Undergrad-level.
TA, USTC

Presentations

2025/08 2025 Joint Statistical Meetings, *"Learning from Similar Linear Representations: Adaptivity, Minimaxity, and Robustness", SLDS Student Paper Award, topic-contributed talk*

2025/03 2025 ENAR Spring Meeting, *"Federated Transfer Learning with Differential Privacy", invited talk*

Department of Statistics, University of Illinois Urbana-Champaign, *"Transfer and Multi-task Learning: Statistical Insights for Modern Data Challenges", job talk*

2025/02 Department of Statistics, Pennsylvania State University, *"Transfer and Multi-task Learning: Statistical Insights for Modern Data Challenges", job talk*

Department of Statistics and Data Science, National Singapore University, *"Transfer and Multi-task Learning: Statistical Insights for Modern Data Challenges", job talk*

Department of Statistics, University of British Columbia, *"Transfer and Multi-task Learning: Statistical Insights for Modern Data Challenges", job talk*

Department of Statistics, University of Michigan, *"Transfer and Multi-task Learning: Statistical Insights for Modern Data Challenges", job talk*

Department of Statistics, University of Virginia, *"Transfer and Multi-task Learning: Statistical Insights for Modern Data Challenges", job talk*

Presentations (continued)

Department of Statistics, Ohio State University, *"Transfer and Multi-task Learning: Statistical Insights for Modern Data Challenges", job talk*

2025/01 Department of Statistics, University of California, Riverside, *"Transfer and Multi-task Learning: Statistical Insights for Modern Data Challenges", job talk*

Department of Statistics, University of California, Irvine, *"Transfer and Multi-task Learning: Statistical Insights for Modern Data Challenges", job talk*

Department of Statistics and Data Science, University of California, Los Angeles, *"Transfer and Multi-task Learning: Statistical Insights for Modern Data Challenges", job talk*

Department of Mathematics and Statistics, Boston University, *"Transfer and Multi-task Learning: Statistical Insights for Modern Data Challenges", job talk*

Department of Biostatistics, University of North Carolina - Chapel Hill, *"Transfer and Multi-task Learning: Statistical Insights for Modern Data Challenges", job talk*

Department of Statistics and Operations Research, University of North Carolina - Chapel Hill, *"Transfer and Multi-task Learning: Statistical Insights for Modern Data Challenges", job talk*

2024/12 Department of Applied and Computational Mathematics and Statistics, University of Notre Dame, *"Transfer and Multi-task Learning: Statistical Insights for Modern Data Challenges", job talk*

Division of Biostatistics and Health Data Sciences, University of Minnesota - Twin Cities, *"Transfer and Multi-task Learning: Statistical Insights for Modern Data Challenges", job talk*

Department of Statistics, Florida State University, *"Transfer and Multi-task Learning: Statistical Insights for Modern Data Challenges", job talk*

2024/11 Department of Statistics, Iowa State University, *"Regularized Fine-tuning in Representation Multi-task Learning: Adaptivity and Robustness", job talk*

2024/10 SIAM Conference on Mathematics of Data Science (MDS24), *"Neyman-Pearson Multi-class Classification via Cost-sensitive Learning", SIAM Student Travel Award, poster*

2024/08 2024 Joint Statistical Meetings, *"Learning from Similar Linear Representations: Adaptivity, Minimaxity, and Robustness", contributed talk*

2024/07 2024 IMS-China International Conference on Statistics and Probability, Ningxia University, China, *"Learning from Similar Linear Representations: Adaptivity, Minimaxity, and Robustness", invited talk*

2024/05 The 37th New England Statistics Symposium (NESS 2024), University of Connecticut, *"Neyman-Pearson Multi-class Classification via Cost-sensitive Learning",contributed talk*

2024/04 Foundations of Data Science Center Workshop, Data Science Institute, Columbia University, *"Neyman-Pearson Multi-class Classification via Cost-sensitive Learning", poster*

2024 Minghui Yu Memorial Conference, Department of Statistics, Columbia University, *"Neyman-Pearson Multi-class Classification via Cost-sensitive Learning", contributed talk*

Student seminar, Department of Statistics, Columbia University, *"Robust Unsupervised Multi-task and Transfer Learning on Gaussian Mixture Models", invited talk*

2024/01 Department of Statistics and Finance, School of Management, University of Science and Technology of China, *"Robust Unsupervised Multi-task and Transfer Learning on Gaussian Mixture Models", invited talk*

2023/12 Young Statistician Workshop, School of Mathematical Sciences, Shenzhen University, China, *"Robust Unsupervised Multi-task and Transfer Learning on Gaussian Mixture Models", invited talk*

2023/11 IMS Young Mathematical Scientists Forum – Statistics and Data Science, Department of Statistics and Data Science, National University of Singapore, *"Robust Unsupervised Multi-task and Transfer Learning on Gaussian Mixture Models", invited talk*

2023/08 2023 Joint Statistical Meetings, *"Neyman-Pearson Multi-class Classification via Cost-sensitive Learning", contributed talk*

Presentations (continued)

2023/06 The 36th New England Statistics Symposium (NESS 2023), Boston, *Student Research Award*, *"Learning from Similar Linear Representations: Adaptivity, Minimaxity, and Robustness"*, *invited talk*

2023/05 Statistical Foundations of Data Science and their Applications: A conference in celebration of Jianqing Fan's 60th Birthday, *"Robust Unsupervised Multi-task and Transfer Learning on Gaussian Mixture Models"*, *poster*

2023/04 2023 Minghui Yu Memorial Conference, Department of Statistics, Columbia University, *"Learning from Similar Linear Representations: Adaptivity, Minimaxity, and Robustness"*, *contributed talk*

Berkeley–Columbia Meeting in Engineering and Statistics, *"Learning from Similar Linear Representations: Adaptivity, Minimaxity, and Robustness"*, *invited talk*

Columbia Statistical Machine Learning Symposium, *"Learning from Similar Linear Representations: Adaptivity, Minimaxity, and Robustness"*, *poster and lightning talk*

2022/12 2022 IMS International Conference on Statistics and Data Science (ICSDS), Florence, Italy, *"Robust Unsupervised Multi-task and Transfer Learning on Gaussian Mixture Models"*, *Student paper award*, *invited talk*

2022/10 International Conference on Advances in Interdisciplinary Statistics and Combinatorics (AISC 2022), The University of North Carolina at Greensboro, *"Robust Unsupervised Multi-task and Transfer Learning on Gaussian Mixture Models"*, *contributed talk*

2022/07 The Alzheimer's Association International Conference (AAIC 2022), San Diego, *"Prediction of Dementia Subtypes by Machine Learning: Feasibility of Error Control"*, *poster*

2022/05 2nd Annual Graduate Student Research Conference (online), National Institute of Statistical Sciences (NISS), *contributed talk*

2022/04 2022 Minghui Yu Memorial Conference, Department of Statistics, Columbia University, *"Transfer Learning under High-dimensional Generalized Linear Models"*, *contributed talk*

Conference on Advances in Bayesian and Frequentist Statistics with a Celebration of the 80th Birthday of Professor William E. Strawderman, Department of Statistics, Rutgers University, *"Transfer Learning under High-dimensional Generalized Linear Models"*, *poster*

2021/04 2021 Minghui Yu Memorial Conference, Department of Statistics, Columbia University, *"RaSE: Random Subspace Ensemble Classification"*, *contributed talk*

2021/03 2021 Women in Science at Columbia Graduate Research Symposium, Columbia University, *"RaSE: Random Subspace Ensemble Classification"*, *poster*

Services

Journal Reviewer

2020-present Annals of Statistics (1), Journal of the Royal Statistical Society (Series B) (5), Technometrics (2), Bernoulli (1), Statistical Papers (2), Journal of the American Statistical Association (9), Journal of Machine Learning Research (3), Statistics and Computing (2), Biometrics (1), Stat (3), Neural Networks (2), Biometrika (1), Journal of Econometrics (1), Scientific Reports (1), Neurocomputing (1), Journal of Business and Economic Statistics (1), Computational Statistics and Data Analysis (1), Annals of Applied Statistics (2), Bioinformatics (1), Journal of Computational and Graphical Statistics (1)

Conference Reviewer

2020-present FODS 2020 (3), ICLR 2023 (1), AISTATS 2024 (6)

Services (continued)

Conference/Seminar Organizer

2024/08	Chair of the session "Applications in Statistical Learning and Data Science" at JSM 2024
2023/08	Chair of the session "Statistical Learning Theory" at JSM 2023
2021/09-2022/04	Co-organizer of Columbia Statistics Student Academic Job Search Workshop (with Arnab Auddy)
2021/08-2022/05	Co-organizer of Columbia Statistics Ph.D. Student Seminar (with Arnab Auddy)

References

Yang Feng
Professor
Department of Biostatistics,
School of Global Public Health,
New York University,
New York, NY, 10003, United States.
yang.feng@nyu.edu

Zhiliang Ying
Professor
Department of Statistics,
Columbia University,
New York, NY, 10027, United States.
zy9@columbia.edu

Yi Yu
Professor
Department of Statistics,
University of Warwick,
Coventry, CV4 7AL, United Kingdom.
yi.yu.2@warwick.ac.uk

Marco Avella Medina
Assistant Professor
Department of Statistics,
Columbia University,
New York, NY, 10027, United States.
marco.avella@columbia.edu