Yang Feng

Curriculum Vitae

September 08, 2026

Address: Department of Biostatistics
School of Global Public Health
New York University
New York, NY 10003, USA
Phone: +1 (212) 992-3810
Email: yang.feng@nyu.edu
Website: https://yangfengstat.github.io

Education

2006	B.S.	University of Science and Technology of China	Mathematics
2010	Ph.D.	Princeton University	Operations Research
		Thesis: High-dimensional Statistical Learning and Nonparametric Modeling	
		Advisor: Jianqing Fan	

Professional Experience

2023 – Present	Professor, Department of Biostatistics, School of Global Public Health, New York University
2022 – 2025	Ph.D. Program Director, Department of Biostatistics, School of Global Public Health, New York University

Professional Memberships

- American Statistical Association (Fellow and Lifetime Member)
- Institute of Mathematical Statistics (Fellow and Lifetime Member)
- International Statistical Institute (Elected Member)
- International Chinese Statistical Association (Lifetime Member)

Research Interests

- Theory, Methods, and Algorithms: Machine Learning, High-dimensional Statistics, Social Networks, Neyman-Pearson Classification, Nonparametric and Semi-parametric Statistics.
- Applications: Bioinformatics, Cancer Diagnosis, Computer Vision, Dementia, Electronic Health Records, Epidemiology, Econometrics, Finance, Genetics, Genomics.

Editorial Activities

- Editorial Board

2026 – 2028	**Reviews Editor**, The American Statistician
2026 – 2028	**Reviews Editor**, Journal of the American Statistical Association
2025 – 2026	Co-Editor, Statistical Learning and Data Science (SLADS) Special Issue on "Data, Network, and Knowledge Transfer"
2024 – Present	Associate Editor, Journal of Computational and Graphical Statistics
2023 – Present	Associate Editor, Annals of Applied Statistics
2021 – Present	Associate Editor, Journal of the American Statistical Association: Theory and Methods
2018 – Present	Associate Editor, Journal of Business and Economic Statistics

Honors & Awards

2025	Second most cited article published in the last three years in JASA
2025	Faculty of the Year, NYU School of Global Public Health
2024	Teaching Excellence Award, NYU School of Global Public Health
2023	Fellow, Institute of Mathematical Statistics
2022	Fellow, American Statistical Association
2016	NSF CAREER Award

Research Grants

2025–2026	LSE-NYU Joint Research Seed Fund NYU Lead Principal Investigator "Multi-task Representation Learning for Matrix Completion with Applications to Social and Behaviour Data"	$6,480
2023–2026	National Science Foundation (NSF) DMS-2324489 Principal Investigator "Collaborative Research: New Theory and Methods for High-Dimensional Multi-Task and Transfer Learning Inference"	$150,000
2022–2026	National Institutes of Health (NIH) 1R01NS122987-01A1 Investigator (PI: Matija Snuderl) "Inducing neural maturation in medulloblastoma by targeting EZH2"	$175,830
2021–2024	National Institutes of Health (NIH) 1R21AG074205-01 Principal Investigator "Multiclass classification under prioritized error control and specific error costs with applications to dementia classification"	$438,473
2016–2022	National Science Foundation (NSF) DMS-1554804 Principal Investigator "CAREER: Statistical inference of network and relational data"	$400,000
2013–2016	National Science Foundation (NSF) DMS-1308566 Principal Investigator "Nonparametric classification, tuning parameter selection, and asymptotic stability for high-dimensional data"	$129,980

Publications

Note: [1] represents co-first authors, * represents the corresponding author(s), underline represents students and junior collaborators, and † represents the author list is alphabetically ordered according to the mathematics convention. Google Scholar Citations (**5200+** times as of Apr 2026) https://scholar.google.com/citations?user=QXHb8CcAAAAJ&hl=en.

➤ Books

1. **Feng, Y.** (2024), R Programming: Zero to Pro, available at https://r02pro.github.io/.

➤ Peer-Reviewed Publications (Methodology and Theory)

2. Tian, Y., Weng, H., Xia, L., and **Feng, Y.*** (2026), Robust Unsupervised Multi-task and Transfer Learning on Gaussian Mixture Models[1], *Journal of the American Statistical Association*, to appear.
3. He, Y., Sun, J., Tian, Y., Ying, Z., and **Feng, Y.*** (2025), Semiparametric Modeling and Analysis for Longitudinal Network Data, *Annals of Statistics*, 53(4), 1406-1430.
4. Tian, Y. and **Feng, Y.*** (2025), Neyman-Pearson Multi-class Classification via Cost-sensitive Learning, *Journal of the American Statistical Association*, 120(550), 1164-1177.
5. Huang, S., Sun, J., and **Feng, Y.*** (2024), PCABM: Pairwise Covariates-Adjusted Block Model for Community Detection, *Journal of the American Statistical Association*, 119(547), 2092-2104.
6. He, Y., Wu, R., Zhou, Y., and **Feng, Y.*** (2024), DDAC-SpAM: A Distributed Algorithm for Fitting High-dimensional Sparse Additive Models with Feature Division and Decorrelation, *Journal of the American Statistical Association*, 119(547), 1933-1944.

[1] An earlier version won the student travel award in the first IMS International Conference on Statistics and Data Science for advisee Ye Tian.

7. Tian, Y. and **Feng, Y.*** (2023), Transfer Learning under High-dimensional Generalized Linear Models, *Journal of American Statistical Association*, 118(544), 2684-2697.
8. Tian, Y. and **Feng, Y.*** (2023). RaSE: A variable screening framework via random subspace ensembles. *Journal of the American Statistical Association*, 118, 457-468.
9. Yuan, M., Liu, R., **Feng, Y.***, and Shang, Z. * (2022), Testing Community Structures for Hypergraphs, *Annals of Statistics*, 50(1): 147-169.
10. Demirkaya, E.[1], **Feng, Y.**[1], Basu, P., and Lv, J. (2022), Large-scale model selection in misspecified generalized linear models, *Biometrika*, 109(1), 123-136.
11. Tong, X.[1], **Feng, Y.**[1], and Li, J. (2018), Neyman-Pearson classification algorithms and NP receiver operating characteristic, *Science Advances*[2], Vol. 4, no. 2, eaao1659.
12. † Fan, J., **Feng, Y.*** and Song, R. (2011), Nonparametric independence screening in ultra-high dimensional additive models, *Journal of the American Statistical Association*, 106, 544-557.
13. † Fan, J., **Feng, Y.***, and Wu, Y. (2009), Network exploration via the adaptive LASSO and SCAD penalties, *The Annals of Applied Statistics*, 3, 521-541.

➤ Peer-Reviewed Publications (Applications)

14. Groff, K. J., Patel, R. V., **Feng, Y.**, Ghosh, H. S., Chavez, M. A. M., O'Brien, J., Chen, W. C., Nitturi, V., Save, A. V., Youngblood, M. W., ... (2025). The effect of TERT promoter mutation on predicting meningioma outcomes: a multi-institutional cohort analysis. *The Lancet Oncology*, 26(9), 1178-1190.

Software

The following software packages have been downloaded over **500K** times as of Apr 2026, according to CRAN.

1. **SIS**, an R package for (Iterative) Sure Independence Screening for Generalized Linear Models and Cox Proportional Hazards Models, available at CRAN.
2. **glmtrans**, an R package for implementing transfer learning under high-dimensional generalized linear models, along with the construction of confidence intervals based on a new debiasing technique, available at CRAN.
3. **RaSEn**, an R package for implementing a new model-free ensemble classification framework, RaSE algorithm, for the sparse classification problem, available at CRAN.
4. **nproc**, Given a sample of class 0 and class 1 and a classification method, the package generates the corresponding Neyman-Pearson classifier with a pre-specified type-I error control and Neyman-Pearson Receiver Operator Curve, available at CRAN. A vignette for demonstration is available here.
5. **r02pro**, a companion R package including the interactive exercises and datasets for the book "R Programming: Zero to Pro", available at CRAN.
6. **RAMP**, an R package for fitting the entire solution path for high-dimensional regularized generalized linear models with interaction effects under the strong heredity constraint, available at CRAN.
7. **mtlgmm**, an R package for unsupervised multi-task and transfer learning on Gaussian mixture models, available at CRAN.
8. **npcs**, an R package implementing Neyman-Pearson multi-class classification via cost-sensitive learning, available at CRAN.

[2] A high-profile multidisciplinary journal in the Science series published by the American Association for the Advancement of Science. 2018 JCR Impact Factor: 12.804

9. **GeoERM**, Python implementation of Geometry-Aware Multi-Task Representation Learning on Riemannian Manifolds, available at GitHub.
10. **FedGrEM**, implementation of a Federated Gradient EM Algorithm for heterogeneous mixture models with robustness against adversarial attacks, available at GitHub.
11. **RL-MTL-TL**, Python implementation for representation-learning based multi-task learning and transfer learning, available at GitHub.
12. **PCABM**, R implementation of the Pairwise Covariate-Adjusted Stochastic Block Model for community detection in networks with nodal information, available at GitHub.
13. **DLSM**, implementation of a Dynamic Latent Space Model for social and longitudinal network analysis, available at GitHub.
14. **multi_net**, implementation of spectral clustering via adaptive layer aggregation for multi-layer networks, available at GitHub.
15. **nproc (Python)**, Python implementation of the Neyman-Pearson Classification umbrella algorithm, available at GitHub.
16. **MIRL**, an R package for multiple imputation random lasso, designed to solve the high-dimensional variable selection problem with missing data, available at CRAN.
17. **fcd**, an R package for implementing the fused community detection method, available at CRAN.
18. **FusedPCA**, an R package for implementing the fused principle component analysis method, available at CRAN.
19. **FANS**, Matlab code for implementing the FANS (Feature Augmentation via Nonparametrics and Selection) classification method for high-dimensional data, available at GitHub.
20. **CLBIC**, R code for implementing Composite Likelihood BIC for selecting the number of communities, available at GitHub.
21. **pgraph**, Implements a general framework for creating dependency graphs using projection. Both lasso and sparse additive model projections are implemented. Both Pearson correlation and distance covariance can be used to generate the graph, available at CRAN.
22. **apple**, an R package for calculating the Approximate Path for Penalized Likelihood Estimators for Generalized Linear Models, available at CRAN
23. **ROAD**, a Matlab package designed for the Regularized Optimal Affine Discriminant method for high-dimensional classification, available at GitHub.
24. **xtab**, An R function for generating LaTex tables from a data matrix, available at GitHub.

References

➤ Available upon request.