Yang Feng

Curriculum Vitae

November 2024

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://yangfeng.hosting.nyu.edu/

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
2019 – 2023	Associate Professor, Department of Biostatistics, School of Global Public Health, New York University
2022 – Present	Ph.D. Program Director, Department of Biostatistics, School of Global Public Health, New York University
2022 – Present	Affiliate Faculty, Center for Data Science (CDS), New York University
2021 – Present	Affiliate Faculty, Center for Practice and Research at the Intersection of Information, Society, and Methodology (PRIISM), New York University
2016 – 2019	Associate Professor, Department of Statistics, Columbia University
2014 – 2016	Howard Levene Assistant Professor, Department of Statistics, Columbia University
2010 – 2014	Assistant Professor, Department of Statistics, Columbia 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.

Editorial Activities

- Editorial Board

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
2023 – Present	Editorial Board of Reviewers, Journal of Machine Learning Research
2014 – 2023	Associate Editor, Statistica Sinica
2020 – 2023	Associate Editor, Stat
2013 – 2022	Associate Editor, Statistical Analysis and Data Mining, The ASA Data Science Journal
2015 – 2018	Associate Editor, Computational Statistics and Data Analysis

Honors & Awards

2024	NYU School of Global Public Health Teaching Excellence Award
2023	Fellow, Institute of Mathematical Statistics
2022	Fellow, American Statistical Association
2021	NYU University Research Challenge Fund
2020	NYU Curriculum Development Challenge Fund Award
2017	Elected Member, International Statistical Institute
2016	NSF CAREER Award
2015	Lenfest Junior Faculty Development Award
2012	New World Mathematics Award (Silver Prize)
2010	Wallace Memorial Honorific Fellowship (the highest award for a Princeton graduate student)
2009	Laha Award from the Institute of Mathematical Statistics (IMS)
2007	The Gordon B. and Nancy R. Stewart, Jr. Fellowship, Princeton University
2005	Samsung Scholarship (11 out of about 7200 undergraduate USTC students)

Research Grants

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
2022–2027	National Institutes of Health (NIH) 1R01CA268932-01A1 Investigator (PI: Jennifer Cantrell) "Using Multiphase Optimization Strategy (MOST) to Optimize a Cost-effective, Sustainable and Scalable Smoking Cessation Package for Smokers in HIV Clinical Care"	$148,438
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
2022–2023	National Institutes of Health (NIH) 1R56NS122987-01 Investigator (PI: Matija Snuderl) "Inducing neural maturation in medulloblastoma by targeting EZH2"	$53,300
2021–2022	NYU University Research Challenge Fund	$12,000

	Principal Investigator "Prioritized Multiclass Classification with Applications to Brain Tumor Diagnosis"	
2020–2021	NYU Curriculum Development Challenge Fund Principal Investigator "Interactive Teaching and Learning of Statistical Programming and Machine Learning in R"	$4,500
2020-2021	National Science Foundation (NSF) Grant DEB-2034022 Co-Principal Investigator (PI: Joshua Epstein) "RAPID: Behavioral Epidemic Modeling For COVID-19 Containment"	$131,954
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 (**4000+** times as of June 2024) 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)

1. Tian, Y. and **Feng, Y.*** (2024) Neyman-Pearson Multi-class Classification via Cost-sensitive Learning, *Journal of American Statistical Association*, to appear.
2. Tian, Y., Rusinek, H., Masurkar, A., and **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*, to appear.
3. Bi, F., Zhu, J., and **Feng, Y.*** (2024) Multi-label Random Subspace Ensemble Classification, *Journal of Computational and Graphical Statistics*, to appear.
4. Tian, Y., Weng, H., and **Feng, Y.*** (2024). Unsupervised Federated Learning: A Federated Gradient EM Algorithm for Heterogeneous Mixture Models with Robustness against Adversarial Attacks. *ICML 2024*.
5. Liu, Q. and **Feng, Y.** (2024), Machine Collaboration, *Stat*, vol. 13, no 1, p. e661.
6. Huang, S., Sun, J., and **Feng, Y.*** (2023), Pairwise covariates-adjusted block model for community detection, *Journal of American Statistical Association*, to appear.
7. He, Y., Wu, R., Zhou, Y., and **Feng, Y.*** (2023), Distributed Feature Selection for High-dimensional Additive Models, *Journal of American Statistical Association*, to appear.
8. Tian, Y. and **Feng, Y.*** (2023), Transfer Learning under High-dimensional Generalized Linear Models, *Journal of American Statistical Association*, 118(544), 2684-2697.
9. Tian, Y. and **Feng, Y.*** (2023). RaSE: A variable screening framework via random subspace ensembles. *Journal of the American Statistical Association*, 118, 457-468.
10. He, Y., **Feng, Y.**, and Song, X. (2023), Variable selection for high dimensional generalized linear model with block-missing data, *Scandinavian Journal of Statistics*, to appear.

11. Shi, Y., Li, H., Wang, C., Chen, J., Jiang, H., Shih, Y. C. T., ..., **Feng, Y.**, and Liu, L. (2023). A flexible quasi-likelihood model for microbiome abundance count data. *Statistics in medicine*, 42(25), 4632-4643.
12. Yuan, M., Liu, R., **Feng, Y.***, and Shang, Z. * (2022), Testing Community Structures for Hypergraphs, *Annals of Statistics*, 50(1): 147-169.
13. 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.
14. Yousuf, K. and **Feng, Y.*** (2022), Targeting predictors via partial distance correlation with applications to financial forecasting, *Journal of Business of Economic Statistics*, 40(3), 1007-1019
15. Huang, S., Weng, H., and **Feng, Y.*** (2022), Spectral clustering via adaptive layer aggregation for multi-layer networks, *Journal of Computational and Graphical Statistics*, to appear.
16. **Feng, Y.**, Liu, Q., Yao, Q., and Zhao, G. (2022), Model Averaging Estimation for Nonlinear Regression Models, *Journal of Business of Economic Statistics*, 40(2), 785-798.
17. Yuan, M., **Feng, Y.**, and Shang, Z. (2022), A Likelihood-Ratio Type Test for Stochastic Block Models with Bounded Degrees, *Journal of Statistical Planning and Inference*, 219, 98-119.
18. Weng, H. and **Feng, Y.*** (2022), Community detection with nodal information: likelihood and its variational approximation, *Stat*, 11(1), e428.
19. Tian, Y. and **Feng, Y.*** (2021). RaSE: Random subspace ensemble classification, *Journal of Machine Learning Research*, 22(45): 1-93.
20. Zhu, J. and **Feng, Y.*** (2021), Super RaSE: Super Random Subspace Ensemble Classification, *Journal of Risk and Financial Management*, 14(12), 612.
21. **Feng, Y.**, Zhou, M., and Tong, X. (2021). Imbalanced classification: A paradigm-based review. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, 14(5), 383-406.
22. Zhang, H., Chen, J., **Feng, Y.**, Wang, C., Li, H., and Liu, L. (2021). Mediation Effect Selection in High-Dimensional and Compositional Microbiome Data. *Statistics in Medicine*, 40(4):885-896
23. Tong, X., Xia, L., Wang, J., and **Feng, Y.*** (2020), Neyman-Pearson classification: parametrics and sample size requirement. *Journal of Machine Learning Research*, 21(12), 1-48.
24. †Fan, J., **Feng, Y.***, and Xia, L.* (2020), A Projection Based Conditional Dependence Measure with Applications to High-dimensional Undirected Graphical Models, *Journal of Econometrics*, 218, 119-139.
25. Weng, H. and **Feng, Y.*** (2020). On the estimation of correlation in a binary sequence model. *Journal of Statistical Planning and Inference*, 207, 123-137.
26. **Feng, Y.*** and Liu, Q. (2020). Nested Model Averaging on Solution Path for High-dimensional Linear Regression. *Stat*, 9(1), e317.
27. **Feng, Y.**, Liu, Q., and Okui, R. (2020), On the Sparsity of Mallows' Model Averaging Estimator, *Economic Letters*, 187, 108916.
28. Weng, H. and **Feng, Y.*** (2020), On the estimation of correlation in a binary sequence model, *Journal of Statistical Planning and Inference*, 207, 123-137.
29. Liu, J., Psarakis, E., **Feng, Y.***, and Stamos, I. (2019), A Kronecker product model for repeated pattern detection on 2D urban images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*[1], 41, 2266-2272.
30. **Feng, Y.*** and Yu, Y. (2019), The restricted consistency property of leave-n_v-out cross-validation for high-dimensional variable selection, *Statistica Sinica*, 29, 1607-1630.

[1] A leading journal in computer vision. 2018 Impact Factor: 17.73.

31. Weng, H., **Feng, Y.*** and Qiao, X. (2019), Regularization after retention in ultrahigh dimensional linear regression models[2], *Statistica Sinica*, 29, 387-407.
32. Hao, N.[1], **Feng, Y.**[1] and Zhang, H.H. (2018), Model selection for high dimensional quadratic regression via regularization, *Journal of the American Statistical Association*, 113, 615-625.
33. Tong, X.[1], **Feng, Y.**[1], and Li, J. (2018), Neyman-Pearson classification algorithms and NP receiver operating characteristic, *Science Advances*[3], Vol. 4, no. 2, eaao1659.
34. Saldaña, D. and **Feng, Y.*** (2018), SIS: an R package for sure independence screening in ultrahigh dimensional statistical models, *Journal of Statistical Software*[4], 83, 2, 1-25.
35. **Feng, Y.***, Li, T., and Ying, Z. (2018), Likelihood adaptive modified penalties, *Applied Stochastic Models in Business and Industry*, 35, 330-353.
36. **Feng, Y.**, Wu, Y., and Stefanski, L. A. (2018), Nonparametric independence screening via favored smoothing bandwidth, *Journal of Statistical Planning and Inference*, 197, 1-14.
37. Gao, X., **Feng, Y.***, (2018), Penalized weighted least absolute deviation regression, *Statistics and Its Interface*, 11, 79-89.
38. Saldaña, D., Yu, Y. and **Feng, Y.*** (2017), How many communities are there? *Journal of Computational and Graphical Statistics*, 26, 171-181.
39. Ji, J., He, D., **Feng, Y.**, He, Y., Xue, F., Xie, L. (2017), JDINAC: joint density-based nonparametric differential interaction network analysis and classification using high-dimensional sparse omics data, *Bioinformatics*, 33, 3080-3087.
40. Li, T., Chen, K., **Feng, Y.***, and Ying, Z. (2017), Binary switch portfolio, *Quantitative Finance*, 17, 763-780.
41. Gao, X., Ahmed, E., and **Feng, Y.** (2017), Post selection shrinkage estimation for high dimensional data analysis (with discussion), *Applied Stochastic Models in Business and Industry*, 33, 97-135.
42. Zhao, A., **Feng, Y.**, Wang, L., and Tong, X. (2016), Neyman-Pearson classification under high-dimensional settings, *Journal of Machine Learning Research*, 17, 1-39.
43. Liu, Y., Wang, Y. *, **Feng, Y.** *, and Wall, Melanie M. (2016), Variable selection and prediction with incomplete high-dimensional data, *The Annals of Applied Statistics*, 10, 418-450.
44. † Fan, J., **Feng, Y.***, Jiang, J., and Tong, X. (2016), A classification rule of feature augmentation via nonparametrics and selection (FANS) in high dimensional space, *Journal of the American Statistical Association*, 111, 275-287
45. Fang, Y., Wang, B. and **Feng, Y.** (2016), Tuning parameter selection in regularized estimations of large covariance matrices, *Journal of Statistical Computation and Simulation*, 86, 494-509.
46. Tong, X., **Feng, Y.**, and Zhao, A. (2016), A survey on Neyman-Pearson classification and suggestions for future research[5], *WIREs Computational Statistics*, 8, 64-81.
47. Ma, W., **Feng, Y.***, Chen, K. and Ying, Z. (2015), Functional and parametric estimation in a semi- and nonparametric model with application to mass-spectrometry data, *International Journal of Biostatistics*, 11, 285-303.
48. Yu, Y. and **Feng, Y.*** (2014), Modified cross-validation for lasso penalized high-dimensional linear models, *Journal of Computational and Graphical Statistics*, 23, 1009-1027.
49. Yu, Y. and **Feng, Y.*** (2014), APPLE: approximate path for penalized likelihood estimators, *Statistics and Computing*, 24, 803-819.

[2] An earlier version of this paper won the 2014 ICSA travel award for advisee Haolei Weng

[3] 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

[4] 2018 SJR Impact Factor: 22.7

[5] Top 10 accessed 2016-2017 articles from WIREs Computational Statistics, http://wires.wiley.com/WileyCDA/WiresCollection/id-51.html

50. † Fang, Y., **Feng, Y.** and Yuan, M. (2014), Regularized principal components of heritability, *Computational Statistics*, 29, 455-465.
51. † Fan, J., **Feng, Y.**, and Tong, X. (2012), A road to classification in high dimensional space: the regularized optimal affine discriminant, *Journal of the Royal Statistical Society Series B*, 74, 745-771.
52. † 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.
53. † Fan, J., **Feng, Y.** and Niu, Y. (2010), Nonparametric estimation of genewise variance for microarray data, *The Annals of Statistics*, 38, 2723-2750.
54. Fan, J., Wu, Y., and **Feng, Y.** (2009), Local quasi-likelihood with a parametric guide, *The Annals of Statistics*, 37, 4153-4183.
55. † 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.
56. **Feng, Y.**, Ma, W., Wang, Z., Ying, Z. and Yang, Y. (2009), Alignment of protein mass spectrometry data by semiparametric random shifting models, *Statistics and Its Interface*, 2, 329-340.

➤ Peer-Reviewed Publications (Applications)

57. Fang, C. S., Wang, W., Schroff, C., Movahed-Ezazi, M., Vasudevaraja, V., Serrano, J., ... **Feng, Y.**, & Snuderl, M. (2024). Racial Distribution of Molecularly Classified Brain Tumors. Neuro-Oncology Advances, vdae135.
58. Domingo-Relloso, A., **Feng, Y.**, Rodriguez-Hernandez, Z., Haack, K., Cole, S. A., Navas-Acien, A., ... & Bermudez, J. D. (2024). Omics feature selection with the extended SIS R package: identification of a body mass index epigenetic multi-marker in the Strong Heart Study. *American Journal of Epidemiology*, kwae006.
59. Galbraith, K., Garcia, M., Wei, S., Chen, A., Schroff, C., Serrano, J., ... **Feng, Y.**, & Snuderl, M. (2024). Prognostic value of DNA methylation subclassification, aneuploidy, and CDKN2A/B homozygous deletion in predicting clinical outcome of IDH mutant astrocytomas. *Neuro-oncology*, noae009.
60. Walsh, B. C., Zhu, J., **Feng, Y.**, Berkowitz, K. A., Betensky, R. A., Nunnally, M. E., & Pradhan, D. R. (2023). Simulation of New York City's Ventilator Allocation Guideline During the Spring 2020 COVID-19 Surge. *JAMA network open*, 6(10), e2336736-e2336736.
61. Liu, E. K., Vasudevaraja, V., Sviderskiy, V. O., **Feng, Y.**, Tran, I., Serrano, J., ... & Snuderl, M. (2022). Association of hyperglycemia and molecular subclass on survival in IDH-wildtype glioblastoma. *Neuro-Oncology Advances*, 4(1), vdac163.
62. Garcia, M. R., **Feng, Y.**, Vasudevaraja, V., Galbraith, K., Serrano, J., Thomas, C., ... & Snuderl, M. (2022). Clinical, Pathological, and Molecular Characteristics of Diffuse Spinal Cord Gliomas. *Journal of Neuropathology & Experimental Neurology*, 81(11), 865-872.
63. Tang, F.[1], **Feng, Y.***[1], Chiheb, H., Fan, J. (2021), The Interplay of Demographic Variables and Social Distancing Scores in Deep Prediction of U.S. COVID-19 Cases, *Journal of the American Statistical Association*, 116(534), 492-506.
64. Chandarana, H., Pisuchpen, N., Krieger, R., Dane, B., Mikheev, A., **Feng, Y.**, Kambadakone, A., and Rusinek, H. (2021). Association of body composition parameters measured on CT with risk of hospitalization in patients with Covid-19. *European Journal of Radiology*, 145, 110031.
65. Lin, L.H., Allison, D.H., **Feng, Y.**, Jour, G., Park, K., Zhou, F., Moreira, A.L., Shen, G., Feng, X., Sabari, J. and Velcheti, V. (2021). Comparison of solid tissue sequencing and liquid biopsy accuracy in identification of clinically relevant gene mutations and rearrangements in lung adenocarcinomas. *Modern Pathology*, 34(12), pp.2168-2174.

66. Black, M.A., Shen, G., Feng, X., Beltran, W.F.G., **Feng, Y.**, Vasudevaraja, V., Allison, D., Lin, L.H., Gindin, T., Astudillo, M. and Yang, D. (2021). Analytical performance of lateral flow immunoassay for SARS-CoV-2 exposure screening on venous and capillary blood samples. *Journal of Immunological Methods*, 489, p.112909.
67. Chandarana, H., Dane, B., Mikheev, A., Taffel, M. T., **Feng, Y.**, and Rusinek, H. (2021). Visceral adipose tissue in patients with COVID-19: risk stratification for severity. *Abdominal Radiology*, 46(2):818-825.
68. Betensky, R. A. and **Feng, Y.** (2020). Accounting for incomplete testing in the estimation of epidemic parameters. *International Journal of Epidemiology*, 49(5), 1419-1426.
69. Fourati, S., Talla, A., Mahmoudian, M., Burkhart, J. G., Klen, R., Henao, R., ... & Sieberts, S. K. (2018). A crowdsourced analysis to identify ab initio molecular signatures predictive of susceptibility to viral infection. *Nature Communication*, 9(1), 1-11.
70. MAQC-II Consortium (2010), MAQC-II Project: A comprehensive survey of common practices for the development and validation of microarray-based predictive models, *Nature Biotechnology*, 28, 827-841.

➤ Book Chapters

71. **Feng, Y.** and Yu, M., (2017), Regularization after marginal learning for ultra-high dimensional regression models, *Big and Complex Data Analysis*, 3-28.
72. † Fan, J., **Feng, Y.** and Wu, Y. (2010), Ultrahigh dimensional variable selection for Cox's proportional hazards model, *IMS Collection, Borrowing Strength: Theory Powering Applications - A Festschrift for Lawrence D. Brown*, 6, 70-86.

➤ Invited Discussions

73. Sun, J. and **Feng, Y.*** (2024). Discussion of "Root and community inference on the latent growth process of a network" by Crane and Xu. *Journal of the Royal Statistical Society Series B*.
74. Tian, Y. and **Feng, Y.*** (2023). Comments on: Statistical inference and large-scale multiple testing for high-dimensional regression models. Test, 32(4), 1172-1176.
75. Weng, H. and **Feng, Y.*** (2022), Discussion of "Cocitation and Coauthorship Networks of Statisticians", *Journal of Business and Economic Statistics*, 40 (2), 486-490.
76. **Feng, Y.** (2017), Discussion of "Random-projection ensemble classification" by T. Cannings and R. Samworth, *Journal of the Royal Statistical Society Series B*, 79, 1011.
77. † Fan, J. and **Feng, Y.*** (2009), Discussion of "Nonparametric prediction in measurement error models" by R. J. Carroll, A. Delaigle, and P. Hall, *Journal of the American Statistical Association*, 104, 1003-1007.

➤ Media

78. **Feng, Y.** and Zhang, X. (2020), Statistical evidence social distancing is working: Look at the effect on new coronavirus cases over time, *New York Daily News*, Apr 2, 2020.

➤ Manuscripts

79. Heng, S., Zhang, J., and **Feng, Y.*** (2023). Design-Based Causal Inference with Missing Outcomes: Missingness Mechanisms, Imputation-Assisted Randomization Tests, and Covariate Adjustment. manuscript.
80. Li, M., Tian, Y., Feng, Y., & Yu, Y. (2024). Federated Transfer Learning with Differential Privacy. arXiv preprint arXiv:2403.11343.
81. He, Y., Sun, J., Tian, Y., Ying, Z., and **Feng, Y.*** (2023), Semiparametric Modeling and Analysis for Longitudinal Network Data, manuscript.

82. Tian, Y., Gu, Y., and **Feng, Y.*** (2023), Learning from Similar Linear Representations: Adaptivity, Minimaxity, and Robustness, manuscript.
83. Tian, Y., Weng, H., Xia, L., and **Feng, Y.*** (2022), Unsupervised Multi-task and Transfer Learning on Gaussian Mixture Models[6], under review.
84. **Feng, Y.**, Tong, X., and Xin, W. (2022), Targeted Crisis Risk Control: Neyman-Pearson Approach, under review.

➤ Ph.D. Thesis

85. **Feng, Y.** (2010) High-dimensional statistical learning and nonparametric modeling, Princeton University.

Software

The following software packages have been downloaded over **390,000** times as of May 2024, 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. **MIRL**, an R package for multiple imputation random lasso, designed to solve the high-dimensional variable selection problem with missing data, available at CRAN.
8. **fcd**, an R package for implementing the fused community detection method, available at CRAN.
9. **fusedPCA**, an R package for implementing the fused principle component analysis method, available at CRAN.
10. **FANS**, Matlab code for implementing the FANS (Feature Augmentation via Nonparametrics and Selection) classification method for high-dimensional data, available at GitHub.
11. **CLBIC**, R code for implementing Composite Likelihood BIC for selecting the number of communities, available at GitHub.
12. **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.
13. **apple**, an R package for calculating the Approximate Path for Penalized Likelihood Estimators for Generalized Linear Models, available at CRAN

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

14. **ROAD**, a Matlab package designed for the Regularized Optimal Affine Discriminant method for high-dimensional classification, available at GitHub.
15. **xtab**, An R function for generating LaTex tables from a data matrix, available at GitHub.

References

➤ Available upon request.