Guo Yu

University of Washington
Department of Statistics   Phone: (607) 279-4296
Padelford Hall B-316   Email: gy63@uw.edu
Seattle, WA 98105

Research Interests

Statistical machine learning; high-dimensional statistics; multiple testing; multi-view data analysis

Academic Position

Postdoctoral Research Associate, Department of Statistics, University of Washington, July, 2018 -

Advisors: Daniela Witten, University of Washington

Jacob Bien, University of Southern California

Education

Ph.D. Statistics, Cornell University, 2018.

Advisor: Jacob Bien   GPA:4.12/4.30

M.S. Statistics, Cornell University, 2016.

B.S. Mathematics, Zhejiang University, 2013.

Publications

Statistical methodology:

Yu, G., Witten, D. M., and Bien, J. (2019) **Controlling Costs: Feature Selection on a Budget**. *Submitted*. https://arxiv.org/abs/1910.03627.

Yu, G., Bien, J., and Tibshirani, R. J. (2019) **Reluctant Interaction Modeling**. *Submitted to Journal of the American Statistical Association: Theory and Methods*. http://arxiv.org/abs/1907.08414

Yu, G., Bien, J., and Witten, D. M. (2019) **Discussion of "CARS: Covariate assisted ranking and screening for large-scale two-sample inference" by Cai et al**. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 81(2), 187-234*.

Yu, G., Bien, J. (2019) **Estimating the Error Variance in a High-dimensional Linear Model**. *Biometrika 106(3): 533-546*.

Yu, G., Bien, J. (2017) **Learning Local Dependence In Ordered Data**. *Journal of Machine Learning Research (JMLR). 18(42), 1-60*.

Applications:

Gao, M., Zhang, Z., **Yu, G.**, Arik, S., Davis, L. S. , and Pfister,. T. (2019) **Consistency-based semi-supervised active learning: Towards minimizing labeling budget**. *Submitted to International Conference on Learning Representations (ICLR)*.

Zhang, Q., **Yu, G.**, Guo, CX., Dang, Y., Swanson, N., Yang, X., Yao, R., Chintalapati, M., Krishnamurthy, A., Anderson, T. (2018). **Deepview: Virtual Disk Failure Diagnosis and Pattern Detection for the IaaS Service**. *15th USENIX Symposium on Networked Systems Design and Implementation (NSDI'18)*.

R Packages on CRAN

`sprintr`: R package for Sparse Reluctant Interaction Modeling (2019).

`natural`: R package for Estimating the Error Variance in a High-Dimensional Linear Model (2017).

`varband`: R package for Variable Banding of Large Precision Matrices (2016).

Experience

Research Intern at Microsoft AI & Research, Redmond, Washington (Summer 2017).

Award

2018 Outstanding Graduate Teaching Assistant Award, Cornell University

2016 Graduate Research Travel Award, Cornell University

Teaching

University of Washington

Guest Lecture

STAT 435, Introduction to Statistical Machine Learning (March 4th, 2019)

BIOSTAT 546, Machine Learning for Biomedical and Public Health Big Data (March 4th, 2019)

Cornell University

Teaching Assistant

STSCI 4090, Theory of Statistics (Spring 2018) (2018 Outstanding Graduate Teaching Assistant Award)

STSCI 4550, Applied Time Series (Spring 2017, Spring 2016)

STSCI 4740, Data Mining and Machine Learning (Fall 2016)

STSCI 4030, Linear Models with Matrices (Fall 2015)

ILRST 5100, Statistical Methods for the Social Sciences I (Fall 2014, Spring 2015)

ILRST 2100, Introductory Statistics (Spring 2014)

MATH 1910, Calculus for Engineers (Fall 2013)

Presentation

Joint Statistical Meetings, Denver, CO (July 31, 2019)

Symposium on Data Science and Statistics 2019, Bellevue, WA (May 31, 2019)

University of Washington Machine Learning Seminar, Seattle, WA (January 22, 2019)

Joint Statistical Meetings, Vancouver, BC, Canada (August 2, 2018)

Joint Statistical Meetings, Chicago, IL (August 3, 2016)

Cornell Department of Statistics Student Seminar, Ithaca, NY (December 7, 2015)

Cornell Department of Statistics Student Seminar, Ithaca, NY (April 15, 2015)

Professional Service

Referee for *Journal of the American Statistical Association: Theory and Methods*, *Biometrika*, *Machine Learning*, *American Statistician*, *Journal of Computational and Graphical Statistics*, *Statistics in Medicine*.