Yuting Ye

Biostatistics at UC Berkeley
Evan 439, UC Berkeley

E-mail: yeyt@berkeley.edu
Phone: 1(310)913-8862

Research interests

- **Machie Learning**: Theory of Nonconvex Learning, Graph Neural Network, Multi-label Calssification
- **Statistics**: Multiple Hypothesis Testing, Bayesian Model, Mixture Model, Non-parametric Statistics
- **Computational Biology**: Pharmocogenomics, Bioinformatics

Education

PhD: Biostatistics, UC Berkeley — Aug 2017 - Feb 2021
GPA: 4.00/4.00 (PhD Candidate) Advisers: Peter J. Bickel and Haiyan Huang

Master: Biostatistics, UC Berkeley — Aug 2015 - May 2017
GPA: 4.00/4.00

Bachelor: Mathematical Sciences, Tsinghua University — Aug 2011 - July 2015
GPA: 3.86/4.00 (Rank 1st in the Division of Probability and Statistics)
Honor: Excellent Graduate of Class 2015

Work Experience

Amazon Search, Applied Scientist Intern. Research on widget ranking with multi-objective learning. — May - Nov 2020

Alibaba Taobao, Software Engineer Intern. Develop algorithms that incorporate prior knowledge into graph embeddings. — Nov 2018 - Feb 2019

Microsoft Azure, Data Scientist Intern. Churn analysis on the customers of Azure Storage. — May - Aug 2018

Honors

The **Genentech** Fellowship Award	2019-2020
The **Mayhew & Helen Derryberry** Fellowship	2018-2019
National Scholarship	2014-2015
Zheng ZongCheng Scholarship	2013-2014
Zheng GeRu Scholarship	2012-2013

Selected Works

Machine Learning

Publication

[1] Da Xu, **Yuting Ye**, Chuanwei Ruan, Understanding the role of importance weighting for deep learning, accepted by ICLR for a spotlight talk, 2021.
keywords: Importance Sampling, Deep Learning, Theory of Nonconvex Learning

[2] **Yuting Ye**, Xuwu Wang, Kunyang Jia, Jingren Zhou, Yanghua Xiao, Hongxia Yang, Bayes EMbedding (BEM): Refining Representation by Integrating Knowledge Graphs and Behavior-specific Networks. In Proceedings of the 28th ACM International Conference on Information and Knowledge Management. ACM, 2019, pp. 679–688.
keywords: Graph Neural Network, Bayesian Model, Knowledge Graph, E-commerce

[3] **Yuting Ye**, Lihua Lei, Cheng Ju, HONES: A Fast and Tuning-free Homotopy Method For Online Newton Step. In Proceedings of the Twenty-First International Conference on Artificial Intelligence and Statistics (AISTATS-18), pages 2008–2017, 2018.
keywords: Online Learning, Netwon Method

Statistics

Publication

[1] Wenpin Tang, **Yuting Ye**, The existence of maximum likelihood estimator in high-dimensional generalized linear models. Electron. J. Statist. 14 (2020), no. 2, 4028–4053. doi:10.1214/20-EJS1766.
keywords: High Dimensional Statistics, Generalized Linear Models, Theory of Nonconvex Learning

[2] **Yuting Ye**, Yin Xia, Lexin Li. Paired Test of Matrix Graphs and Brain Connectivity Analysis. Biostatistics, kxz037, 2019.
keywords: Precision Matrix, Brain Image, Multiple Hypothesis Testing

Preprint

[3] **Yuting Ye**, Christine Ho, Ci-Ren Jiang, Wayne Tai Lee, Haiyan Huang, Hierarchical multi-label classification with local precision rate, submitted to Journal of the American Statistical Association (JASA).
keywords: Hierarchical Multi-label Classification, Local Precision Rate

[4] **Yuting Ye**, Peter J. Bickel, Non-parametric density estimation and inference under the U-shape constraint, with an application to the Gene Fishing method, manascript.
keywords: Binomial Mixture Model, Shape Constraint, Non-parametric Density Estimation

[5] **Yuting Ye**, Lihua Lei, Michael I. Jordan, Bayesian Inference On Gaussian Graphical Models With Soft G-Wishart Distributions, manuscript.
keywords: Bayesian Model, Markov Chain Monte Carlo, Whishart Distribution, Precision Matrix

Computational Biology

Publication

[1] Calvin Chi, **Yuting Ye**, Bin Chen, and Haiyan Huang, Bipartite graph-based approach for clustering of cell lines by gene expression-drug response associations, accepted by Bioinformatics.
keywords: Pharmacogenomics, Bi-clustering, Canonical Correlation Analysis

[2] Zhiyue Hu, **Yuting Ye**, Nwbury Patrick, Haiyan Huang, Bin Chen. AICM: A Genuine Framework Targeting Inconsistency Between Large Pharmacogenomics Datasets. In Proceedings of Pacific Symposium on Biocomputing (PSB), pp. 248-259. 2019.
keywords: Pharmacogenoimcs, Drug Sensitivity, Matrix Completion,

[3] Miao Chang, Elliot K. Edmiston, Fay Y. Womer, Qian Zhou, Shengnan Wei, Xiaowei Jiang, Yifang Zhou, **Yuting Ye**, Haiyan Huang, Xi-Nian Zuo, Ke Xu, Yanqing Tang, Fei Wang. Spontaneous low-frequency fluctuations in the neural system for emotional perception in major psychiatric disorders: amplitude similarities and differences across frequency bands. Journal of psychiatry & neuroscience: JPN. 2019 Mar; 44(2):132.
keywords: fMRI, Bipolar, Major Depressive Disorder, Schizophrenia

[4] Miao Chang, Fay Y. Womer, Elliot K. Edmiston, Chuan Bai, Qian Zhou, Xiaowei Jiang, Shengnan Wei, Yange Wei, **Yuting Ye**, Haiyan Huang, Yong He, Ke Xu, Yanqing Tang, Fei Wang. Neurobiological Commonalities and Distinctions Among Three Major Psychiatric Diagnostic Categories: A Structural MRI Study. Schizophrenia Bulletin. 2017 Jun 13.
keywords: Strutural MRI, Bipolar, Major Depressive Disorder, Schizophrenia

[5] **Yuting Ye**, Jingyi Jessica Li, NMFP: a non-negative matrix factorization based preselection method to increase accuracy of identifying mRNA isoforms from RNA-seq data. BMC Genomics 17, 11 (2016).
keywords: Bioinformatics, Isoform Selection, Pre-screening, Non-negative Matrix Factorization

Preprint

[6] Zoe Vernon, **Yuting Ye**, Bin Chen, Haiyan Huang, Reversal Correlation for drug discovery and biomaker identification, manuscript.
keywords: Pharmacogenomics, Drug Discovery, Rank Testing, Multiple Hypothesis Testing

Presentations

Conference

Conference on Information and Knowledge Management, Beijing, China	11/2019
Artificial Intelligence and Statistics, Playa Blanca, Lanzarote, Canary Islands	04/2018
Joint Statistical Meetings, Chicago, IL, US	08/2016
Asia Pacific Bioinformatics Conference, San Francisco, CA, US	01/2016

Invited Talks

JD AI Research, Mountain View, CA, US	11/2019
Baidu Research, Sunnyvale, CA, US	03/2019

Seminar Talks

UC Berkeley, Biostatistics student seminar	11/2019
UC Berkeley, SGSA Student Seminar	04/2019
UC Berkeley, Biostatistics student seminar	02/2019

Teaching Experience

Teaching assistant, STAT 134: Concepts of Probability	Fall 2020, Fall 2019
Teaching assistant, STAT 230: Linear Models	Spring 2019, Spring 2017
Teaching assistant, STAT 201B: Introduction to Statistics at an Advanced Level	Fall 2017
Teaching assistant, STAT 135: Concepts of Statistics	Fall 2016
Teaching assistant, PH 142: Introduction to Probability and Statistics in Biology and Public Health	Fall 2015

Professional Services

Referee

- The Canadian Journal of Statistics
- Conference on Neural Information Processing Systems, 2017
- Intelligent Systems for Molecular Biology, 2017