

WEICHI YAO

Curriculum Vitae

44 W 4th St., 8-154, New York, NY 10012

☎ (+1) (718)316-7377

✉ weichi.yao@gmail.com

🌐 www.weichiyao.com

🐙 github.com/weichiyao

EDUCATION

- Sept. 2017 - **Ph.D., Stern School of Business, New York University**,
present *Department of Technology, Operations, and Statistics*,
Advisor: Professor Halina Frydman. Email: hfydman@stern.nyu.edu.
- Sept. 2015 - **M.A., Graduate School of Arts and Sciences, Columbia University**,
Dec. 2016 *Department of Statistics*, GPA: 3.8/4.0.
- Sept. 2011 - **B.A., Economics & Management School, Wuhan University**,
June 2015 *Department of Finance*, GPA: 3.8/4.0.
- Sept. 2012 - **B.A., Mathematics & Statistics School, Wuhan University**,
June 2015 *Department of Applied Mathematics*, GPA: 3.9/4.0.

RESEARCH

My research interests lie in developing machine learning methodology and algorithms to tackle problems in various fields. I have worked in *healthcare* related fields. Recently I focus on *deep active learning* and *active optimization* for more accurate and reliable deep learning algorithms at a reduced cost of time and human resources for science projects. I also work on designing *equivariant neural networks* that respect physical symmetries and application in dynamical systems..

ONGOING PROJECTS

- **W. Yao** and Y. Wang. Deep active learning and active optimization.
- **W. Yao**, K. Storey-Fisher, D. W. Hogg and S. Villar. Equivariant machine learning methods for dynamics based on scalars.
- **W. Yao** and J. Loftus. Is the best good enough? The direction of selection bias for high-dimensional goodness-of-fit tests.

PREPRINTS

- [1] **W. Yao**, H. Frydman, D. Laroque and J. S. Simonoff. Ensemble methods for survival data with time-varying covariates. *arXiv:2006.00567*, 2020.

[LINK](#) [CODE](#)

CONFERENCES

- [2] **W. Yao**, K. Storey-Fisher, D. W. Hogg and S. Villar. A simple equivariant machine learning method for dynamics based on scalars. *In the Advances in Neural Information Processing Systems (NeurIPS) 2021 Workshop on Machine Learning and the Physical Sciences (to appear)*, 2021.

[LINK](#) [CODE](#)

- [3] S. Villar, D. W. Hogg, K. Storey-Fisher, **W. Yao** and B. Blum-Smith. Scalars are universal: Equivariant machine learning, structured like classical physics. *In Proceedings of the Advances in Neural Information Processing Systems (NeurIPS) (to appear)*, 2021.

[LINK](#) [CODE](#)

- [4] **W. Yao**, A. S. Bandeira and S. Villar. Experimental performance of graph neural networks on random instances of max-cut. *In Proceedings of the Society of Photographic Instrumentation Engineers*, 2019.

[LINK](#) [CODE](#)

- [5] J. H. Lee, D. E. Carlson, H. S. Razaghi, **W. Yao**, G. A. Goetz, E. Hagen, E. Batty, E. J. Chichilnisky, G. T. Einevoll and L. Paninski. YASS: Yet Another Spike Sorter. *In Proceedings of the Advances in Neural Information Processing Systems (NeurIPS)*, 4005-4015, 2017.

[LINK](#) [CODE](#)

JOURNALS

- [6] H. Moradian, **W. Yao**, D. Larocque, J. S. Simonoff and H. Frydman. Dynamic estimation with random forests for discrete-time survival data. *The Canadian Journal of Statistics (to appear)*, 2021.

[LINK](#) [CODE](#)

- [7] **W. Yao**, H. Frydman and J. S. Simonoff. An ensemble method for interval-censored time-to-event data. *Biostatistics*, 22(1):198-213, 2021.

[LINK](#) [CODE](#)

INTERNSHIP

2016–2017 **Research Intern**, Grossman Center for the Statistics of Mind, Columbia University, Professor Liam Paninski.

As a research assistant working in the neuroscience lab on projects that develop statistical methodology for understanding how neurons encode information.

TEACHING

2020 **Course Instructor**, STAT-UB Statistics for Business Control.

2019–2021 **Teaching Fellow**, XBA1-GB: Operations Analytics.

2021 **Teaching Fellow**, STAT-GB: Analytics & Machine Learning for Managers.



2021 **Teaching Fellow**, STAT-GB: Introduction to Stochastic Processes.

2020 **Teaching Fellow**, STAT-UB: Statistics for Business Control Regress & Forecasting Models.

2018 **Teaching Fellow**, COR1-GB: Statistics and Data Analysis.

SKILLS

Programming Python, R, Matlab, SAS

- Highlights
- Implemented equivariant neural networks based on scalars in Python; see **ScalarEMLP** .
 - Implemented graph neural networks in Python for Max-Cut; see **maxCut**  (contributed to 2020 GitHub Archive Program).
 - Contributed two public available R packages, **ICcforest** and **LTRCforests**, on CRAN.