

ZHUOYAN XU

1402 Regent St, Apt 428, Madison, WI 53711 ♦ (608)960-1191 ♦ zxu444@wisc.edu

EDUCATION

Total 5 years. Spent 3 years in Wuhan University, 1 year in UW-Madison as an undergraduate student, 1 year in UW-Madison as a graduate student.

M.S. in Statistics

University of Wisconsin-Madison, WI

- **GPA:** 3.83/4.00 09/2018 – 05/2020
- **Relevant Major Coursework:** Statistical Inference (A)/ Statistical Learning Theory (A)/ Linear Regression Theory and Methods (A)/ Experimental Design (A)

B.S. in Statistics

Wuhan University, China

- **Major Percentage:** 87.2/100 09/2015 – 06/2018
- **Relevant Major Coursework:** Mathematical analysis (95/100)/ Advanced Algebra (96/100)/ Sampling Survey (96/100)/ Regression analysis (93/100)/ Stochastic Process (90/100)/ Mathematical Statistics (97/100)/ Statistical computing (90/100)

PUBLICATION

Xu, Z., Hu, J. and Wang, M., 2019. Generalized tensor regression with covariates on multiple modes. arXiv preprint [arXiv:1910.09499](https://arxiv.org/abs/1910.09499). Under review for Journal of Machine Learning Research W&CP (AISTATS track).

Xu, Z., Hu, J. & Wang, M., 2019. R package *Tensorregress*: Generalized tensor regression with covariates on multiple modes. Published on The [Comprehensive R Archive Network](#).

RESEARCH EXPERIENCE

Generalized Tensor Regression with Covariates on Multiple Modes

UW-Madison

Research Assistant to **Prof. Miaoyan Wang**

02/2019 – present

- Proposed a tensor response regression model incorporating covariates on multiple modes.
- Extended proposed model to generalized tensor decomposition for observations in exponential family.
- Proved the theoretical accuracy guarantees of the proposed model.
- Proposed efficient alternating updating algorithm robust to outliers. Evaluated on both simulations and two real dataset(Human Connectome Project (HCP) & Nations data).
- Developed the R package to implement the algorithm.
- Co-author Professor Wang will deliver this work on the invited talk at Columbia University, Purdue University and 2020 ENAR (Eastern North American Region, International Biometric Society).

Uncertainty Quantification Seminar (repo: [UQ-Seminar](#))

UW-Madison

Presenter & Discussant in seminar held by **Prof. Peter Chien**

01/2019 – 05/2019

- Discussed state-of-art statistical machine learning methods for quantifying uncertainties in complex systems and applications in engineering, medical, finance and other fields.
- Presented papers in optimization, convolutional neural network and deep reinforcement learning to PhD and professor audiences.

Optimal Transport Project (repo: [OT-project](#))

UW-Madison

Student Researcher supervised by **Prof. Nicolas Garcia Trillos**

09/2019 – present

- Constructed the Optimal Transport map over geodesic metric space. Characterized the distribution of pixel values of nuclei images. Computed Wasserstein distance using Sinkhorn's algorithm proposed by Marco Cuturi.
- Implemented shape interpolation between nuclei images using Convolutional Wasserstein Distances proposed by Justin Solomon et al. 2015.

Presenter in Academic Poster Session

UW-Madison

Computation and Informatics in Biology and Medicine/Bio-Data Science Training Programs 10/2019

- Presented work in Training Programs to the professor and PhD audiences in Statistics/Computer Science/ Biostatistics and Medical Informatics department.
- Applied the tensor regression model to identify functional brain connectivity patterns related to individual attributes.

Statistical Machine Learning Seminar

UW-Madison

Presenter in Statistical Seminar held by **Prof. Miaoyan Wang**

7/2019 – 9/2019

- Presented papers in *International Conference on Machine Learning*.
- Led a discussion about the papers by Tony Cai, Quentin Berthet and Nicolai Baldin.

Boosting Method Implementation on Machine Learning Task

UW-Madison

Major Researcher for Graduation Project

04/2019 – 06/2019

- Investigated the classical papers of boosting methods including Adaboost, Gradient Boosting, XGBoost.
- Discussed the relationship between boosting method and non-parametric approximation from a statistical perspective. Presented the relevant papers from Jerome Friedman, Trevor Hastie, and Robert Tibshirani.

Big data innovation contest held by Peking University

Wuhan University

Leader of the group supervised by Research Fellow **Dr. Shirong Deng**

1/2018 – 2/2018

- Proposed a semiparametric regression model. The final fitting results were consistent with the economic theory.
- Predicted the tendency of the designed problem.

PROJECT EXPERIENCE

Natural Language Processing Project: Yelp Data Review Analysis

UW-Madison

Leader of the Group of supervised by **Prof. Hyunseung Kang**

9/2019 – 12/2019

- Implemented NLP preprocessing on review text. Implemented word counting and statistical analysis.
- Constructed Ordered Logit Regression model predicting the rating of each business.
- Constructed LSTM with other techniques identifying topics in review text. Conducted sentiment analysis.

TECHNICAL SKILLS

- Python, R, Linux, HPC, HTC, C/C++, Git, Matlab, SQL, Latex