Mabel Qianqian Yao

• mabel.qq.yao@gmail.com • LinkedIn • Github • Google Scholar • X • ME

EDUCATION

North Dakota State University *2026*
PhD, Statistics
focusing on Statistical Inference, Statistical Machine Learning, Data Science
Overall GPA: 4.00/4.00

North Dakota State University *2022*
Master, Computer Science
focusing on Machine Learning, Data Mining, Data Science
Overall GPA: 4.00/4.00

Tohoku University *2016*
Master in Engineering, Structural Engineering
Overall GPA: 3.68/4.00

Dalian Jiaotong University *2013*
Bachelor in Engineering, Civil Engineering & Software Engineering
Overall GPA: 87/100

Research

- **Interests**

- Statistical Inference, Statistical Machine Learning, Data Science

- **Research Projects by Topics**

* ***Statistical Inference on Graphs*** *2022-present*

- Statistical Methods in Network Analysis

- Graph Representation/ Embedding: Learned and Non-Learned

* ***Algorithmic and Data Modeling, Theoretical Statistics, Machine Learning*** *2022-present*

- How algorithms learn? using statistical models to understand when and why machine learning algorithms work.

- Comparison of Modeld of Parametric, Semiparametric, Nonparametric; to understand how to do model selection for real world applications.

* ***Applications in Tech/IT Company, Econometric, Financial Economics, Medical Data Science, Drug Discovery/Precision Medicine, etc*** *2022-present*

- Explore applications of statistical methods in real world scenarios, model implementation, evaluation, selection, and validation
- Application of Graphs or Networks in Financial Markets/ Economics, etc.
- Outlier Analysis/ Anomaly Detection in Financial Fraud
- Feature Selection, Feature Extraction, and Analysis and Modeling on Diverse Data: multivariate, time series, survival analysis, etc.

* ***Graph Representation Learning*** *2019-2022*

- Network embedding
- graph-level representation learning: Learned and Non-Learned Representations
- Graph-Level Representation Learning for Chemical Screening for Catalyst Discovery/ Material Discoveries
- Molecular Property Prediction by Descriptor based Statistical Modeling, Sequential Modeling, Graph Modeling

* ***Recommendation Systems*** *2019-2022*

- Recommendation in e-commerce and healthcare
- Graph Learning & graph neural networks for Recommendation Systems
- Statistical methods for Recommendations

* ***Machine Learning, Deep Learning, Data Mining, Data Science*** *2019-2022*

- Algorithms, Frameworks, Infrastructures, Implementations
- Data Processing and Applications in Applied Domains

PUBLICATION

- **Peer Reviewed Papers**

- Qianqian Yao, Mingao Yuan. Empirical Likelihood test for common invariant subspace of multilayer networks. **to be finished.** (2025)
- Yuan, M., Yao, Q. Testing common degree-correction parameters of multilayer networks. Stat Comput 35, 38 (2025). https://doi.org/10.1007/s11222-025-10576-z
- Mingao Yuan, Qianqian Yao. Testing common invariant subspace of multilayer networks. arXiv preprint arXiv:2406.05010 (2024). (Preprint)
- Qianqian Yao, Yoshihiro Ito, Yusuke Suzuki, Maeda Masaki. (2015). "Effect of Earthquake Response Spectrum Characteristics on Residual Seismic Performance", Japan Concrete Institute, vol.37 NO.2, 685-690.

- **Conference**

* **2024 Red River Valley Statistics Conference** *Fargo, ND*

Presentations *May, 2024*

- Presentation on time series: Analysis of Time Series Models for Electricity Consumption Forecasting

- Presentation on survival analysis: Survival Analysis for Colorectal Cancer Considering Effects of Staging and Treatment Variables

* **2023 Red River Valley Statistics Conference** *Fargo, ND*

Presentation *May, 2023*

- Presentation on biostatistics: Statistical Learning for Virtual Screening in Drug Discovery

* **2015 Architectural Institute of Japan Annual Conference (Kanto)** *Hiratsuka, JP*

- Linfei Hao, Qianqian Yao, Yusuke Suzuki, Masaki Maeda, "Comparison Between Residual Seismic Capacity Evaluation Method Based on Seismic Capacity Ratio and Recommended by Current Guideline", published in the Architectural Institute of Japan, page 403-404, 2015.9.

* **2014 Architectural Institute of Japan Annual Conference (Kinki)** *Kobe, JP*

- Qianqian Yao, Linfei Hao, Yoshihiro Ito, Yusuke Suzuki, Maeda Masaki, "Seismic Evaluation of Damaged RC buildings Basing on Earthquake Response Spectrum, part 1 Basic Evaluation Basing on Seismic Capacity Index and Effect of Building Capacity Reduction", published in the Architectural Institute of Japan meeting in Koube, page195-196, 2014.8.

- Linfei Hao, Qianqian Yao, Yoshihiro Ito, Yusuke Suzuki, Masaki Maeda, "Residual Seismic Capacity Assessment of Damaged RC Buildings Based on Response Spectrum Part 2. Assessment of residual seismic capacity ratio based on ideal model and comparison with seismic capacity reduction factor", published in the Architectural Institute of Japan, page 197-198, 2014.8.

- **Thesis**

- Qianqian Yao: topic about statistical inference on graphs/ learning in networks **(to be finished.)** (2026). (PhD in Statistics, North Dakota State University)

- Qianqian Yao: Comparison of non-learned and learned molecule representations for catalyst discovery (2022). (Master's in Computer Science, North Dakota State University)

- Qianqian Yao: Residual Seismic Capacity Evaluation for Reinforced Concrete Buildings Considering Effects of Characteristics of Structure and Earthquake Response Spectrum (2016). (Master's in Architecture and Building Science, Tohoku University)

Working Experience

- **Teaching Assistantship in North Dakota State University** *Fargo, US*

Department of Statistics *2022-present*

duties: help with computer labs, tutoring, grading, review sessions

STAT 725 Applied Statistics

STAT 726 Applied Regression and Analysis of Variance

STAT 330 Introductory Statistics

Computer Science Department *2019-2022*

duties: help with computer labs, tutoring, grading, review sessions

CSCI 160 Computer Science I

CSCI 161 Computer Science II

CSCI 213 Modern Software Development

CSCI 114 Computer Applications

CSCI 122 Visual BASIC

- **Part-Time Work in North Dakota State University** *Fargo, US*

  Lab Assistant in Plant Sciences Department *2019-2023*

  Implementation of Experimental Design, including seeding, planting, harvesting, data collection and entry

  Data Analysis

- **Industry Working Experience**

  Shenzhen Yuanlizhu Engineering Consultants Co.,Ltd *Shenzhen, China*

  Structural Engineer *2017-2019*

  Using computer aided engineering tools to design and analyze building structures.

  Communicate with clients including investors, constructors, designers to optimize the structural design.

  Shanghai Saiyo Construction Technology Co.,Ltd *Shanghai, China*

  Project Assistant *2016-2017*

  Participated in a Japanese project of Shopping Mall Construction in Ningbo, and applied Building Information Modeling (BIM) to construct a virtual model of the building for design and clash detection.

  Yamashita Sekkei INC. Tohoku Branch *Sendai, Japan*

  Intern *9/2015-10/2015*

  Analyze structures with SNAP, created building model, considered seismic isolators and seismic control devices, analyzed seismic response controlled structure and seismic isolation structure to get seismic performance, created animation; Drew construction drawings with AutoCAD.

Technical Skills

- **Programming Languages**

  R, Python, Julia, SAS, Java, etc. (able to get hands on new languages very quick)

- **Computing and Editing Softwares**

  Latex, RStudio, Anaconda Environment, Matlab, Octave, Minitab, JMP, Visual Studio, etc. (able to get hands on new tools very quick)

- **Domain Knowledge**

  Experienced and Gained Comprehensive Knowledge in Interdisciplinary Applied Sciences and Engineering

- **Speaking Human Languages**

  English, professional level

  Japanese, professional level

  Chinese, Native level

COURSE

- **Related Courses from North Dakota State University** *in-person*
- *STAT886 Advanced Inference* *2025 Spring*
- *STAT851 Bayesian Statistical Inference* *2025 Spring*
- *STAT874 Generalized Linear Models* *2024 Fall*

  Implementation: Generalized Linear Models for credit card expenditure analysis by considering variance functions

- *MATH650 Real Analysis I* *2024 Fall*

  Presentation: Proof of Weierstrass Approximation Theorem

- *STAT672 Time Series* *2024 Spring*

  Implementation: Analysis of Time Series Models for Electricity Consumption Forecasting, R

- *STAT770 Survival Analysis* *2024 Spring*

  Implementation: Regression Models for Survival Analysis in CRC Considering Staging Groups, R

- *STAT768 Mathematical Statistics II* *2024 Spring*
- *STAT767 Mathematical Statistics I* *2023 Fall*
- *STAT764 Multivariate Methods* *2023 Fall*

  Implementation: Multivariate Analysis for Discrimination of Carcinogenesis Staging, SAS

- *STAT661 Applied Linear Models* *2023 Fall*

  Implementation: Detection and Evaluation of Outliers by Linear Models, R

- *STAT669 Introduction to Biostatistics* *2023 Spring*

  Implementation: Descriptor based multiple linear regression model for molecule property prediction, python
- *STAT662 Introduction to Experimental Design* *2023 Spring*
- *STAT663 Nonparametric Statistics* *2022 Fall*
- *STAT726 Applied Regression and Variance Analysis* *2022 Fall*
- *STAT860 Statistical Machine Learning* *2022 Spring*

  Implementation: Statistical Methods for Recommender System, python
- *STAT725 Applied Statistics* *2021 Fall*
- *CSCI859 Computational Methods in Bioinformatics* *2021 Fall*
- *CSCI848 Empirical Methods for Software Engineering* *2021 Spring*
- *CSCI702 Survey of Cybersecurity* *2021 Spring*
- *CSCI 717 Software Construction* *2020 Fall*

  Implementation: Natural Language Processing: text classification, python
- *CE793 Machine Learning for Engineers* *2020 Spring*

  Implementation: Multi-label classification based on image similarity, python
- *CSCI846 Distributed Systems* *2020 Spring*

  Implementation: Distributed database built on client-server architecture, java
- *CSCI679 Introduction to Data Mining* *2019 Fall*

  Implementation of recommender system based on different models, python
- *CSCI736 Advanced Intelligent Systems* *2019 Fall*

  Implementation of expert system for real estate recommnendation by drools, java
- *CSCI713 Software Development Processes* *2019 Fall*
- *CSCI741 Algorithm Analysis* *2019 Summer*
- *CSCI879 Advanced Data Mining* *2019 Spring*

  Implementation: Network Mining and analysis using deepwalk, line, and node2vec, python
- *CSCI724 Introduction to Artificial Intelligence* *2019 Spring*

  Implementation: Large scale study of programming languages and code quality in github, python
- *CSCI765 Introduction to Database Systems* *2019 Spring*

  Implementation: Evaluation of real estate market using deep learning, python

- **Free Open Source: Theories in Mathematical Statistics and Machine Learning** *online*

CMU 36-708 Statistical Methods for Machine Learning, Dr. Larry Wasserman — *Syllabus*

CMU 36-709 Advanced Statistical Theory I, Dr. Sivaraman Balakrishnan — *Syllabus*

CMU 36-710 Advanced Statistical Theory II, Dr.Alessandro Rinaldo — *Syllabus*

UCLA STATS 100C Linear Models, Dr. Arash A. Amini — *Syllabus*

UCLA STATS 200B Theoretical Statistics, Dr. Arash A. Amini — *Syllabus*

UCLA STATS 200C High-dimensional Statistics, Dr. Arash A. Amini — *Syllabus*

UCLA STATS 231C Theories of Machine Learning, Dr. Arash A. Amini — *Syllabus*

UCSD MATH 181A Mathematical Statistics, Dr. David Quarfoot — *Podcast*

UCSD MATH 181B Mathematical Statistics, Dr. David Quarfoot — *Podcast*

UCSD MATH 180A Introduction to Probability for Data Science, Dr. Todd Kemp — *Podcast*

UManchester MATH38161 Multivariate and Machine Learning, Dr. Korbinian Strimmer — *Podcast*

UCLA STATS 100C Linear Models, Arash A. Amini — *Syllabus*

MIT OpenCourseWare: Topics In Mathematics With Applications In Finance — *Syllabus*

DavidsonX: Drug Discovery & Medicinal Chemistry, edX — *Syllabus*

- **Verified Courses with Certificates on Coursera** — *online*

IBM Data Science Specialization — *May, 2023*

Machine Learning — *May, 2023*

Deep Learning Specialization — *March, 2021*

Dear Hiring Team,

Thank you for considering my application for the position related to statistical inference, machine learning and data science. I am actively seeking opportunities that align with my career aspirations for professional growth.

My academic path highlights my passion for interdisciplinary learning and innovation in sciences and engineerings. I pursued dual majors in Civil Engineering and Software Engineering during my bachelor's at Dalian Jiaotong University in China, driven by the evolving role of computer science in industrialization and intellectualization. Then I advanced my graduate studies at Tohoku University in Japan, specializing in Structural Engineering. This experience not only strengthened my expertise in structural analysis and design but also inspired me to explore the transformative potential of technology in engineering through computer-aided solutions. Afterwards, two years working in industry empowered me with hands on experiences about applications of technologies in real world. Fueled by a deep interest in machine learning and its diverse applications, I pursued Master's in Computer Science and PhD in Statistics at North Dakota State University in US. My academic journey has illuminated the intricate connections between computer science, statistics, and domain knowledge, and has empowered me expertise in mathematical, statistical, and algorithmic theories behind the curtain of this physical world.

My research focuses on statistical inference, machine learning, data science, emphasizing on theories, also exploring some real-world applications. I devote in theoretical statistics and statistical learning, which is an umbrella of statistical methods including Likelihood, Bayesian, Regression, Multivariate analysis, Nonparametric, Hypothesis Testing, Machine Learning, etc. Meanwhile I am also drawn to applications in Tech/IT Company, Econometric, Financial Economics, Medical Data Science, Drug Discovery/Precision Medicine, etc.

I am highly motivated by the opportunities to achieve my career goals by meaningful contribution to your organization's mission. I appreciate all who have ignited lights through my life journey, and I am inspired to become the one who are able to light up the paths for other people. With a resilient mindset and a dedication to continuous learning, I am confident in my ability to deliver impactful results as part of your team.

I appreciate of your time to have more discussion on how my skills and expertise align with your team's needs. Please feel free to contact me at qianqian.yao@ndsu.edu or mabel.qq.yao@gmail.com to arrange a conversation at your convenience. Thank you for considering my application.

Sincerely,
Y.