MABEL QIANQIAN YAO

LINKS

Email: qianqian6.yao@gmail.com; qianqian.yao@ndsu.edu

LinkedIn

Github

Google Scholar

Personal Webpage for Outputs

EDUCATION

North Dakota State University *2026*

PhD, Statistics, focusing on Mathematical Statistics, Statistical Machine Learning, Data Science

Overall GPA: 4.00/4.00

North Dakota State University *2022*

Master, Computer Science, focusing on Machine Learning, 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**

- Mathematical Statistics, Machine Learning, Data Science, Statistical Inference

- **Research Projects by Topics**

- *Statistical Inference on Graphs*

- Mathematical Modeling for Graphs, Statistical Inference, Hypothesis Testing

- Network Analysis and Embedding

- Graph Representation/ Embedding: Learned and Non-Learned

- *Algorithmic, Mathematical Modeling, Statistics, Machine Learning*

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

- Comparison of Parametric, Semiparametric, Nonparametric Models, to understand how to select models for real world applications

- *Outlier Analysis, Anomaly Detection in Computational Finance*

- Time Series Data and Multivariate Data Analysis
- Anomaly Detection in Financial Fraud.

- *Medical Data Science, Drug Discovery, Precision Medicine*

- Chemometrics: Molecular Profiling, Feature Selection & Feature Extraction
- Molecular Property Prediction: Descriptor based Statistical Modeling, Sequential Modeling, Graph Modeling
- Graph-Level Representation Learning for Chemical Screening for Catalyst Discovery/ Material Discoveries
- Application to other types of data

- *Graph Representation Learning*

- Network embedding
- graph-level representation learning: Learned and Non-Learned Representations

- *Recommendation Systems*

- 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*

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

PUBLICATION

- **Peer Reviewed Papers**

- 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 Papers**

- 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.
- 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.

• Thesis

- 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)

PROFESSIONAL ACTIVITY

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

Presentation — *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

• PhD Program Training in Computer Science — *Fargo, ND*

PhD training — *2022-2024*

registered in PhD program in Computer Science from 2022 Spring to 2024 Summer, conducted research credits focusing on machine learning and graph representation learning. finished 18 research credits and 3 seminars. Degree is incomplete but continued on as PhD in Statistics with same focus on mathematical statistics, machine learning, data science.

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

Presentation — *May, 2023*

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

• Intelligent Ground Vehicle Competition — *Rochester, Michigan*

Competition — *June, 2019*

placed fourth in the Intelligent Ground Vehicle Competition's Autonav Challenge,
held June 7-10 at Oakland University in Rochester, Michigan

• North Dakota Science Olympiad — *Fargo, ND*

Volunteer — *April, 2019*

helped to manage the competition of Bridge Structure *webpage*

• Exchange Program, University of California, Davis — *Davis, US*

Cooperative Laboratory Study — *02/2015-03/2015*

Academic English Program for Science and Technology; Stayed in structural engineering lab in UC Davis for field study;

- **Inter-Graduate School program, Tohoku University** *Sendai, Japan*

  Project of Doctoral Degree Program on Science for global safety *2014-2016*

  Finished all corresponding curriculum including training of specialized seminar of science for natural disasters, Multi-disciplinary and Specialized Basic Subjects about natural disaster generation mechanism, and fundamental subjects of philosophy and sociology; Activities with other Universities about natural disasters; Concentrated lectures about development of human culture; English skill training;

- **Scholarships and Awards, Dalian Jiaotong University** *Dalian, China*

  Undergraduate scholarships *2009-2013*

  received scholarships semesterly due to outstanding performance.

WORKING EXPERIENCE

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

  Department of Statistics *2024-present*

  STAT 725 Applied Statistics

  STAT 726 Applied Regression and Analysis of Variance

  STAT 330 Introductory Statistics

  Computer Science Department *2019-2023*

  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, etc.

- **Computing Softwares**

  Matlab, Octave, Minitab, JMP, etc

- **Writing Editors and Tools**

  Latex, RStudio, etc

- **Applied Sciences**

  Experienced and Gained Deep Knowledge in Interdisciplinary Sciences and Engineering

- **Speaking Human Languages**

  English, professional level

  Japanese, professional level

  Chinese, Native level

COURSE

- **Courses from North Dakota State University**

- *STAT874 Generalized Linear Models* *2024 Fall*

  Implementation: Dispersion Analysis (to be finished)

- *MATH650 Real Analysis I* *2024 Fall*

  Proof: Power Series (to be finished)

- *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*

  theories on mathematical statistics

- *STAT767 Mathematical Statistics I* *2023 Fall*

theories on mathematical statistics

- *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*

  about theories on experimental design models

- *STAT663 Nonparametric Statistics* *2022 Fall*

  about theories on nonparametric models

- *STAT726 Applied Regression and Variance Analysis* *2022 Fall*

  about theories on linear regression models

- *STAT860 Statistical Machine Learning* *2022 Spring*

  Implementation: Statistical Methods for Recommender System, python

- *STAT725 Applied Statistics* *2021 Fall*

  about theories on statistical inference

- *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

- **Courses about Theories in Mathematical Statistics and Machine Learning** *Winter, 2024*

  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 200B Theoretical Statistics, Dr. Arash A. Amini *Link, Videos*

  UCLA STATS 200C High-dimensional Statistics, Dr. Arash A. Amini *Link, Videos*

  UCLA STATS 231C Theories of Machine Learning, Dr. Arash A. Amini *Link, Videos*

- **Courses about Theories in Mathematics and Statistics** *Summer, 2024*

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

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

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

  UManchester MATH38161 Multivariate and Machine Learning, Dr. Korbinian Strimmer *Syllabus and Videos*

  UCLA STATS 100C Linear Models, Arash A. Amini *Link, Videos*

- **Courses about Applied Sciences in Economics** *Spring, 2024*

  MIT OpenCourseWare: Development Economics *Videos*

  MIT OpenCourseWare: Topics In Mathematics With Applications In Finance *Videos*

- **Courses about Applied Sciences in Drug Discovery** *Spring, 2023*

  DavidsonX: Drug Discovery & Medicinal Chemistry, edX *Link*

- **Verified Courses with Certificates on Coursera**

  IBM Data Science Specialization *May, 2023 Certificate*

  Machine Learning, Coursera *May, 2023 Certificate*

  Deep Learning Specialization, Coursera *March, 2021 Certificate*

REFERENCE

Information on Department of Statistics

Dr. Mingao Yuan, Associate Professor, Department of Statistics
North Dakota State University
Mingao.Yuan@ndsu.edu

Dr. Rhonda Magel, Professor and Chair, Department of Statistics
North Dakota State University
Rhonda.Magel@ndsu.edu

Dr. Megan Orr. Associate Professor, Department of Statistics
North Dakota State University
Megan.Orr@ndsu.edu

Dr. Ron Degges, Associate Professor of Practice, Department of Statistics
North Dakota State University
Ronald.Degges@ndsu.edu

Dr. Gang Shen, Associate Professor, Department of Statistics
North Dakota State University
Gang.Shen@ndsu.edu