MABEL QIANQIAN YAO

(701) · 491 · 2632 ◇ qianqian6.yao@gmail.com

EDUCATION

North Dakota State University *2026*

PhD, Statistics, focusing on Machine Learning and Data Science

Overall GPA: 4.00/4.00

North Dakota State University *2022*

Master in Science, Computer Science, focusing on Machine Learning and Data Science

Overall GPA: 4.00/4.00

Tohoku University *2016*

Master in Engineering, Architecture and Building Science, focusing on Structural Engineering

Overall GPA: 3.68/4.00

Dalian Jiaotong University *2013*

Bachelor in Engineering, Civil Engineering & Software Engineering

Overall GPA: 87/100

RESEARCH

• Research Interests

- Machine Learning, Statistical Inference, Data Science
- Statistical Inference on Graphs

• Publications

Google Scholar: https://scholar.google.com/citations?user=S7k_gdkAAAAJ&hl=en

- Yuan, M., Yao, Q.: Testing common invariant subspace of multilayer networks. arXiv preprint arXiv:2406.05010 (2024)
- Yao, Q.: Comparison of non-learned and learned molecule representations for catalyst discovery (2022).

• Current Research Projects

Statistical Inference on Graphs *2024*

- Community Detection, Graph Modeling, Hypothesis Testing

• Previous Related Research Projects

• *Medical Data Science, Healthcare, Drug Discovery, Precision Medicine* *2023*

- 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.

• *Outlier Analysis & Anomaly Detection* *2023*

- Time Series Data, Multivariate Data.
- Anomaly Detection in Financial Fraud.

- *Algorithms, Models, Statistical and Mathematical Modeling* *2022, 2023*
  - How algorithms learn?
  - Comparison of Parametric and Nonparametric Models
  - Frameworks of deep neural networks
  - Frameworks of graph neural networks
- *Recommendation Systems* *2021, 2022*
  - Recommendation in e-commerce,
  - Statistical methods, machine learning, deep learning, graph neural networks

Related Course Projects

- *STAT874 Generalized Linear Models* *2024 Fall*
  - Dispersion Analysis
- *MATH650 Real Analysis* *2024 Fall*
  - Proof: Power Series
- *STAT672 Time Series* *2024 Spring*
  - Analysis of Time Series Models for Electricity Consumption Forecasting, R
- *STAT770 Survival Analysis* *2024 Spring*
  - Regression Models for Survival Analysis in CRC Considering Staging Groups, R
- *STAT764 Multivariate Methods* *2023 Fall*
  - Multivariate Analysis for Discrimination of Carcinogenesis Staging, SAS
- *STAT661 Applied Linear Models* *2023 Fall*
  - Detection and Evaluation of Outliers by Linear Models, R
- *STAT669 Introduction to Biostatistics* *2023 Spring*
  - Descriptor based multiple linear regression model for molecule property prediction, python
- *STAT860 Statistical Machine Learning* *2022 Spring*
  - Statistical Methods for Recommender System, python
- *CE793 Machine Learning for Engineers* *2020 Spring*
  - Multi-label classification based on image similarity, python
- *CSCI846 Distributed Systems* *2020 Spring*
  - 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

- *CSCI879 Advanced Data Mining* *2019 Spring*

- Network Mining and analysis using deepwalk, line, and node2vec, python

- *CSCI724 Introduction to Artificial Intelligence* *2019 Spring*

- Large scale study of programming languages and code quality in github, python

- *CSCI765 Introduction to Database Systems* *2019 Spring*

- Evaluation of real estate market using deep learning, python

SKILLS

- *Programming Languages*

  R, Python, etc.

- *Computing Softwares*

  Latex, Matlab, Octave, Minitab, JMP, etc

- *Applied Sciences*

  Experienced and Gained Deep Knowledge in Interdisciplinary Sciences and Engineerings

WORKING EXPERIENCE

- **North Dakota State University** *Fargo, US*

  Graduate Teaching Assistant in Department of Statistics *2024-present*

  stat 725 Applied Statistics

  stat 726 Applied Regression and Analysis of Variance

  Graduate Teaching Assistant in Computer Science Department *2019-2023*

  csci 160 Computer Science I

  csci 160 Computer Science II

  Lab Assistant in Plant Science Department *2019-2023*

  Experimental Design (seeding, planting, harvesting, data collection and entry)

  Data Analysis

- **Previously**

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

  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, CN*

  Project Assistant *2016 - 2017*

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

SELF IMPROVEMENT

- *Unverified Courses about Theories in Math and Stat* *Winter, 2024*

  UCLA STATS 200B Theoretical Statistics, Arash A. Amini
  http://www.stat.ucla.edu/~arashamini/teaching/

  UCLA STATS 200C High-dimensional Statistics, Arash A. Amini
  http://www.stat.ucla.edu/~arashamini/teaching/

- *Unverified Courses about Theories in Math and Stat* *Summer, 2024*

  UCSD MATH 181A Mathematical Statistics, David Quarfoot
  https://podcast.ucsd.edu/watch/sp22-b/math181a_a00

  UCSD MATH 181B Mathematical Statistics, David Quarfoot
  https://podcast.ucsd.edu/watch/sp22-b/math181b_a00

  UCSD MATH 180A Introduction to Probability for Data Science, Todd Kemp
  https://mathweb.ucsd.edu/~tkemp/180A/
  https://podcast.ucsd.edu/watch/fa19/math180a_b00/

  MATH 38161 Multivariate Statistics and Machine Learning, Korbinian Strimmer
  https://strimmerlab.github.io/korbinian.html

  UCLA STATS 100C Linear Models, Arash A. Amini
  http://www.stat.ucla.edu/~arashamini/teaching/

- *Unverified Courses about Applied Economics* *Spring, 2024*

  MIT OpenCourseWare: Development Economics
  https://ocw.mit.edu/courses/14-771-development-economics-fall-2021/

  MIT OpenCourseWare: Topics In Mathematics With Applications In Finance
  https://ocw.mit.edu/courses/18-s096-topics-in-mathematics-with-applications-in-finance-fall-2013/

- *Unverified Courses about Drug Discovery* *Spring, 2023*

  DavidsonX: Drug Discovery & Medicinal Chemistry, edX
  https://www.edx.org/learn/drugs/davidson-college-drug-discovery-medicinal-chemistry

- *Verified Courses with Certificates on Coursera*

  IBM Data Science Specialization *May, 2023*
  https://www.coursera.org/account/accomplishments/professional-cert/NJ2PYKM3KYDU

  Machine Learning, Coursera *May, 2023*
  https://www.coursera.org/account/accomplishments/certificate/HSNF9PYJVDUW

  Deep Learning Specialization, Coursera *March, 2021*
  https://www.coursera.org/account/accomplishments/specialization/certificate/FFBNKVM82AXS

VARIOUS TYPES OF OUTPUTS

- *Writing Papers*

- Proof

- Graph Mathematical Models

- *Implementations*

- Use R Markdown in RStudio Build Personal Website
  https://blog-4-ds-ml-stat.netlify.app/