

MABEL YAO

<https://blog-4-ds-ml-stat.netlify.app>

www.linkedin.com/in/mabelqqy

qianqian.yao@ndsuo.edu | qianqian6.yao@gmail.com

EXPECTED POSITION & CAREER

Data Scientist/ Machine Learning Engineer

EDUCATION & TRAINING

North Dakota State University, Fargo, US

2024-2026

Doctor of Philosophy in Applied Statistics

GPA: 4.0/4.0

Concentration: Machine Learning, Data Science, Representation Learning, Graphs

North Dakota State University, Fargo, US

2022-2026

Doctor of Philosophy in Computer Science

GPA: 4.0/4.0

Concentration: Machine Learning, Data Science, Representation Learning, Graphs

North Dakota State University, Fargo, US

2019-2022

Master of Science in Computer Science

GPA: 4.0/4.0

Concentration: Machine Learning, Data Science, Representation Learning, Graphs

Tohoku University, Sendai, JP

2013-2016

Master of Science in Architectural and Building Science

GPA: 3.68/4.0

Concentration: Seismic Evaluation & Damage Assessment, Structure Analysis

University of California, DAVIS, US

March, 2015

Cooperative Laboratory Study Program (training)

Academic English Program for Science and Technology

Concentration: Structural Engineering

Dalian Jiaotong University, Dalian, CN

2009-2013

Bachelor of Engineering in Civil Engineering & Software Engineering (double majors)

GPA: 87/100

Concentration: Building Structural Design, Software Development

INTEREST & SKILL

Machine Learning, Statistics, Data Science, Embedding/Representation Learning, Graphs

Applied Domains:

Molecule Representations/Molecular Property Prediction, Cheminformatics, Bioinformatics, Drug Discovery, Medicinal Chemistry, Biochemistry;

FinTech/Quant/Econometrics, Outlier Analysis;

Climate Tech/Climate Change;

Risk Analysis in Seismicity;

Programming Languages

Often Used: Python, R, SAS, Latex,

Scientific Computing: Julia, Matlab, Octave,

Computer Science: Java, SQL, HTML/CSS/JavaScript, Git, Anaconda,

ONGOING RESEARCH

Molecule Representations for Drug Discovery

Chemometrics,

Descriptor based Statistical Modeling,

Sequential Modeling,

Graph Modeling,

Molecular Property Prediction

Graph-Level Representation Learning for Chemical Screening.

Catalyst Discovery/ Material Discoveries.

Adverse Drug Reaction Detection

Knowledge Graph Embedding, Drug-Drug Interaction Network.

Adverse Drug Reaction in Clinical Trials in Drug Discovery.

Graph Neural Networks

How algorithms learn?

Frameworks of graph neural networks

Outlier Analysis

Time Series Data, Multivariate Data.

Anomaly Detection in Financial Fraud.

Recommendation Systems

Recommendation in e-commerce,

Statistical methods, machine learning, deep learning, graph neural networks

WORK EXPERIENCE

Teaching Assistant, North Dakota State University

Fargo, US

Department of Statistics

2024-2026

stat725 Applied Statistics

stat726 Applied Regression and Analysis of Variance

Department of Computer Science

2019-2023

csci160 Computer Science I

csci161 Computer Science II

Laboratory Assistant, North Dakota State University

Fargo, US

Department of Plant Science

2019-2023

(2019fall, 2020spring, 2020summer, 2022summer, 2023summer)

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

Data Analysis

Structural Engineer

Shenzhen, CN

Shenzhen Yuanlizhu Engineering Consultants Co.,Ltd

2017-2019

using computer aided engineering tools to design and analyze building structure,

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

Project Assistant

Shanghai, CN

Shanghai Saiyo Construction Technology Co.,Ltd

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;

also took the role of translator between Japanese and Chinese during the meetings.

Intern

Sendai, JP

Yamashita Sekkei INC. Tohoku Branch

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.

RELATED COURSE PROJECT & RESEARCH PROJECT

Molecule Representations, Multivariate Analysis, Outlier Analysis

2023

Multivariate Analysis for Discrimination of Carcinogenesis Stages, SAS

Fall

Detection and Evaluation of Outliers by Linear Models, R

Molecule Representation Learning for Virtual Screening in Drug Discovery, python

Spring

Descriptor based multiple linear regression model for molecule property prediction, python

Graph Representation Learning, Molecule Representation Learning

2022

Knowledge graph embedding for drug discovery

Fall

Comparison of Non-Learned and Learned Molecule Representations for Catalyst Discovery

Spring

Statistical Methods for Recommender System, python

Graph Representation Learning, Molecule Representation Learning

2021

Molecular Representation Learning for Catalyst Discovery, python	<i>Fall</i>
Graph Representation Learning: a survey on graph convolutional neural network, python	<i>Spring</i>
Computer Vision, Natural Language Processing	<i>2020</i>
Natural Language Processing: text classification, python	<i>Fall</i>
Natural language Processing: chatbot as virtual assistant, python	
Distributed database built on client-server architecture, java	<i>Spring</i>
Multi-label classification based on image similarity, python	
Network Mining, Deep Learning, Recommendation	<i>2019</i>
Implementation of recommender system based on different models, python	<i>Fall</i>
Implementation of expert system for real estate recommendation by drools, java	
Large scale study of programming languages and code quality in github, python	<i>Spring</i>
Network Mining and analysis using deepwalk, line, and node2vec, python	
Evaluaitoion of real estate market using deep learning, python	

CERTIFICATE

IBM Data Science Specialization , Coursera	<i>May, 2023</i>
https://www.coursera.org/account/accomplishments/professional-cert/NJ2PYKM3KYDU	
Machine Learning , Coursera	<i>May, 2023</i>
https://www.coursera.org/account/accomplishments/certificate/HSNF9PYJVDUW	
Deep Learning Specialization , Coursera	<i>March, 2021</i>
https://www.coursera.org/account/accomplishments/specialization/certificate/FFBNKVM82AXS	

PUBLICATION

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