

Qianyi Xue

qianyi.xue@duke.edu | [Github](#) | [LinkedIn](#) | Durham, NC

EDUCATION

DUKE UNIVERSITY

Master of Science in Electrical & Computer Engineering

Expected 08/2023 - 05/2025

SICHUAN UNIVERSITY

Bachelor of Engineering in Computer Science

09/2019 - 06/2023

SKILLS

- **Languages:** Python (advanced), Java (intermediate), JavaScript(intermediate), SQL (intermediate), C++, C, dart, Shell.
- **Frameworks:** Spring MVC, NodeJS, React, Node.js, MongoDB, MyBatis, PyTorch & Hugging Face (AI Framework).
- **Developer Tools:** Linux, Git, Relational Databases, MySQL, Redis.
- **Software Applications:** Anaconda, SPSS, Matlab, Arduino.
- **Certifications & Training:** Introduction to Computation and Optimization for Statistics (UCLA summer online course 2021), Web Mining (NUS School of Computing Summer Workshop 2022).

WORK EXPERIENCES

Sichuan University-Pittsburgh Institute Teaching Assistant

Chengdu, China

Data Structures (CS 0445 2 sections)

02/2023 - 06/2023

- **Led weekly class discussions** on the following topics: Java basics--Object-oriented languages, Array-based implementation of Deque, Implementation of Deque using Doubly Linked List and Circular Array, Efficient StringBuilder based on Circular Linked List, Implementation of Double Linked List using Recursion, Backtrace, Implementation of Merge Sort and Quick Sort.

West China Biomedical Big Data Center Research Intern

Chengdu, China

Optimize the mRNA vaccine sequence algorithm for the covid-19

04/2022-02/2023

- Developed a **machine translation (MT) model** between the UTR and ORF of the RNA sequence and enhanced the **BLEU-4 score** from **9.46** to **30.21** to optimize the RNA sequence.
- Redesigned the pre-training task of **DNABERT** to enable **RNABERT** to hold up to **2551** tokens compared to the original **512** tokens of BERT. Pre-trained RNABERT with a total of **152,000** collected UTRs and ORFs data.
- Constructed a novel RNABERT2RNABERT model using **BERT2BERT** and pre-trained RNABERT, which introduces pre-trained gene sequence models to the translation task of gene languages for **the first time**.
- Applied MT model to optimize mRNA sequence which confirmed to produce of **highly expressed proteins** and be applicable in vaccine development.

Yillion Bank Data Analyst Intern

Beijing, China

Intelligent Risk Control Department

12/2021 - 02/2022

- Used **xlwings** package in **Python** to process weekly, daily and annual data in Excel data (VBA), and **built an automated data process workflow** that greatly **reduced time costs**. Saved time manual filling time and is highly praised by the team leader.
- Built complex **SQL** to extract data from the data lake and completed **up to 300 lines** of SQL statements in a single query.
- Performed **binning method** to extract data and conducted customer portrait analysis.

PROJECTS

Full Stack MERN IOS/Android App

06/2023 - 07/2023

- Developed a full-stack Inventory Management App with the MERN stack from scratch, with **React** and **SCSS** on the front-end and **NodeJS** and **Express** on the back-end and **Axios** to connect the front-end and back-end.
- Built **RESTful APIs** using **Node.js (Express)** and **MongoDB** for handling HTTP requests and responses, along with performing CRUD (create, read, update, delete) operations.
- Utilized **Redux Toolkit** to manage application states and deployed the app to **Heroku** and **Render**.

Essay: Breast Cancer Classification Based on Various CNNs and Classifiers

05/2022 - 07/2022

- Extracted image features for breast cancers classification on the BreCaKHis dataset and tested different backbones, including **VGG-16, VGG-19, Xception, ResNet50, Inception-V3, and Inception-Resnet-V2**.
- Compared different network performance, including **fully-connected layer, logistic regression (LR), and SVM** resulting a 93.9% accuracy rate on eight-classification with the ResNet50-SVM model based on cyclical learning rate policy.

Mathematical Contest in Modeling (MCM) | Meritorious Award

02/2021 - 02/2021

- Compared several machine learning algorithms (**LR+word2vec, LR+TF-IDF, xgboost+word2vec, xgboost+TF-IDF, xgboost+LR+word2vec**) with Bert on text classification task.
- Extracted keywords by word frequency statistics and text clustering.
- Implemented difference processing on time series data and used **ARIMA** to make predictions based on time series to determine ARIMA model parameters.