

CV for Pei Tian

Tel: +1-6464699928 E-mail: tpatrix29@outlook.com

Academic Education

TONGJI UNIVERSITY, Shanghai, China

BS in Bioinformatics

Sep 2019 - Jun 2023

- ♦ GPA: 4.58/5.0, 90.81/100 Rank: 2th/33

Minor in Software Engineering

Sep 2021 - Jun 2023

- ♦ GPA: 4.81/5.0, 96.25/100

COLUMBIA UNIVERSITY, New York, NY

MS in Biostatistics, Public Health Data Science Track

Sep 2023 - Expected in May 2025

- ♦ GPA: Not released yet

Research Experience

Evaluation of Reference Datasets Oriented to Single-cell Annotations

Jul 2022 - Jun 2023

Advisor: Prof. Qi Liu

Undergraduate Thesis

- Designed the sampling method based on the determinantal point process and oversampling algorithm, which is able to generate high-quality reference data while keeping the original data amount.
- Implemented the method by Python and packaged code into a library with an MIT open-source license.
- Evaluated the method by comparing the automatic annotation results between original and refined datasets via several models, including Seurat, SciBet, scmap, etc.

PhageMAP (Phage-Microbiome Assist Phagotherapy) for iGEM 2021

Jan 2021 - Nov 2022

Advisor: Prof. Qi Liu

Group Research & International Competition

- Achieved BEST Award in Software Track, iGEM 2021
- Developed phagotherapy assisting method for analyzing bacterium-phage interactions by using algorithms including CRISPRDetect, BLAST, PHP, WisH and statistical methods including the kernel.
- Designed and implemented a relational database that stored information of bacterium, phage and bacterium-phage interactions via MySQL platform.
- Offered a user-friendly usage of analysis procedure by deploying the software environment and packages in a Docker container and uploading the container to DockerHub with instructional documentation.
- Implemented the front-end webpages using HTML/CSS/JavaScript collaboratively.

Prediction for Solid Waste Composition of Different Neighborhoods

May 2022 - Oct 2022

Advisor: Prof. Dongjie Niu

College Students' Innovative Entrepreneurial Training Plan Program

- Implemented data cleaning, preprocessing and feature engineering for raw datasets.
- Predicted solid waste composition using neural network model while utilizing many machine learning technologies including L2 regularization, Adam optimizer and dropout layer via PyTorch library.
- Visualized data features and model evaluation results via matplotlib and Tensorboard library.

Internships Experience

Shanghai Foxhub Network Technology Company

Aug 2022 - Oct 2022

R&D Department, Database Engineer

- Designed relational database architecture (ER diagrams) and unstructured data source (OSS, Object Storage Service) on the Alibaba Cloud platform to support front-end webpage access.
- Managed database access permission and backup operation by writing scripts to ensure production and development environment stability.
- Wrote working log and instructional documentation about database usage to smooth collaboration.

Course Projects

Course Management System

Oct 2022 - Jan 2023

Advisor: Prof. Jie Wang

Course Design of Software Engineering

- Specified requirements, system design and implementation technologies under the view of micro-service.
- Implemented backend service by MySQL, MongoDB, Docker, SpringMVC and front-end web page by React, Webpack, Axios, Node.js, etc.
- Wrote development documentation for collaboration and user manual for friendly usage.

PlantDB (Plant Information Application)

May 2022 - Jun 2022

Advisor: Prof. Keping Yuan

Course Design of Foundation of Database

- Designed and implemented relational database on SQL Server platform to store data of plants and users.
- Built a desktop application on VS.NET platform to deliver data visualization and user interaction.

Neurodegenerative Diseases Onset Prediction

Jul 2022

Advisor: Prof. Lulu Gong

Course Design of Practice of Medical Information

- Implemented data cleaning, preprocessing and feature engineering for open-source patient data about the onset of Alzheimer's disease and Parkinson's disease.
- Trained the prediction models by classical machine learning algorithms including SVM, decision tree, etc.
- Demonstrated project details and delivered an interactive prediction interface on a webpage constructed by the Flask framework.

Skills & Interests

Core Courses:

- **Mathematics:** Advanced Mathematics (Calculus, Differential Equation, etc), Linear Algebra, Probability, Statistics, Discrete Math, Numerical Methods and Computer Algorithms
- **Computer Science:** C/C++ Programming, Foundation of Data Structure, Software Engineering, Software Test, Database and Data Warehouse Technology, Introduction to Embedded Systems, SOA and Web Service
- **Data Science:** Machine Learning Theory and Algorithms, MIT xPRO Machine Learning Course
- **Bioinformatics:** Modern Bio-computing Environments, Computational Genomics, Bioinformatics

Programming Language: C/C++, Python, SQL, Java, R, shell, HTML5/CSS3/Javascript

Development Suite: PyTorch, SpringMVC, React, Docker

Interests: Machine Learning, Deep Learning, Omics Data Analysis

Awards

First Class Scholarship in Tongji University, Second Prize in the Mathematics Competition of Tongji University, Third Prize in the National College Students Mathematical Modeling Competition