Yanjie Xu

📞 19895808812   ✉ 19895808812@163.com

🎓 Education

- **Nanyang Technological University (QS 15)** — Aug 2024 – Dec 2025
  - *M.S. in Biomedical Data Science*   Major GPA: 4.67/5.0

- **Xi'an Jiaotong-Liverpool University** — Sep 2020 – Jul 2024
  - *B.S. in Information and Computing Science*   Major GPA: 3.9/4.0   First Class Honors
  - **Awards and Scholarship:** 2022-23 University Academic Achievement Award (top 10%)

Seeking full-time, on-site internship from March to September 2025.

📄 Internship Experience

Software Development Intern | Shanghai Yunkuo Information Technology Co., Ltd   Jun 2024 - Jul 2024

○ Optimized database interactions using MyBatisPlus by auto-generating CRUD interfaces, reducing redundant code, and improving batch processing performance.

○ Optimized SQL queries for the user permissions database by analyzing execution plans and adding indexes to high-frequency fields, reducing query response times by 30%.

🏛 Project Experience

Data Science in Educational AI | Individual Project   Sep 2024 - Jan 2025

○ Aimed at investigating the role of artificial intelligence in supporting educators at different stages of teaching.

○ Designed a questionnaire and conducted Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) using SPSS to develop and validate research models.

○ Conducted **correlation analysis**, **ANOVA**, and **MANOVA** to identify significant variable relationships and differences.

★ **This work is under review at the Journal of Behavior and Information Technology.**

Data Mining of Differential Gene Expression | Group Project (Team Leader)   Aug 2024 - Oct 2024

○ Aimed at analyzing gene expression differences in leukemia patients by mining and visualizing data to identify key factors influencing disease progression.

○ Collected and cleaned datasets from public sources, addressing outliers and sample size imbalances with bootstrap methods. Used Kolmogorov-Smirnov tests and fuzzy scoring to identify differentially expressed genes (DEGs).

○ Visualized gene expression with volcano, density, box, scatter, and heatmaps, showing no genes controlling both relapse status and timing.

Disease Prediction Modeling with Environmental Data | Individual Project   Feb 2024 - May 2024

○ Aimed at building models using environmental data provided by official sources to predict the weekly dengue fever cases in two specific cities.

○ Used Python for data preprocessing tasks, including dimensionality reduction, interpolation, and normalization.

○ Built predictive models using **Neural Networks**, **Gradient Boosting Regressor**, and **LSTM**, applying feature selection and grid search for optimization. The Neural Network achieved the best performance with an MAE of 6.13.

🔧 Skills

- IELTS Score: 6.5
- Programming & Tools: Python, Java, SQL, HTML
- AI & Algorithms: Machine Learning, Neural Networks, Dynamic Programming, Cryptography, Sorting, Graph Theory