JIAYUE QIN

Email: qinjiayuekkk@163.com Mobile: +65 8084 4215 https://www.linkedin.com/in/jiayue-qin-data/

EDUCATION

National University of Singapore **Aug 2024 – Dec 2025**

- MSc. in Data Science and Machine Learning

East China Normal University **Sep 2020 – Jun 2024**

- B.Sc. in Statistics (GPA: TOP 20%)
- Key Courses: Probability Theory and Mathematical Statistics (98), Machine Learning (99), Time Series Analysis (92), Corporation Finance (92), Bayesian Statistics (100)
- Awards: University-level Scholarship (Second and Third Prize)

SKILLS AND LANGUAGES

- Programming Skills: Python(pandas, numpy, sklearn, keras), SQL, R, QuickBI, SPSS, Thinkcell
- Analytical Skills: Machine Learning, Experimental Design, Casual Inference, Data Visualization, Data Crawling
- Languages: Fluent in English and Mandarin (both spoken and written)

WORK EXPERIENCE

LVMH - Data Modelling Intern **Feb 2024 – Jul 2024**

- Transformed large, complex data into user and product features using SQL, and helped my mentor building an XGBoost classification model to predict customer purchasing behaviour, achieving a high accuracy (AUC of 0.8) and increasing purchase rates by 350%-440%.
- Developed dashboards to monitor model performance (AUC, F1 score, etc.), purchase rate, and data source anomalies in order to evaluate and optimize models by using QuickBI.
- Conducted feature quality checks, implemented Python scripts for automation, saving time.

Zhongyan Technology - Research Intern **Jul 2023 – Nov 2023**

- Visualized survey data through bar charts, word clouds, etc., to analyse feedback from McDonald's events, and presented enhancements in reports to improve customer experience.
- Executed significance tests on satisfaction metrics across different demographics and regions, aiding in product positioning.
- Self-taught web scraping techniques to extract store information and implemented automated processes to extract and analyse review keywords.

PROJECTS

Prediction of Chinese NEV Sales and Evaluation of Regional Policy Effects Jan 2024 – May 2024

- Collected data from macroeconomic, policy, and product perspectives in sources like WIND, government websites; cleaned raw data using numpy and pandas, resulting in 11 features.
- Built SARIMA and LSTM models to predict NEV sales, with final multivariate LSTM model achieving a higher accuracy (12.91% MAPE).
- Utilized Synthetic Control Method to prove two non-subsidy policies increased NEV sales in China.

Prediction of Elderly Health Status Using Stacking (Kaggle Top 15%) **May 2023 – Jun 2023**

- Visualized data distribution using violin plots, heatmaps, etc., to identify features distributions.
- Compared models using balanced log loss and 5-fold cross-validation. Optimized parameters for CatBoost , LightGBM with Optuna. Evaluated variable importance with Gini and SHAP. Implemented a Stacking model, reaching a balanced log loss of 0.405.

Digital Technology Empowering Chinese Cultural Theme Park Construction Dec 2022 – May 2023

- Collaborated with four teammates to direct on-site interviews at Fangte Theme Park, developing ideas for a questionnaire and distributing 2,723 questionnaires.
- Clustered populations and calculated gray correlation using SPSS Pro, providing differentiated recommendations for various demographics.