



Education

Applied Statistics, Shanghai Lixin University of Accounting and Finance; 2020-2024

- Grade: 88.83/100
- Modules: Mathematical Statistics(85);Applied Time Series Analysis(92);Applied Regression Analysis(81);Applied Multivariate Statistical Analysis(93);Statistical Computing(python)(94);Statistical Machine Learning(97)

Applied Statistical Modelling And Health informatics, King's College London; 2024-present

- Modules: Causal Modelling and Evaluation;Introduction to Health Informatics;Multilevel and Longitudinal Modelling ;Bioinformatics, Interpretation and Data Quality in Genome Analysis;Multilevel and Longitudinal Modelling

Internships

Customer data analysis, Atlas Copco; Apr 2024 - Jun 2024

- Expanded the SharePoint platform, organized AI Brain tools files, and implemented automated dialogue.
- Organized over 200 documents in SharePoint, enhancing accessibility and reducing retrieval time by 40%.
- Visualized 120,000 customer data points with Power BI, contributing to sales productivity.

Mathematical Modelling Experience

MSc Dissertation Project , Large Language Models (LLMs) for Electronic Health Records,Nov 2024 -Jun 2025

- Built a Retrieval-Augmented Generation (RAG) clinical Q&A system for patient record summarization and retrieval, achieving 60% accuracy on the MIMIC-IV dataset (14,000+ patients)
- Applied MiniLM embeddings, Chroma/FAISS vector databases, and LangChain pipelines to enhance clinical text retrieval & generation
- Developed interactive interfaces (Streamlit/Gradio) for real-world clinical query and summary display

Bachelor's thesis, Semantic Context Analysis Based on Machine Learning -An Example of Sina Weibo Rumor Detection Research

- Collected 6000+ raw data entries using Python web scraping; cleaned the data to retain 3516 valid entries.
- Analyzed text characteristics using the Word2Vec model to extract text features and applied Stacking for model ensembling

2022 National University Student Data Analysis Competition, 1st Prize; Dec 2022 - Jan2023

- Independently completed the pharmaceutical e-commerce sales data analysis report based on 6,000 Taobao sales data.
- Used Python and R to clean data, process regular expressions, and establish ARIMA, seasonal ARIMA (multiplicative and additive), Holt-Winters, three-parameter exponential smoothing models, and KNN time series models.
- From the perspective of the e-commerce platform, conducted a brand SWOT analysis, developed a 4Ps marketing strategy, and constructed indicators for pharmaceutical e-commerce operations to formulate strategies for new brands.

Group leader, 2022 "Asia and Pacific Mathematical Contest in Modelling"(AMPCM), 2nd Prize; Nov 2022

- Analyzed global temperature data from 1722 to 2022 to predict the temperature for 2024 and examined the factors affecting changes in global temperature data.
- Constructed machine learning and time series models such as ARIMA, KNN, Lasso, and XGBoost using R and Python, providing suggestions for addressing global warming.
- Served as team leader, managing task allocation, establishing major models, and writing the thesis.

Group leader, 2023 Mathematical Contest In Modelling, Honorable Mention; Feb 2023

- Analyzed 360 data points from the Wordle game and used a time series model to predict player numbers for the next three years.
- Utilized Python to analyze data by extracting the information entropy of English words and the number of vowels. Employed K-means and machine learning methods for classification and prediction.

Campus Experience

Head of Reporting Department – Lixin University Media Center, Jun 2021 – Jul 2022

- Directed daily operations of reporting team, including interviews, features, and event coverage
- Edited and reviewed submissions to ensure quality and accuracy
- Coordinated major campus events such as press conferences and anniversary activities

Award

- 2024 Shanghai Outstanding Graduates
- 2021-2022 Merit Student
- 2020-2021 Outstanding Student Leader
- 2021-2022; 2022-2023 Merit Student Scholarship 2nd price
- 2020-2021 Merit Student Scholarship 1st price

Skills & Volunteering

Technical Skills: Python, R, Spss, Stata, Power BI, Galaxy

Languages: English (IELTS 6.5; CET-6: 518; CET-4: 530) and Mandarin

Volunteering: Outstanding Volunteer – Robotex Asia 2021; China International Consumer Goods Expo 2021