CHEN QIAN

399 Crocker Lane, One Shields Avenue, Davis, CA 95616
Phone: (+1)5712670414
Email: mrcqian@ucdavis.edu
Homepage: https://chen20191020.github.io/ChenQian

OBJECTIVE

I am actively applying for doctoral programs in Statistics, Data science, and Econ & Stats for Fall 2023.

EDUCATION

University of California, Davis — **Davis, CA**
M.S. Statistics (*Data* Science Track) — *Sep.2021 - Jun.2023*

Complutense University of Madrid — **Madrid, MAD**
B.S. Applied Statistics — *Sep.2016 - Jul.2020*
Advisor: Enrique Gonzalez Aranguena
Thesis: A comparison of Box-Jenkins ARIMA models and Artificial Neural Network models for financial time series.

RESEARCH INTEREST

- Time Series Analysis (Stationary, Non-stationary, Nonlinear, Wavelets)
- Machine Learning(RNN, LSTM, BERT)
- Bootstrapping
- Functional Data Analysis

RESEARCH EXPERIENCE

University of California, Davis — *Davis, CA*
Independent Research — *Jun.2022 - Present*
Supervisor: Xiucai Ding

- Created '*Sie2nts*' R package for auto-regressive approximations to non-stationary time series with estimation via the method of sieves and 2 high dimensional test basis on multiplier bootstrap procedure. Available at [*Link*].
- Provided function to generate wavelet table by Cascade algorithm and tables of scaling function for Daubechies 1-20, Coiflet 1-5. Wrote a user's guide for '*Sie2nts*'. Available at [*Link*].

Complutense University of Madrid — *Madrid, MAD*
Thesis Research — *Jan.2020 - Jul.2020*
Supervisor: Enrique Gonzalez Aranguena

- Compared the forecasting ability of ARIMA and 2 layers LSTM models for financial time series in long and short periods based on MAPE and MdAPE.[*Link*]

TEACHING EXPERIENCE

University of California, Davis *Davis, CA*
Teaching Assistant *Jan.2023 - Mar.2023*
Supervisor: Shizhe Chen

- *STA207*: Statistical Methods for Research II (Winter 2023). Topics are related to linear and nonlinear statistical models.

INDUSTRY EXPERIENCE

Zhongyu Tech (mobile game) *Wuhan, China*
Data Scientist *Sep.2020 - May.2021*

- Promoted the company to establish the data center team, formulated data-driven solutions to help the company's product decision-making, and created technology platform for data analysis.
- Designed metrics and analyzed experiments to evaluate features for top100(Spring 2021) game "Wheel Offroad 3D"; Improved retention rate by cohort analysis and monetized the fitness of revenue.
- Constructed ETL procedure with API functions provided by Sensor-Tower and Performed EDA about mobile game market.
- Wrote weekly reports and delivered on the dashboard to aid other partners in design and engineering using.
- Optimized features of RFM model and K-means algorithm to create user cohort for "Wheel Offroad 3D".

PROJECTS

Predicting Acute Aquatic Toxicity by QSAR model
Supervisor: Debashis Paul

- A dataset of 908 chemicals was used to develop a multi-linear regression model and kNN method to predict the LC50 96 hours for the fathead minnow(Cooperated with *Le Chen*).[*Report*, *Slides*]

Forecasting of COVID-19 endpoints in US, Spain and Canada
Supervisor: Shizhe Chen

- Built features for COVID-19 data to predict pandemic endpoints in 3 different countries using Bidirectional LSTM and compared results based on experiment design (Keras, Statsmodels).[*Report*]

Causal study in mobile game industry

- Wrote annual report with content of the seasonality analysis of game market(trough in Feb) and causal study between big events and mobile game downloads (Language: Mandarin, Cooperated with *ChengJun Zhang*).[*Slides*]

COURSES

University of California, Davis

- **Stats, DS, CS**: Probability Theory (Miles Lopes), Statistical Inference (Can Le), Statistical Methods and Research (Jie Peng, Shizhe Chen), Applied Time Series (Xiucai Ding), Data Web Technologies for Data Analysis (Duncan Temple Lang), Multivariate Data Analysis, Statistical Data Science

- **Planned**: Optimization for Big Data Analytics (Mina Karzand), Applied Statistics (D. Paul), Generalized Linear Models (H. Mueller), Independent Research (Xiucai Ding)

Complutense University of Madrid

- **Stats, Math**: Probability Theory and Stochastic Processes, Statistical Estimation(Inference), Optimization Techniques, Survey Sampling, Design of Experiments, Multidimensional Analysis, Simulation and Queuing Systems, Lineal Linear Prediction Methods, Industrial Applied Statistics, Time Series, Survival Analysis
- **DS, CS**: Programming I & II(C++), Statistical Software I(SAS), Statistical Software II(R), Computational Methods for Mathematics(Matlab), Database Design, Data Cleaning, Data Structures and Algorithms

HONORS

The Honours(The top student in 10 courses at Complutense) *2016-2020*

Top performance(in M.S. Comprehensive Exam at UC Davis) *2022*

SKILLS

programming: Python, R, MySQL, C++, SAS, Matlab
Languages: English(fluent), Spanish(advanced), Mandarin(native)
Certificates: Machine Learning (Coursera)