Zhong-Yi Zheng

+65-86432457 • e0536234@u.nus.edu • 21 Lower Kent Ridge Road, Singapore 119077

EDUCATION

National University of Singapore — **Singapore**
M.Sc. in Financial Engineering — *Aug 2020 - Jun 2022*
- Coursework: Financial Econometrics, Derivatives and Fixed Income, Stochastic Calculus and Quantitative Method
- Opencourse: Stanford CS229 Machine Learning

Shanghai JiaoTong University — **Shanghai**
B.S. in Business Administration — *Sep 2015 - Jun 2018*
Microelectronics Science and Engineering (changed major) — *Sep 2013 - Jun 2015*
- Coursework: Probability and Statistics (A+), Linear Algebra (A), C++ Programming (A-)
- Awards: C-level Scholarship, 2014 (Top 20%)

EXPERIENCE

Asian Institute of Digital Finance (NUS) — **Singapore**
Model development intern, Credit Research Initiative — *Nov 2020 - Apr 2021*
- Translated 2k lines of Matlab production code to Julia language with parallel feature.
- Used Pandas to map 2.4 billion CRI history credit rating data with S&P's and Moody's.
- Improved likelihood calculation speed by 50 times by gpu computing module CuPy.
- Developed automation tools such as rebooting distributed Julia API system with email alarming.

Jiyan Information Technology — **Shanghai**
Co-Founder, Lead Product development — *Nov 2014 - Mar 2020*
- Built up a company with annual revenue 3.5 million RMB and social platform Wechat 25k subscribers .
- Conducted strategic management to maintaining company's expansion feasibility.
- Optimized routine workload by Numpy and Pandas for calculation of expenses and sales data.

China Merchants Bank — **Shanghai**
Project Collaborator, Retail finance division — *Mar 2019 - Dec 2019*
- Cooperated with head of retail finance division on electronic transactions subsidies project aiming new generation.
- Made deal with clients to open 2k cards in total, and attracted 3k new registered users for official application.

PROJECT

Sentiment analysis on credit news for default probability prediction (CRI)
Alternative Data, NLP, PyTorch, Python, C++ — *Nov 2020 - Mar 2021*
- Programmed parallel web crawler to fetch credit news from Reuters/FT/WSJ by Selenium.
- Modified C++ source code of semi-supervised SRC-LDA to make model could predict on new article.
- Pre-processed 56k articles by LDA to calculate credit relating score and NER to extract company names.
- Fine-tuned 24-layer BERT model with out-of-sample accuracy as 75% in 3 class and 66% in 5 class.
- Implemented logistic regression on default event prediction and obtained p-value as 0.002.

Ensemble of neural network and boosting decision tree for distance to default regression (CRI)
Time Series, TensorFlow Keras, Python — *Dec 2020 - Feb 2021*
- Designed proper feature engineering pipeline, including missing values processing and skew features transforming.
- Alleviated over-fitting by tuning proper dropout rate and L2 regularization based on k-fold cross-validation.
- Trained MLP and XGBoost model with out-of-sample R-2 both outperformed 2nd polynomial's by 15%.
- Stacked trained model by proper weight and achieved additional 2% R-2 improvement.

SKILLS

Programming: Proficient in C++ (STL, OOP), Python (Numpy, Pandas, Matplotlib, ...); Intermediate SQL, Julia, Matlab, R.

Compute Science: Enjoy coding and experienced in Algorithms and Data structure; Familiar with Linux, Shell, Vim.

Statistics & Machine Learning: Econometrics, Regression (LASSO, Ridge), Logistic Regression, Random Forest, XGBoost, K-Mean, KNN, PCA, SVD, Neural Network