Xueyan Hu (Shwayan)

Durham, NC, U.S. | +1 (919) 491-0113 | xueyanhu1231@gmail.com | Personal Webpage

EDUCATIONAL BACKGROUND

Duke University, Trinity College of Arts and Sciences **Aug. 2024 - Jul. 2026**
M.S., Statistical Science, *Cumulative GPA: 3.9/4.0*
Central University of Finance and Economics, School of Finance, China **Sep. 2020 - Jul. 2024**
Top 1 finance university in China
B.S., Financial Technology, *Overall GPA: 95.01/100 (# rank 1/244)*

RELEVANT COURSEWORK

Maths: Statistical Inference, Predictive Modeling, Advanced Algebra, Advanced Mathematics, Probability and Statistics, Discrete Mathematics, Analysis of Financial Time Series, Statistics, Ordinary Differential Equations, Stochastic Processes
Computer: Statistical Programming, C++ for Programming, Data Structure and Algorithms, Big Data and Finance, Principles and Applications of Artificial Intelligence, Deep Learning and Natural Language Processing, Cloud Computing and Big Data Technology

WORK EXPERIENCE

Quantitative Researcher, **Huatai Securities Co., Ltd., China** **Jun. 2023 - Sep. 2023**

- Developed a report-driven stock-selection strategy based on research of financial analysts on the Chinese stock market; reduced the turnover ratio from 18.46 to 12.84 by adding a component adjustment mechanism of market indexes; achieved an excess return of 38.32% and an Information Ratio (IR) of 4.00; implemented with Python
- Constructed a factor based on the jump in price with an IC (Barra Neutralized) of 2.40%; filtered stock pool based on Syntax Analysis of industry research reports; the strategy achieved an excess return of 32.37% and IR of 3.01; implemented with Jieba and Gensim
- Reproduced market indexes, including CSI932000, based on compilation requirements of index companies with 100% accuracy; tracked market index performance and conducted corresponding statistical work
- Tested I/O efficiency of different data types, including Parquet, Feather, Mat (MatLab), CSV, Pickle, and Jay
- Refactored projects with packages including Numba, Threading, and Collections; reduced the backtest time of the report-driven stock-selection strategy from 6 hours to 5 minutes

Quantitative Researcher, **Tongxinyuan (Sanya) Fund Management Co., Ltd., China**
Jan. 2023 - May 2023

- Designed and developed a backtesting framework for time-series strategies in the Chinese Future market with Python and MySQL
- Optimized profit and loss (P&L) decomposition models from Brinson and Barra with accuracy for each investment process, including selection and combination of factors
- Established cross-sectional selection model of Future market with Arrow-Debrew two-phase market model; achieved a Rank-Information Coefficient (Rank IC) of 0.013 and a winning rate of 52.43%
- Conducted sensitivity analysis to test the stability of factors, demonstrating the robustness and effectiveness of factors

RESEARCH EXPERIENCE

First Author, "How Policy Texts Affect Financial Markets: A Market State Recognition Perspective" (submitted to Journal of Empirical Finance)

Jul. 2023 – May 2024

- Built and preprocessed corpus of Chinese government policies by crawling websites; implemented with Shell
- Tokenized and eliminated stopwords with both StanfordNLTK and Jieba; improved time efficiency with multi-threading; implemented with Threading
- Established Doc2Vec-Kmeans Model and Deep Temporal Clustering (DTC) to recognize market states; improved results with LDA topic model
- Evaluated results with K-S test to determine the distribution of the difference return from portfolio analysis of FF3 factors in Chinese stock market
- Built a directed graph of the order of policy impact on industries; analyzed the centrality of each node, and found systemically important industries; examined the ability of the state of the market to predict the average value of the industry's rate of return

First Author, "Investor Sentiment Index Based on Prospect Theory: Evidence from China" (submitted to Applied Economics Letters, attended 2023 China Fintech Research Conference)

Mar. 2023 - May 2023

- Originated an Investor Sentiment Index based on Prospect Theory and Technical Analysis; implemented machine learning algorithms, including Principal Component Analysis (PCA) and Partial Least Squares (PLS)
- Varified the significance of the index with Fama-Macbeth Regression, Portfolio Analysis, and in-and-out-of-sample prediction; implemented under contexts including FF-3, FF-5, and different industries
- Analyzed economic mechanisms of indexes, including Campbell-Shiller Decomposition and Merton's Market Volatility Risk Channel

ACTIVITIES & PROJECTS

Big Data Analysis of Taobao Shopping Record **May 2023 - May 2023**

- Prepared data with Hive; transferred data between Hive and MySQL with Sqoop; operated under Linux
- Constructed the Logistic Regression Model and the Decision Tree Model to predict returned customers with an accuracy of 58%; implemented with Spark; visualized result with Tableau and Matplotlib

Spam Email Real-time Detector **Apr. 2023 - Apr. 2023**

- Monitored mailbox with Zmail; quantified email with TFIDF and BOW
- Eliminated stop words, and constructed Naïve Bayes model to do classification

Construction of Database of Hydrogen Energy by CUFE **Dec. 2021 - Mar. 2022**

- Collected policy targets on hydrogen energy from Chinese governments and sorted out the data
- Transported the data from the previous database to the new one

HONORS & AWARDS

- Hongru Scholarship (5 out of 10105 students)
- Second-Class Prize in China Undergraduate Mathematical Contest in Modeling (top 2.65%)
- First-Class Prize in the 32nd College Student Mathematics Competition (top 5%)
- Academic Outstanding Scholarship of CUFE (2021, 2022, and 2023) (top 10%)
- Honorable Mention in American Mathematical Contest in Modeling (top 30%)

PROGRAMMING SKILLS & INTERESTS

Computer: C++ (proficient); Python (proficient); R (proficient); SQL (proficient); Spark (proficient)
Interests: Marathon; Clarinet