HENGXU (HUGH) LIN

(646) 236-8558 | helin23@gsb.columbia.edu | [Homepage] | [LinkedIn]

EDUCATION

COLUMBIA BUSINESS SCHOOL — New York, NY
MS, Financial Economics GRE: 329 08/2021 - 05/2023

A highly selective program (7.3% acc. rate) which combines PhD and MBA courses in quantitative econometrics and finance
Coursework (PhD-level)*: Optimization, Statistical Inference, Econometric & Time Series Analysis, Continuous Time Finance*
Programming*: Python (PyTorch/Tensorflow), Linux/Bash, C++, Java/C#, SQL, 50k+ lines programming experience*

SUN YAT-SEN UNIVERSITY — Guangzhou, China
BS, Mathematics and Applied Mathematics GPA: 3.9/4 08/2015 - 06/2020

Coursework*: Neural Networks, Machine Learning, Probability & Statistics, Operations Research, Algebra, Real Analysis, ODE/PDE*
Honors*: Outstanding Graduate (highest distinct,Top 0.5%), First-place Scholarship, Microsoft Stars of Tomorrow*

EXPERIENCE

MICROSOFT RESEARCH — Beijing, China
Research Intern (Full-time), Machine Learning Group, Advisor: Dr. Weiqing Liu & Dong Zhou 06/2020 - 09/2021

- **Learning Multiple Trading Patterns (ACM-SIGKDD 2021)**
  - Proposed a lightweight extensive module, temporal routing adaptor (TRA), to automatically dispatch samples into multiple domains and select a best predictor, applied optimal transport restrict balance assignments while keeping lowest overall loss;
  - Researched ablation studies on influence of number of domains, performance of hidden states and memory mapping combination, surpassed state-of-the-art baselines' RankIC by 1%, Annual Return by 3.1%.
- **Deep Risk Model (ACM-ICAIF 2021):**
  - Formulated risk mining as a supervised learning task and overcame fundamental and statistical risk models' s deficiency;
  - Put forward a deep learning solution (GAT-GRU) for mining risk latent factors to improve covariance matrix estimation;
  - Advanced cutting-edge performance in R^2 by 1.9%, manifested stability and explainability in model and objective design.
- **Representation Learning of Stock Data**
  - **Deep Clustering**: Plugged reconstruction loss in auto-encoder with clustering based pseudo labels, projected data into a linear separable hidden space, accomplished 3% enhancement to baseline (analogue to manifold clustering);
  - **Contrastive Learning**: Applied a contrastive method in representing learning with optimal transport, exceeded benchmarks on downstream with 1st online clustering algorithm, achieved oracle accuracy on synthetic data (99%).
- **Qlib (1st open-source AI platform for Quant Finance):** Released the TRA model and its baselines. (github.com/microsoft/qlib)

SUNSHINE QUANT INVESTMENT CONSULTING (Prop Trading) — Shenzhen, China
Research Intern, Quantitative Research 02/2020 - 05/2020

- **Portfolio Optimization**: Built Black-Litterman model with risk parity strategy as prior and adversarial learning predictions as posterior; enhanced index's return by approx. 4%, max drawdown by 30%. (github.com/linhx25/BlackLittermanModel)
- **Adversarial Learning**: Constructed a GAN to empower LSTM's baseline by 2%, attained accuracy of 58% and MSE of 0.44%.

MORGAN STANLEY HUAXIN FUNDS (Mutual Fund) — Shenzhen, China
Quantitative Analyst, Risk Management 05/2019 - 08/2019

- **Performance Attribution**: Cracked funds' return into security selection and industry allocation based on Brinson model, implemented program in C# hence improved attribution algorithm thereby hastening original process by 5 times;
- **Fair Trading Analysis**: Computed profit discrepancies in trading window according to regulations in analyzing unfair trades.

SOUTHERN STATISTICAL SCIENCE LAB — Guangzhou, China
Research Intern, Financial Engineering 09/2017 - 02/2019

- **Numpy.NET**: Constructed a data structure emulating Numpy in C# and allowed efficient functions in time series analysis model;
- **Hidden Markov Model**: Utilized the EM algorithm to estimate MSVAR's parameters. (github.com/linhx25/MarkovSwitching)

PROJECTS

Undergraduate Thesis (Best Paper) – "*Readability, Opaqueness and Crash Risk*", Advisor: Prof. Keming Wang
- Developed a Python package to process firms' 10-K fillings (217G) and applied NLP methods for annual report textual analysis;
- Carried robustness test to eliminate endogeneity: searched alternative variables, tested fixed effect of panel data;
- Proved that 10-K fillings of high-risk firms are of low readability and high opaqueness. (github.com/linhx25/FReader).

Kaggle's PUBG Machine Learning Prediction (Ranked Top 1%)
- Conducted feature engineering and tested linear/non-linear models with regularization in pre-experiment;
- Implemented Random Forest, Genetic Algorithm, Particle Swarm and GBDT in Python for prediction, with 1.95% MAE;
- Ranked Top 1% in Kaggle's Project competition (kaggle.com/pubg-prediction , Team SYSU).

ADDITIONAL INFORMATION

Languages: Cantonese (Fluent), Mandarin (Native) | **Interests**: Hip-hop dance, Climbing (elevation 6,000 m), Cooking