HENGXU(HUGH) LIN

(86) 137-1934-4881 | 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% acc. rate) which combines PhD and MBA courses in quantitative econometrics and finance
Anticipated Coursework (All PhD-level)*:* Optimization, Applied Multivariate Statistics, High Dimensional Statistics, Statistical Inference, Assets Pricing, Econometric & Time Series Analysis, Microeconomics Analysis
Programming*:* Python (Torch/Tensorflow), Linux/Bash, C++, Java/C#, SQL, 50k+ lines programming experience

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

Coursework*:* Neural Networks, Machine Learning, Probability & Statistics, Optimization, Calculus, Algebra, Real Analysis, ODE, Algorithms, Financial Engineering, Investment, Corporate Finance, Econometrics, Micro/Macro-Economics
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 - Present

- **Learning Multiple Trading Patterns (ACM-SIGKDD 2021)**
  - **Temporal Routing Adaptor**: Proposed a lightweight extensive module to automatically dispatch samples into multiple domains and select a best predictor, applied optimal transport keep balance assignments while keeping lowest overall loss;
  - **Ablation Study**: Researched on influence of number of domains, performance of hidden states and memory mapping combination, and improvement on different SOTA models, enhanced top baselines' RankIC by 1%, Annual Return by 3.1%.
- **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%).
- **Deep Risk Attribution (Ongoing):** Formulated risk factor mining as representation learning; designed a risk factor mining model with efficient, adaptive and non-linear interpretability on the data sets.
- **Qlib (1st open-source AI platform for Quant Finance):** Released the TRA model and baseline tests. (github.com/microsoft/qlib)

SUNSHINE QUANT INVESTMENT CONSULTING (Prop Trading) — Shenzhen, China
Research Intern, Quantitative Research, Advisor: Dr. Guang Yang — 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 an architecture of GAN with a MLP as discriminator and an LSTM as generator for forecasting return; empowered single LSTM's accuracy by 2%, attained average accuracy of 58% and MSE of 0.44%.

MORGAN STANLEY HUAXIN FUND (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, developed the 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**: Developed data structure emulating Numpy in C#, allowing 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 package (Python) to process firms' 10-K fillings (217G), applied NLP method for annual report textual analysis;
- Conducted robustness test to eliminate endogeneity: searched alternative variables, tested fixed effect of panel data;
- Proved 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%)
- Feature Engineering: extracted, combined features 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