SHITIAN (STEVE) LIN

646.645.9323 www.linkedin.com/in/shitian-steve-lin shitianl@andrew.cmu.edu

EDUCATION

CARNEGIE MELLON UNIVERSITY, TEPPER SCHOOL OF BUSINESS — New York, NY
Master of Science in Computational Finance – MSCF — ***QPA: 3.84/4.33*** — 12/24

- **Coursework:** Data Science, Machine Learning, Financial Computing, Stochastic Calculus, Fixed Income, Time Series, Simulation Methods for Option Pricing, Advanced Derivative Models, Market Microstructure and Algorithmic Trading
- **Programming Skills:** Python (NumPy, Pandas, PySpark, PyTorch), C++, SQL
- **Honors:** MSCF Scholarship Recipient; 3rd Place in Poker AI Competition

RENSSELAER POLYTECHNIC INSTITUTE — Troy, NY
Bachelor of Science in Applied Mathematics and Computer Science — ***GPA: 3.97/4.0 (Summa Cum Laude)*** — 12/22

QUANTITATIVE TRADING PROJECTS

IMC – 2024 GLOBAL TRADING CHALLENGE (18th of 9,139 Teams Worldwide, 6th in the U.S) — 4/24 – 5/24

- **Alpha Research:** Performed analysis on order book and market trade data to find alpha signals for various tradable products
- **Algorithmic Trading:** Developed Python scripts for algorithmic trading, implementing strategies such as market making, delta hedging for options, pair trading for ETF, considering inventory risk; backtested trading strategies on historical data
- **Commodity Trading:** Modeled commodity prices using weather and temperature; executed cross-counter arbitrage across domestic and foreign exchanges, optimizing profitability by accounting for tariffs, storage fees and transportation fees

QUANTBOT – CMU DATATHON (1st Place Winner) — 4/24

- **Alpha Mining:** Constructed new alpha signals using order book factors, price momentum, macro indicators, statistical metrics, and liquidity factors to enhance the model predictive power; implemented a LightGBM model to predict daily stock returns, enhancing stability by training on cumulative returns across various horizons to reduce noise
- **Model Optimization:** Leveraged model insights to construct a long-short equity portfolio, periodically refitting to mitigate drift; backtested and optimized hyperparameters, yielding an average Sharpe Ratio of 2.52 by aligning with alpha signals

OPTIVER – TRADING AT THE CLOSE (54th out of 4,436 Teams) — 9/23 – 3/24

- **Data Analysis:** Performed analysis to assess correlations, distributions, and autocorrelation on 17 auction book and order book features with visualization; conducted data cleaning to ensure data integrity for modeling using Scikit-learn
- **Feature Engineering:** Constructed 170+ features encompassing order book imbalances, price momentum, rolling features, shifted historical features, statistical aggregations, and market microstructure to enhance the model predictive power
- **Ensemble Modeling:** Implemented ensemble models with LightGBM and XGBoost using stacking and voting to predict closing price movements for 200 stocks; optimized the model via time-series cross-validation and hyperparameter tuning

EXPERIENCE

WELLS FARGO — Charlotte, NC
Quantitative Analytics Intern — 6/24 – 8/24

- **Data Preprocessing:** Preprocessed 10 billion transaction data utilizing PySpark; created model label to identify recurring transactions for classification modeling; employed stratified sampling and oversampling to mitigate class imbalance
- **Classification Modeling:** Implemented Logistic Regression, Decision Tree and XGBoost models to identify recurring transactions for future balance forecast; compared performance with the existing vendor model for potential replacement, selecting the optimal model that achieved an average F1 Score of 0.77 and AUC-PR of 0.86, delivering a 20% improvement
- **Model Optimization:** Optimized the model via cross-validation and hyperparameter tuning; conducted sensitivity analysis using Partial Dependency Plots and Shapley Values for feature selection; assessed robustness across three seasonal samples

JPMORGAN CHASE & CO — Columbus, OH
AI & Data Science Summer Analyst — 6/22 – 8/22

- **Big Data Analysis:** Analyzed approximately 4 billion digital transaction data utilizing SQL and implemented advanced data manipulation and analytics strategies to discover causes of rent hikes and identify income sources of credit-invisible renters
- **Feature Engineering:** Constructed 20+ new rent features using PySpark, resulting in an average Area Under the Curve (AUC) of 0.80; conducted sensitivity analysis to assess the marginal benefit of each feature, enhancing model performance
- **Credit Risk Modeling:** Created model label via ground truth labeling process to accurately identify customer delinquency status; implemented and fine-tuned a Logistic Regression model for credit risk assessment

ADDITIONAL INFORMATION

- **Interests:** Competitive Gaming, Poker, Trading, Table Tennis, Swimming
- **Languages:** English, Mandarin, Cantonese