Volume 2025, Issue 2 - May 2025

Quant Quarterly

NC STATE UNIVERSITY

Financial Mathematics

Contents

Dr. Tao Pang
Ph.D., CFA, FRM
Professor and Director

Four new students joined our program in Spring 2025, enrolling either as full-time or part-time students. They come from diverse backgrounds and hold undergraduate degrees in computer science, economics, and finance. Our applicant pool for Fall 2025 is also exceptionally strong. We have received more applications than last year and anticipate welcoming even more students this fall.

During Spring 2025, students worked on eight group projects covering topics such as credit risk, interest rate modeling, option hedging, machine learning, and cryptocurrencies. These projects were led by FM Ambassadors and the teams presented their work to invited industry guests on April 18. Through this hands-on training, students applied classroom knowledge to real-world challenges. They also had the opportunity to collaborate with peers from diverse backgrounds, further enhancing both their technical expertise and their teamwork and communication skills.

Summer internship placements for 2025 are very strong. As of April 25, over 90% of students seeking internships have received offers. For those not pursuing internships, we will provide an industry-sponsored project during the summer to ensure continued professional development.

In Spring 2025, we launched a new website for the program:
https://financialmath.sciences.ncsu.edu/
The new site provides expanded information about our curriculum, faculty, student opportunities, and alumni success stories.

We also created an official Instagram account for the program:
https://www.instagram.com/ncstatefinmath/
Going forward, we will regularly update our website, LinkedIn page, and Instagram account with news, events, and highlights from both current students and alumni.

To support incoming students, we will offer the FM Preparation Workshop from July 28 to August 15. Topics will include mathematics, statistics, econometrics, risk management, Python, and machine learning. All sessions will be led by FM faculty members and are designed to help students prepare for a successful first semester. More information can be found at https://financialmath.sciences.ncsu.edu/summer-workshop/.

Wishing everyone a wonderful and productive summer!

Sincerely yours,

Tao Pang, PhD, CFA, FRM
Professor and Director
Financial Mathematics Graduate Program.

Patrick Roberts
Director of Career Services

Maximizing LinkedIn: Engaging with Industry Experts

LinkedIn is considered the top social media outlet for professional networking and is a powerful platform for students seeking career advice, job opportunities, or developing valuable industry connections. The major challenge when navigating this resource however is getting responses from industry professionals, many of which have busy schedules both professionally and personally therefore limiting their responsiveness and engagement on the site. This article will discuss effective strategies to increase your chances of receiving replies and further enhancing the usefulness of digital networking through LinkedIn.

First, you should always keep your profile up to date with the most current relevant and recent information that would appeal to potential professional connections. Your headline and about section should highlight your skills, interests, academic focus and career goals. Also make sure to include relevant experiences, projects, coursework, and certifications to increase the likelihood that professionals will be interested in engaging with you and perceive you as person that could bring value to their organization or industry.

Second, when sending connection requests, since the standard version of LinkedIn limits your customized messaging to 5 per month, this should be utilized on a limited basis. However, after sending a connection request, once the contact has agreed to connect, you should follow up immediately with a personalized message informing them about who you are and why you would like to connect. This message should be clear, respectful, and action oriented. Be sure to include relevant information such as your current degree, career goals, and why this person is of interest. Here is an example message:

> "Hi [Name], thank you for accepting my connection! I am currently studying Financial Mathematics at NC State and came across your work at [company]. I would be thrilled to schedule a brief call to discuss your experience and any advice you have for students entering this field. Would you be open to a brief chat? Thanks in advance!"

Following this message, if you do not receive a response, wait 1-2 weeks, set a reminder to follow up but make sure your message is courteous and concise. After this second follow up, you should move on to the next connection and stop attempting to message until they respond since this can be viewed as harassment and create a bad impression. Here is an example follow up message:

> "Hi [Name], I appreciate your time, and I know that you are busy. I wanted to follow up on my previous message. I would really appreciate any insights you can share when you have time. Thanks again!"

When someone responds, always thank them for their time and show appreciation for the advice they give. Be flexible and try to adhere to their schedule if possible. When given suggestions on skill development, keep the contact updated as you find strategies to apply their advice. Create reminders every 2-3 months to provide brief updates and follow up checkpoints to maintain the relationship, this will keep the door open to future opportunities. These checkpoints can be as simple as letting the contact know that you have recently completed an interesting project or course to updates on your career goals or securing employment.

In the end, building professional connections on LinkedIn can be seen as a numbers game with strategic planning and execution involved. Targeting specific industries, professionals and relevant networks like alumni, locations, organizations, or positions can help narrow the vast landscape of professionals and increase the chances of response. By following these strategies, students can successfully utilize LinkedIn to build valuable industry connections.

Students' Corner

George Armentrout
Motivations and Implications of Cryptocurrency Research

Divija Balasankula
Trade Policy Turbulence: The VIX Rollercoaster of 2025

Daniel Chertock
Bridging Financial Mathematics and Algorithmic Trading

Thaddeus Creech
The Power of Attention

Haozhe Cui
PCA Statistical Arbitrage Mean Reversion

Jacob Dolan
Modeling Mortgage Default Probability

Sophia Feng
Prediction of BTC Price with Onchain Data Using Bayesian Neural Network

Isaac Gohn
Pursuing a Career in Quantitative Finance

Zhen He
Navigating Risk Management with NC State's MFM Program

Chance Humiston
Experimenting With Support Vector Machine Trading Models

Vidyul Jain
Navigating The Journey In Financial Mathematics

Joseph Jonasson
When Optimal Transport meets Quantitative Finance

Sarang Kansara
Trading On Market Mood

Dev Kewlani
Deep RL for Portfolio Optimization

William Lanzoni
A Decision Without Regret

Tharun Mandadi
Markowitz Model With Large Simulation And Sensitivity Analysis

Chelsea Niles
Enhanced Modern Portfolio Optimization

Jinjia Peng
Monte Carlo Stock Price Simulation with Jump Diffusion GBM

Adharsh Rajagopal
Beyond The Numbers: Charting My Path To Quantitative Research

Sanith Rao
Exploring Sentiment Analysis For Stock Prediction

Vismit Rekhan
Bridging Traditional Finance with the Future

Partha Sanampudi
Levelling Up In Modelling

Eric Wang
Program Observations And Career Goals

Jingxing Wang
Mortgage Loan Delinquency Prediction

Yu Wang
A Rewarding Journey, and More to Come

Jinyi Yang
Navigating Financial Math: From Theory to Practice

Nick Zehnle
Regime-Adaptive PCA Mean Reversion Strategy -- Understanding the Theory

George Armentrout

Motivations and Implications of Cryptocurrency Research

Humble Beginnings & Bad Reputations

Cryptocurrencies, what once were shunned in the financial sector due to their lack of regulation, high volatility, and illicit uses, have made quite the comeback.

In the aftermath of the 2008 global financial crisis, many people sought a way to store and exchange wealth in ways that were not reliant upon the institutions that much of finance and economics are built upon (i.e. banks, government regulators, credit agencies). This led to the birth of cryptocurrency, a frankly impressive realization of software development, encryption, and financial theory.

This was a system that enabled for specified ownership of a fully digital asset and allowed for safe transfer of said assets from one party to another. This system was devoid of any form of real regulation, and all value in these digital assets was based solely upon the belief of the asset holding value. Naturally, no regulatory oversight allowed for cryptocurrency to be used as a payment for illicit activities, and value on belief implies incredibly high volatility. These factors, combined with the incredibly small market cap of the cryptocurrency market in its infancy (< $1billion prior to 2013), led many financial institutions to completely disregard crypto as a viable investment tool.

Shift in the Industry

While many other tech fads or financial schemes faded, cryptocurrency did not go anywhere. On the contrary, over time crypto's market cap increased dramatically, demonstrating a sustained and growing demand for a digital asset. Several countries' financial regulators started actually investigating potential uses and impacts of cryptocurrency on their country's economy. Some investors, both commercial and retail, started considering the potential of investing in the highly volatile asset. Many of these commercial investors were startups intentionally targeting cryptocurrencies, but it marks a distinct shift in the financial industry's opinion of crypto.

With this increased interest from the financial sector, much more scrutiny and regulation has been implemented towards cryptocurrencies, making the digital asset more appealing to many financial institutions and retail investors. Now, there is enough of a backlog of historical data that analysis can effectively be conducted on the crypto market, and as a result these currencies are viewed more and more as a viable and useful financial tool.

Research & Implications

Currently, the financial industry in its entirety is working to embrace cryptocurrency, and with that many institutions and companies are conducting analysis on the prospects of this market and its potential uses in the context of their specific business practices.

For example, brokerage firms are now looking to offer different cryptos. As such, research and development must be done to properly implement a cryptocurrency framework within the context of their other services. Additionally, work must be done to discern which currencies should be listed in the first place, and models need to be designed to optimize and automate this kind of process.

Another example would be more traditional proprietary trading firms and hedge funds. Now, these institutions have access to an entirely new class of financial tools they can utilize in hedging and portfolio management. Analysis needs to be done to determine the usefulness of cryptos in this context as well as how these new formats of financial management integrate into their current services and systems.

Through all of this, it is certain that there is a growing enthusiasm for embracing the crypto market within the financial sector; continued research and development will only foster more innovation within finance.

.

Divija Balasankula

Trade Policy Turbulence: The VIX Rollercoaster of 2025

The first quarter of 2025 has been a stark reminder of how trade policy can send shockwaves through financial markets, with the VIX index—Wall Street's "fear gauge"—serving as a barometer for market anxiety. The VIX, or Volatility Index, measures expected volatility in the S&P 500 over the next 30 days, providing investors with a quantitative measure of market fear and uncertainty. As President Trump's aggressive tariff strategy unfolded, the VIX reached levels not seen since the pandemic-induced market crash of 2020.

The VIX Surge

From a relatively calm level of 14 in mid-February, the VIX surged to 24 by early March—a 71% increase in just weeks. This spike reflected growing investor concern about the economic implications of the tariffs and the potential for a full-blown trade war.

Timeline of Tariff Turmoil and VIX Reactions

- February 1: Trump signs executive orders to impose tariffs on Canada, Mexico, and China. VIX begins to climb.
- February 4: 10% tariff on Chinese imports takes effect. VIX continues its upward trajectory.
- March 4: 25% tariff implemented on most imports from Canada and Mexico; Chinese tariffs increased to 20%. VIX spikes, reaching its highest level of the year at 26.35 intraday before settling at 23.51 at market close.Later that day, Commerce Secretary hints at possible compromise.
- March 5: One-month exemption announced for goods qualifying under the USMCA. VIX falls 6.72% to 21.93.
- March 7: Trump's inconsistent messaging reignites market jitters. VIX surges 13.4% to 24.87.

VIX Fluctuations in Detail

The VIX's behavior during this period was particularly noteworthy:

- Pre-tariff announcement: VIX at a relatively calm level of 14.
- Post-initial tariff announcement: VIX jumped to 15.82, a 13% increase.
- March 4: VIX spikes to 26.35 intraday (highest level of the year) following the implementation of new tariffs, before settling at 23.51 at market close.
- March 5–7: VIX experienced significant daily swings, reflecting the market's sensitivity to each policy shift and statement:
  - March 5: VIX falls 6.72% to 21.93
  - March 7: VIX surges 13.4% to 24.87

This sustained elevation above 20 post-March 4 suggests a fundamental shift in market risk perception, with investors pricing in long-term uncertainty.

Investment Strategies Centered on VIX

As the VIX rollercoaster unfolded, savvy investors employed various strategies to capitalize on or hedge against the heightened volatility.

- **Long Volatility Strategy**: Going long on VIX futures or call options when volatility was low in mid-February would have yielded significant profits. Traders could have employed more advanced strategies like VIX calendar spreads, selling near-term options and buying longer-term options with the same strike price to benefit from time decay while positioning for increased longer-term volatility.
- **Short-Term Trading on VIX-related ETFs**: Products like ProShares VIX Short-Term Futures ETF (VIXY) and iPath Series B S&P 500 VIX Short-Term Futures ETN (VXX) saw significant gains, soaring to 2025 highs. The ProShares Ultra VIX Short-Term Futures ETF (UVXY), a leveraged product offering 1.5x exposure to VIX futures, also experienced substantial increases during this volatile period.
- **Hedging with VIX-based instruments**: Investors bought put options on major indices or vulnerable sectors to protect their portfolios from downside risk. Additionally, they used inverse VIX-based ETFs like the ProShares Short VIX Short-Term Futures ETF (SVXY) to hedge against sudden spikes in volatility. When the VIX surged, these inverse ETFs declined, offsetting potential losses in long stock positions.

Continue Next Page

Divija Balasankula

Trade Policy Turbulence: The VIX Rollercoaster of 2025

Continued From Previous Page

For investors anticipating further increases in VIX, additional strategies could be considered, albeit with higher costs due to elevated option premiums:

- **Long Straddle**: When VIX is high, options become more expensive, but a long straddle can still be profitable if volatility increases further or if there's a significant move in the underlying asset. Traders could buy both a call and a put at the same strike price, profiting from large moves in either direction.
- **Long Strangle**: Similar to a long straddle, but with different strike prices for the call and put. This strategy can be less expensive than a straddle and still profit from large moves or increased volatility. This method is particularly effective when traders expect a big move but are unsure of the direction.

The events of early 2025 serve as a stark reminder of how closely the VIX tracks trade policy developments in our interconnected global economy. As UBS strategists noted, elevated risk premia are expected to persist, making a swift drop in volatility unlikely in the near term. Looking ahead, investors should remain vigilant, as the ongoing trade negotiations and potential policy shifts could continue to drive VIX fluctuations. In this environment, a thorough understanding of volatility dynamics and associated trading strategies will be crucial for navigating the markets effectively.

In the world of international trade, the only certainty seems to be uncertainty itself—a reality clearly reflected in the VIX's recent rollercoaster ride.

UBS. Trade war fears spark volatility https://www.ubs.com/global/en/wealthmanagement/insights/chief-investment-office/house-view/daily/2025/latest-05032025.html

Daniel Chertock

Bridging Financial Mathematics and Algorithmic Trading

Financial markets exhibit complex behaviors that challenge traders and quantitative analysts to develop effective strategies for identifying opportunities. In our FIM 601 project, our team has been researching mean reversion techniques using Bollinger Bands and the Relative Strength Index (RSI) to trade stocks. The objective of our project is to determine whether these indicators can consistently identify mean reversion opportunities, leading to profitable trading strategies.

The underlying premise of mean reversion is that asset prices tend to oscillate around a fair value, often reverting to their historical average after deviating. Bollinger Bands help identify overbought and oversold conditions by plotting standard deviations around a moving average, while RSI quantifies momentum and potential reversals. By combining these two indicators, we aim to enhance signal reliability and reduce false positives in trading decisions.

To conduct this research, our team has been leveraging Python and QuantConnect's algorithmic trading framework, which allows for back-testing strategies using historical stock market data. QuantConnect provides access to a wide range of market data, including equities and alternative datasets, enabling a more comprehensive evaluation of the mean reversion approach. Our project involves extensive quantitative analysis, including testing different parameter combinations for Bollinger Bands and RSI to determine their effectiveness across various market conditions.

One of the key challenges I encountered was overfitting—where a strategy performs well on historical data but struggles in live trading. To address this, I implemented various validation techniques to improve robustness and generalizability. I also had to adjust my approach to account for different market conditions, as I found that mean reversion strategies are significantly less effective in strongly trending markets. Identifying the right market environments for this approach was a challenge, requiring careful analysis of volatility regimes and historical price behavior.

Through this project, I have gained valuable insights into the strengths and limitations of mean reversion strategies. I have also improved my ability to structure quantitative research, analyze data, and refine trading algorithms based on statistical evidence. While the strategy has shown promise in range-bound markets, it requires further optimization to ensure consistent performance across different market cycles.

My decision to pursue the Master's in Financial Mathematics program at NC State stems from my long-standing passion for financial markets, algorithmic trading, and quantitative research. I have been actively involved in financial markets as both a trader and investor, specializing in technical analysis and algorithm-driven trading strategies. While I have developed strong intuition and experience in trading, I recognized that a deeper mathematical and statistical foundation was necessary to advance my understanding of quantitative finance. The program provides the rigorous training needed to refine my analytical skills and enhance my ability to design, test, and optimize trading strategies.

Through my coursework, I have gained exposure to advanced topics such as stochastic calculus, time-series analysis, and machine learning applications in finance. These concepts are critical for refining trading models and understanding market dynamics beyond traditional technical indicators. My experience in financial markets—particularly in cryptocurrency, futures, and equities trading—has provided me with a strong intuition for market movements. However, translating intuition into quantitative models requires a deeper mathematical framework, which the MFM program is helping me develop.

One of the key skills I seek to improve is my ability to construct and analyze mathematical models that can be applied to algorithmic trading. While I have a solid understanding of technical analysis and trading strategies, formal training in probability theory, optimization, and statistical inference will enable me to make data-driven decisions with greater confidence. Additionally, learning programming techniques for quantitative finance, such as Monte Carlo simulations and machine learning-based pattern recognition, will be invaluable for both my trading and professional pursuits.

To further enhance my skills, I plan to engage in research projects that bridge financial mathematics with real-world trading applications. The MFM program is equipping me with the tools necessary to elevate my trading strategies, optimize risk management, and expand my capabilities as a quantitative trader. As I progress in my studies, I look forward to applying these skills to both my personal trading endeavors and the development of sophisticated algorithms that can adapt to changing market conditions.

Thaddeus Creech

The Power of Attention

Ever since Vaswani et al. (2017) introduced the transformer model in their seminal paper Attention Is All You Need transformer architecture has been taking over the world. With the release of Large Language Models (LLMs), almost everyone has seen the power of this architecture in analyzing vast amounts of text data and building some of the world's first truly human-like chat bots. However, the power of this architecture does not end there.

In our semester project, our group has aimed to take this architecture, known for its remarkable abilities in capturing long-term dependencies, and apply it to financial data. In an attempt to mimic the results of a more recent paper AlphaPortfolio (Cong et al., 2021). While this endeavor has been a challenge, this experience has brought the power and scope of this development to the forefront of our minds. Primarily, this method's ability to develop models that are able to optimize almost any problem with the same underlying architecture.

While our project is focusing on digesting monthly stock data and aiming to optimize the returns of a portfolio through its Sharpe ratio, in an effort to overcome some of the common bottlenecks in traditional portfolio optimization, it has not escaped the group that this architecture has application in many other forecasting problems. As long as enough quality data exists and there is some metric to be optimized, simply altering the training data and the loss function of the model allows it to be applied to almost any forecasting application in the financial industry. Along with many other issues that are faced by financial professionals in modern day such as anomaly detection and risk management.

In short, transformer architecture has the power to completely reshape how we analyze financial data by offering a more nuanced and comprehensive understanding of the dynamics of the markets. With their many applications, these models are soon to become an indispensable tool in the practitioners' toolbox in the more data-driven financial landscape that we face today.

Cong, L., Tang, K., Wang, J., & Zhang, Y. (2021). AlphaPortfolio: Direct construction through deep reinforcement learning and interpretable AI [Working Paper]. SSRN Electronic Journal. http://dx.doi.org/10.2139/ssrn.3554486

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. Advances in Neural Information Processing Systems, 30, 5998–6008.

Haozhe Cui

PCA Statistical Arbitrage Mean Reversion

This project focuses on developing a statistical arbitrage strategy that leverages principal component analysis (PCA) and linear regression to capture mean-reversion opportunities in US equities. The goal is to exploit temporary mispricings by identifying deviations between the predicted and actual prices of securities, based on the notion that prices will eventually revert to their long-term averages. In this approach, PCA is applied to standardized log returns of the asset prices to extract the most significant factors driving market movements, typically using the first three principal components. These components capture over 90% of the variance in the data, making them a robust basis for further analysis.

The primary technology used in this project is Python, along with its rich ecosystem of libraries for data analysis and machine learning. Key libraries include scikit-learn for performing PCA, statsmodels for executing ordinary least squares (OLS) regression, and pandas for efficient data manipulation. Additionally, visualization tools such as matplotlib are used to generate performance plots and equity curves, which are crucial for understanding the behavior of the strategy over time. The project also integrates with a quantitative trading platform that provides real-time and historical market data, ensuring that the analysis is based on accurate, adjusted price information (taking into account corporate actions like splits and dividends).

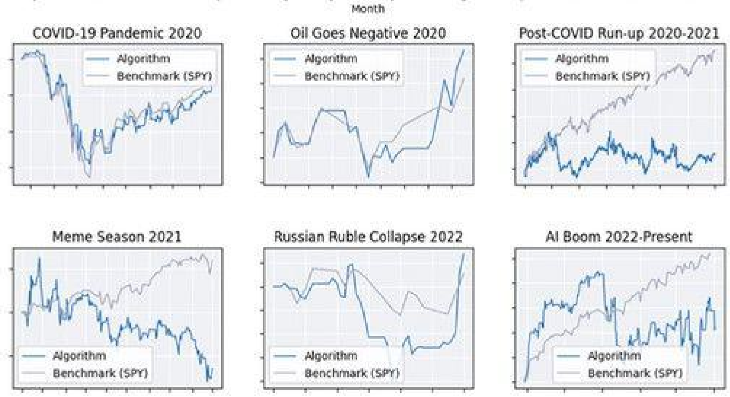
Data sources for this project are centered on US equities, specifically targeting the 100 most liquid stocks with prices above $5. This selection criterion ensures that the strategy is implemented on assets with sufficient trading volume and price stability, which are critical factors for both model accuracy and the feasibility of executing trades in real-market conditions. Historical daily price data is used, and the analysis considers various lookback periods—from 21 days (approximately one month) to 126 days (roughly six months)—to determine the optimal window for model training and performance assessment.

Quantitatively, the project employs a multi-step analytical process. Initially, PCA is conducted on the standardized log returns of the selected stocks to identify the key market factors. The first few principal components, which collectively explain the majority of the variance, are then used as independent variables in an OLS regression model for each stock. This regression predicts the expected price based on these underlying factors. The difference between the predicted and actual prices (the residuals) is then standardized using z-scores. Stocks with z-scores significantly below a predetermined threshold (e.g., −1.5) are identified as potential candidates for mean reversion; that is, they are expected to appreciate as their prices revert to the mean.

Several challenges have been encountered during the project. One significant hurdle is selecting the optimal number of principal components and determining the appropriate lookback period. Although the first three components generally capture most of the variance, adding too many may introduce noise, while too few may omit critical information. Moreover, ensuring the quality and consistency of historical data—particularly after adjustments for splits and dividends—requires careful data preprocessing. The model’s sensitivity to outliers and the dynamic nature of market behavior further complicate the process, necessitating rigorous backtesting and parameter tuning. Overall, while the strategy is conceptually straightforward, its successful implementation demands meticulous attention to detail in data handling, model selection, and continuous performance evaluation.

Monthly Returns (%)

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2019	2.80	7.44	1.36	3.05	-10.87	2.66	5.00	10.74	1.70	4.64	0.39	-2.28
2020	2.61	-7.49	-10.93	8.45	2.37	6.72	-17.51	4.63	-5.69	4.25	10.67	1.19
2021	-1.96	-0.67	-2.63	-1.26	-0.53	5.91	-4.67	-5.06	-2.95	6.86	0.92	3.75
2022	-11.87	1.59	3.99	-13.70	1.43	-1.99	13.26	-2.29	-8.08	9.49	-0.21	-3.43
2023	14.07	-2.69	3.93	-3.46	1.64	2.65	15.44	-3.99	-3.23	-2.37	13.42	5.95
2024	-11.80	3.36	3.04	-2.97								



Jacob Dolan
Modeling Mortgage Default Probability

In the Fall 2024 NCSU Financial Mathematics Career Development course, I completed a group project called Loan Application Scorecard Models: Logistic Regression and Random Forest. In our project, we developed predictive models to assess mortgage default risk using the Freddie Mac Single-Family Home Loan data. The general objective was to assist lenders in screening risky loans and improving loan approval decisions, thereby minimizing financial losses. With logistic regression and random forest techniques, we sought to model significant borrower characteristics and loan attributes to forecast the likelihood of mortgage default and compare the performance of both models.

To achieve our goals, we employed a number of technologies, with the primary focus being Python programming. Data manipulation was handled through libraries such as Pandas and NumPy, while model implementation and performance measures utilized Scikit-learn. For data visualization, we leveraged Matplotlib and Seaborn.

Our project centered on the Freddie Mac Single-Family Home Loan data set, which comprised historical mortgage data including borrower credit scores, loan-to-value ratios, debt-to-income ratios, and loan performance. This data set was the foundation for both training and testing of our models, with real-world applicability in predictive analytics for finance. A core challenge in working with this dataset was its enormous class imbalance—96% of the observations were non-default cases, and thus it was tricky to build models that could successfully predict default events.

To address these challenges, we implemented various quantitative and analytical techniques. We began with data preprocessing, ensuring data quality by handling missing values and identifying outliers. Since certain variables exhibited high multicollinearity, careful feature selection was necessary. Synthetic Minority Oversampling Technique (SMOTE) was applied to balance the data set by generating artificial minority class samples to allow the model to learn patterns for loan default more effectively. In the logistic regression model, we also applied the Weight of Evidence (WOE) transformation to enable interpretability as well as increased performance. Model training involved cross-validation and hyperparameter tuning to optimize both the random forest and logistic regression models.

One of the biggest obstacles was selecting the most relevant features for our models. With as many variables as we had available, deciding which ones contributed the most to the accuracy of our predictions was a combination of domain knowledge, statistical techniques, and machine learning algorithms. To overcome this barrier, we implemented several feature selection methods, including forward selection and feature importance scoring from our models. These methods enabled us to refine our dataset and improve model efficiency.

One challenge was ensuring our models accurately distinguished between default and non-default loans while minimizing false positives. The random forest model provided us with strong predictive strength since it was capable of capturing complex relationships in the data. Logistic regression, while less predictive, was more interpretable but performed slightly worse. Through hyperparameter tuning and model performance validation on independent test sets, we could achieve a trade-off between recall, precision, and overall model performance.
Overall, the project was a success and allowed us to develop practical skills. This experience helped us enrich our understanding of machine learning pipelines, data preprocessing, feature engineering, and model evaluation. We gained hands-on experience in handling imbalanced datasets using methods such as SMOTE and WOE, which are highly relevant in predictive modeling in finance applications. Additionally, this project strengthened our ability to apply theoretical concepts to real-world problems, reinforcing our expertise in quantitative finance and risk management.

Our results were positive, indicating the effectiveness of our approach. On all the key measures, the random forest model performed better than logistic regression with a recall of 0.84, precision of 0.83, F1 of 0.83, and ROC-AUC of 0.92. This suggests that the model is highly effective in identifying high-risk loans while minimizing false positives. The logistic regression model, while slightly underperforming, still provided valuable observations and a helpful baseline. Overall, our project validated the importance of thorough data preparation, feature selection, and model tuning in predictive analytics. This project has significantly enhanced our analytical and technical skills, reaffirming our potential contribution to decision making in data-driven financial services.

Sophia Feng

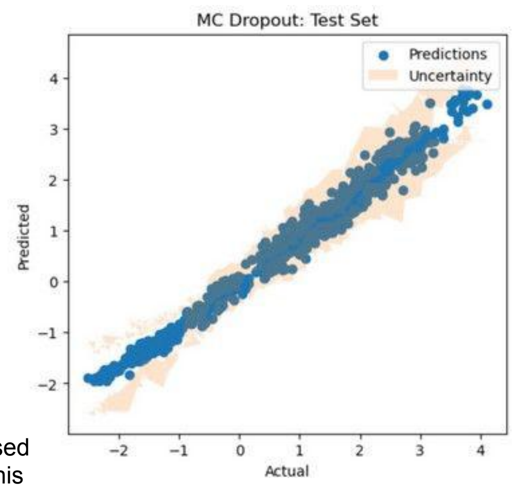
Prediction of BTC Price with Onchain Data Using Bayesian Neural Network

The goal of this project is to predict Bitcoin (BTC) prices using on-chain data and macroeconomic indicators through a Bayesian Neural Network (BNN). By leveraging blockchain transaction data, wallet activity, network health, and supply metrics, we aim to develop a robust predictive model that can provide reliable price forecasts. The Bayesian approach is chosen to quantify uncertainty in predictions, offering a probabilistic perspective on BTC price movements.

This project integrates various technologies, including Python for data processing and model development, TensorFlow/PyTorch for building and training Bayesian Neural Networks, and APIs from Blockchain.com and Investing.com for retrieving relevant financial data. Additional tools such as Pandas, NumPy, and Scikit-learn are utilized for data preprocessing and analysis, while Matplotlib and Seaborn aid in visualization.

The data is sourced from Blockchain.com and Investing.com, covering the period from January 1, 2009, to January 1, 2025. Blockchain-specific data includes transaction volume, wallet activity, and network statistics, while macroeconomic variables encompass stock indices, gold prices, volatility indices, and exchange rates. The historical BTC price data is used as the response variable, sourced from financial and blockchain data repositories.

In terms of data collection, cleaning, and preprocessing, we ensure the dataset is comprehensive and consistent. Missing values are handled through imputation techniques, and normalization is applied to make the features comparable. Feature selection is a critical step in the process. Initially, 26 features were considered, but after performing a multicollinearity check using the Variance Inflation Factor (VIF), the feature set was reduced to 17 variables by removing highly correlated predictors. This step helps prevent overfitting and improves the model's efficiency.



The predictive model is based on a Bayesian Neural Network (BNN), which differs from traditional deep learning models by incorporating probability distributions over weights. The choice of prior distribution is crucial, with Gaussian, Laplace, and Horseshoe priors considered. The posterior distribution is approximated using three methods: Monte Carlo (MC) Dropout, Markov Chain Monte Carlo (MCMC), and Variational Inference (VI). After extensive evaluation, MC Dropout was found to deliver the best performance, achieving an R^2 score of 0.9746, indicating strong predictive capability. Additionally, Bayesian inference provides credible intervals, which allow for uncertainty quantification in BTC price predictions.

Several challenges were encountered during the project. Data quality and availability posed a significant hurdle, as blockchain data often contains irregularities and missing values. This was addressed by applying imputation techniques and ensuring data consistency through cross-validation. Feature multicollinearity was another issue, as high correlation among blockchain metrics led to redundant information. To overcome this, a VIF analysis was conducted, and highly correlated predictors were removed. Furthermore, Bayesian methods, particularly MCMC, require significant computational power, which posed a challenge. To mitigate this, MC Dropout was chosen as an efficient approximation method, striking a balance between accuracy and computational efficiency.

Through this project, several key learning outcomes were achieved. We gained proficiency in on-chain data analysis and its impact on BTC price movements, developed expertise in Bayesian deep learning techniques, and understood the importance of uncertainty quantification in financial modeling. Additionally, the project enhanced our skills in data preprocessing and feature engineering, which are critical for improving model efficiency and accuracy.

The results demonstrate the effectiveness of a Bayesian Neural Network for BTC price prediction. The model successfully integrates blockchain metrics and macroeconomic indicators to deliver a high R^2 score of 0.9746. The superiority of MC Dropout over other approximation methods was evident, and the ability to provide probabilistic forecasts with credible intervals is a significant advantage. This project highlights the potential of Bayesian deep learning techniques in financial time-series prediction, particularly for volatile assets like Bitcoin.

In conclusion, this research underscores the effectiveness of Bayesian Neural Networks in financial modeling. The combination of blockchain and macroeconomic data significantly enhances predictive accuracy, while the Bayesian approach allows for better uncertainty estimation. Future work will explore alternative deep learning architectures, such as Bayesian LSTMs, to capture temporal dependencies in BTC price movements more effectively.

Isaac Gohn

Pursuing a Career in Quantitative Finance

My goal is to build a career in quantitative finance within commercial banking and I chose the Master of Financial Mathematics degree at NC State because I knew this program would give me the best shot at achieving this. Some of the aspects that stood out to me was the fact the program offers a strong combination of advanced mathematics, programming, and real-world finance applications, making it an ideal fit for someone like me looking to apply quantitative skills in a banking environment. Coming from a background in electrical engineering and physics, I wanted a program that was not just theoretical but also provided the practical skills needed to solve real problems in finance. The coursework in stochastic modeling, programming, and risk analysis have been instrumental in preparing me for the industry, while NC State's strong industry connections and emphasis on applied learning made this a natural choice for my graduate studies.

Through the program, I have come to understand that quantitative finance in commercial banking is more than just advanced math and coding. It requires a deep understanding of financial markets, risk management, and data-driven decision-making while also being able to analyze complex patterns in large datasets efficiently. The field is constantly evolving, especially as predictive modeling and risk assessment play a growing role in financial institutions, so staying up to date with new techniques is crucial. Beyond technical skills, I have realized that interpreting results, collaborating with teams, and communicating insights effectively are just as important as building models. These soft skills are essential for ensuring that data-driven strategies align with a bank's broader financial objectives.

In addition to the technical coursework, one of the most valuable aspects of the MFM program has been its career development support, particularly in interviewing, LinkedIn networking, and job searching. I have always felt unsure about the interview process, but the career development class has helped me understand how to prepare effectively. This has boosted my confidence and performance in interviews, especially as I navigate the hiring process for quantitative finance roles.

Throughout the semester, I have focused on strengthening skills and expertise to make me a better candidate in quantitative analysis. There are still several areas where I need to improve. Financial theory is one of them—while I understand core concepts, I need to deepen my knowledge in areas like portfolio management and market microstructure, which are highly relevant to commercial banking fields. Additionally, while I can code, I want to focus on writing more efficient and scalable algorithms for handling large datasets, a skill that is critical for building real-time financial applications. Another key area I want to develop is communicating complex ideas to non-technical stakeholders—learning that it is one thing to build a model, but being able to explain the implications clearly to decision-makers is just as important.

To address these gaps, I plan to fully leverage the MFM program's coursework, and work on projects that apply what I am learning to real banking scenarios. I also aim to network with industry professionals, learn from their experiences, and seek out internship opportunities that provide hands-on exposure to the commercial banking sector. By combining formal education with practical experience, I am confident that I will develop the skills needed to excel as a quantitative analyst.

Zhen He

Navigating Risk Management with NC State's MFM Program

My decision to enroll in the Master of Financial Mathematics program at North Carolina State University was driven by my clear and focused career ambition: specializing in risk management. The program emerged as my top choice because of its exceptional career services and highly customizable curriculum, allowing me to strategically select elective courses from various colleges that align directly with my professional goals.

My current career goal revolves around risk management, a field that continually captivates me due to its dynamic and multifaceted nature. Through my coursework and projects, I have enhanced my understanding and practical skills in financial risks.
For instance, my recent project on Market Risk Analysis involved modeling techniques such as volatility forecasting using GARCH models. Through the historical method, I effectively estimated critical risk metrics, including Value at Risk (VaR) and Expected Shortfall (ES). These experiences have improved my quantitative analytical skills, providing me with essential insights into real-world risk assessment and management.

Furthermore, my work in Loss Severity Modeling of Single-Family Residential Mortgage Loans, leveraging Fannie Mae datasets, enhanced my capabilities. By employing models including Linear Regression and XGBoost, I successfully integrated macroeconomic variables to enhance model robustness and accuracy. The meticulous management of data integrity and outlier detection, coupled with strategic feature engineering, solidified my analytical skills in credit risk modeling.

While these experiences have equipped me with solid technical proficiency in Python, SQL, R, Tableau, and Bloomberg Terminal, I recognize that technical expertise alone is insufficient. Risk management professionals must possess outstanding communication skills to translate complex analyses into clear and strategic business insights.

Acknowledging this, I am committed to improving my communication and interpersonal skills through the resources provided by the MFM program's career services. Actively participating in professional communication workshops, industry networking events, and targeted electives from various academic departments will enable me to further develop my skills in advanced programming and data management.

Ultimately, my experience in the MFM program at NC State has reinforced my passion for risk management, providing both theoretical knowledge and practical tools. With ongoing dedication to continuous learning and skill development, I am confident in achieving my career goal of becoming a proficient risk management professional capable of delivering strategic value within the financial industry.

Chance Humiston

Experimenting With Support Vector Machine Trading Models

Introduction

Financial mathematics is a highly interdisciplinary pursuit and my coursework in the program thus far has covered math, statistics, financial theory, and computer science. One of the more prominent applications of this material is quantitative trading, which relies on algorithmic methods to buy and sell assets. This article will detail a project that I have been working on to develop a quantitative trading strategy for leading tech stocks using a Support Vector Machine (SVM). The project utilized daily open and closing prices for eight popular tech stocks from 2021 through 2024. Everything was coded in Python using NumPy and Pandas for data manipulation, and Scikit-Learn for the machine learning components. The basic framework for the project was taken from a third-party source that implemented a similar strategy for predicting moves in the EUR/USD exchange rate. The output of the model built here was a daily signal predicting whether the stock would move up or down in the next trading day and the objective was to maximize returns following such a strategy.

Feature Engineering

The two explanatory variables fed into the model were Trend and Relative Strength Index (RSI). This choice was inspired by the EUR/USD model this project was based on. Trend was defined as the percentage difference between the closing price and the moving average. The RSI is a technical indicator that measures the speed and magnitude of recent price changes. Both of these metrics are defined based on a variable lookback period which was determined as part of the optimization process. The purpose of predicting next-day moves constitutes short-term trading. As such, the periods considered were also relatively short. During optimization, lookback periods of 7, 14, 21, and 28 days were evaluated for each metric.

Modeling

The Support Vector Machine is a machine learning algorithm that seeks to optimally separate classes of data using boundaries, or support vectors. The input data is numerical as discussed above. The output is a categorical response variable, in this case whether the stock would move up or down during the next trading day. All SVMs tested for this project used a Radial Basis Function (RBF) kernel. The two optimized parameters were the c and gamma values. In SVM, c is the regularization parameter, with a higher c value encouraging a more complex model. The gamma parameter controls the influence of individual data points on the decision boundary. Higher gamma values also lead to more complex, less generalizable models. Five values each for c and gamma were tested as a part of the optimization phase.

Optimization

The premise for the optimization phase was to conduct a grid search and test every combination of the parameters discussed above for each stock being considered. The models were trained using a time-series cross validation procedure on data from 2022 and 2023. The objective function considered was a profitability-oriented modification of Root Mean Squared Error (RMSE). The function could be called the RMS-Return as it substituted the percent return of holding or shorting the stock each day for error in the traditional RMSE formula. This method was meant to reward predictions that were directionally accurate and also consider the magnitude of the movements the stock prices made. Final tests for the optimal models were conducted on 2024 trading data.

Results

Overall, the outcomes for the final models were quite poor. In only one case did trading outperform investing. Most of the trading models yielded decent returns, but a simple buy-and-hold strategy would have been better. In most cases, the trading strategy came in about 10% lower than the 2024 return for the stock. One reason for this was the stock selection. These names benefited greatly from AI tailwinds and saw substantial gains during 2024, so the overall return benchmark was already a high bar to meet. Another issue with this strategy is the binary output that yielded either a buy or short signal each day. Perhaps a more dynamic output could have been put to better use. But probably the biggest reason why this strategy underperformed was the choice of explanatory variables. At least for these stocks, there are no discernable patterns that could have been used to separate the data into “next-day-up” and “next-day-down” categories. The scatter plots showed no groupings, and each of these variables were essentially normally distributed with respect to next-day returns.

Conclusion

Although the trading strategy is underperforming so far, this process was still a positive experience. For one, there is still a lot of room to experiment with different stocks, data, and outputs. I think the immediate next steps would be to bring in additional technical indicators as inputs and modify the output to be more dynamic. Perhaps the response classes could be changed to include a third “sit-out” category, where the strategy doesn’t hold or short the stock and instead leaves that capital available for other purposes. The other reason this was a worthwhile exercise is that I learned a lot while refining the process trying to improve profitability. Specifically, the objective function was an interesting concept that I believe could be reused elsewhere even if this model is otherwise discarded. And that is not something I would have thought about if the model had immediately outperformed its benchmarks. Just because this experiment did not yield a profitable trading strategy does not mean that the results were not valuable.

Vidyul Jain

Navigating The Journey In Financial Mathematics: A Reflection On Projects And Career Aspirations

Interest Rate Forecasting Project (FIM 500)

Project Goals: Last semester, I undertook an interest rate forecasting project, aiming to predict future interest rate movements using historical data and various forecasting models. The primary goal was to develop a robust model that could provide accurate short-term interest rate predictions, which is crucial for investment strategies and risk management.

Technology and Data Sources: This project utilized Python for data analysis and modeling. I sourced data from Bloomberg Terminal and other financial databases, ensuring a comprehensive dataset for analysis.

Challenges and Solutions: One significant challenge was handling data volatility and missing values. I addressed this by employing data cleaning methods and robust statistical techniques to fill gaps. Another challenge was model overfitting, which I mitigated by using cross-validation and adjusting model parameters.

Learning Outcomes and Results: Through this project, I gained a deeper understanding of time-series forecasting and the complexities of financial data. The final model achieved a reasonable level of accuracy, providing valuable insights into interest rate trends.

2-Step Loss Given Default (LGD) Modeling Project (FIM 601)

Project Goals: This semester, my focus has shifted to a 2-step LGD modeling project, aiming to predict the recovery rate of defaulted loans. The project's goal is to develop a predictive model that can be used for risk assessment and management in financial institutions.

Technology and Data Sources: I am using Python for statistical modeling and analysis. The data is sourced from internal banking datasets and publicly available financial information, providing a robust basis for model development.

Quantitative and Analytic Techniques: The project involves logistic regression and linear regression models in a two-step approach. The first step predicts the probability of default, and the second step estimates the recovery rate. I am also exploring the use of ensemble methods to improve prediction accuracy.

Challenges and Solutions: A key challenge has been ensuring data quality and dealing with imbalanced datasets. To overcome this, I have implemented data augmentation techniques and employed stratified sampling methods. Additionally, I am focusing on model validation and backtesting to ensure reliability.

Learning Outcomes and Results: This project has enhanced my skills in risk modeling and statistical analysis. Although still in progress, the preliminary results indicate a promising model that could significantly contribute to effective risk management practices.

Career Goals and the MFM Program at NC State Why I Chose the MFM Program
I chose to pursue the Master of Financial Mathematics (MFM) program at NC State because of its strong emphasis on quantitative finance and the practical application of mathematical models in financial decision-making. The program's curriculum aligns perfectly with my career aspirations in quantitative analysis and risk management.

Learning and Skills Development: Through the program, I have learned about advanced financial models, risk assessment techniques, and the integration of technology in finance. The projects I have undertaken have provided hands-on experience that directly relates to my career goals in financial risk management and quantitative finance.

Skills to Improve and Plan: While I have developed a strong foundation in quantitative analysis, I recognize the need to enhance my coding skills, particularly in data-driven techniques such as machine learning and AI. I plan to achieve this through elective courses, self-study, and practical applications in ongoing projects.

Conclusion: The MFM program has been instrumental in shaping my career trajectory. The projects I have worked on have not only reinforced my interest in finance but also equipped me with the necessary skills to excel in the field. I am committed to continuous learning and skill enhancement to achieve my career goals.

Joseph Jonasson

When Optimal Transport meets Quantitative Finance

As part of the North Carolina State University Master of Financial Mathematics curriculum, multiple courses in Probability and Statistical Inference equip students with a deep understanding of probability distributions. In popular finance careers like quantitative finance, actuarial science, risk management, and many more, these concepts are essential to know how to view markets using the same statistical framework used in these courses. Often, problems presented in probability and statistics are not always obviously applicable to finance. However, with some creativity, and a good understanding of probability theory, financial mathematicians have found ways to apply many famous statistics problems to finance. This article focuses on one such topic: Monge-Kantorovich Optimal Transport.

The original optimal transport problem dates back to 1781 France. Mathematician Gaspard Monge wanted to answer a simple question: given a pile of sand, how can one move the sand into a prescribed shape, like a sandcastle, in the most efficient way? This question was then further explored and formalized by Nobel Prize winner Leonid Kantorovich in the 1940s. Kantorovich's formulation is the following:

> Let (X, μ) and (Y, ν) be two probability spaces and let $c: X \times Y \rightarrow R \cup \{+\infty\}$ be the cost function associated with moving from the first probability space to the other. Use probability measures π to model the transference plans from $x \in X$ to $y \in Y$. The goal is to minimize $\int c(x,y) d\pi(x,y)$ for all probability spaces π. In this formulation, we can think of $d\pi$ informally as measuring the amount of sand being moved from x to y.

Finding the least costly way of moving materials from one location to another has broad applications across many fields of science, mathematics, and economics. In finance, these ideas can be applied to asset pricing to build a model-independent framework, providing an alternative to other popular models. For example, the Black-Scholes model, and numerous other similar models, rely on parameters that may require model calibration. However, this process of calibrating theoretical option prices to what is seen in the market can yield varying results and disagreements on option prices. The goal of using optimal transport for option pricing avoids the model calibration step, aiming to find prices that are entirely independent of any predetermined model.

To fully apply optimal transport to asset pricing, the martingale constraint is an important addition, which states the following:

> For a sequence of random variables $X_1, X_2, X_3, \ldots$ and for any time n with $E[|X_n|] < \infty$, then $E[X_{n+1} | X_1, \ldots, X_n] = X_n$.

This now requires that the transportation from one probability space to another reflects the "no-arbitrage condition" often used in asset pricing. In martingale optimal transport, we consider the probability distributions of an asset's price at an initial time and a future time and seek the most efficient way to "transport mass" between them while ensuring that the asset's expected value at any intermediate time matches its current price. Mathematically, this means that if X_t represents the asset price at time t, then any transport plan π between the initial distribution and terminal distribution must satisfy the martingale constraint.

To see how martingale optimal transport has been used, consider the following paraphrased problem on pricing forward start options from Hobson and Neuberger's paper "Robust Bounds for Forward Start Options":

Considering the current vanilla call and put options prices, we want to use martingale optimal transport to search over all possible price processes, finding the one that yields the highest price for the derivative while respecting the martingale constraint. The price process is the "transportation plan" from the original Monge-Kantorovich problem. For a derivative with payoff $|P_{T_1} - P_{T_2}|$, Hobson and Neuberger find bounds on the price without any of the assumptions of a predetermined stochastic model.

Papers like this one demonstrate how mathematicians interested in finance can use martingale optimal transport to gain deeper insights into option pricing beyond traditional stochastic models. By focusing on observable market data and enforcing arbitrage-free constraints, martingale optimal transport allows for the derivation of model-independent price bounds that remain valid even in the face of uncertainty about volatility and price dynamics. This framework enhances theoretical understanding and provides tools for traders and risk managers seeking more reliable hedging strategies in incomplete markets.

Reference: Hobson, D., & Neuberger, A. (2010). Robust bounds for forward start options. Mathematical Finance, 22(1), 31–56. https://doi.org/10.1111/j.1467-9965.2010.00473.x

Sarang Kansara

Trading On Market Mood: Integrating Sentiment Analysis Into Algorithmic Strategies

Financial markets are influenced by investor sentiment, yet traditional trading strategies primarily rely on price, volume, and fundamental indicators. With advancements in natural language processing (NLP) and the increasing availability of alternative data, sentiment analysis is becoming a powerful tool for traders. As part of my FIM 601 project, I am exploring how sentiment analysis can be integrated into quantitative trading strategies to enhance predictive accuracy and improve risk-adjusted returns.

Project Overview

The goal of this project is to extract sentiment from social media, financial news, and online discussions, quantify its impact on asset prices, and incorporate this data into algorithmic trading models. By leveraging NLP techniques, I aim to determine whether shifts in sentiment can serve as leading indicators for price movements and volatility changes.

Technology and Data Sources

I am implementing this project using Python and leveraging several key technologies:

- FinBERT: A financial-domain adaptation of BERT, used for sentiment classification of financial news and online discussions.
- NLTK and VADER: Used for sentiment analysis of short-form text such as tweets and Reddit comments.
- Pandas and SQL: For data preprocessing and structured storage.
- Backtrader: For backtesting trading strategies.

The primary data sources include:

- Social Media: Twitter (X) and Reddit, which provide real-time market sentiment from retail traders, institutional analysts, and financial influencers.
- Financial News: News articles from financial media outlets to gauge sentiment shifts in professional reporting.
- Macroeconomic Indicators: Data from Yahoo Finance and the Federal Reserve Economic Data (FRED) to assess broader economic conditions.

Quantitative and Analytical Techniques

To extract insights from sentiment data and integrate them into a trading model, I am employing several analytical methods:

- Sentiment Scoring with FinBERT: News articles and Reddit discussions are processed using FinBERT to classify sentiment as positive, neutral, or negative.
- Feature Engineering: Rolling sentiment scores are computed and combined with stock price, volume, and macroeconomic indicators to construct predictive signals.
- Machine Learning Models: Regression techniques, including logistic regression and random forests, are used to analyze the relationship between sentiment and asset price movements.
- Backtesting Performance: A sentiment-based long-short trading strategy is being backtested, with performance measured through Sharpe ratio, maximum drawdown, and alpha generation.

Challenges and Solutions

One major challenge is handling the noise and biases inherent in social media sentiment data. Unlike structured financial reports, tweets and Reddit posts contain slang, sarcasm, and exaggerated emotions. To mitigate this, I am applying text preprocessing techniques such as tokenization, stop-word removal, and sentiment smoothing.

Another challenge is determining the optimal weighting of sentiment indicators relative to traditional price-based signals. To address this, I am conducting lag analysis and optimizing feature selection to ensure sentiment signals provide meaningful predictive value.

Key Learnings and Expected Outcomes

This project is helping me deepen my understanding of how alternative data sources can be integrated into quantitative finance. I expect the sentiment-enhanced trading strategy to demonstrate improved risk-adjusted returns compared to traditional technical and fundamental strategies. If successful, the model could provide valuable insights into market trends, especially during periods of high volatility driven by news and social media narratives.

Next Steps

I plan to refine the model by experimenting with different NLP techniques, such as transformers and topic modeling, to improve sentiment classification accuracy. Additionally, I aim to explore reinforcement learning approaches to dynamically adjust trading strategies based on sentiment fluctuations.

By incorporating sentiment analysis into trading models, I hope to bridge the gap between qualitative market sentiment and quantitative decision-making, creating a more comprehensive approach to systematic trading.

Dev Kewlani

AlphaPortfolio: Direct Portfolio Optimization using Deep Reinforcement Learning

In today's fast-paced financial landscape, quantitative innovation is key. At NC State's Master of Financial Mathematics program, I have led a dynamic team of three to build AlphaPortfolio—a project that integrates deep reinforcement learning into portfolio optimization. As a program ambassador, I lead a team of three on this initiative which I conceptualized following comprehensive research conducted during my winter break. This semester, we have successfully transitioned from theoretical exploration to practical implementation, bridging the gap between academic concepts and real-world financial applications. Our approach redefines traditional portfolio construction, aiming to maximize risk-adjusted returns by directly learning optimal investment decisions in real time.

Project Overview and Objectives

AlphaPortfolio represents a paradigm shift in portfolio construction by directly optimizing investment objectives through deep reinforcement learning rather than following the conventional two-step approach of first estimating return distributions and then constructing portfolios. Our primary goal is to maximize the Sharpe ratio over 12-month returns while creating a flexible framework that can accommodate an investor's investment constraints and objectives.

The model architecture consists of three components:

- Sequence Representation Extraction Module (SREM): Utilizes a transformer encoder to process sequential historical asset features, effectively capturing temporal dependencies in financial time series. Unlike traditional recurrent neural networks, this approach better handles long-range dependencies by reducing network path length.
- Cross-Asset Attention Network (CAAN): Computes attention across different assets, modeling cross-sectional relationships in the investment universe. This component generates "winner scores" that rank assets for portfolio inclusion.
- Portfolio Generator (PG): Transforms the "winner scores" into optimal portfolio weights by implementing a long-short strategy of selecting the top X stocks for long positions and bottom X stocks for short positions, while maintaining dollar neutrality to mitigate market risk exposure.

Technology Stack and Data Sources

Our implementation leverages Python's robust ecosystem, with PyTorch serving as our deep learning framework due to its flexibility and automatic differentiation capabilities essential for reinforcement learning. We maintain a modular codebase with clear separation between data processing, model architecture, and training loops.

For data, we access WRDS platforms to obtain comprehensive financial information from CRSP, Compustat and Financial Firm Ratios databases. Our dataset spans the top 500 companies each year from 1975 to 2024, incorporating over 50 firm characteristics including price-based indicators, investment and profitability metrics, valuation ratios, and liquidity measures.

Quantitative Techniques

The main idea of AlphaPortfolio lies in embedding portfolio construction into a reinforcement learning framework where:

- States represent historical market conditions and asset features
- Actions correspond to portfolio weights
- Rewards align with the Sharpe ratio computed over specified time horizons

This formulation enables the model to learn optimal asset allocation directly from market experience, capturing non-linear relationships and temporal dependencies that traditional methods might miss.

Implementation Challenges

The project has presented several complex challenges:

- Data Preprocessing: Ensuring proper alignment of time-series data to prevent look-ahead bias has been particularly demanding. Financial data features varying reporting frequencies and delayed availability, requiring careful as-of joins and point-in-time construction.
- Computational Efficiency: Training deep reinforcement learning models with large financial datasets demands significant computational resources. We implemented batch processing and gradient accumulation to manage memory constraints.
- Model Interpretability: Unlike traditional factor models, deep learning approaches can be "black boxes." We're trying to incorporate gradient-based sensitivity analysis to identify which features most significantly impact the model's decisions.
- Hyperparameter Optimization: The model's performance is sensitive to numerous hyperparameters, including learning rates, network architecture, and reinforcement learning parameters.

Continue Next Page

Dev Kewlani

AlphaPortfolio: Direct Portfolio Optimization using Deep Reinforcement Learning

Continued From Previous Page

Preliminary Results

Initial back testing results from 2000-2021 are promising, with our implementation achieving a Sharpe ratio of 1.5 compared to 1.2 for equal-weighted portfolios (which is validated in walk-forward manner (3 years of validation for every 6 years of training).

Learning Outcomes and Future Directions

Leading this project has significantly enhanced my technical expertise in deep learning applications for finance, particularly in designing attention mechanisms for cross-sectional financial data. Additionally, I've developed leadership skills in coordinating research efforts, managing implementation timelines, and synthesizing theoretical knowledge into practical applications.

As we refine our implementation, we plan to extend AlphaPortfolio to incorporate transaction costs, market impact models, and alternative objectives beyond Sharpe ratio optimization, while also trying to make the project more interpretable. The framework's flexibility allows for potentially groundbreaking applications in goal-based wealth management and personalized portfolio construction.

William Lanzoni
A Decision Without Regret

During my senior year at Siena College, I had a conversation with my advisor about my career goals and post-graduation plans. I had always known I wanted to pursue an advanced degree, but I was debating whether enrolling immediately after undergrad was the right choice. Looking back, I have no regrets.

During our discussion, my advisor suggested that a Master's in Financial Mathematics would be the best fit for my background, as well as my career aspirations, as a Finance major with a Mathematics minor. He provided me with a list of top programs in the field. After conducting my own research and speaking with alumni from various programs, I ultimately decided on NC State.

What initially attracted me to the program was not only its strong reputation but also the collaborative environment where professors and students push each other toward academic success. Looking deeper, I was particularly drawn to the opportunity to work on industry-level projects and the ability to present findings to a panel of professionals, something unique to NC State. These projects allow students to bridge the gap between theory and practice, applying classroom knowledge to real-world scenarios. This hands-on experience is invaluable for early career success, providing both practical skills and relevant industry exposure.

Coming into a program with a strong focus on machine learning, while having a background in pure Finance and Mathematics, was initially daunting. However, this pushed me to develop the ability to adapt to challenging situations. Over the summer, I took advantage of free online courses to build a strong foundation in various programming languages. This preparation helped me succeed when applying these skills to industry projects, particularly in financial modeling and projections.

One unexpected but invaluable resource has been my classmates. Collaborating with them, exchanging ideas, and learning from their strong backgrounds in computer science have significantly enhanced my capabilities in financial modeling. Their insights have helped me bridge the gap between finance, mathematics, and programming, making me a more well-rounded problem solver.

I aspire to pursue a career in wealth management, and concentrating in Portfolio Management has allowed me to refine and enhance my skills in this field. Additionally, enrolling in a few MBA courses has further strengthened my expertise in equity valuation.

Overall, combining a strong quantitative background with a Finance degree has given me a unique and competitive edge in the industry, one that I plan to continue building upon. The program’s emphasis on financial modeling through programming languages ensures that students stay ahead of the curve and are well prepared for the evolving landscape of the financial industry.

Tharun Mandadi

Markowitz Model With Large Simulation And Sensitivity Analysis

Portfolio optimization remains a fundamental aspect of modern finance, and the Markowitz model continues to be a cornerstone for constructing efficient investment portfolios. This project applied the Markowitz model to develop a minimum variance portfolio and a maximum Sharpe ratio portfolio, leveraging financial data, simulation techniques, and optimization algorithms. By simulating thousands of portfolio combinations, we aimed to identify the most efficient asset allocations while conducting sensitivity analysis to assess how varying data conditions impact portfolio performance.

The project was executed using Python, integrating libraries such as yfinance for data acquisition, Pandas and NumPy for data preprocessing and numerical computations, and Matplotlib for visualization. Optimization was carried out using PyPortfolioOpt and SciPy's minimize function, while QuantStats was used for performance evaluation metrics, including the Sharpe ratio, Sortino ratio, Alpha, and Beta.

Our dataset comprised historical adjusted close prices of six selected stocks—Apple (AAPL), Adobe (ADBE), Microsoft (MSFT), Google (GOOGL), Cisco (CSCO), and Intel (INTC)—from 2014 to 2024. Data preprocessing included handling missing values, ensuring alignment across assets, and eliminating erroneous price points to maintain data integrity.The methodology involved simulating 25,000 portfolios with randomly assigned weights summing to one. Annualized returns and volatility were computed for each portfolio, and the Markowitz model was applied to determine the minimum variance and maximum Sharpe ratio portfolios. The Efficient Frontier was plotted, visually illustrating the trade-off between risk and return, with optimal portfolios highlighted (image1). Cumulative returns of the maximum Sharpe ratio portfolio were compared against individual stocks to assess performance (image2).

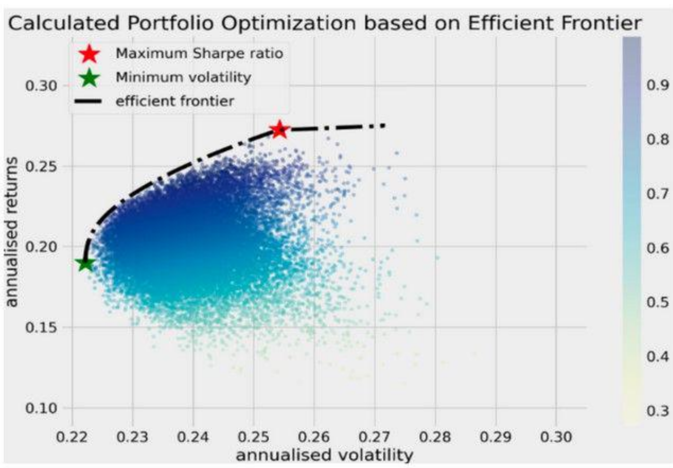


Image1: Calculated Portfolio Optimization based on Efficient Frontier

Image2: Cumulative Returns of the Max Sharpe ratio Portfolio and Individual Stocks

Challenges encountered included ensuring data consistency, optimizing computational efficiency for large-scale simulations, and estimating a stable covariance matrix crucial for optimization. These were addressed by refining data cleaning processes, optimizing code for efficiency, and leveraging advanced numerical techniques to enhance stability.
Key learning outcomes included proficiency in specialized financial libraries, a deeper understanding of portfolio optimization and risk management, and hands-on experience in large-scale simulations and sensitivity analysis. The project demonstrated the robustness of optimized portfolios under varying conditions, providing actionable insights into investment strategies.

Future extensions could include incorporating multi-objective optimization to balance additional criteria such as ESG factors and transaction costs, applying machine learning for risk prediction, and scaling simulations to larger datasets with real-time data integration. Additionally, incorporating clustering techniques such as K-means and classification models like Support Vector Machines could enhance portfolio construction methods.

The results of this study provided strong evidence for the effectiveness of portfolio optimization techniques in real-world applications. The minimum variance and maximum Sharpe ratio portfolios consistently outperformed individual stocks in risk-adjusted returns. Sensitivity analysis further confirmed the robustness of these portfolios under different market conditions, making the methodology highly relevant for institutional and retail investors alike.

Overall, this project reinforced essential quantitative finance skills, equipping us with the expertise necessary for roles in investment management and financial engineering. The knowledge gained will be instrumental in tackling future challenges in financial modeling, risk assessment, and algorithmic trading strategies.

Chelsea Niles

Momentum Enhanced Modern Portfolio Optimization

With financial markets constantly evolving, investors need strategies that not only manage risk and return but also adapt as conditions change. As part of the Master of Financial Mathematics (MFM) program, I had the opportunity to work on a team project where we enhanced Modern Portfolio Theory (MPT) using momentum factor and rebalancing techniques to address this.

Project Goals and Overview

The project aimed to build and identify optimal portfolios by enhancing Harry Markowitz's MPT, which identifies optimal portfolios offering the highest expected return for a given risk level. However, MPT comes with some limitations, such as its reliance on long-term historical averages, the assumption of normally distributed returns, and high sensitivity to input estimates. These factors can lead to suboptimal performance in dynamic markets. To address these challenges, we introduced Momentum Factor and Rebalancing Triggers.

Our goal was to construct efficient portfolios and analyze the performance at three different risk levels—Aggressive, Moderate, and Conservative—using risk-adjusted performance metrics such as Sharpe Ratio, Maximum Drawdown, and Annualized Return. Ultimately, we aimed to generate Alpha—excess return relative to the benchmark—across all risk levels.

Data and Technology

Data, from 2002 to 2024, was sourced from Morningstar for 15 asset classes (for diversification). Each asset class represented a unique market segment, from US equities and international stocks to commodities and corporate bonds. For each asset class, we selected mutual funds or ETFs that closely track its performance, allowing for realistic backtesting.

We conducted all data processing, analysis, and modeling in Python, utilizing libraries such as Pandas for data manipulation, NumPy for mathematical calculations, and Matplotlib for visualizations. PyPortfolioOpt was essential for performing the core optimization routines.

Quantitative Techniques

Our process began with calculating monthly expected returns, standard deviations, and the covariance and correlation matrix for all asset classes to analyse the relationship between them.

$$\mathrm{Cov}(X,Y) = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{n-1} \qquad r(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_Y}$$

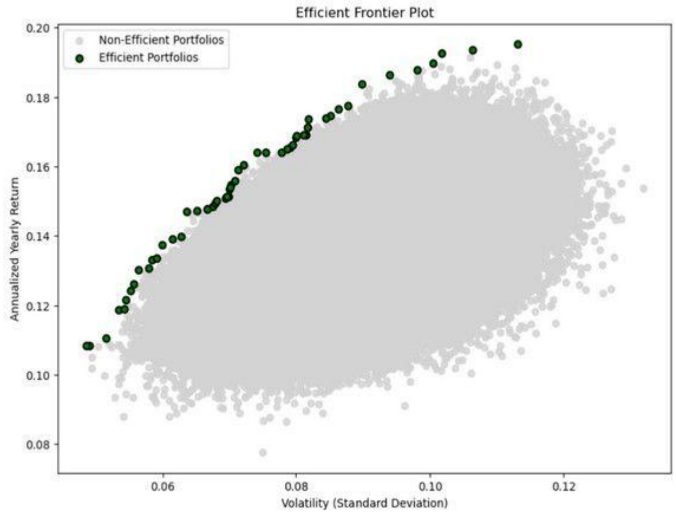
These inputs were then used to identify the efficient frontier—portfolios providing the optimal balance between risk and return.

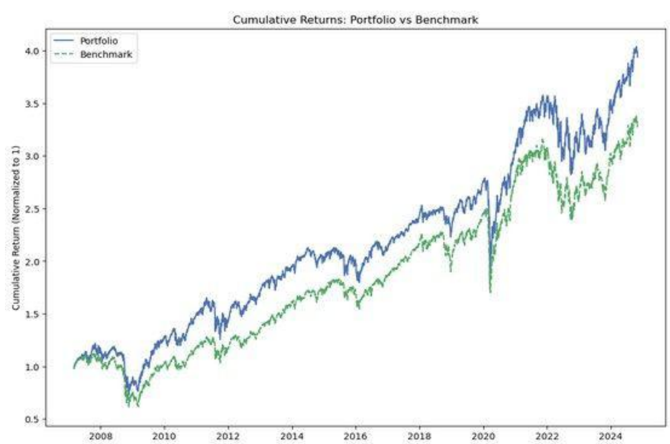
Next, we incorporated the Momentum Factor, recalculating expected returns based on the most recent 12-month period, rather than averaging across the entire dataset. This adjustment allowed our portfolios to capture and capitalize on current price trends, unlike the previous more static model.

To maintain the target risk level for each of the generated portfolios, rebalancing was introduced. A relative tolerance band was implemented, and if any asset's weight drifted outside a ±20% range around its target allocation, the portfolio was automatically rebalanced. This ensured that risk exposures remained consistent over time.

Results

We saw promising results across all risk profiles. The optimized portfolios outperformed their respective benchmarks, achieving cumulative returns between 38.89% and 69.47% higher over the backtested period (2007-2024). The Aggressive portfolio generated annualized Alpha of 1.57%, while Moderate and Conservative portfolios produced Alpha of 1.02% and 0.80%, respectively.





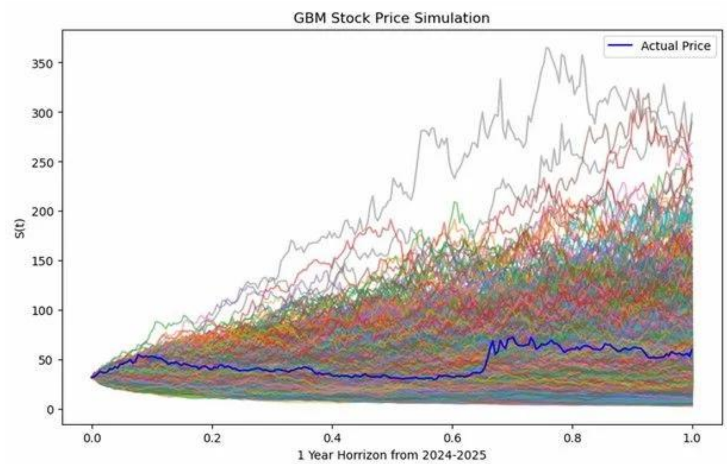
Jinjia Peng

Monte Carlo Stock Price Simulation with Jump Diffusion GBM: A Quantitative Approach

Quantitative finance relies on mathematical models to analyze and predict market movements. As a graduate student in the Master of Financial Mathematics (MFM) program at North Carolina State University, I explored modeling techniques that bridge theory with real-world financial applications. My project, Stock Price Monte Carlo Simulation with Jump Diffusion Geometric Brownian Motion, applied these techniques to analyze iSoftStone Information Technology (Group) Co., Ltd.'s stock prices. By combining Monte Carlo simulations with a Jump Diffusion model, the project provided insights into price trends, risk assessment, and trading strategies.

The goal was to simulate iSoftStone's stock price movements over 14 days using the Jump Diffusion Geometric Brownian Motion (JDGBM) model. Traditional Geometric Brownian Motion (GBM) assumes continuous price evolution but fails to capture sudden market jumps. The JDGBM model improves this by integrating a Poisson-distributed jump process, making it more realistic. Using 487 historical stock price observations from February 6, 2023, to February 6, 2025, I estimated key parameters such as drift, volatility, and jump frequency. These were used in 10,000 Monte Carlo simulations to generate possible price trajectories, offering insights into future price distributions and risks.

The methodology began with identifying jump events, defined as log returns exceeding the mean plus three standard deviations. These jumps were modeled using a Poisson process to estimate the probability of sudden price changes. The Monte Carlo simulation combined Brownian motion for continuous movements and Poisson-distributed jumps for sudden spikes or drops. The results showed that the mean simulated price trajectory closely followed actual movements, validating the model's accuracy.

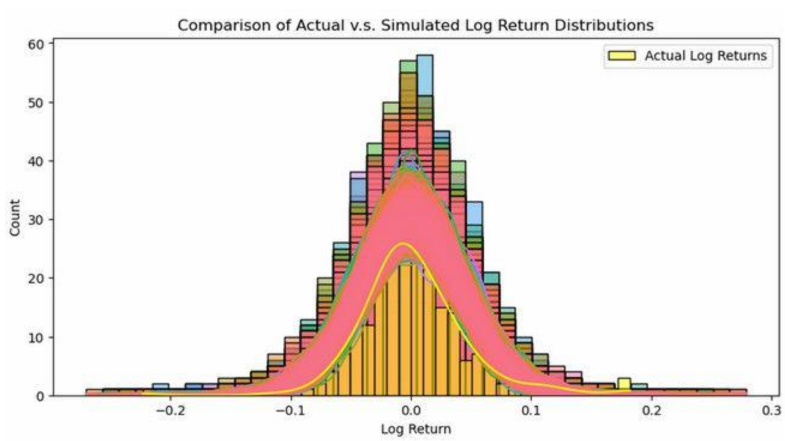


10,000 simulation paths and the actual stock price of iSoftStone from 2024 to 2025

From a risk perspective, I conducted a Value at Risk (VaR) analysis at 95% and 99% confidence levels. The analysis revealed a 5% probability of a loss greater than 53.39% and a 1% probability of a loss exceeding 66.40%, highlighting the stock's high-risk nature. A Kolmogorov-Smirnov (KS) test confirmed that the simulated log returns closely matched historical data. Additionally, the annualized Sharpe ratio of 0.31 indicated low risk-adjusted returns, emphasizing the need for portfolio diversification. Given the stock's bullish trend, I concluded that a long position with options-based hedging strategies might be better to mitigate the downside risks.

One challenge for me during this project was accurately estimating jump parameters due to unpredictable market factors. To refine these, I used Kernel Density Estimation (KDE), ensuring the simulated distribution matched historical trends and enabling the model to correctly capture the movements. Another challenge was the computational complexity of running 10,000 simulations. I optimized efficiency using Python libraries like NumPy, Pandas, and SciPy for faster processing and better memory management. For future work, I plan to incorporate technical indicators like Bollinger Bands and MACD to enhance predictive accuracy by accounting for market behavior.

This project aligns with my career goal of working in quantitative finance, particularly in derivatives pricing, algorithmic trading, and risk management. The MFM program has equipped me with skills in stochastic calculus, machine learning, and financial modeling. Moving forward, I aim to explore machine learning techniques to uncover deeper patterns in asset price movements and improve model accuracy. I also plan to integrate Markowitz Portfolio Theory for portfolio optimization and analyze macroeconomic factors like interest rates and inflation to better understand price dynamics.



Comparison of the Actual V.S. Simulated Log Return Distributions for the KDE test

In conclusion, this project strengthened my expertise in Monte Carlo simulations, risk modeling, and quantitative trading strategies. This experience reinforced my passion for quantitative finance and demonstrated how rigorous quantitative analysis can inform investment and risk management strategies. As I continue my journey, I look forward to refining my skills and contributing to the field through innovative, data-driven approaches.

Adharsh Rajagopal

Beyond The Numbers: Charting My Path To Quantitative Research

Choosing the Master of Financial Mathematics (MFM) program at NC State was a natural extension of my passion for quantitative finance—a field that perfectly blends my interests in mathematics, computer science, and problem-solving. My goal is clear: to become a quantitative researcher who leverages mathematical models and statistical techniques to address complex financial challenges, optimize risk strategies, and drive innovation in financial products.

My background in Computer Science, complemented by a mathematics minor, has equipped me with foundational skills crucial to succeeding in quantitative finance. Through coursework and professional experiences, I have honed my programming skills, particularly in languages such as React, R, and Java. My internship as an application developer at the Bank of Montreal in 2023 and my current role as an Actuarial Quantitative Analysis Associate at Allstate have provided practical exposure to coding-intensive environments. These roles reinforced the critical importance of translating complex mathematical insights into efficient, functional code—a core competency for a quantitative researcher. However, I recognize the need for growth in certain areas essential for my career aspirations.

My mathematical proficiency is robust, yet I acknowledge that advanced probability theory and mathematical assessment techniques demand deeper exploration and practice. The MFM program offers a structured path to strengthen these competencies, and I intend to fully immerse myself in coursework and quantitative projects that emphasize applied probability and statistical modeling.

Additionally, quantitative researchers frequently engage with cutting-edge financial literature. To succeed, I must enhance my ability to critically read, interpret, and synthesize information from academic and industry papers. Participating in research discussions, collaborating with peers on quantitative projects, and actively presenting my findings will significantly improve my analytical and communication skills. Moreover, I aim to refine my visualization and presentation capabilities—essential tools for effectively communicating complex quantitative concepts.

Throughout my journey in the MFM program, my strategy for skill development will be hands-on and proactive. By combining rigorous coursework with practical projects, research activities, and peer collaboration, I am confident in my ability to not only meet but exceed the demands of a career as a quantitative researcher. NC State provides the ideal environment to transform my ambitions into impactful outcomes.

Sanith Rao

Exploring Sentiment Analysis For Stock Prediction: A Quantitative Approach

As part of the Financial Mathematics program at NC State, my team and I embarked on a project for FIM 601 that explores the integration of sentiment analysis into stock price prediction models. Our goal is to assess whether incorporating market sentiment from news and social media can improve predictive accuracy compared to traditional quantitative models.

The primary objective of our project is to evaluate the impact of sentiment analysis on stock price movements, particularly for highly volatile stocks. By analyzing textual data alongside traditional market indicators, we aim to determine whether sentiment-driven signals enhance trading strategies and risk management techniques.

We leverage Python for data processing and machine learning, utilizing key libraries such as Pandas, NumPy, Scikit-learn, and TensorFlow. For natural language processing (NLP), we employ NLTK and the VADER sentiment analysis tool. Our data sources include historical stock prices and financial metrics from Bloomberg and Yahoo Finance, news articles and tweets sourced via APIs from platforms like Twitter and Alpha Vantage, and sentiment analysis tools using pre-trained sentiment models and custom-built classifiers.

Our methodology integrates several techniques. We use sentiment analysis with NLP techniques to assign sentiment scores to news headlines and tweets. Feature engineering is employed to construct sentiment-based indicators, including moving averages of sentiment scores and correlation with price movements. Time series modeling is conducted using LSTMs and ARIMA models to analyze stock price trends. Finally, predictive modeling is implemented to compare the performance of traditional regression models with hybrid models that incorporate sentiment features.

One of the main challenges has been handling noisy and unstructured textual data. We mitigated this by applying advanced NLP preprocessing techniques such as tokenization, stopword removal, and lemmatization. Another challenge was aligning sentiment data with stock movements, as market reactions often exhibit lag effects. To address this, we experimented with different time windows and cross-validation techniques to refine our model's predictive capability.

This project has provided valuable insights into the integration of sentiment analysis in quantitative finance. We observed that sentiment-enhanced models exhibit improved short-term prediction accuracy, particularly for stocks with high retail investor interest. While challenges remain in refining predictive power over longer time horizons, our findings highlight the potential of sentiment-driven analytics in trading strategies.

The intersection of machine learning, NLP, and finance is an exciting frontier with significant implications for quantitative research and trading. By continuing to refine our methodologies and exploring alternative sentiment sources, we aim to develop a more robust framework for sentiment-based market prediction. This project has deepened our understanding of both financial modeling and data science, equipping us with valuable skills for our future careers in quantitative finance.

Vismit Rekhan

Bridging Traditional Finance with the Future

Finance is always evolving, shaped by new trends, technologies, and market shifts. When I started my Master of Financial Mathematics (MFM) program, I saw it as a way to build strong financial modeling and risk management skills. What I didn't expect was how much it would change my understanding of finance, especially its connection to Decentralized Finance (DeFi).

From Foundations to the Future

My background has always leaned toward the qualitative side of finance—market trends, investor behavior, and macroeconomics. But as I moved through courses like FIM 500 and FIM 601, I realized finance is not just about numbers. It's about using data-driven insights to make better decisions. Learning about stochastic calculus, portfolio optimization, and risk management has given me strong technical skills, but what excites me most is applying these ideas in real-world settings.

One key experience in my MFM journey was working on projects, where I got to see finance in action. This hands-on experience deepened my interest in bridging traditional finance (TradFi) with DeFi. These projects showed me how structured financial models could be adapted to decentralized systems, making finance more efficient and transparent.

Career Aspirations: A Hybrid Approach

With the rise of blockchain-based financial products, there is a growing need for professionals who understand both quantitative finance and decentralized systems. My goal is to work at the intersection of portfolio management, risk assessment, and DeFi, using what I've learned in the MFM program to analyze both traditional and digital assets.

During my time at NC State, I have also explored venture capital, investment research, and market trends, which has strengthened my belief that finance is changing in fundamental ways. The ability to combine classic financial models with new, tech-driven solutions is where I see the most opportunity, and that's the space I want to work in.

The Road Ahead

As I move forward in the MFM program, I am focused on expanding my quantitative skills and applying them in real-world financial markets. Whether through internships, research, or industry projects, I want to use my knowledge to keep up with the rapid changes happening in finance.

The next few years will be crucial as finance continues to adapt to digital transformation. Through the MFM program, investment experience, and exposure to DeFi, I want to be at the forefront of this shift—helping connect traditional finance with the future, one model at a time.

Partha Sanampudi

Levelling Up In Modelling

Post pursuing my undergraduate degree in Finance at Kelley School of Business, I had the desire to learn more in depth about the math behind the models and algorithms I was introduced to in my undergraduate. I believed a strong base in mathematics would help me in my goal of becoming an equity analyst. After deciding to join a master program in Financial Mathematics (MFM) I chose NC State in particular, as it is a small program with more hands-on feedback and support from the teaching staff. The emphasis on project work and putting to use what you learn in class, was also another major deciding factor for me to choose NC State in particular.

While I was able to learn applicable skills in Excel for my career during my undergraduate, like building discounted cash flow models, CAPM models and income statement forecasts. After joining the MFM program I was able to understand modeling can be done on a much higher level with larger and more complicated data sets while also integrating aspects such as machine learning and simulations of multiple variable outcomes. The program made me realize that my aspirations of equity analysis using various financial models can be done on a much more technical and accurate basis.

As a finance undergraduate with minimal exposure to technical skills such as coding and data analysis, I realized that these are much needed skills to become a professional in the field of financial analysis. Being enrolled in the MFM program is perfect for me to both acquire and perfect my knowledge in this domain. The program has a large emphasis on quantitative skills by implementing project work and most coursework includes assignments where we can put our learned knowledge to practical use.

During my time in the MFM program I plan to focus mainly on acquiring knowledge around the topics of statistics, mathematical modeling of financial data using coding languages, and the math behind various financial models. I plan to not only learn about these topics but complete projects where I can put these skills to practical use and showcase them to future prospective employers.

Eric Wang

Program Observations And Career Goals

I chose to pursue the Master of Financial Mathematics (MFM) at NC State because I wanted to improve my quantitative background in a way that was conducive to using both my Computer Science (CS) background as well as have directly relevant classes on topics related to the stock market. My end goal is to understand the underlying concepts behind complex trading models, and to secure a job in the trading space that makes profits using innovative models. I believe that NC State's MFM program was a cost-effective option in getting me this background while also allowing me to network and build connections with my potential peers in the space.

I have learned that banks are generally somewhere I do not see myself working long term. The basis of these models is just too simple for my tastes, and I want to work with Machine Learning (ML) state of the art models that are not held back by strict government regulations. I have also learned that my math background is insufficient to get through the top firms currently, and that is something I really want to grasp better as my coding skills seem efficient enough to give myself a fair shot.

Previous to joining the MFM program, I have gained several relevant work experience, having worked for over a year and a half at Wells Fargo. There I saw that banks were restricted and tasks were often mundane. Additionally, I completed an internship at Cerberus Capital Management, where I saw how my models could directly impact portfolio decisions in the housing market, and included more complex models than just simple linear regression. This confirmed my interest and made me feel like it was something I wanted to do long term.

The main skill I need to improve is my overall grasp of mathematics. High Frequency Trading and top firms in general like Optiver, SIG, and Belvedere, all 3 of which I received interview assessments, had significant logic puzzles that were extremely difficult.
I plan to understand mathematics more readily during my time in the MFM program. Fundamentally, if I can understand Monte Carlo and Stochastic Calculus, I believe this will set me up for success in the future. I will also focus on networking and making professional connections within the field of quantitative finance.

Jingxing Wang
Mortgage Loan Delinquency Prediction

In my FIM 500 project, I embarked on an analytical journey to predict mortgage loan delinquency, aiming to provide financial institutions with a reliable tool for early risk detection. The primary goal was to develop a model capable of forecasting delinquency with high accuracy, thereby enabling proactive decision-making in loan management.

Throughout the project, I utilized Python as the backbone of my analysis, leveraging powerful libraries such as scikit-learn and XGBoost. These tools allowed me to build and fine-tune multiple predictive models on a comprehensive mortgage loan dataset that contained detailed historical records of loan performance. This rich data source provided the necessary foundation for understanding borrower behavior and the factors that lead to default.
One of the most significant challenges was the inherent class imbalance within the dataset, where non-delinquent loans far outnumbered delinquent ones. To address this issue, I applied the Synthetic Minority Oversampling Technique (SMOTE), which effectively augmented the minority class and improved the fairness of the model. This approach ensured that the predictive models did not become biased toward the majority class and could accurately identify the risk of delinquency.

In pursuit of a robust solution, I experimented with several models. I trained a Logistic Regression model with Lasso regularization, which not only provided a solid baseline for prediction but also offered the advantage of interpretability through its coefficients. Additionally, I implemented ensemble methods such as Random Forest and XGBoost to capture complex interactions within the data. The combined use of these techniques proved to be highly effective, as evidenced by an impressive AUC of 0.9 on the test set.

A crucial part of the project was ensuring that the model's decisions were transparent and actionable. I conducted model attribution by analyzing the coefficients of the Logistic Regression model and evaluating the feature importance in the Random Forest and XGBoost models. This process provided valuable insights into which factors were most influential in predicting loan delinquency, thereby adding an extra layer of interpretability to the analysis.

Throughout this project, I not only honed my technical skills in machine learning and data analysis but also gained a deeper understanding of the practical challenges involved in financial modeling. I learned the importance of addressing data imbalances, selecting the appropriate model for the task, and ensuring that the results are both statistically sound and easily interpretable. These lessons have been instrumental in shaping my approach to quantitative analysis and have prepared me to tackle similar challenges in the future.

In the end, the project was a resounding success, demonstrating that a thoughtful combination of SMOTE, ensemble modeling, and rigorous attribution techniques can yield a highly accurate predictive model. Achieving a high AUC on the test set validated the effectiveness of my approach, while the interpretability of the results ensured that the findings could be confidently used in real-world financial decision-making. This experience has not only enriched my understanding of predictive analytics in finance but has also equipped me with the tools and insights necessary for future challenges in the field.

Yu Wang

A Rewarding Journey, and More to Come

After spending over a year coaching students for math competitions such as AMC 8 and AMC 10, I became even more intrigued by the rigor of mathematics. This experience deepened my interest in the subject and motivated me to pursue a career that integrates mathematics, programming, and finance to solve practical problems. Working as a teaching assistant in the math competition field also helped me realize that effective learning and skill development rely heavily on timely feedback and close mentorship from experienced professionals. These insights played a significant role in my decision to join the Master of Financial Mathematics (MFM) program at North Carolina State University.

The career services offered by this program are structured methodically and continuously improving. Before the semester began, the program's director of career services reached out for a one-on-one meeting, providing guidance on tailoring my resume and navigating job opportunities. This early support helped me secure internship interviews before classes even started. Beyond that, I benefited from career development lectures and industry guest speakers, which taught me how to network professionally, stay informed about industry trends, and identify valuable skill sets to improve. The career development course assignments not only helped turn goals into feasible plans, and eventually reality, but also instilled a mindset of continuous growth—whether through online resources, hands-on projects, or professional connections. The course also provided opportunities to practice networking and public speaking. I truly appreciate our career services director's dedication—he consistently considers what is best for students and integrates those insights into the curriculum. I am grateful to have joined this program under his leadership and cannot imagine it without him.

With the program's academic and professional support, I secured multiple internship offers within three months of starting my first semester. After that, I was encouraged to continue networking and improving my skills, to prepare for the full-time recruiting season. Through projects and coursework, I gained firsthand experience in model development—collecting and cleaning data, selecting and tuning models, and evaluating performance—which deepened my understanding of machine learning algorithms and statistical modeling. My experiences in this program also led me to another field that is equally, if not more, fascinating: quantitative development. Thanks to the program's interdisciplinary flexibility, I took an algorithm class in the computer science department this semester. At the same time, I worked on an independent project to develop a cryptocurrency trading platform in C++, simulating real exchange functions such as matching orders in the order book and allowing participants to make bids and asks. As a math student, analyzing algorithm complexity and designing efficient algorithms came naturally to me. However, I found myself particularly drawn to the nuances of C++ and leveraging these features to build high-performance systems. While algorithmic knowledge enables me to implement more efficient functions, being able to interact with the computers at a low level is crucial for developing low-latency trading systems—an area I now plan to focus on in my full-time job search.

Looking back on my journey in this program, I have grown more than I expected. My initial goal was to secure a summer internship, which I achieved as a short-term milestone. However, I also gained something more valuable—clarity about my long-term career aspirations. If I could go back to the moment I received my admission offer, but were still worried about the competitive and seemingly uncertain job market, I would tell myself: "Don't worry. A rewarding journey lies ahead."

Jinyi Yang

Navigating Financial Math: From Theory to Practice

During the Spring 2025 semester, I had the opportunity to collaborate on a team project focused on Deep Hedging for Options. The primary objective of our project was to develop a deep hedging framework in Python using Seq2Seq Neural Networks to enhance risk management for Snowball Options and compare its performance to traditional dynamic delta hedging.

To achieve this, we began by utilizing S&P 500 daily bar data to fit a returns model. This allowed us to use Monte Carlo (MC) simulations to model stock prices based on Cauchy Geometric Brownian Motion. For the classic delta hedging approach, I employed the MC method to estimate the price of the Snowball Option and computed the delta using finite difference quotients. For the deep hedging approach, we built an LSTM-based Seq2Seq model, using historical data of underlying assets and options as input features. The model's loss function was designed to minimize the variance of the portfolio's PnL. By inputting a sequence of historical data, the model generates a corresponding sequence of optimal hedging ratios to effectively minimize risk.

Throughout the project, I encountered several engineering challenges, with one of the most significant being the high time complexity of MC simulations. For traditional delta hedging, simulating numerous price paths for each day was computationally expensive, requiring an impractical amount of time to complete. To address this, we implemented vectorized operations and parallel computing techniques to optimize simulation performance. These optimizations reduced the runtime from over ten days to just two, significantly improving efficiency.
Although the project is still ongoing, I have already gained valuable insights and skills. This experience deepened my understanding of both pricing theory and deep learning models, while also enhancing my practical abilities in areas such as MC methods and vectorized coding.

Currently, I am halfway through the MFM program at NC State, which I chose because of my passion for quantitative modeling and my aspiration to pursue a career in quantitative finance. Over the past months, I have gained a broader understanding of the various roles within the quant field, which has helped me solidify my goal of becoming a quantitative researcher and analyst. To further enhance my competitiveness, I plan to continue improving my coding skills through advanced coursework and regular practice on platforms like LeetCode. This journey has been both challenging and rewarding, and I am excited to continue building my expertise in quantitative finance.

Nick Zehnle

Regime-Adaptive PCA Mean Reversion Strategy – Understanding the Theory Strategy Overview

This project explores a statistical arbitrage strategy rooted in the existence of mean-reverting relationships among assets. Principal Component Analysis (PCA) is used to gauge the underlying structure of price movements and to obtain the so-called "alpha factors." Hence, PCA mean reversion assumes that if an asset deviates significantly from this structure it will tend to revert back in the near future. The general algorithm of the strategy is to perform PCA on the price data for a set of assets, regress each asset against the desired number of principal components, and trade the assets proportionally to their deviation so long as they cross a specified threshold. The deviation is measured by computing the Z-scores of residuals and a Z-score threshold is typically set as a global constant.

That said, a couple of shortcomings are glaring in this simple trading algorithm. First, the noise in price movements cannot be ignored as it could introduce faulty Z-score interpretations. Second, a constant Z-score threshold does not account for different market regimes such that high-volatility and low-volatility regimes are traded in the same manner. Ergo, in this project a Kalman filter is applied to the residuals and an Exponential GARCH (EGARCH) model is used to estimate the current market regime in terms of conditional volatility. A sigmoid function is then implemented to produce fluid Z-score thresholds according to market regimes. Below are the mathematical representations of these improvements – Kalman filtering, EGARCH, and the applied sigmoid function.

Kalman Filtering for Noise Reduction

To reduce noise in the residuals that this strategy ultimately trades upon, a Kalman filter is applied. The filter is defined by the subsequent state-space equations:

$$x_t = Ax_{t-1} + w_t, \quad w_t \sim N(0, Q)$$
$$z_t = Hx_t + v_t, \quad v_t \sim N(0, R)$$

Where:
- x_t represents the hidden state (true residual process),
- A is the state transition matrix,
- Q is the process noise covariance,
- z_t is the observed noisy residual,
- H is the observation matrix,
- R is the measurement noise covariance.

The Kalman update step refines the state estimate $\hat{x}_t$ using prior information and the latest observation:

$$K_t = P_t H^T (H P_t H^T + R)^{-1}$$
$$\hat{x}_t = \hat{x}_{t|t-1} + K_t (z_t - H\hat{x}_{t|t-1})$$

Where K_t is the Kalman gain and P_t is the covariance matrix of the state estimate.

EGARCH for Market Regime Detection

To ensure that this strategy is regime-adaptive, the EGARCH(1,1) model is fitted to each asset and the average of the last conditional volatility estimate is taken to represent the market regime. The EGARCH(1,1) model is defined as:

$$\log(\sigma_t^2) = \omega + \beta \log(\sigma_{t-1}^2) + \alpha \left(\frac{|\epsilon_{t-1}|}{\sigma_{t-1}} - E\left[\frac{|\epsilon_{t-1}|}{\sigma_{t-1}} \right] \right) + \gamma \frac{\epsilon_{t-1}}{\sigma_{t-1}}$$

Where:
- σ_t^2 is the conditional variance,
- ω, β, α, γ are model parameters,
- ϵ_t is the innovation term.

Continue Next Page

Nick Zehnle

Regime-Adaptive PCA Mean Reversion Strategy – Understanding the Theory Strategy Overview

Continued From Previous Page

Sigmoid Function for Adaptive Z-Score Threshold

To modify the trading threshold based on the market regime estimated by EGARCH, the following sigmoid function is employed:

$$\tau_t = \tau_0 + (\tau_{max} - \tau_0)\frac{1}{1 + e^{-k(\sigma_t - \sigma_{base})}}$$

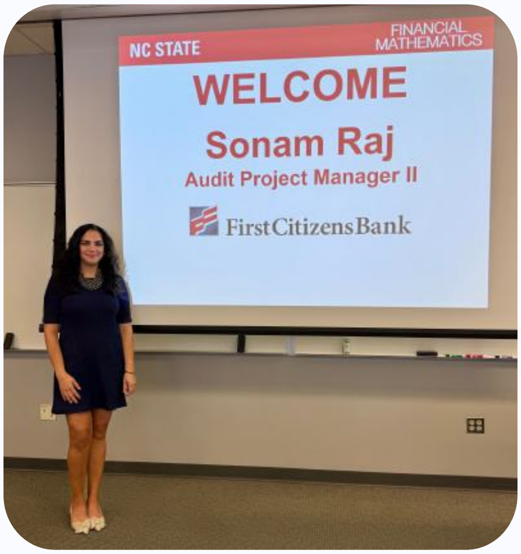
Where:

- τ_0 is the lowest possible threshold,
- τ_{max} is the maximum possible threshold,
- k is a scaling parameter,
- σ_t is current volatility,
- σ_{base} is the chosen baseline volatility.

Conclusion

This statistical arbitrage strategy improves upon the typical PCA mean reversion strategy by implementing a Kalman filter and an EGARCH model paired with a sigmoid function. In doing so, the strategy becomes regime-adaptive as well as less susceptible to faulty trade triggers. When backtested from January 2022 to 2025, a compounding annual return of 30.492% with an alpha and beta of 0.192 and 0.534, respectively, was achieved.


NC STATE UNIVERSITY



NC STATE
FINANCIAL MATHEMATICS
WELCOME
Sonam Raj
Audit Project Manager II
FirstCitizensBank

Why?

Arch | MI
Introduction to Arch Mortgage Insurance

SAS Hall at NCSU's Main Campus

NC STATE UNIVERSITY

Financial Mathematics Graduate Program

Ranked # 5 by Risk.net Ranked #12 by QuantNet

Financial-mathematics@ncsu.edu
https://financialmath.sciences.ncsu.edu