Risk Forecasting with Machine Learning

Candidate Number: 37279

2024-12-17

Introduction

Risk forecasting plays a fundamental role in financial risk management, helping institutions optimize portfolios, anticipate market volatility, and comply with regulatory frameworks. While traditional approaches like the GARCH model offer a strong theoretical foundation for volatility prediction (Bollerslev, 1986), their reliance on linearity and stationarity assumptions often limits their effectiveness in dynamic and complex financial markets. In contrast, machine learning models such as Random Forest and XGBoost provide a more flexible and powerful alternative by capturing non-linear relationships and processing high-dimensional data efficiently (Breiman, 2001; Chen & Guestrin, 2016). The incorporation of engineered features, including lagged returns and technical indicators like MACD and RSI, further enhances the predictive capabilities of these models (Hull, 2018). Machine learning's ability to uncover intricate patterns and improve forecast accuracy has been highlighted in empirical studies (Gu, Kelly & Xiu, 2020). This project evaluates the performance of Random Forest and XGBoost models using historical data from Citigroup (C), Ford (F), and Apple (AAPL), providing a diverse sectoral testbed. By synthesizing course concepts—volatility modeling, risk measures, and backtesting—this study constructs a robust and dynamic framework for market risk analysis.

Data

Metric	Price (PRC)	Trading Volume (VOL)	Returns (RET)	BID Price	ASK Price
Count	17,355	17,351	17,355	17,355	17,355
Mean	78.28	51,672,300	0.00064	78.26	78.30
Standard Dev.	113.69	110,122,300	0.0265	113.68	113.70
Minimum	1.02	719,436	-0.3902	1.01	1.02
25th Percentile	13.06	14,512,810	-0.0105	13.06	13.07
Median	43.21	26,085,860	0.0001	43.21	43.24
75th Percentile	80.76	48,541,190	0.0115	80.74	80.77
Maximum	702.10	1,897,900,000	0.5782	702.11	702.18

The dataset, spanning January 1, 2000, to December 23, 2023, sourced from WRDS, includes daily trading data for Apple (AAPL), Citigroup (C), and Ford (F), representing technology, financial services, and automotive sectors, respectively. Key variables include BID, ASK, and closing prices (PRC), daily returns (RET), and trading volumes (VOL). Preprocessing involved standardizing dates, filtering tickers, calculating log returns, and rolling volatility, while engineered features like MACD, RSI, and lagged returns were included to enhance model inputs. The dataset reflects significant market events, such as Apple's 2007 iPhone launch, Citigroup's turbulence during the 2008 Global Financial Crisis, and Ford's cyclical trends tied to economic conditions and EV innovations. Descriptive statistics highlight long-term price growth, particularly in technology stocks, with prices ranging from 1.02 to 702.10 and returns varying between -39.02% and 57.82%, aligning with periods of market volatility. Trading volumes peaked at 1.89 billion shares, reflecting heightened activity during market shocks. These diverse sectoral dynamics and advanced preprocessing ensure a robust dataset for evaluating predictive models like Random Forest and XGBoost in financial risk forecasting.

Empirical Analysis

This analysis forecasts stock volatility using Random Forest (RF) and XGBoost (XGB), machine learning models adept at capturing non-linear relationships. Implemented in R with libraries like randomForest, xgboost, and caret, RF utilized 500 trees and predictors set to the square root of total features for optimal performance (Breiman, 2001). XGBoost, configured with a depth of 6, learning rate of 0.1, and 100 boosting rounds, effectively modeled complex patterns (Chen & Guestrin, 2016). Models were trained on an 80/20 train-test split, using 20-day rolling standard deviation of log returns as the volatility target. Predictors included lagged returns, bid-ask prices, and technical indicators such as MACD and RSI (Hull, 2018). Evaluation metrics—MSE, R^2, VaR, ES, and violation rates—provided a comprehensive assessment. While RF delivered consistent baseline predictions, XGB excelled in accuracy and capturing extreme values, confirming its robustness for financial applications (Gu, Kelly & Xiu, 2020). Traditional methods like GARCH, despite parametric rigor, were excluded due to assumptions of linearity and stationarity (Bollerslev, 1986). Simpler statistical models lacked the sophistication required for capturing complex interactions. The inclusion of rolling volatility and technical indicators ensures methodological rigor, aligning with best practices in financial forecasting (Makridakis et al., 2020).

Results

Model Comparison Summary

Ticker	RF MSE	RF R^2	RF VaR	RF ES	RF Violation Rate	XGB MSE	XGB R^2	XGB VaR	XGB ES	XGB Violation Rate
AAPL	8.60e-06	0.991493	0.009431	0.008304	0.062718	4.45e-08	0.999956	0.009163	0.007806	0.051394
C	2.72e-05	0.980804	0.007484	0.006715	0.056620	8.28e-08	0.999942	0.007366	0.006352	0.053136
F	1.18e-06	0.991239	0.009820	0.008828	0.061847	5.44e-08	0.999598	0.009469	0.008378	0.052265

The **empirical analysis** evaluates the predictive capabilities of **Random Forest (RF)** and **XGBoost (XGB)** for forecasting financial volatility and assessing key risk metrics across three assets: **Apple (AAPL)**, **Citigroup (C)**, and **Ford (F)**. The performance of each model is summarized in the tables below, focusing on **Mean Squared Error (MSE)**, **R^2**, **Value at Risk (VaR)**, **Expected Shortfall (ES)**, and **Violation Rates**. **Random Forest (RF)** demonstrated robust performance, achieving **R^2 values exceeding 0.98** across all stocks, indicating its effectiveness in capturing the majority of variance. For instance, RF recorded **R^2 values of 0.9912 for Ford** and 0.9915 for Apple. However, RF's precision, as measured by **MSE**, varied across stocks; for example, Citigroup had a relatively higher **MSE (2.71 × 10)**, suggesting challenges in modeling more complex market dynamics. In terms of risk metrics, RF's **Violation Rates**, such as 0.0566 for Citigroup and 0.0627 for Apple, reflect its conservative approach but indicate room for improvement in adhering to confidence thresholds. **XGBoost (XGB)** consistently outperformed RF across all metrics. With near-perfect **R^2 values**, such as `0.9996` for Ford and `0.9999` for Citigroup, XGB demonstrated superior ability to model non-linear relationships and complex dynamics. Its **MSE** values, such as `4.45 × 10` for Apple and `5.44 × 10` for Ford, were significantly lower than RF's, reflecting greater precision. Additionally, XGB exhibited better adherence to confidence thresholds, with lower **Violation Rates**, such as `0.0531` for Citigroup and `0.0514` for Apple, suggesting its predictions were more reliable. These results align with findings in the literature, such as Gu, Kelly, and Xiu (2020), which highlight machine learning models' advantages in capturing intricate financial patterns. While RF serves as a robust baseline for variance modeling, XGB's superior performance underscores its potential as a leading tool for financial risk forecasting. The findings illustrate the trade-offs between machine learning models. RF provides strong baseline performance, particularly in capturing variance, but struggles with precision in some scenarios. XGB excels in both precision and risk metric adherence, making it well-suited for financial applications requiring high accuracy. Future research could investigate combining machine learning models with methods like **GARCH**, exploring the balance between interpretability and adaptability.

Analysis

This study demonstrates the practical applications of machine learning models, particularly XGBoost (XGB), in financial risk forecasting. XGB consistently outperformed Random Forest (RF) in predictive accuracy, achieving lower MSE, higher R^2, and more reliable risk metrics such as VaR and ES. These results have direct implications for various financial applications. In trading and risk management, XGB's lower violation rates enable improved loss prediction and dynamic hedging strategies. For investment fund risk management, its capacity to model non-linear relationships supports enhanced asset-level risk assessment and portfolio stability. In the context of regulatory compliance, XGB's conservative and accurate risk estimates align with Basel III standards, facilitating stress testing and automated risk monitoring. While machine learning models like RF and XGB excel in capturing non-linear dependencies and high-dimensional data, methods such as GARCH remain valuable for their interpretability and parametric rigor. However, the limitations of GARCH in dynamic and non-linear markets highlight the advantages of machine learning in modern financial applications. These findings align with prior research (Gu, Kelly, & Xiu, 2020) emphasizing machine learning's superior performance in financial forecasting while suggesting opportunities for hybrid approaches that combine the strengths of both traditional and machine learning models. Future research could refine these methodologies by integrating macroeconomic indicators, exploring advanced hybrid models, and conducting comparative analyses to clarify the optimal use of these approaches in trading, fund management, and regulatory compliance. These results offer a robust foundation for tailoring risk forecasting strategies to meet diverse financial objectives.

Conclusion

This project examined the effectiveness of Random Forest (RF) and XGBoost (XGB) in forecasting market volatility and assessing financial risks using historical data from Citigroup, Ford, and Apple. Both models demonstrated strong predictive capabilities, with XGB outperforming RF in terms of accuracy, evidenced by lower MSE, higher R^2, and superior risk metrics like VaR and ES. While RF proved reliable, its higher MSE indicated lower precision compared to XGB. The study highlighted the importance of tailored approaches, leveraging asset-specific characteristics and feature engineering, such as lagged returns and technical indicators, to enhance model performance. These findings reaffirm the strength of machine learning models in capturing non-linear relationships and handling dynamic market conditions. Future research could explore hybrid approaches that integrate traditional methods like GARCH with machine learning models, balancing interpretability with predictive power. Additionally, investigating advanced models such as LSTM networks or incorporating macroeconomic indicators could further enhance forecasting accuracy. This work underscores the value of machine learning in financial risk management, providing a foundation for developing adaptable methodologies for dynamic financial markets.

References

- Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31(3), 307–327.
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5–32.
- Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794.
- Gu, S., Kelly, B., & Xiu, D. (2020). Empirical Asset Pricing via Machine Learning. *The Review of Financial Studies*, 33(5), 2223–2273.
- Hull, J. C. (2018). *Options, Futures, and Other Derivatives* (10th ed.). Pearson Education.
- Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2020). The M4 Competition: 100,000 Time Series and 61 Forecasting Methods. *International Journal of Forecasting*, 36(1), 54–74.