Stock Volatility Prediction: A Comparative Analysis of Machine Learning Models

Moussaab OUALIJANE

September 4, 2024

1 Introduction

While watching the "Financial Markets" course on Coursera, provided by Yale University, a statement by the instructor piqued my curiosity. He mentioned how the high volatility of a stock like Apple made predicting its growth challenging and noted that a $40,000 investment could have grown to $15 million by 2021. This made me wonder: if I had been there, what should I have known, and how should I have treated market information to predict such volatility?

2 Apple Stock Overview

Apple Inc. (AAPL) is one of the most valuable and traded companies in the world, known for its innovation in technology and consumer electronics. Over the years, Apple's stock has experienced significant growth, but with that growth has come high volatility. This volatility makes predicting stock price movements particularly challenging.

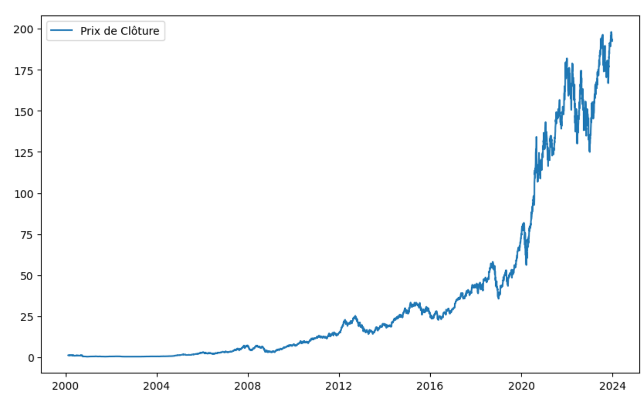

Figure 1: Historical variation of Apple stock price.

Figure 1 illustrates the historical price variation of Apple stock, highlighting periods of significant growth as well as volatility. This figure provides a visual foundation for understanding the complexities involved in predicting Apple's stock price movements, especially during periods of market instability.

3 Does intrepretable ML models work?

3.1 Model Performance

Model	MSE	R2 Score
Linear Regression	0.3897	-18.7173
Decision Tree	0.0352	-0.7788
Random Forest	0.0305	-0.5414
XGBoost	0.0289	-0.4639

Table 1: Performance comparison of machine learning models.

The performance metrics reveal interesting insights about our models:

- **XGBoost** demonstrates the best performance with the lowest Mean Squared Error (MSE) of 0.0289 and the highest R2 score of -0.4639.
- **Random Forest** follows closely, with an MSE of 0.0305 and R2 score of -0.5414.
- **Decision Tree** shows moderate performance.
- **Linear Regression** performs poorly, with a high MSE and extremely low R2 score, indicating it's not suitable for this complex prediction task.

It's important to note that all models have negative R2 scores, suggesting that they perform worse than a horizontal line (mean of the data). This indicates the challenging nature of predicting stock volatility and the need for further model refinement or feature engineering.

3.2 Prediction Visualization

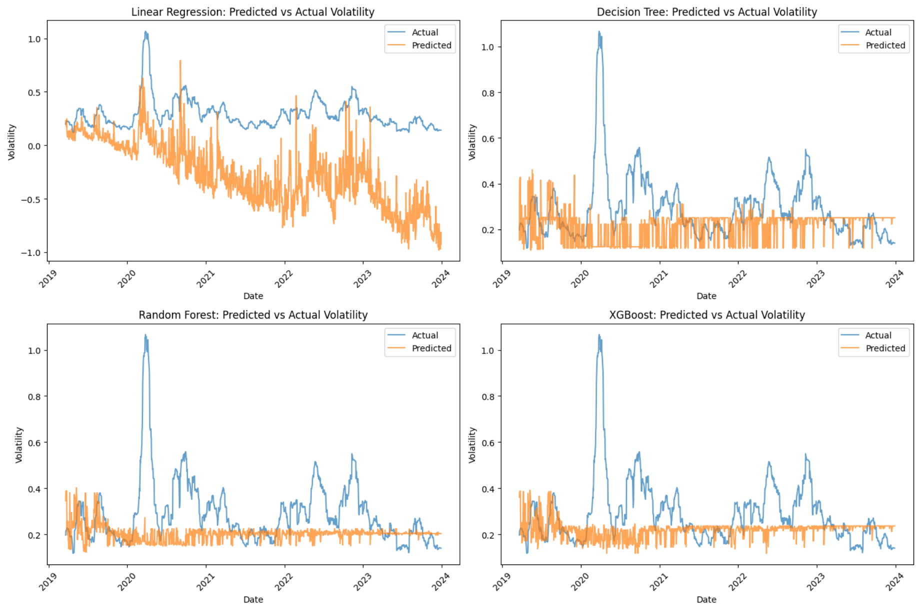

Figure 2: Predicted vs Actual Volatility for different models.

The visualization of predictions reveals the following patterns:

- **Linear Regression** shows poor performance, with predictions diverging significantly from actual values over time.
- **Decision Tree** exhibits a step-like pattern, indicating overfitting to certain value ranges.
- **Random Forest** and **XGBoost** demonstrate similar patterns, capturing the overall trend but missing extreme volatility spikes.
- All models struggle to predict the high volatility period in early 2020, likely corresponding to the COVID-19 market crash.

These visualizations corroborate the numerical performance metrics and highlight the challenges in predicting stock volatility, especially during extreme market events.

3.3 Feature Importance

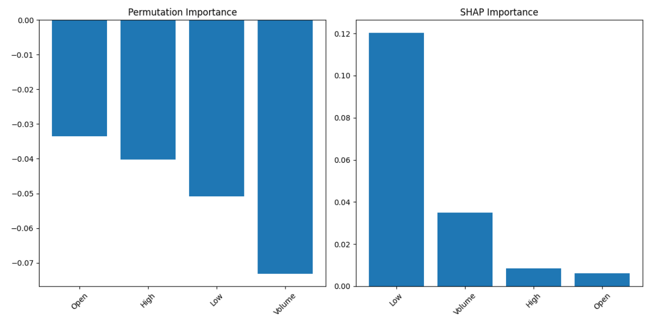

Figure 3: Feature importance analysis using Permutation and SHAP.

Figure 3 presents feature importance analysis using two methods:

- **Permutation Importance**:
  - All features show negative importance, indicating that permuting any feature improves the model's performance.
  - **"Volume"** has the highest negative impact, followed by **"Low"**, **"High"**, and **"Open"**.
- **SHAP Importance**:
  - **"Low"** price is the most important feature.
  - **"Volume"** is the second most important, followed by **"High"** and **"Open"**.

The negative permutation importance suggests potential overfitting or complex feature interactions. The consistency of "Low" price importance across methods indicates its significance in predicting volatility. The divergence in importance rankings between methods highlights the complexity of feature relationships in this prediction task.

Why the Discrepancy Between SHAP and Permutation Importance?

SHAP focuses on the marginal contributions of features across all possible scenarios, which often emphasizes features that consistently have a large impact on individual predictions. This explains why SHAP might identify "Low" price as the most important feature, as it likely plays a significant role across a wide range of predictions, particularly in capturing certain patterns in the stock's volatility.

Permutation Importance assesses the impact of a feature on the overall model accuracy. If "Volume" is identified as most important here, it suggests that changes in volume significantly affect the model's performance, perhaps because "Volume" directly relates to market activity, which drives price volatility.

4 Results

Our analysis reveals the challenges in predicting stock volatility:

- XGBoost and Random Forest models outperform simpler models, but all models struggle with accurate predictions, especially during extreme market events.
- The negative R2 scores across all models indicate the need for further refinement in our approach, possibly including additional features or alternative modeling techniques.
- Feature importance analysis suggests that the "Low" price and trading "Volume" are crucial factors in volatility prediction, but the complex interactions between features warrant further investigation.

5 Conclusion

The analysis highlights the significance of certain market indicators, like the "Low" price and trading "Volume," in predicting stock volatility. However, the intricate relationships between these factors and their non-linear interactions make the prediction process highly complex. This complexity suggests that no single piece of market information can be treated in isolation; instead, a holistic approach is necessary. To answer the question, "How should I have treated the market information?"—it's now evident that predicting volatility requires more than just observing individual metrics.

6 Limitations and Problems with future Work

- Exploring additional features, such as market sentiment or macroeconomic indicators.

  One of the significant challenges in further developing this analysis lies in incorporating market sentiment and macroeconomic indicators. Market sentiment analysis, particularly through social media, has become a powerful tool in recent years. However, applying it retrospectively is problematic. For instance, during the announcement of the first iPhone, social media platforms were either non-existent or in their infancy, making it difficult to gauge public sentiment accurately during that period. Furthermore, while macroeconomic indicators are often used to predict market trends, their application to Apple's performance is somewhat controversial. Notably, Apple managed to thrive during the 2008 financial crisis, a period when many other companies struggled. This counterintuitive success suggests that traditional economic indicators may not fully capture the unique factors driving Apple's stock performance, making it challenging to integrate these variables effectively into our predictive models

- Investigating the use of time series-specific models or hybrid approaches.

  While exploring advanced time series models could potentially improve predictive accuracy, we've already demonstrated the non-linearity in volatility variation, which presents a significant challenge. Even if such models perform well, they often become highly complex and difficult to interpret. This complexity makes it challenging to answer the initial question of how I should have treated market information to predict volatility. The advanced nature of these models may obscure the very insights we're seeking to uncover, limiting their practical utility in providing actionable market analysis.

- Considering different time horizons for volatility prediction to capture both short-term and long-term patterns.

  Adjusting the time horizon for volatility prediction could offer valuable insights, allowing us to differentiate between short-term fluctuations and long-term trends. However, this approach is not without its challenges. For instance, if I had focused solely on the first two years after Apple's release, I would have observed a period of declining stock prices. This short-term perspective might have led to a misleading conclusion about Apple's future performance, missing the substantial long-term growth that followed. Therefore, while varying the time horizon may help, it also introduces the risk of drawing inaccurate conclusions based on limited data, underscoring the need for a balanced and comprehensive approach.

So in your opinion, what's the next Apple that we'll be asking ourselves, 'What should we have seen or understood to predict it right?