Volatility Forecasting in Financial Market

Author: Thao Nguyen

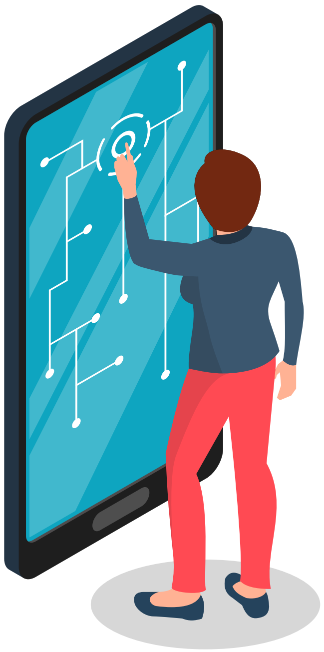


TABLE OF CONTENTS

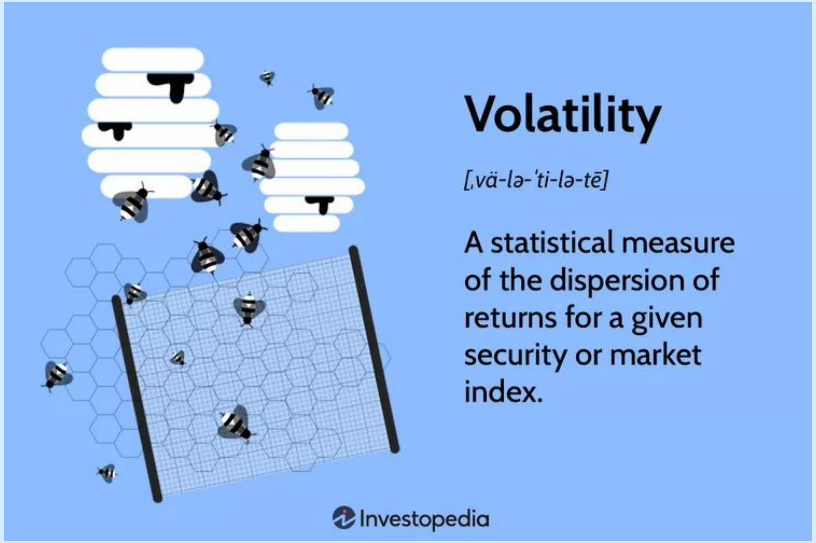
01 Overview

About Project

This project focuses on developing a robust framework for estimating stock market volatility using the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model.

Volatility?

Volatility, the degree of variation in stock prices over time, plays a crucial role in financial analysis, influencing decision-making processes such as risk assessment, portfolio optimization, and option pricing.



Dataset

The dataset comprises the S&P 500 index data spanning a 5-day week, beginning from April 6, 2009, through December 31, 2020

- Open: The price when trading starts for the day
- High: The highest price during the trading day - Low: The lowest price during the trading day
- Close: The final price at the end of the day
- Adj Close (Adjusted Close): The closing price adjusted for any corporate actions that may affect the stock price

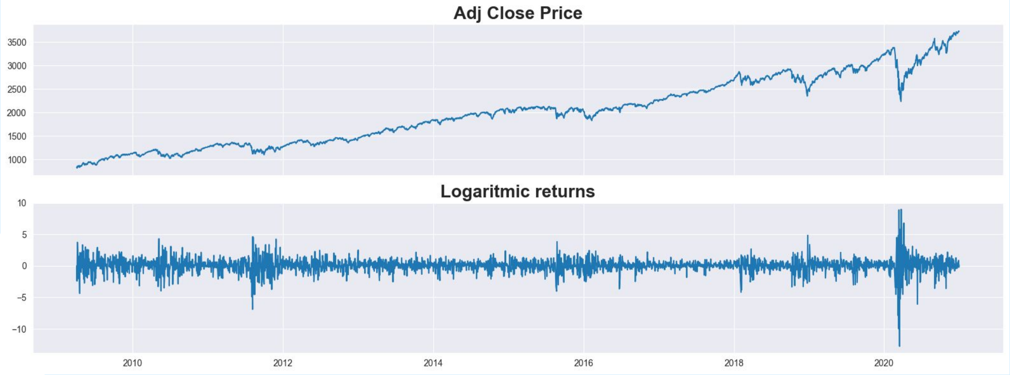
For historical returns or analysing past performance, the Adjusted Close is frequently relied upon. Consequently, in this project, i will consistently incorporate the Adj Close data

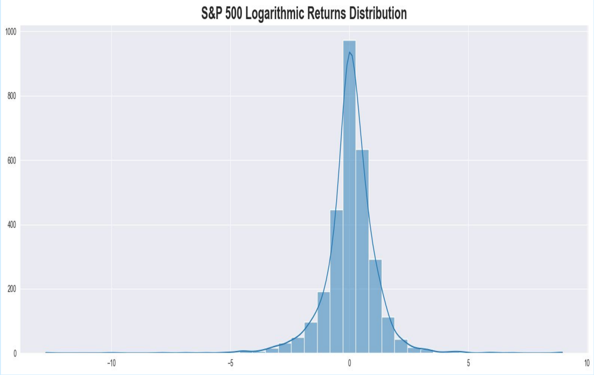

S&P 500: Adjusted Close Price
3500
3000
2500
2000
1500
1000
2010
2012
2014
2016
2018
2020
Time

02 Implementation

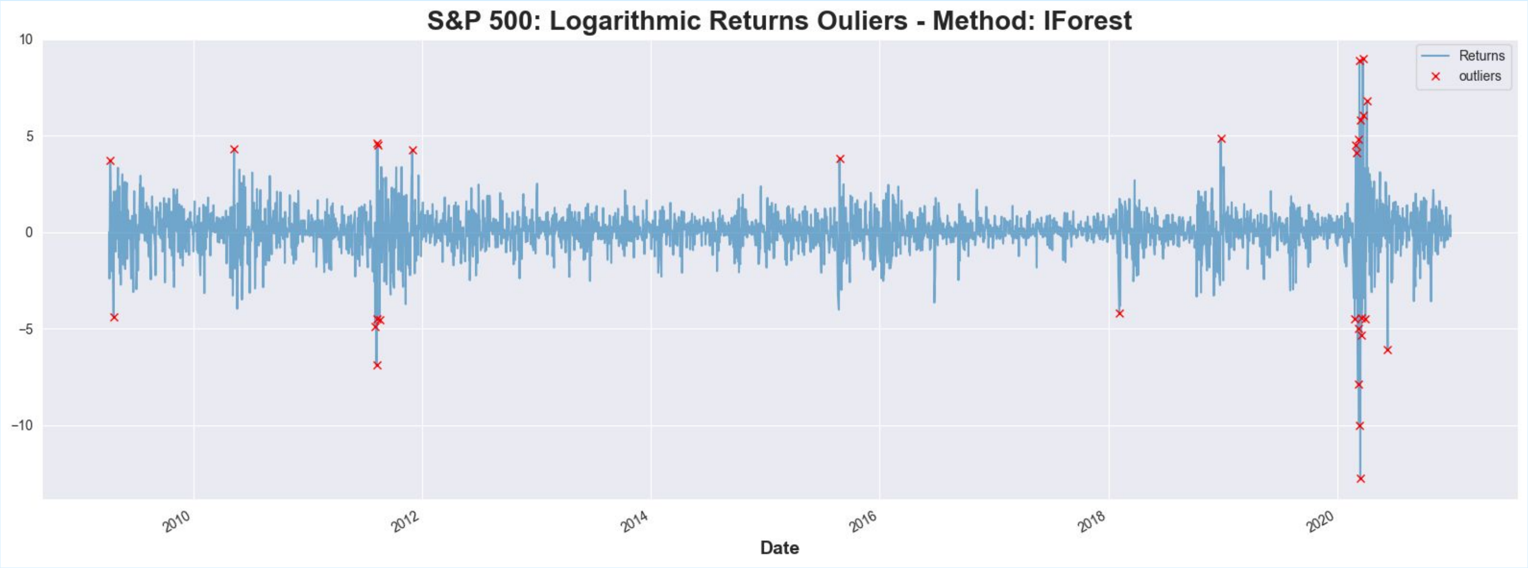
Returns, Logarithmic Returns

- Convert stock price to returns: Returns quantify growth rate or change in the value of a financial asset over time
- Using Daily Logarithmic returns instead of Daily simple returns :
  - Additive
  - Normal Distribution Assumption



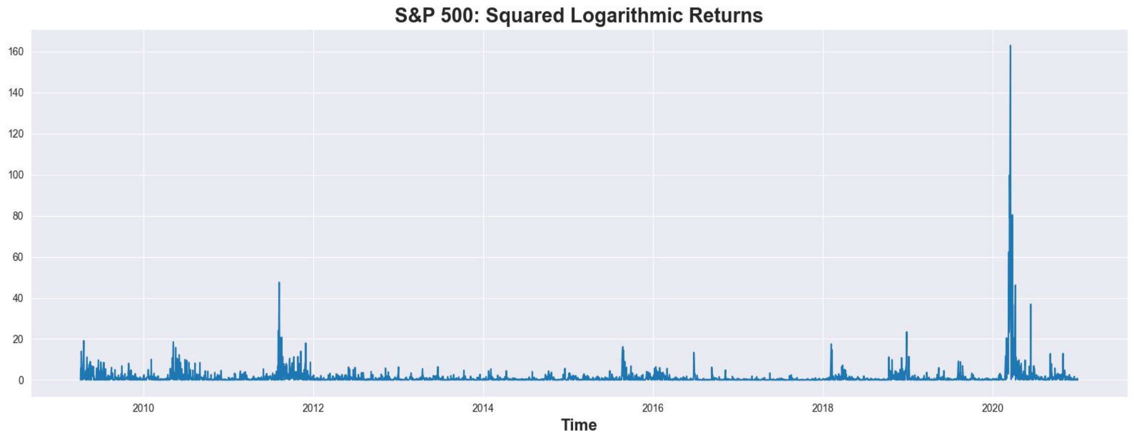


Basic Outlier Detection



Square Returns

Square returns give us the general idea about volatility. High values of squared returns indicate higher volatility or risk in the underlying asset.



The returns exhibit significant volatility particularly in 2020 and 2012

Methods for computing volatility

Close-to-close

Garman Klass Volatility

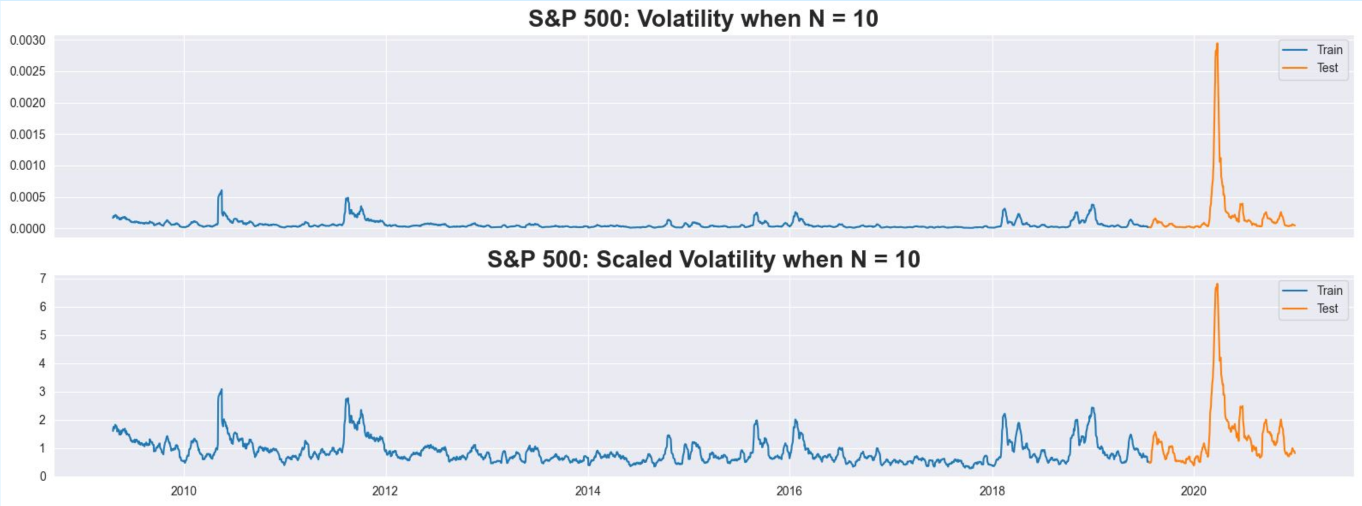
Parkinson Volatility

GKYZ Volatility

This estimator can address most aspects of realized price dynamic and can handle intraday dynamic, as well as overnight jump. Hence, it is considered the most efficient estimator.

The Window Option

$$\sigma_{GKYZ} = \sqrt{\frac{F * 252}{N}} \sqrt{\sum_{i=1}^{N} \left[Ln(\frac{o_i}{c_{i-1}})^2 + \frac{1}{2}\left(Ln\left(\frac{h_i}{l_i}\right)\right)^2 - (2Ln(2) - 1)\left(Ln\left(\frac{c_i}{o_i}\right)\right)^2 \right]}$$

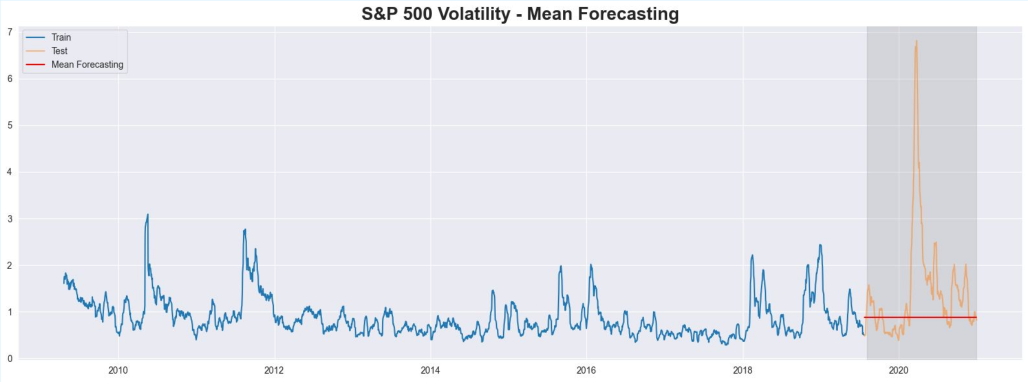


We need to scale the Volatility estimator to match the magnitude of the volatility estimates with the ones produced from Garch model. Here is a formula proposed by Fiszeder (2005,2009):

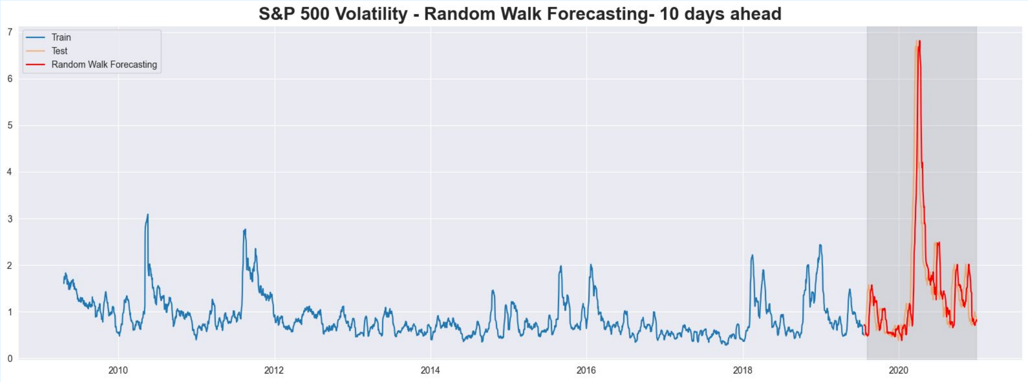
$$Scaled\ \sigma^{GKYZ} = \frac{a}{b} \cdot \sigma^{GKYZ}$$

$$a = \sqrt{\frac{1}{T}\sum_{t=1}^{T} r_t^{\ 2}}, b = \sqrt{\frac{1}{T}\sum_{t=1}^{T} \sigma_t^{GKYZ,2}}$$

Baseline models



	Model	MSE	MAE
0	Mean	1.716823	0.738683



Model	MSE	MAE
Naive Forecasting	0.841333	0.571458

GARCH Model

The GARCH model, which stands for Generalized Autoregressive Conditional Heteroskedasticity. -

- Autoregressive: means that the current value of a variable is influenced by past values of itself at different periods
- Heteroskedasticity: means that the model may have different magnitudes or variability at different time points (variance changes over time).
- Conditional: since the volatility is not fixed, the reference here is on the constant that we place in the model to limit heteroskedasticity and make it conditionally dependent on the previous value or values of the variable.

	Model	MSE	MAE
0	Naive Forecasting	0.8413	0.5715
1	Garch (Constant Mean, Skew Students Distribution)	0.07454	0.1816

GARCH VARIANTS' PERFORMANCE

	Model	MSE	MAE
0	Naive Forecasting	0.8413	0.5715
1	Garch (Constant Mean, Skew Students Distribution)	0.07454	0.1816
2	GJR-GARCH model	0.2117	0.2868
3	EGARCH model	0.4574	0.354
4	APARCH model	0.1916	0.2772

Conclusion

In this project, I have used GKYZ estimator for estimating volatility, which serves as a benchmark for this project.Through the implementation of the GARCH model, we have achieved superior performance regarding predicting stock volatility compared to various other variants within the GARCH model family, including APARCH and EGARCH.

Reflection and Next Steps

- This project signals the initial step towards advancing volatility forecasting methodologies.
- Although the GARCH model has demonstrated effectiveness, opportunities for enhancement and fine-tuning abound.
- Future progress could involve integrating deep learning methods into our volatility estimation framework.
- While currently in the learning phase, I am enthusiastic about exploring how neural networks can enhance our understanding of intricate patterns and nonlinear relationships within financial time series data.

REFERENCES

ATWAN, T. A. (2024). Time series analysis with Python cookbook: Practical recipes for exploratory data analysis... , data preparation, forecasting, and model evaluat. PACKT PUBLISHING LIMITED.

Peixeiro, M. (2022). Time series forecasting in Python. Manning Publications Co. Hyndman, R. J., & Athanasopoulos, G. (2021). Forecasting: Principles and practice. OTexts.

Feasel, K. (2022). Finding ghosts in your data anomaly detection techniques with examples in Python. Apress. Michańków, J., Kwiatkowski, Ł., & Morajda, J. (2023). Combining deep learning and GARCH models for financial volatility and risk forecasting. SSRN Electronic Journal.https://doi.org/10.2139/ssrn.4589950

https://arch.readthedocs.io/en/latest/

https://medium.com/swlh/the-realized-volatility-puzzle-588a74ab3896

https://www.topstep.com/blog/implied-vs-realized-volatility-the-vix/#:~:text=%E2%80%9CImplied%20volatility%20represents%20the%20value,for%20global%20derivatives%20at%20Cboe.

https://dynamiproject.files.wordpress.com/2016/01/measuring_historic_volatility.pdf

https://github.com/chibui191/bitcoin_volatility_forecasting

THANK YOU SO MUCH FOR YOUR TIME !

Further Details:: official Notebook