Historical Perspectives in Volatility Forecasting Methods with Machine Learning

Zhiang Qiu[1], Clemens Kownatzki, Ph.D[2], Fabien Scalzo, Ph.D[1], Eun Sang Cha Ph.D[1]

1. Seaver College, Pepperdine University, Malibu, CA 90263. 2. Graziadio Business School, Pepperdine University, Malibu, CA 90263

Abstract

- Volatility forecasting in the financial market plays a pivotal role across a spectrum of disciplines, such as risk management, option pricing, and market making. However, volatility forecasting is challenging because volatility can only be estimated, and different factors influence volatility, ranging from macroeconomic indicators to investor sentiments. While recent works suggest advances in machine learning and artificial intelligence for volatility forecasting, a comprehensive benchmark of current statistical and learning-based methods for such purposes is lacking. Thus, this paper aims to provide a comprehensive survey of the historical evolution of volatility forecasting with a comparative benchmark of key landmark models. We open-source our benchmark code to further research in learning-based methods for volatility forecasting.

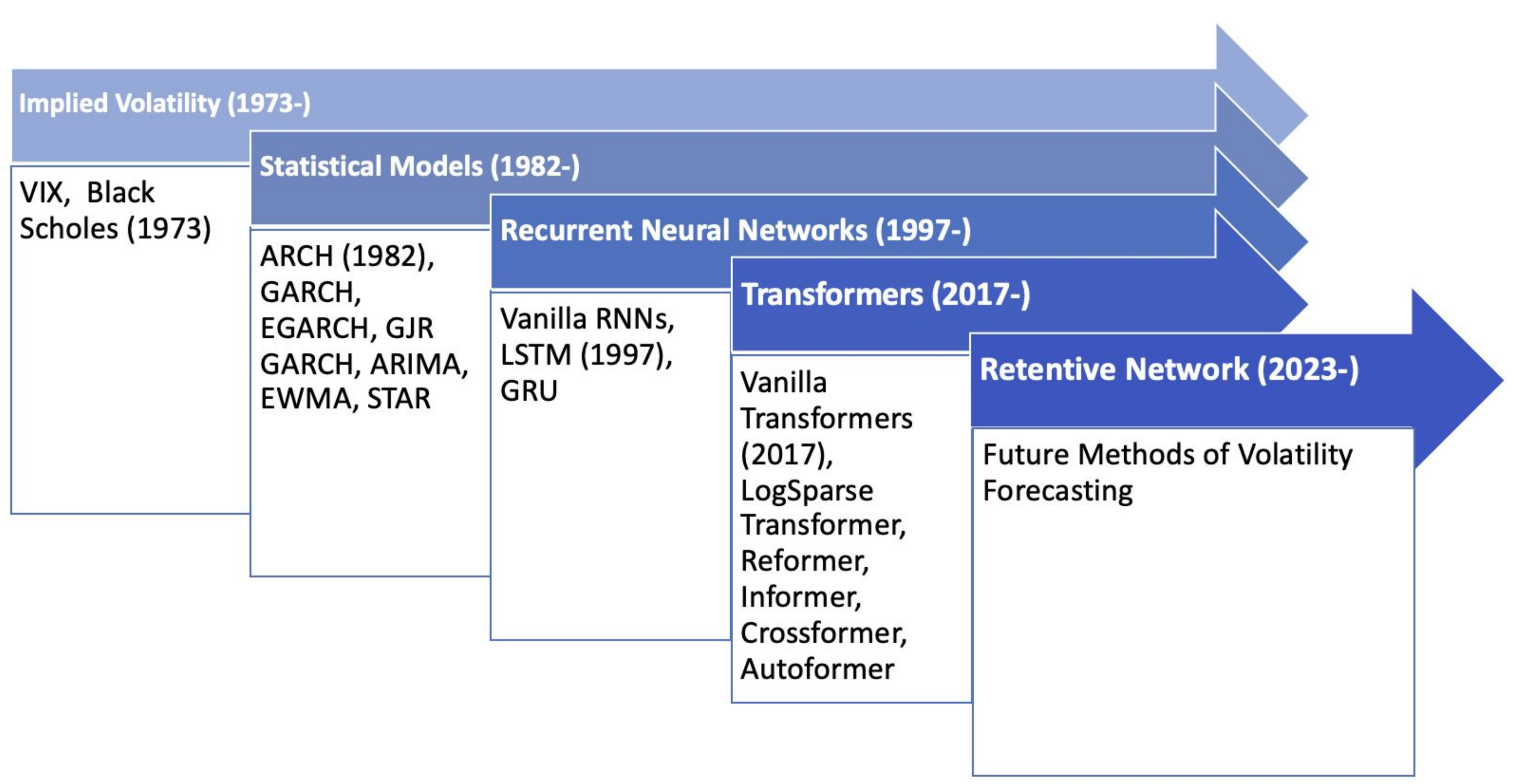
KEY WORDS

- volatility forecasting, risk management, deep learning, time series analysis, GARCH, LSTM, transformer

Contributions

- 1. We survey the evolution of volatility forecasting models, transitioning from traditional AR models to contemporary variations of the Transformer models, which represent the current state-of-the-art.
- 2. We select a representative model from each category and conduct a comparative benchmark to show their respective performances, paving the way for subsequent model developments.
- 3. We open-source our benchmarks and comprehensively analyze the advantages and disadvantages inherent to each model type.

Timeline of Model Evolution



Methods

- We will discuss four milestone models: GARCH, Implied Volatility, Long Short Term Memory, and Transformer.
- We used thirty years of S&P 500 data from Yahoo Finance, October 1, 1993, to October 1, 2023 (the first twenty- seven years for training and the last three years for testing). We calculated the realized volatility as the standard deviation of 22 rolling trading days (as an approximation for one month in time)'s log return. To compute for errors, we used both root mean square error (RMSE) and root mean square percentage error (RMSPE).

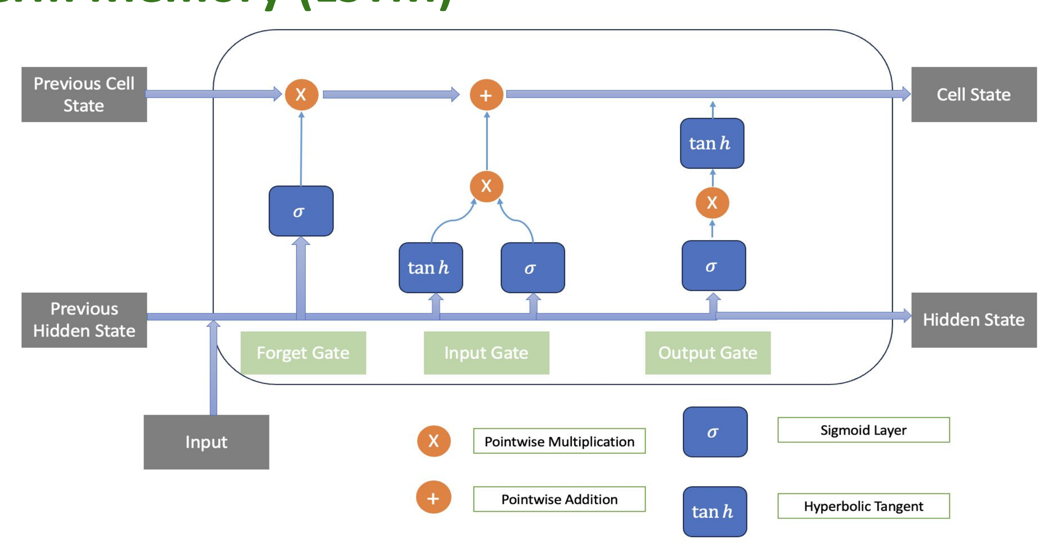
Generalized Autoregressive Conditional Heteroskedasticity (GARCH)

- The ARCH model was introduced prior to the GARCH model for forecasting volatility. The name,ARCH means that volatility depends on time series value in previous periods and some error term.
- GARCH is a variant of the ARCH model that addresses the problem of predictions being bursty, which means the prediction can vary by a huge amount day by day. This enhancement is achieved by incorporating the previous day's volatility into the current day's calculation, alongside the ARCH model's time series value and error term. Therefore, the resulting predictions tend to be more stable, given that today's volatility is likely to mirror the previous day due to its inclusion in the equation.

Implied Volatility

- What distinguishes IV from other models is that it is forward-looking. Implied volatility captures the market's expectation of the volatility for the next 22 days, calculated backward from the option's price using the Black Scholes Merton Formula.

Long Short Term Memory (LSTM)



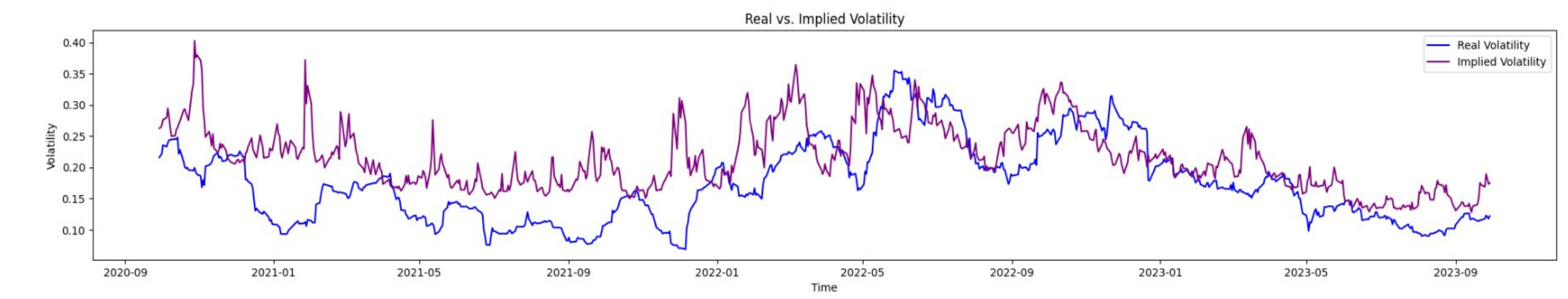
Transformers

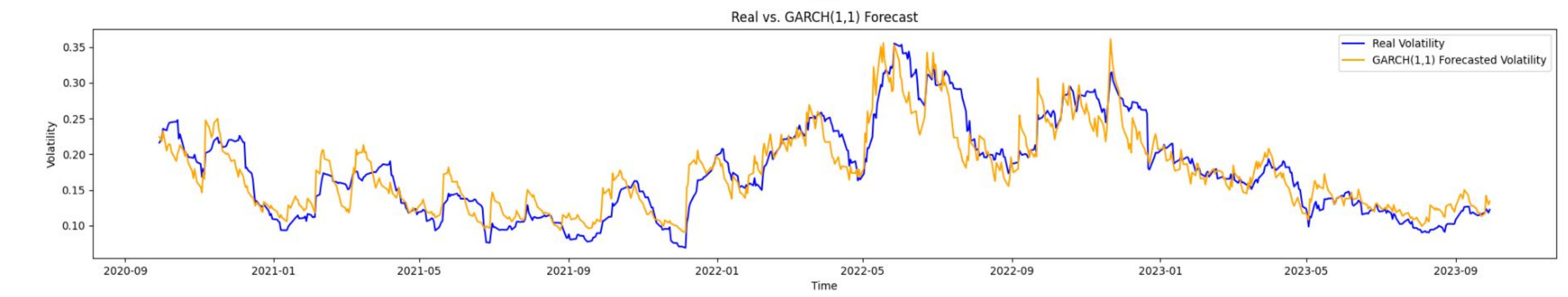
- Transformers abandon recurrence entirely and use attention mechanisms instead. Transformers have an encoder-decoder structure. The encoder maps the input sequence and produces a continuous representation, and the decoder chooses what and how much previously encoded information to access. The encoder's attention mechanism derives attention scores from input vectors of queries, keys, and values. These scores determine the weight of each piece of information in predicting every time step.

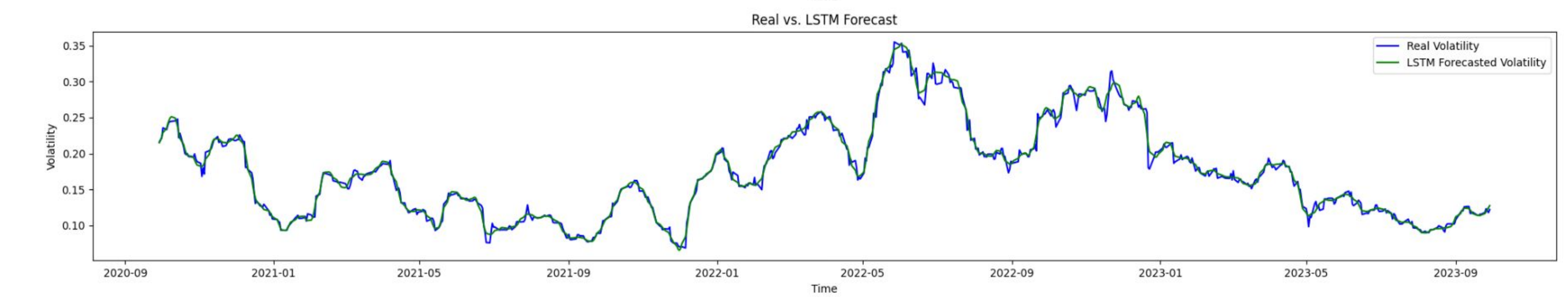
Results

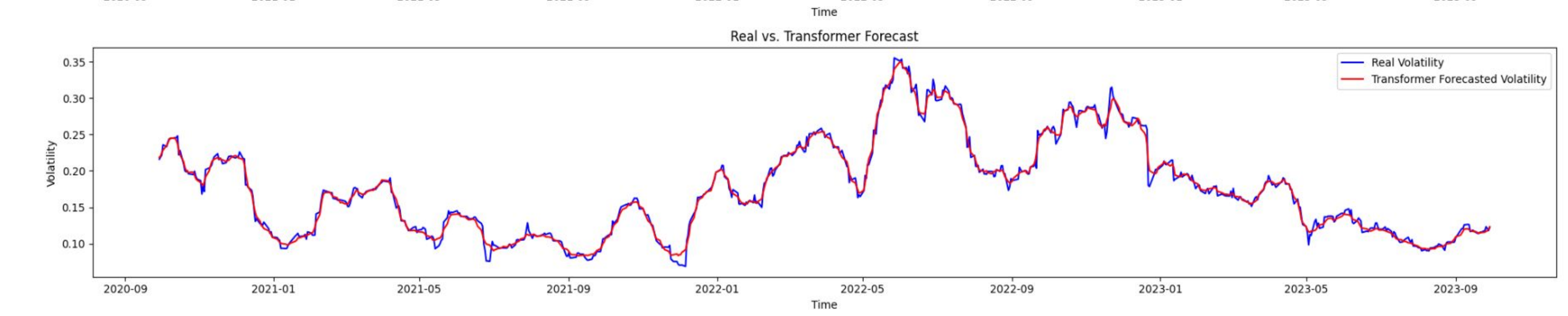
- In the thirty years of S&P 500 data we used to train and test our models, two-layered LSTM performs the best, followed by Transformer.

Model Name	RMSPE	RMSE
GARCH (1,1)	0.151825	0.024772
Implied Volatility	0.599153	0.069205
2-layered LSTM	0.032105	0.005744
Transformer	0.041281	0.005883









Conclusion

- Volatility forecasting is essential for various disciplines, such as risk management, derivatives pricing, and market making. A benchmark comparing the efficacy and performance of different forecasting techniques can be beneficial. While some existing literature reviews focus on particular forecasting models, like GARCH, there remains a gap for a holistic and up-to-date benchmark. Our study offers a comprehensive review of volatility forecasting methods, ranging from traditional models to the current SOTA. We summarize the advantages and disadvantages inherent to each group of models. This research consolidates insights from model-specific reviews, presenting the evolution of volatility forecasting methods in a structured manner.

Acknowledgements

- This work was supported in part by the Keck Institute Undergraduate Research Grant and the Dean's Research Grant at Pepperdine University.