Published by :
http://www.ijert.org
International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 11 Issue 05, May-2022

Volatility Prediction

Muskaan Sinha
Dept. of Computer Science & Engineering
Dayananda Sagar University, India

Muskaan Goel
Dept. of Computer Science & Engineering
Dayananda Sagar University, India

Mohammed Kashif
Dept. of Computer Science & Engineering
Dayananda Sagar University, India

Lysetti Lakshmi Poojitha
Dept. of Computer Science & Engineering
Dayananda Sagar University, India

Sayeeda Umaima Sadafunnisa
Dept. of Computer Science & Engineering
Dayananda Sagar University, India

Geetha A ,
Assistant Professor
Dept. of Computer Science & Engineering
Dayananda Sagar University, India

Abstract: **In finance , volatility refers to the variation degree of the asset price, and it measures the uncertainty of the price. It plays an important role in both academic research and the financial industry. In risk management and performance measurement, volatility is a risk indicator itself and can be a part of some other indicators, like the Sharpe ratio. In portfolio theory, Markowitz used volatility to measure the risks of assets and the overall risk of the portfolio. Volatility is both the input and the optimisation target of the portfolio construction model. In derivative pricing , prices of derivatives can be determined by the volatility of the underlying assets. This paper would aim to help trading markets , stock markets to predict volatility beforehand and take measures with respect to trade. Volatility is a measure of how prices or returns are scattered over time for a particular asset or financial product. It is a key metric because volatility creates profit potential. However , trading on volatility can also create losses , if traders do not learn the appropriate information and strategies.**

INDEX TERMS:- Volatility , prediction, asset price, LSTM ,Neural Networks, Garch Model, Arch Model, finance, stock prices, Market analysis, uncertainty,risk management.

INTRODUCTION

Volatility in stock markets evokes varying responses from market participants. While some perceive it as an opportunity to make money , others perceive it as a threat and start unwinding their positions. While affecting portfolio choice, changes in stock market volatility also gives some idea about the current economic state. In today's globalized environment , increased volatility means global uncertainty.

The effect of stock market volatility can be six fold :
(i) it enhances the chance of profit making opportunities from intraday trading for stock market traders.
(ii) it leads to portfolio rebalancing by fund managers.
(iii) it increases volatility trading in the options market
(iv) it increases hedging activities in the financial market
(v) it does influence policy makers in taking difficult decisions
(vi) it affects capital formation, as volatile markets are not conducive for fresh equity issues in the market.

In today's globalized world, as financial integration and enhanced trade in goods and services,volatility in one country spreads to other countries almost immediately. In India, where foreign institutional investors are large players in the stock market, their fund allocation is shaped by macroeconomic conditions in other economies. Generally, when the stock market becomes volatile, there is a tendency for gold prices to rise. It is considered to be a safe asset and hence there is a tendency to substitute stocks with gold. Thus volatility in gold prices is also a reflection of volatility in stock markets.

Accurate forecasting of market variables is critical to economists, analysts, and investors. This task gets complex as world financial markets get increasingly interconnected and interdependent. This complexity has created opportunities for neural networks which have the ability to explore interrelationships among a large number of market variables. Hence they are gaining popularity.

Neural networks have gained use in economics and finance more recently. The networks have been used in issues like economic prediction, stock picking, portfolio construction, identifying insider trading, analyzing corporate financial health, bond risk assessment, recognizing financial distress, detecting credit card fraud, improving real estate appraisal, identifying good credit or insurance candidates, exchange rate prediction, valuing options, commodity trading.

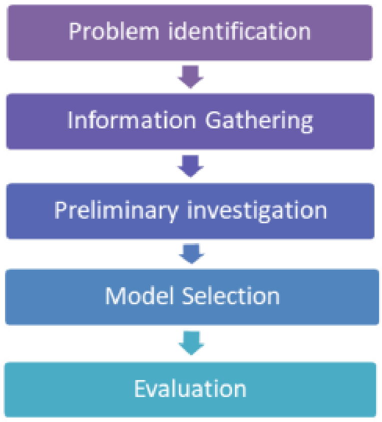
Forecasting is an important problem but with vital importance in all areas of the real world like business and industry, medicine, social science, politics, finance, government, economics, environmental sciences and others. In recent years, with the rise of social media and other promising applications, stock market forecasting has attracted huge interest from people in general and business in particular. Advances in financial sectors are responsible for growth and stability of overall economy

The purpose of this paper is to develop a framework for forecasting volatility in the Indian stock market.

- **THE FORECASTING PROCESS**

A forecast refers to scientific prediction or estimation. Forecasting is a technique where historical data is taken as input to make informed estimates that determine the direction of future trends. .

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 11 Issue 05, May-2022



Step 1: Problem Identification

It is important to identify who needs these predictions which we are going to make and who will benefit from these predictions. We also need to know in what ways these forecasts will be used and in what way the forecasting function fits inside the organization that requires the forecasts.

Step 2: Information Gathering
We determine how the data is assembled. The data can be of two types - Primary and Secondary Data. Primary data does not have any previous existence and is collected directly from the respondents. Primary data is considered very important , though it is not very reliable. Secondary data is historical data.

Step 3: Preliminary Analysis
In this step we check whether the data collected is useful or not. This analysis also helps us in revealing a pattern which ultimately helps us in choosing the model to fit in. Any repeated data is removed in this process.

Step 4: Choosing and Fitting of Models
Once the model is fitted , we need to choose the prediction model which gives us the best results.

Step 5: Usage and Evaluation of Forecasting Model
Once the prediction model is selected and different parameters have been set , the model is then put into practice for making predictions.

- **TYPES OF VOLATILITY AND THEIR MEASUREMENT**

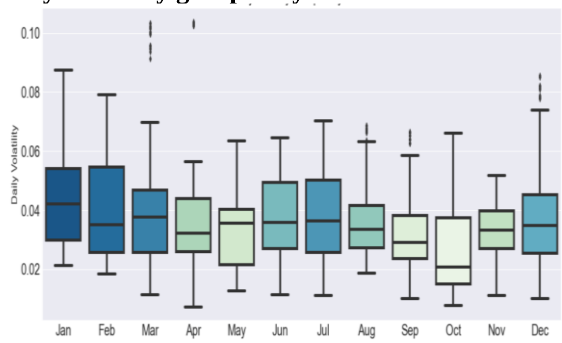
Volatility attempts to measure the magnitude of price movements that a financial instrument experiences over a certain period of time. The more dramatic the price swings are in that instrument, the higher the level of volatility, and vice versa. “Volatility does not measure the direction of price changes, merely their dispersion. This is because when calculating standard deviation (or variance), all differences are squared, so that negative and positive differences are combined into one quantity. Two instruments with different volatilities may have the same expected return, but the instrument with higher volatility will have larger swings in values over a given period of time.” Volatility can either be historical or implied; both are usually expressed in percentage terms.

Historical Volatility (HV) or Realized Volatility is the actual volatility demonstrated by the underlying over a period of time, such as the past month or year. Realized Volatility is commonly calculated as the standard deviation of price returns, which is the dollar change in price as a percentage of previous day's price.

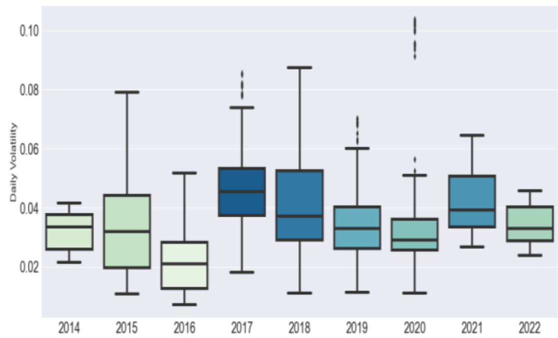
Implied volatility (IV), on the other hand, is the level of volatility of the underlying that is implied by the current option price.

- **EXPLORATORY DATA ANALYSIS**

Daily Volatility grouped by Month



Daily Volatility grouped by Year



- **MODELING**

1) Mathematical Model : In this model we do performance metrics using RMPSE(Root Mean Squared percentage error) and RMSE(Root Mean Squared Error) with RMSPE prioritized.

Usually with financial time series, if we just shift through the historic data trying different methods, parameters and timescales, it's almost certain to find some strategy with in-sample profitability at some point. However the whole purpose of "forecasting" is to predict the future based on currently available information, and a model that performs best on training data might not be the best when it comes to out-of-sample generalization (or overfitting). Avoiding/Minimizing overfitting is even more important in

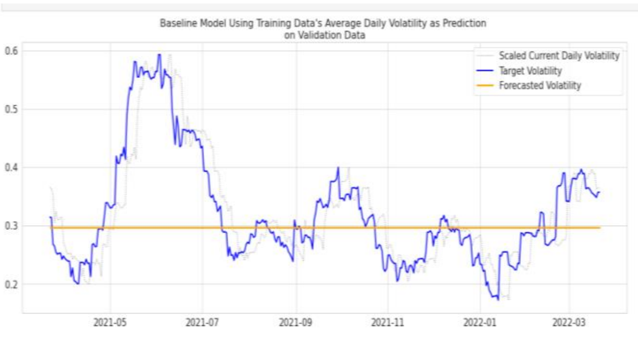
International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 11 Issue 05, May-2022

the constantly evolving financial markets where the stake is high.

2)Time Series Model: Time series analysis includes many categories or variations of data, analysts sometimes must make complex models. However, analysts can't account for all variances, and they can't generalize a specific model to every sample. Models that are too complex or that try to do too many things can lead to a lack of fit. Lack of fit or overfitting models lead to those models not distinguishing between random error and true relationships, leaving analysis skewed and forecasts incorrect.

3)Baseline Models: A baseline model is essentially a simple model that acts as a reference in a machine learning project. Its main function is to contextualize the results of trained models. Baseline models usually lack complexity and may have little predictive power. Regardless, their inclusion is a necessity for many reasons.

a) Mean Baseline Model- One of the essential characteristics of Volatility is its mean-revert over the long term. Therefore my first baseline model his would be a very simple one that only outputs the average current realized volatility of the whole training set as predictions everything. We calculate the mean of the scaled training data. We then create a series of prediction for the baseline model based on the validation data set.



b) Random Walk Naive Forecasting- A commonly known fact about volatility is that it tends to be autocorrelated, and clusters in the short-term. This property can be used to implement a naive model that just "predicts" future volatility by using whatever the daily volatility was at the immediate previous time step. We use previous n_future days volatility and plot predictions vs target values on validation set. We then append the metrics output to dataframe.

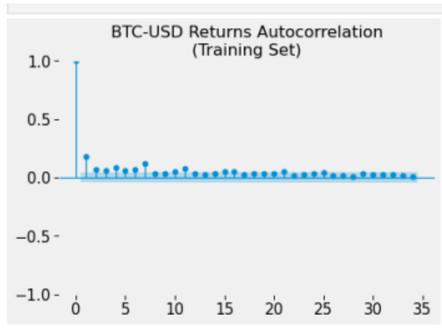
4) Statistical Model: A statistical model is a mathematical model that embodies a set of statistical assumptions concerning the generation of sample data. A statistical model represents, often in considerably idealized form, the data-generating process.

a) Garch Model- It stands for "Generalized Autoregressive Conditional Heteroskedasticity". It is an extension of the ARCH model. It includes lag variance terms with lag residual errors from the mean process calculated in the above models. Mathematically , GARCH can be represented as:

$$\sigma_t^2 = \omega + \sum_i^q \alpha_i \epsilon_{t-i}^2 + \sum_1^p \beta_i \sigma_{t-i}^2$$

i)Basic Garch: We visualize autocorrelation of squared returns and we also visualize partial autocorrelation of squared returns. The plot seems to indicate that there is only significant correlations upto the 7th lags, the ones other than that doesn't seem significant.

ii)



GARCH model with Asymmetric Shock Response- The basic garch model assumes that positive or negative news have similar impact on volatility. In reality the market tends to change in an asymmetric way and negative news impacts the market more than the positive ones. GJR-Garch accounts for all the asymmetry of shock responses.

b) Tarch Model- It is another member in the GARCH family called TARCH(threshold autoregressive conditional heteroskedasticity). TARCH models the volatility using absolute values. This model is specified using power = 1.0, but its default value is 2.0. In this model, we first set the seed for reproducibility. Later we get volatility scaler and scaled conditional volatility from the model result.

There are three parts of TARCH Modeling:

i) Bootstrap-based Forecasting for TARCH

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 11 Issue 05, May-2022

Ii ii)Simulation Based Forecasting for TARCH

iii)Hyperparameter Tuning for TARCH

5) Neural Networks: GARCH model is the gold standard for volatility prediction without traditional financial institutions, there has been an increasing number of professionals and researchers turning to Machine Learning, especially Neural Networks, to gain insights into the market in recent years.

a) LSTM- It is translated for long short term memory and its an artificial neural network used in fields of artificial intelligence and deep learning. LSTM networks are also suitable for classification, processing and making predictions based on time series data.

i) univariate LSTM: A part of neural networks which can predict "the future" is RNN. This works best for time series data, like stock prices.

- Bidirectional LSTM

ii) multivariate LSTM: The first multivariate model would be relatively simple within 2 layers of bidirectional LSTM. However, having more features would mean the model would be prone to overfitting. We set the seed for reproducibility after which multivariate bidirectional LSTM neural network is constructed. The further process is followed by layering, fitting, visualising , forecasting and plotting predictions vs target values on the validation set.

We also try to forecast the predictibility using 3 layers and 4 layers multivariate LSTM.

RESULTS

Final Model : We use the same architecture as multivariate LSTM with 2 layers , a lookback window of 30 days and a batch size of 64 days. We first create the dataset that combines training and validation set. Then a multivariate bidirectional LSTM is created and layers are added. The training is stopped if validation RMSPE is not improved by the end.

CONCLUSION

The main contribution of this paper is to find out the best model for forecasting volatility. This is done by comparing different models, which include mathematical models, time-series models, statistical models and neural network models. On doing so it was found out that multivariate LSTM with 2 layers is considered to be the best model, because the forecasted volatility is almost similar to the targeted volatility. This is certainly an improvement compared to all the previous models.

REFERENCES

[1] Jacopo De Stefani, Olivier Caelen1 , Dalila Hattab2 , and Gianluca Bontempi Machine Learning Group,Departement d'Informatique, Universit´e Libre de Bruxelles, Boulevard du Triomphe CP 212, 1050 Brussels, Belgium Worldline SA/NV R&D, Bruxelles, Belgium 1 Equens Worldline R&D, Lille (Seclin), France2{jacopo.de.stefani,gianluca.bontempi}@ulb.ac.be olivier.caelen@worldline.com dalila.hattab@equensworldline.com

[2] Machine Learning for Realized Volatility Forecasting.Machine Learning for Realized Volatility Forecasting Rahimikia and Ser-Huang PoonThis revision: November 10, 2021.

[3] Christiansen, C., Schmeling, M., & Schrimpf, A. A comprehensive look at financial volatility prediction by economic variables, Journal of Applied Econometrics, Vol 27, no 6, pp 956-977, August 2012

[4] Tamal Datta Chaudhuri Principal, Calcutta Business School, Diamond Harbor Road, Bishnupur – 743503, 24 Parganas (South), West Bengal Indranil Ghosh Assistant Professor, Calcutta Business School, Diamond Harbor Road, Bishnupur – 743503, 24 Parganas (South), West Bengal" Forecasting Volatility in Indian Stock Market using Artificial Neural Network with Multiple Inputs and Outputs" International Journal of Computer Applications (0975 – 8887) Volume 120 – No.8, June 2015

[5] Bhattacharya, S., & Ahmed, A. Forecasting crude oil price volatility in India using a hybrid ANN-GARCH model. International Journal of Business Forecasting and Marketing Intelligence, Vol. 4, pp 446-457, 2018

[6] M. F. Anaghi and Y. Norouzi, ''A model for stock price forecasting based on ARMA systems,'' in Proc. 2nd Int. Conf. Adv. Comput. Tools Eng. Appl. (ACTEA), Dec. 2012, pp. 265–268.

[7] Maqsood, Arfa & Safdar, Suboohi & Shafi, Rafia & Lelit, Ntato., Modeling Stock Market Volatility Using GARCH Models: A Case Study of Nairobi Securities Exchange (NSE). Open Journal of Statistics. Vol 07, pp 369-381, April 2017

[8] F.-M. Tseng, G.-H. Tzeng, H.-C. Yu, and B. J. C. Yuan, ''Fuzzy ARIMA model for forecasting the foreign exchange market,'' Fuzzy Sets Syst., vol. 118, no. 1, pp. 9–19, 2001.

[9] Christiansen, C., Schmeling, M., & Schrimpf, A. A comprehensive look at financial volatility prediction by economic variables, Journal of Applied Econometrics, Vol 27, no 6, pp 956-977, August 2012

[10] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[11] Wang, J. The Impact of Attention on Stock Returns: An Empirical Study of China's Securities Market. Master's Dissertation, Shanghai Jiaotong University, Shanghai, China, 2012.

[12] Sheikh Mohammad Idrees Department of Computer Science and Engineering, Jamia Hamdard, New Delhi, India. M. Afs. Parul Agarwal Department of Computer Science and Engineering, Jamia Hamdard, New Delhi, India har Alam Department of Computer Science and Engineering, Jamia Hamdard, New Delhi, India.

[13] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, ''Distributed representations of words and phrases and their compositionality,'' in Proc. Adv. Neural Inf. Process. Syst., 2013, pp. 3111–3119.

[14] W. Kristjanpoller and M. C. Minutolo, "A hybrid volatility forecasting framework integrating GARCH, Artificial Neural Network, technical analysis and principal components analysis," *Expert Systems with Applications*, vol. 109, pp. 1–11, 2018.

[15] H. Markowitz, "Portfolio selection," *The Journal of Finance*, vol. 7, no. 1, pp. 77–91, 1952.

[16] H. Tong, Threshold Models in Non-Linear Time Series Analysis, vol. 21. New York, NY, USA: Springer, 2012.

[17] G. S. Atsalakis and K. P. Valavanis, ''Surveying stock market forecasting techniques—Part I: Conventional methods,'' J. Comput. Optim. Econ. Finance, vol. 2, no. 1, pp. 45–92, 2010.

[18] R. F. Engle, "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation," *Econometrica*, vol. 50, no. 4, pp. 987–1007, 1982.

[19] Bai, S.; Kolter, J.Z.; Koltun, V. An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling. *arXiv* **2018**, arXiv:1803.01271.

[20] K. K. Suresh and S. R. K. Priya, ''Forecasting sugarcane yield of Tamilnadu using ARIMA models,'' Sugar Tech, vol. 13, no. 1, pp. 23–26, 2011

[21] G. P. Zhang, ''A neural network ensemble method with jittered training data for time series forecasting,'' Inf. Sci., vol. 177, no. 23, pp. 5329–5346, 2007.

[22] L. Zhang, C. Aggarwal, and G.-J. Qi, ''Stock price prediction via discovering multi-frequency trading patterns,'' in Proc. 23rd ACM S

[23] Poon, S.H., Granger, C.W.: Forecasting volatility in financial markets: A review. Journal of economic literature 41(2), 478–539 (2003)

[24] Xiong, R., Nichols, E.P., Shen, Y.: Deep learning stock volatility with google domestic trends. arXiv preprint arXiv:1512.04916 (2015)

[25] Deng, S.M.; Zhang, N.Y.; Zhang, W.; Chen, J.Y.; Pan, J.; Chen, H.J. Knowledge-Driven Stock Trend Prediction and Explanation via Temporal Convolutional Network. In Proceedings of the WWW '19 Companion World Wide Web Conference, San Francisco, CA, USA, 13–17 May 2019.