International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 06, June-2021

Volatility Forecasting Techniques using Neural Networks: A Review

Pranav B. M
Department of Computer Science
R.V College of Engineering
Bengaluru, India

Dr. Vinay Hegde
Department of Computer Science
R. V College of Engineering
Bengaluru, India

***Abstract*—Volatility is one of the key aspects in option pricing and considered as a risk associated with an asset. Because of its noisy, non-stationary, and heteroscedastic nature, predicting volatility for various forms of financial assets is one of the more mathematically challenging issues in time series prediction. Option risk management and trading depend heavily on the evaluation of option prices and implied volatility. The current studies use parametric models as a common strategy. But these models stand on a number of idealistic assumptions. Neural networks are widely used in all fields in recent years and their applicability includes the financial world as well. In this paper, existing neural network techniques in predicting volatility are studied. The paper covers mainly three types of neural networks- Artificial Neural Networks (ANNs), Recurrent Neural Networks (RNNs) and Convolutional Neural Networks (CNNs). These deep neural network models are compared with traditional models such as GARCH and its variants by using Root Squared Mean Error (RMSE) as the main loss function. It is observed that all neural network models perform much better than traditional models. But since most of these models depend only on historical data, more research is needed in considering market sentiment as a variable as it plays an important role in market fluctuations.**

Keywords—Volatility forecasting; Option Pricing; Neural Networks; CNN; LSTM-RNN

I. INTRODUCTION

Volatility in the options market refers to the underlying asset's market price fluctuation. It is a measure of how rapidly and up to what extent the underlying asset prices change. Investors who are aware of various volatility models and have a great knowledge about the financial markets understand the reason behind varying option prices better. The theory that assigns premium value for an option based on the probability of the option expiring In The Money (ITM) where “In the money” refers to the option whose strike price is more than that of its stock price value is called Option Pricing Theory. To state in simple terms, Option Pricing Theory gives the fair value of an asset after considering various parameters. These parameters include volatility, expiration date, rate of interest, stock price and strike price etc. Many models are used to price options and amongst them Black-Scholes model is widely used. Binomial Option Pricing and Monte-Carlo simulations are also the commonly used models.

Volatility is also an important factor in risk management, asset pricing, and portfolio management. As a result, financial econometrics continue to focus on predicting and forecasting stock market volatility. Some researchers have developed new and powerful predictors or parameters to enhance stock market volatility forecasting. Schwert [1] finds little support for correlations between volatility and macroeconomic predictors, but more recent researches such as Christiansen, Schmeling, and Schrimpf [2], Paye [3], Conrad and Loch [4], and Nonejad[5], Mohsen and Sujata [6] create macroeconomic and financial variables and arrive at somewhat more promising results.

Finance is highly nonlinear in nature, and stock price data can appear to follow no particular pattern. Traditional time series approaches like ARIMA and GARCH models are only useful when the series is stationary, which is a constraining assumption that necessitates pre-processing the data by taking log returns or any other similar transformations. However, the biggest difficulty is faced when implementing these models in a real-market system, because stationarity cannot be guaranteed as new data is introduced at regular intervals. This is countered by the use of neural networks, which do not require stationarity. Furthermore, by their very nature, neural networks are good at discovering correlations between data and using them to predict and classify new data.

Neural networks and deep learning techniques are used extensively in the field of Computer Science and technology industry. They provide an efficient solution to many real-world problems such as image processing, speech recognition, natural language processing, building 3D models and the list is growing every day. Artificial Neural Networks, as opposed to expert systems, are better at handling uncertainty, which is why they have gained popularity in recent years in dealing with financial applications. Financial applications are generally concerned with forecasting future events based on historical data. One such application is forecasting volatility. There have been numerous researches going on in the field and many models are developed for the purpose. This paper reviews all neural network techniques to analyse the effectiveness and efficiency.

The paper is organized as follows: Section 2 describes the importance of volatility in option pricing along with the basic definition of basic terms used in the field of finance. Section 3 analyses volatility forecasting techniques with ANNs. Section 4 and 5 studies volatility predictions using RNNs and CNNs respectively. Section 6 summarizes the review and concludes the paper.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 06, June-2021

II. IMPLIED VOLATILITY AND OPTION PRICING

As discussed earlier, volatility can be considered as the risk associated with the option. When volatility is higher, it indicates that the risk in investing is higher. Hence such options will have greater premium. Similarly, when the risk decreases, demand for assets declines as well. Therefore, prices of the option highly depend on variation in Implied Volatility, which can affect the outcome of an investment.

A. Definitions

Here are the definitions of important terms used throughout the paper:

(i) *Implied volatility (IV)*: Prices of an option are driven by various events. The forecast of such a movement is defined as Implied Volatility. As discussed, Options prices are directly proportional to Implied Volatility.

(ii) *Options*: An option can be defined as a contract between a buyer and a seller wherein the buyer is provided with an opportunity, but has no obligation to buy or sell an asset for a certain maturity date at a fixed price, known as the strike price. There are two types of options, viz., call option, put option. The holder of a call option can buy an asset at a specific price before or at a specific time. Put options are the total opposites of call options in that they allow the holder to sell an asset at a certain price before or at a predetermined period.

(iii) *Black-Scholes Model (BSM)*: Although volatility impacts option prices, it is not the only one that decides. There are multiple factors that contribute to the pricing of an option. To consider all these variables into account a mathematical model named Black-Scholes model is used. It is also known as the Black-Scholes-Meteron model. This model is leveraged to calculate the price of a European call option and takes into account six variables. These are volatility, option type, price of the underlying stock, time to maturity, strike price and rate in a risk-free investment. to estimate the fair price or theoretical value of a call or put option.

B. Volatility Prediction Model

A volatility model should be able to predict the volatility of an option in the near future. The most observed financial application of volatility prediction models is anticipating future returns. To predict magnitudes of returns, such volatility prediction models are used in most of the cases, but it can also be utilized to predict quantiles.

There are two types of volatility models that are widely used [7]. The first type formulates the conditional variance directly as a function of observed variables. The ARCH and GARCH models are the most basic examples here. Models of volatility in the second general category are not purely observable functions. These models are known as latent or stochastic volatility models.

1) GARCH Model

The generalised autoregressive conditional heteroskedasticity (GARCH) process is a method for predicting financial market volatility [8,9]. In many scenarios, the market movements are unpredictable. In a statistical model, heteroskedasticity refers to the irregular pattern of movements of any variable. To simplify, whenever observations do not follow any particular linear pattern, heteroskedasticity is said to exist. In such cases, observations appear to cluster. Therefore, the model's predicted value will be uncertain in time.

Various kinds of financial data, especially the macroeconomics data is examined using the GARCH model. It is a statistical model incorporated by many financial institutions to predict volatility of underlying assets such as stocks, bonds, and even market indices return. The result obtained from the GARCH model is extremely useful for institutional investors to allocate the assets, managing the risk, hedging, optimizing the portfolio decisions by determining price, judging which assets may potentially generate higher returns, and forecasting the returns of present investments. The GARCH model can be implemented in three steps. The first one is to create an autoregressive model in such a way that the data is fit to the best. After this step, autocorrelations between error terms are determined. Last step is to perform the significance test with appropriate significance levels.

The studies [10,11] have indicated that GARCH gives an efficient result compared to other models including ARCH as the model has only three parameters that allow for an infinite number of square roots to influence the conditional variance. However, in some cases there are aspects of the model which can be improved so that it can better detect the features and dynamics of a particular time series which are discussed in [12]

2) Stochastic (latent) Volatility Model

A model in which the variance is specified to follow a latent stochastic process provides an alternative to the ARCH framework. In the theoretical finance literature on option pricing, such models are referred to as stochastic volatility (SV) models. It is also popularly known as, Heston-Hull-White model.

With structural breaks at random times and amplitudes, multiple factors, jumps and fat-tailed shocks, fractals and multifractals, and general sorts of nonlinearities, latent volatility models can be arbitrarily elaborated. Models like this can usually be simulated, but they're hard to predict and forecast [7]

The Hull-White model takes into account the influence of correlation and volatility on option prices. The tails of the risk-neutral distribution of returns are wider when the volatility of volatility is higher. Extreme positive and negative returns are more frequent in this environment than in the Black-Scholes universe, where asset prices follow a lognormal distribution.

Many studies including [14, 15] have been conducted in comparing GARCH models and SV models for their accuracy. It is observed that no model can be declared to dominate another.

III. VOLATILITY FORECASTING USING NEURAL NETWORKS

After understanding the basic definitions and concepts, the forecasting techniques used with the help of neural networks are studied in this section. As the main objective of the paper is to review various methodologies, different models and their performances are compared. Artificial Neural Networks

(ANNs), Recurrent Neural Networks (RNNs) and Convolutional Neural Networks (CNNs) are the three major types of neural networks that are studied.

A. Volatility forecasting with ANNs

Artificial Neural Network (ANN) is a collection of numerous neurons at each layer. Each of these neurons are described by Logistic regression. This network is popularly known as Feed-forward network as inputs are processed in forward direction only. The network can be described as the combination of three layers namely, input layer which receives the data, hidden layer which processes the data and output layer which generates the output. In essence, each layer is attempting to learn specific weights. The most exploited feature of these networks is that it can learn any nonlinear functions. Hence these are also known as Universal Approximator. Forecasting stock market volatility is an important part of assessing market risk. In this section, various state-of-the-art approaches preferred by the traders are investigated.

Most of the recent research integrates ANN with the GARCH model to predict volatility of Commodities, especially metals such as Gold, Silver and Copper. Objective of one such study [23] is to analyze if there is any improvement in predicting the volatility based on returns. This method compares traditional methods such as GARCH and its variants with the proposed hybrid neural network. There are mainly six variables used by this model. They include component indices such as SZSE index (Chinese stock market), FTSE100 (British stock market), Sensex (Indian Stock market). Exchange rate between US dollar - Euro, variation in USD-Yen exchange rate and Oil price volatility. The low interconnectedness of variables indicates that the model will utilize more information to improve the fit. A Multi-Layered Perceptron (MLP) is used to model an ANN. In the base case, the ANN consists of two hidden layers and each layer contains five neurons. The training algorithm used in this study is known as the Levenber-Marquardt Backpropagation where each network is denoted by Ann(l,n). In this notation, the letter “l” indicates the number of hidden layers present in the network architecture and “n” indicates the number of neurons in each layer. As discussed before, the initial model is denoted by Ann(2,5). This configuration was initialized based on the work [24]. This initial model is used as a baseline for volatility forecasting which predicts volatility for 3 weeks ahead. The results obtained for metals Gold, Silver and Copper which are considered high-risk commodities in the market showed reduction in the error as compared to the traditional GARCH model by 2.2%. The loss function considered in the study is Heteroskedasticity-adjusted MSE (HMSE). From observation it can be seen that this base case of Ann(2,5) does not outperform the traditional GARCH model for copper metal. After the base case, many other network architectures were experimented as well. It is observed that the model with 5 layers gave best results for gold. Similarly, best results for silver and copper are obtained at four and six layers respectively. Another observation made indicates that more number of neurons in each layer helps in modelling non-linearity in volatility. Table 1 shows network architecture in (lxn) form and % variation in HMSE as compared to the GARCH model.

TABLE I. BEST-SUITED ANN(L,N) MODEL AND % VARIATION OF LOSS FUNCTION

Commodity	*(L x N) BEST*	*HMSE*
Gold	5 x 20	-16.7%
Silver	4 x 20	-93.2%
Copper	6 x 20	-8.9 %

Similarly, [25] employs ANN-GARCH hybrid model to forecast the volatility of Crude oil in Indian market. To anticipate volatility, the GARCH model and its variants such as Exponential GARCH (EGARCH) and Integrated GARCH (IGARCH) are used and the output thus obtained are in terms of return vectors. These vectors are used as input for a neural network. The loss functions Root Mean Square Error (RMSE) of GARCH family models in return forecasting is compared to the GARCH-ANN models. This study uses the same model used in [23], but differs in using the different loss function and in extending the concept by hybridizing ANN with EGARCH and IGARCH. The result thus obtained is as shown in the Table 2

TABLE II. % VARIATION OF LOSS FUNCTIONS FOR CRUDE OIL VOLATILITY

Loss function	*ANN-GARCH*	*ANN-EGARCH*	*ANN-IGARCH*
MSE	-12.6%	-19.4%	-8.15%
RMSE	-6.5%	-10.2%	-4.2%

B. Volatility forecasting with RNNs

A large number of studies earlier had suggested that Support Vector Machines (SVMs) are the best predictors of volatility. But recent research has shown that RNNs outperform SVMs. Recurrent Neural Networks (RNNs) are artificial neural networks in which the connections between the units form a directed cycle. This enables RNNs to save relevant information from prior inputs and utilise it to adjust the current output. This indicates that it has memory, which enhances its intelligence. The Long Short Term Memory (LSTM) architecture is an RNN architecture designed for long-term training and memory retention.

Yang Liu [16] evaluates Long Short Term Memory (LSTM) RNNs and v-SVR models and compares the result with the popular GARCH. During training, 11 years of historical data of financial stock indices was used. Density Estimation (DE) and Logarithmic returns were calculated for S&P 500 and AAPL. The DE calibration results for AAPL data, with ideal and consistent parameters determined using the p value test with significance 0.05 for 20 separate iterations were obtained and all the constraints were met.

From the analysis it can be seen that LSTM RNNs and v-SVR can forecast long time intervals better than the GARCH model. The v-SVR model trades off accuracy of predicting the performance of a particular option, to a minimal training set. The model requires a small size of

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 06, June-2021

training set due to its robust approach. The LSTM RNNs method is an efficient approach to learn big raw data. In order to increase the speed of training, GPUs can be used as an accelerator.

Most of the papers use historical data for predicting volatility. Although they perform much better than traditional models, in the real-world market the scenario is a bit different. Along with historical data, current events and sentiment drive changes in volatility. Therefore, the need arises to leverage social media news in predicting option prices and volatility. One such study is observed in [17] where authors have designed a novel hybrid neural network named RNN-boost. The news content crawled from Sina Weibo - a major social media platform in China and the data is analysed by extracting sentiment components and Latent Dirichlet allocation (LDA) characteristics. These characteristics are then fed into a novel hybrid model named RNN-boost as input, along with technical indicators, to forecast stock volatility in the Chinese stock market. In this study, the Shanghai-Shenzhen 300 Stock Index (HS300) was used.

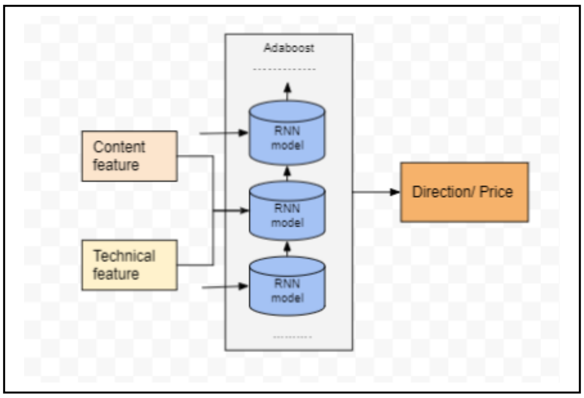

Fig. 1. RNN-Boost model proposed in [17]

The stock market's technical features are calculated using a historical dataset. Sentiment features and LDA features are two types of content features. The term Adaboost stands for "Adaptive boosting" which is a machine learning approach suggested by Y. Freund et. al [18]. AdaBoost is best used to improve decision tree performance on binary classification problems. The RNN-boost model is compared with Single RNN and observed to outperform. The proposed model in [17] reduces loss functions such as Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Root Mean Squared Error (RMSE) by 8.3%, 7.2% and 5.5% respectively compared to Single RNN.

C. Volatility forecasting with CNNs

Convolutional Neural Network (CNN) is a deep neural network that was created with image analysis in consideration. CNN has recently been discovered to have exceptional capabilities in sequential data analysis, such as natural language processing. Convolution and pooling are two basic operations that are always included in CNN. Multiple filters are used in the convolution operation to extract features (feature map) from the data collection, preserving their related spatial information.

The main objective of the study [20] is to improve the volatility forecasting of Gold by combining the concepts LSTMs and CNNs. Since CNN models are mostly used in case of image analysis, the hybrid LSTM-CNN model takes image as the input and provides useful information from the financial time series. But since the series is not capable of extracting data from images and feature matrices, it is difficult to detect CNN architecture. [22] developed a new framework for encoding time series as several forms of pictures, including Gramian Angular Fields (GAF) and Markov Transition Fields (MTF). Because of these techniques, the application can perform classification based on computer vision. The suggested volatility prediction model has two phases:

Firstly, the RGB picture series is generated using [22]. The second stage is model training, in which the image embedding is created into a new feature space on the one hand, and the LSTM layers that include squares of logarithmic returns on the other.

The two image embeddings corresponding to each of the previous layers are chained and joined by two dense layers. The initial dense layer consists of output neurons to allow for dimensionality reduction via concatenation. The final layer contains only one neuron that provides the output. It can be observed from the experimental analysis that this approach reduces the MSE by 37% in comparison to the traditional GARCH approach. When compared with the LSTM model, a reduction of 18% is observed in the loss function.

TABLE III. % VARIATION OF LOSS FUNCTIONS WHEN COMPARED LSTM-CNN MODEL WITH OTHER MODELS [20]

Model	*MSE*	*% Var*
LSTM-CNN	1.9840E08	0.00%
CNN	2.6909E08	-26.27%
ANN-GARCH	2.6807E08	-25.99%
SVR	3.5787E08	-44.56%
GARCH	3.1866E08	-37.74%

The study [21] used Chinese SSE ETF 50 Options to demonstrate the model developed using CNN in order to forecast the implied volatility of the market and thus the pricing of options. The customised non-parametric learning techniques are used to forecast implied volatility. To estimate option prices, numerous traditional parametric models are used. Then, based on various input sets, convolutional neural networks are employed to estimate pricing. Option co-movements and time-invariance problems in options markets are addressed using CNN's pattern recognition method. The similar pattern may not endure long in the finance market, particularly for newly established instruments. If the learning window is set too large, the learning may overlook the market's true trend, resulting in inaccurate forecasts. As a result, the model is built on a 10-day training scheme. This model outperforms the traditional methods such as Black-Scholes (BS) ad-hoc Black-Scholes (AHBS), Jump Diffusion

(JD), Stochastic Volatility (SV) with the minimum of 40.11% reduction in RMSE.

IV. CONCLUSION

The objective of this review is to explore more Neural Network Techniques used in Volatility Prediction. On understanding basic terminologies used in the finance world, many studies were analyzed and it is observed that ANN-GARCH model is highly used for predicting commodities with higher volatility which exhibit more non-linearity. Therefore, more research is yet to be done in generalizing the model. RNNs and CNNs which are widely used in Computer Science are also extensively used in forecasting of volatility for various options. Large portions of the studies show that these methods are generally combined with LSTM. Although these techniques perform much better than traditional methods, they use data from previous observations only. The historical prices and return sequences will be used as input information when predicting expected stock returns and volatility. But unfortunately in the real-world, asset prices can be driven by unforeseen events and market sentiment. Therefore, the need arises to include a sentiment analysis of the market and thus introducing a new variable to the model.

REFERENCES

[1] Schwert, G. W., "Why does stock market volatility change over time?", *The journal of finance*, Vol 44, no 5, pp 1115-1153, December 1989

[2] 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

[3] Paye, B.S, "Deja Vol: Predictive Regressions for Aggregate Stock Market Volatility Using Macroeconomic Variables", *Journal of Financial Economics*, Vol 106, pp 527-546, December 2011

[4] Conrad, C., & Loch, K., "Anticipating Long-Term Stock Market Volatility", *Journal of Applied Econometrics*, Vol 30, pp 1090-1114, August 2014

[5] Nima Nonejad, "Forecasting aggregate stock market volatility using financial and macroeconomic predictors: Which models forecast best, when and why?", *Journal of Empirical Finance*, Vol 42, June 2017

[6] Bahmani-Oskooee, M., & Saha, S, "On the effects of policy uncertainty on stock prices: an asymmetric analysis". *Quant Financ Econ*, Vol 3, pp 412-424, June 2019

[7] Engle, R., & Patton, A, "What good is a volatility model?" *Quantitative Finance*, Vol 1, pp 237 - 245, February 2001

[8] Bollerslev, T., Chou, R. Y., & Kroner, K. F. "ARCH modeling in finance: A review of the theory and empirical evidence", *Journal of Econometrics*, Vol 52, no. 1-2, pp 5–59, April-May 1992

[9] Bollerslev, T., "Generalized autoregressive conditional heteroskedasticity", *Journal of Econometrics*, Vol 31, Issue 3, April 1986

[10] Brooks, C., Burke, S., & Persand, G., "Benchmarks and the accuracy of GARCH model estimation". International *Journal of Forecasting*, Vol 17,pp 45-56, January 2001

[11] 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

[12] Zivot, Eric. (2008). "Practical Issues in the Analysis of Univariate GARCH Models". *Handbook of Financial Time Series*, May 2009

[13] Kim, S., Shephard, N., & Chib, S., "Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models", *The Review of Economic Studies*, Vol. 65, no 3, pp 361-393, July 1998

[14] Bragoudakis, Zacharias & Voulgarakis, Rafail., "A comparison of GARCH type models with the SV model on the financial markets of BRICS". *SSRN Electronic Journal*. May 2019.

[15] Bedowska-Sójka, B., & Kliber, A. "Realized volatility versus garch and stochastic volatility models. the evidence from the wig20 index and the eur/pln foreign exchange market.", *Statistical Review*, Vol. 57, no 4, pp 105-127, January 2010

[16] Liu, Y. "Novel volatility forecasting using deep learning–long short term memory recurrent neural networks". *Expert Systems with Applications*, Vol. 132,pp 99-109, October 2019

[17] Chen, W., Yeo, C., Lau, C., & Lee, B., "Leveraging social media news to predict stock index movement using RNN-boost". *Data Knowl. Eng*., Vol 118,pp 14-24, November 2018

[18] Y. Freund, R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting", *Journal of computer and system sciences*, Vol. 55, no. 1,pp 119–139, August 1997

[19] R. Luo, W. Zhang, X. Xu, J. Wang, "A neural stochastic volatility model", *Proceedings of AAAI*, Vol. 32, no. 1, pp. 6401–6408, November 2017

[20] Vidal, A., & Kristjanpoller, W., "Gold volatility prediction using a CNN-LSTM approach". *Expert Syst. Appl*., Vol. 157, November 2020

[21] Wei, X., Xie, Z., Cheng, R., & Li, Q., "A CNN Based System for Predicting the Implied Volatility and Option Prices". *HICSS,* January 2020

[22] Wang, Z., & Oates, T, "Encoding time series as images for visual inspection and classification using tiled convolutional neural networks". *In Workshops at the twenty-ninth AAAI conference on artificial intelligence*, Vol. 1, January 2015

[23] Werner Kristjanpoller R. , Esteban Hernandez P. , "Volatility of main metals forecasted by a hybrid ANN-GARCH model with regressors"., *Expert Systems With Applications*, 2017

[24] Kristjanpoller, W., Fadic, A., & Minutolo, M. C. "Volatility forecast using hybrid Neural Network models". *Expert Systems with Applications*, Vol. 41, no. 5,pp 2437-2442, 2014

[25] 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