International Journal of Science and Research (IJSR)
ISSN: 2319-7064
Index Copernicus Value (2016): 79.57 | Impact Factor (2017): 7.296

Forecasting Volatility with LSTM Techniques

Hemanth Kumar P[1], S. Basavaraj Patil[2]

[1]Ph.D Scholar, Department of CSE, VTURRC, Belgaum, Karnataka, India

[2]Professor, Department of CSE, AMCEC, Bangalore, Karnataka India

Abstract: ***Volatility forecasting is most searched topic in recent times, from past fears there has been tremendous research in this field of finance. This paper aims at forecasting volatility of stock index with high accuracy. The historical volatility was calculated from daily prices using Yang-Zhang method. Deep learning techniques have evolved over the years and have been successfully applied in time series forecasting problems. In this paper LSTM techniques are applied to forecasting volatility 10 days ahead. The performance of the techniques were measured with mean square error and mean absolute error. The performance of LSTM techniques has outperformed Arima, Arfima and Neural network based techniques.***

Keywords: Volatility, Forecasting, LSTM, Time series

1. Introduction

Volatility forecasting is an interdisciplinary research topic which has some helpful applications in a several of other research fields. In the finance, volatility forecasting can enable individuals to make an insightful arrangement and choice in order to reduce the risk of trading. In volatility, time period of data considered is generally an essential variable utilized for making decisions and forecasts. We can utilize time series property of volatility to forecast the volatility pattern, by making use of historical data over certain interval to time. Specialists generally utilize historical data to forecast different future values, for example, the forecast of stock prices and changes in movement.

In this paper, it is expected to make contributions in volatility forecasting deep learning LSTM techniques. To begin with, the data is stock index data is downloaded from internet, which contains open, high, low and close prices. These columns represents the prices at different durations. The volatility can be represented from group of prices, precisely we say that it is standard deviation of close prices for a longer period. The volatility is computed using volatility computing standard formula. The forecasting is implemented with deep learning LSTM technique. With the advancement in technology there is development of advanced algorithms that can take care of huge data and complex computations. Hence a LSTM approach is employed to forecast volatility. LSTM can capture temporal values pattern and produce most accurate forecast in majority cases.

In further section of the paper is sorted out as follows. Segment II gives the background of the volatility forecasting frameworks. Section III leads to explorative discussion on deep learning and LSTM techniques. SectionIV presents the methodology followed in this work for forecasting. The results summary are discussed in Section Vand VIconcludes the paper.

2. Background

Artificial neural network are good nonlinear capacity approximators [1], so they are a characteristic way to deal with consider with displaying volatility series which are suspected to have nonlinear dependence on inputs. It is without a doubt not new to forecast financial volatility series utilizing machine learning strategies and recurrent neural network are appropriate to this task. For example, this mid (1990) work [2] is among the first of a few which utilize recurrent neural nets to forecast stock prices, Ref. [3] rather detailed a volatility forecasting model, and Ref. [4] fused sentiment analysis information in directional forecast of the Dow Jones Industrial Average.

Time series based neural systems models incorporate RNNs with neuron structures, for example, long short term memory (LSTM)[5], bidirectional RNN (BRNN) [6], gated recurrent unit (GRU) [7] and consideration system [8]. Ongoing outcomes demonstrate that RNNs exceed expectations for sequence modeling and generation in different applications [9, 10]. Nonetheless, its ability as non-linear widespread approximator, one of the disadvantages of neural systems is its deterministic nature. Including dormant factors and their procedures into neural systems would effortlessly make the posteriori computationally intractable. Ongoing work demonstrates that inference can be found when hidden consistent factors are incorporated into the neural network structure [11]. Some early work has begun to investigate the utilization of variational interference to make RNNs stochastic [12]. Characterized a sequential framework with complex communicating elements of coupling noticeable also, inactive factors though [13] used heterogeneous in back propagating layers in derivation organize as per its Markovian properties.

Marino et al. made the first attempt towards a similar forecasting issue by utilizing LSTM [14] exhibit practically identical outcomes as in [15]. Be that as it may, the viability of the two spearheading works was just confirmed on the metric of root mean square error (RMSE) rather than the more typical metric of MAPE, which makes it difficult to complexity to different works. Likewise, the comparable work of [16] come up short on the legitimate dialog of the explanation behind picking LSTM. In addition, all the diverse methodologies in [15, 16, 17] were just checked with a solitary private family unit, so the totaled impact of the client data forecasts cannot be assessed.

a) Volatility

Volatility forecasting in financial world is of extraordinary reasonable and hypothetical premium topic. Volatility plays essential role in finance, for example, in option pricing, portfolio management, and hedging techniques. Hence, it is demanding to discover a decent technique to forecast volatility with high accuracy. Modeling and forecasting volatility is imperative for econometricians, analysts and professionals, and therefore it has increased much enthusiasm for the finance and literature writing. A comprehension of the diverse methodologies utilized to forecast volatility and the implifications of their presumptions and conditions gives a strong system to the procedure of risk elimination.

In this paper we are forecasting historical volatility. The historical volatility is calculated from the closing prices. Historical volatility is not a direct measurement of loss, it's a measure of how far does security prices drifts away from average prices. This is how a strongly trending but smooth market can have low volatility despite the fact that prices change dramatically over the years. Historical volatility is also used in all types of risk valuations. Stocks with an excessive historic volatility generally require a higher risk tolerance. And high volatility markets additionally require wider stop-loss ranges and probable better margin necessities.

b) LSTM

LSTM (Long Short-Term Memory Networks) is an extraordinary subset of RNN that can bargain with memory unit for any longer timeframes. The thought behind a LSTM is utilizing every node as a memory cell that can store other data as opposed to being just a node with a singleactivation function. In particular, it keeps up its own cell state. Typical RNNs take in their past concealed state and the present input, and output another hidden state. A LSTM does likewise, but it likewise takes in its old cell state and outputs its new cell state. This property encourages LSTMs to address the vanishing gradient issue from the past time-steps. LSTM has three gates: input gate, forget gate and output gate as envisioned in Figure 1. All gates are produced by a sigmoid function over the outfit of information and the former covered up state. Keeping in mind the end goal to produce the hidden state at current advance, it initially creates a transitory outcome by a tanh non-linearity over the assemble of input and the former hidden state, at that point consolidates this brief outcome with history by input gate and forget gate, separately, to get a refreshed history, at long last uses output gate over this refreshed history to get the last hidden state.

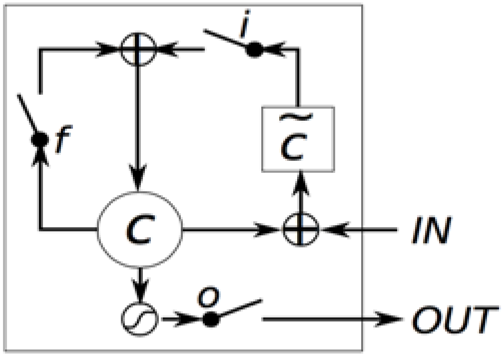

Figure 1: Representation of LSTM Architecture

3. Methodology

The methodology followed in this paper is discussed in this section. The methodology is simple and straight as this paper uses deep learning technique. Deep learning techniques are capable of feature extraction, selection or important variables and other forecasting modules. The methodology consists of database containing stock index, A historical volatility computing module, a data preprocessing module, LSTM for forecasting followed by Results. The block diagram of methodology is shown in figure.

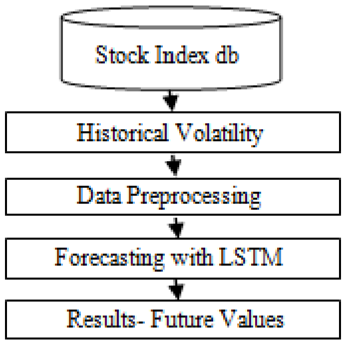

Figure 2: Methodology diagram of volatility forecasting

Data

In this paper we are using stock index data S&P 500 from 03 March 2005 till 03 March 2015. The data was collected from yahoo finance, an open source data provider with 15 minutes delay. The S&P 500 is an American stock index which comprises of 500 Companies whose common stocks are also listed in NASDAQ. It is one of the trusted indices and it may be considered as best representation American Stock Market. The index values are updated for every 15 seconds of the trading sessions.

The downloaded data consists of open, high, low, close, adjusted close and volume. This paper considers EOD close prices for volatility calculations. The data is arranged in the increasing order of the dates, 10 years of data contains 2517 records. The closing prices are very important in financial markets and specially while calculating volatility. The closing price of SP 500 for a period of 10 years is shown in figure 2.

Figure 3: SP 500 closing prices for a period 10 years

Historical Volatility Calculation

There exists several methods for estimating volatility from stock data, in this paper we have used Yang method for estimating volatility. It is derived from Rogers-Satchell method; it is weighed average of the open-close and close-open volatility. The experiments have shown that it is 14 times accurate compared to close method. The plot of estimated volatility for S&P 500 using Yang-Zhang method is as shown in figure 3.

$$\text{Volatility}_{\text{Yang-Zhang}}=\sqrt{\sigma_{YZ}}=F\,\sigma^2_{\text{overnight Volatility}}+k\,\sigma^2_{\text{open to close Volatility}}+(1-k)\sigma^2_{RS} \quad (1)$$

$$\text{Where } k=\frac{0.34}{1.34+\frac{N+1}{N-1}} \quad (2)$$

$$\sigma_{\text{overnight Volatility}}=\frac{1}{N-1}\sum_{i=1}^{N}\left[\ln\left(\frac{oi}{ci}\right)-\text{bar}\left(\ln\left(\frac{oi}{ci}\right)\right)\right]^2 \quad (3)$$

$$\sigma_{\text{open to close Volatility}}=\frac{1}{N-1}\sum_{i=1}^{N}\left[\ln\left(\frac{ci}{oi}\right)-\text{bar}\left(\ln\left(\frac{ci}{oi}\right)\right)\right]^2 \quad (4)$$

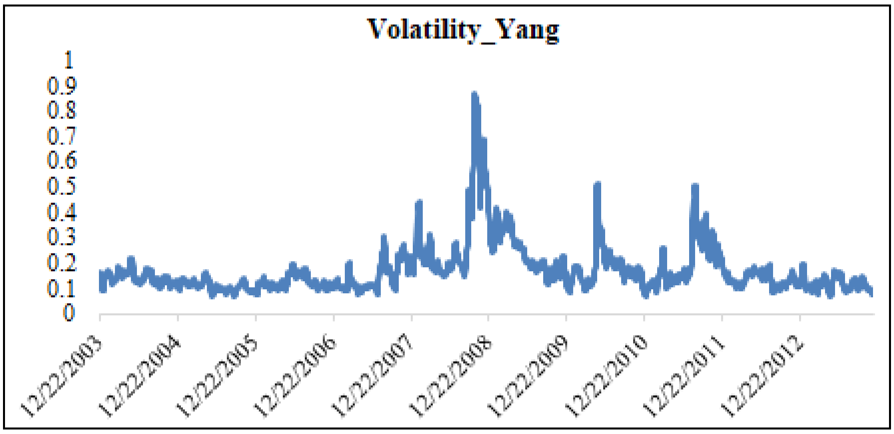

Figure 4: Volatility estimation by yang method

Data Preprocessing

The stock prices consists of outliers, missing and other noisy data. The noise data may be generated by human errors while typing or calculating. The outliers in the data used for forecasting may induce errors by deviation in average values. The missing data are generated with internet and server problem where the connection is lost.

The measures taken in this paper to avoid or eliminate these errors are as follows. The missing data is filled with the average of previous and next values. The outliers are handled by identifying the data greater than 95 percentile or lesser than 5 percentile of the data. These data are either eliminated or replaced with 95th and 5th percentile values respectively.

Forecasting with LSTM

The main aim of the paper was to forecast volatility 10 days ahead> The LSTM network is designed to forecast volatility after 10 periods. The input and output layer consists of single nodes and 10 hidden layers with linear activation function. The network was trained and tested for several epoch experimentations. However for 1000 epochs the forecasting model has shown good results.

4. Results

In this experiment S&P 500 data downloaded for a period for 10 years was used as input. The data was preprocessed and trained with LSTM technique. The mean absolute error and mean square errorof LSTM forecasts are 0.015595 and 0.00502 respectively. The LSTM algorithm was trained for 1000 epochs and results obtained are shown below.

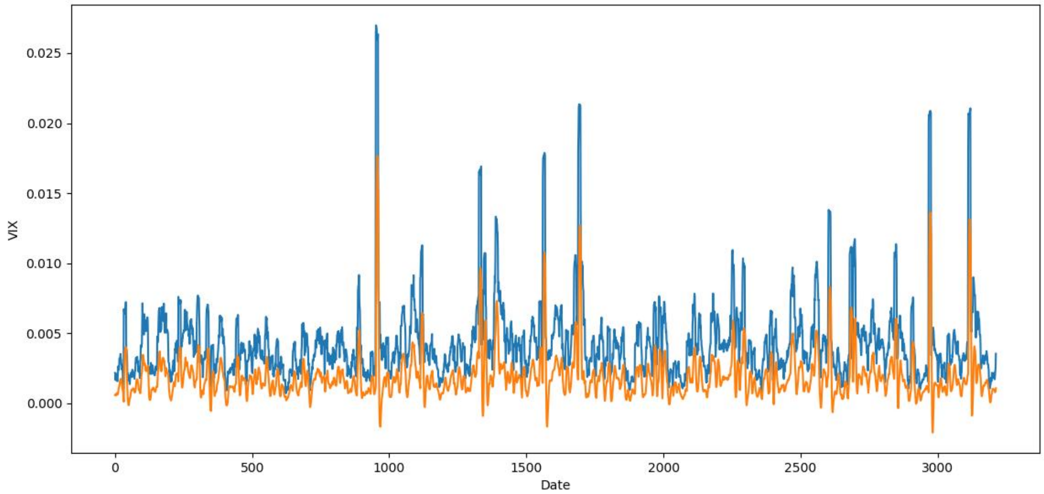

Figure 5: Actual versus forecasted results

Mean Square Error (MSE) is measure of the differences between values predicted by a trained model and the actually values.

$$MSE = 1/N(\sum_{i=1}^{N}(\sigma_t^{\wedge} - \sigma_t)^2) \qquad (5)$$

Mean Absolute Error (MAE) is an error measuring quantity used to measure the closeness of forecasts and the actual outcomes.

$$MAE = \frac{1}{N}\sum_{i=1}^{N} |\sigma_t^{\wedge} - \sigma_t| (6)$$

5. Results Comparison

In the table, the performance of Arima, Arfima, Neural Network and LSTM used for forecasting volatility are tabulated. It can be observed from table that LSTM has least MAE and MSE compared with Arima, Arfima and Neural Networks.

Table 1: Performance measure comparison of Arima, Arfima, Neural Network and LSTM

Techniques	MAE	MSE
Arfima	0.016034	0.026589
Arima	0.016043	0.024775
Neural Networks	0.024259	0.026
LSTM	**0.015595**	**0.00502**

6. Conclusion & Future Scope

The aim of this paper is to forecast volatility with high accuracy. This paper used SP 500 index data with a volume of 10 years recorded at end of the day. Experimentations were conducted on LSTM techniquesto forecast10 days volatility in advance. Deep learning based LSTM technique was used as forecasting model to obtain forecasted values. The historical volatility was computed using Yang-Zhang method from SP 500 open, low, high and close prices. The error measuring parameters mean error and mean square error are used for measuring accurate technique. From the experiments LSTM techniques outperformed Neural networks, Arima and Arfima techniques.

The future on this research is handling of outliers, a hybrid techniques to remove outliers can reduce the errors in forecasts. The LSTM techniques are part of deep learning techniques, there are other techniques available such as CNN techniques. The future works in using different frequency of data set, to study the forecasting behavior of LSTM techniques.

References

[1] K. Hornik, M. Stinchcombe and H. White, Multilayer Feedforward Networks Are Universal Approximators, Neural Networks 2, 359 (1989).

[2] K. Kamijo and T. Tanigawa, Stock price pattern recognition—a recurrent neural network approach, IJCNN 1, 215 (1990).

[3] S. A. Hamid and Z. Iqbal, Using Neural Networks for Forecasting Volatility of S&P 500 Index Futures Prices, J. Bus. Res. 10, 1116 (2004).

[4] J. Bollen, H. Mao, and X. Zeng, Twitter Mood Predicts the Stock Market, J. Comput. Sci. 1, 1 (2011).

[5] Hochreiter, S., and Schmidhuber, J. 1997. Long short-term memory. Neural Computation 9(8):1735–1780.

[6] Schuster, M., and Paliwal, K. K. 1997. Bidirectional recurrent neural networks. IEEE Trans. Signal Processing 45(11):2673–2681.

[7] Cho, K., van Merrienboer B., Gulcehre C, Bahdanau D., Bougares F., Schwenk H., and Bengio Y. 2014. Learning phrase representations using RNN encoder-decoder for statistical machine translation. In Moschitti, A.; Pang, B.; and Daelemans, W., eds., Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, EMNLP 2014, October 25-29, 2014, Doha, Qatar, A meeting of SIGDAT, a Special Interest Group of the ACL, 1724–1734. ACL.

[8] Xu, K., Ba, J.; Kiros R., Cho, K., Courville A. C., Salakhutdinov R., Zemel, R. S. and Bengio Y. 2015. Show, attend and tell: Neural image caption generation with visual attention. In Bach, F. R., and Blei, D. M., eds., Proceedings of the 32nd International Conference on Machine Learning, ICML 2015, Lille, France, 6-11 July 2015, volume 37 of JMLR Workshop and Conference Proceedings, 2048–2057. JMLR.org.

[9] van den Oord A., Kalchbrenner N.; and Kavukcuoglu K. 2016. Pixel recurrent neural networks. In Balcan, M. and Weinberger, K. Q., eds., Proceedings of the 33nd International Conference on Machine Learning, ICML 2016, New York City, NY, USA, June 19- 24, 2016, volume 48 of JMLR Workshop and Conference Proceedings, 1747–1756. JMLR.org.
[10] Kingma D. P., and Welling M. 2013. Auto-encoding variational bayes. CoRR abs/1312.6114.
[11] Chung, J.; Kastner, K.; Dinh, L.; Goel, K.; Courville, A. C.; and Bengio, Y. 2015. A recurrent latent variable model for sequential data. In Cortes et al. (2015), 2980–2988.
[12] Rezende, D. J.; Mohamed, S.; and Wierstra, D. 2014. Stochastic backpropagation and approximate inference in deep generative models. In Proceedings of the 31th International Conference on Machine Learning, ICML 2014, Beijing, China, 21-26 June 2014, volume 32 of JMLRWorkshop and Conference Proceedings, 1278–1286. JMLR.org.
[13] Fraccaro, M., Sønderby, S. K., Paquet, U., and Winther, O. 2016. Sequential neural models with stochastic layers. In Lee, D. D., Sugiyama M., von Luxburg, U., Guyon I. and Garnett, R., eds., Advances in Neural Information Processing Systems 29: Annual Conference on Neural Information Processing Systems 2016, December 5-10, 2016, Barcelona, Spain, 2199–2207.
[14] M. Chaouch, "Clustering-Based Improvement of Nonparametric Functional Time Series Forecasting: Application to Intra-Day Household-Level Load Curves," IEEE Transactions on Smart Grid, vol. 5, pp. 411-419, 2014.
[15] M. Ghofrani, M. Hassanzadeh, M. Etezadi-Amoli, and M. S. Fadali, "Smart meter based short-term load forecasting for residential customers," in North American Power Symposium (NAPS), 2011, pp. 1- 5.
[16] S. Ryu, J. Noh, and H. Kim, "Deep neural network based demand side short term load forecasting," in 2016 IEEE International Conference on Smart Grid Communications (SmartGridComm), 2016, pp. 308-313.
[17] E. Mocanu, P. H. Nguyen, M. Gibescu, and W. L. Kling, "Deep learning for estimating building energy consumption," Sustainable Energy, Grids and Networks, vol. 6, pp. 91-99, 6// 2016.