2019 4th International Electrical Engineering Conference (IEEC 2019)
Jan, 2019 at IEP Centre, Karachi, Pakistan

Stock market prediction with the integration of daily news sentiments using Long Short Term Memory in Artificial Neural Network with the ensemble approach

Zain Ul Mustafa[1*], Arish Yaseen[1*], Hassan Naeem[1], Maria Aftab[1]
[1] Department of Electrical Engineering, National University of Sciences and Technology, Karachi, 75350, Pakistan (pnec@nust.edu.pk) *Corresponding author

Abstract: This paper explains the use of Supervised Artificial Neural Network and proves it to be massively powerful in determining the daily trend in the financial market. Several back-propagation feed-forward neural networks using multiple techniques are mentioned in many studies for prediction over a certain range of time. This study, however, focuses on day to day prediction of stock data using data from the preceding day. As this is a time-series problem, the Recurrent Neural Network with Long Short Term Memory is taken into consideration and sentiment of news is calculated using natural language processing. The results produced using ANN LSTM and ANN LSTM ensemble are also compared. All the results produced in this study are carried out using Python 3 programming language whose code is made open-source.

Keywords: news, sentiments, LSTM, ANN, stocks, ensemble, NLP.

I. Introduction

In this modern era of booming international and local market, increasing wealth can be crucial unless one knows the proper way of doing it. One of the most trending topics in the sector of finance is the prediction of stock market price. There is a vast variety of researches and studies on this topic all of which having an aim to accurately predict the market trend. Financial analysts, investors, and businesses have their interest in the financial gain which is why stock prediction has become really important. But the stock market is considered to be a constantly swinging pendulum which means markets do not follow a linear path which makes it difficult to predict what might happen in the future. It is essential for an individual to know the best time to invest and sell the stock but it is difficult to know that best time. Accurate prediction of that best time in the stock exchange can provide huge opportunities for the investors to gain profit and could make trading extremely profitable [1]. Leung, Daouk [2] suggests financial analysts and traders should put their focus on to minimize the estimates' deviations from the actual observed values in the stock index. Mostafa [3] is also of the view that predicting the direction of the movement of the stock market index is extremely essential for an investor. The behavior of the stock market depends on many factors such as opening price, closing price, trading volume, day's high, day's low, economic conditions, and political events.

With the rapid development in the digital media sector, digital data acquisition is not a thing of the past. Online news services available on the World Wide Web provide easy means of accessing daily news. Accessibility to 24/7 digital news, the latest developments in Natural Language Processing (NLP), and the availability of powerful processors allow extraction of more information from news articles and reports, and predict their effect on the stock market trends. Financers, brokers, and trader are the main contenders in the stock market and are not usually aware of the current trends. They are usually puzzled as to which stock share they should buy and which to sell to gain profits. What they do know is that stock prices depend highly on relevant daily news and they need to keep themselves up-to-date with all the news available in digital or print media. It is not easy to handle the inflow of data from such vast sources of information, let alone analyzing them is beyond human capabilities. With developments in data science, text mining techniques are now available through which useful information can be automatically extracted from multiple sources, relieving humans from this vicious task.

There are two standard analyses for stock market trend prediction which are technical analysis and fundamental analysis. Analysis of statistics gathered from price movement and volume in order to evaluate securities and to identify opportunities in trading is known as technical analysis [4]. Analysis of the factors affecting the well-being of industries, companies, and the economy is known as fundamental analysis [5].

II. Background

Various models have been created for the prediction of the stock market price using various algorithms and learning schemes. Şenol and Özturan [6] proposed seven different prediction models and applied them to the stock exchange in Turkey. The authors concluded that the Artificial Neural Network (ANN) could be one of the most robust algorithms for forecasting. Huang, Nakamori [7] used the Support Vector Machine (SVM) to forecast the movement in the stock market in which the authors concluded that SVM works quite well in predicting the direction. Kara, Boyacioglu [8] proposed the ANN and SVM models and applied them on the Istanbul Stock Exchange in which the authors concluded that ANN is significantly better than SVM.

A useful and powerful technique that has the ability to capture the subtle relationships among empirical data which are hard to define is known to be ANN [9, 10]. ANN has become a popular machine learning technique that is known to exceed the performance capabilities of the various conventional method [11–17]. Tsaih [18] describes that markets walk in a veil of randomness which makes it totally nonlinear and dynamic. The author also mentions that using simple techniques for prediction do not stand near a chance and a supervised learning technique such as Back-propagation neural network (BPNN) should be used.

Xiotian, Zhu [19] proposed a four three-layered feedforward neural network models for the prediction of the stock index using neural networks in which authors concluded that with different epoch cycle neural network gave different prediction accuracies on NASDAQ data with the highest accuracy achieved to be 74%. Amin [20] proposed using Multi-Layer Perceptron for the prediction of stock index in which the authors used TANSIG and LOGSIG as the transfer function giving extremely good accuracy on NASDAQ data with the highest accuracy achieved to be 96%. Erkam [21] proposed using GARCH Multi-Layer Perceptron on NASDAQ data in which the model gave unexpected good accuracies with MSE on test data to be 0.03665. Xiao Ding [22] proposed an event-driven model with deep learning for stock prediction and applied it on S&P500. The author concluded that the sole purpose of the study was to predict better trading strategy. Mathew [23] proposed a classification based deep neural network with five fully connected layers with one thousand neurons in four layers. The author applied his model on CME Commodity and concluded classification accuracy and strategy performance measurement can be tied together.

III. METHODOLOGY

The algorithms used for this research are described in detail as below.

A. Artificial Neural Network

When it comes to automating a task, a computer system needs to be built that mimics what the human brain can do. A human brain learns a new thing, sees what an outcome turns out to be, and remembers it. An Artificial Neural Network (ANN) mimics the way neurons work in the human brain where they process and retain the information for computational problems and complex tasks.

Biological neurons are the most important part of the brain that forms a complete nervous system. Every neuron carries information through electric impulses. In computer science, a neuron is known as a node that receives and transmits information to and from other nodes and processes the information to compute an output. The connection between the nodes is called edges. This whole node connection is known as a computational graph and the system is called the ANN.

An ANN is simply a network of nodes that are connected to each other via directed edges. Every node has an associated weight with it which shows its importance in computation compared to other nodes in the system. Each edge connecting the nodes also has an associated numerical weight which indicates the influence of the tail node on the head node, where positive weights represent strength and negative weights representing inhibition [24].

Generally, the initial connection weights are randomly selected. A neural network consists of layers and the simplest form of ANN consists of an input layer, a hidden layer, and an output layer. A hidden layer is where the complex computations are performed and passed through an activation function to add non-linearity in the result. This is because of the fact that information and computations performed in real life are non-linear and there is a need to make a neural network that does the job in similar criteria. Connectionist model, such as ANNs, is fitted for machine learning where connection weights are adjusted in order to improve the performance of a network.

In our neural network, the data flows forward to the output continuously without any feedback. A typical four-layer neural network is used for prediction of the closing price of various stock companies. Technical variables are used in the input layer as nodes, while predicted result based on true value (actual closing price) and sentiment analysis is the output. Hidden layer nodes with a non-linear activation function are used to process the information from input nodes. This is represented by Eq. (1).

$$X_j = \sum w_{ij} I_i + \theta_j \tag{1}$$

Where i is the number of neuron nodes, j is the number of hidden layer neurons, X is the output of the node, w is the weight, I is the input neuron, θ is the bias of hidden nodes.

An activation function is used to calculate the output of the neuron of the hidden layer or output layer. From available sigmoid function namely, logistic, hyperbolic tangent (tanh), and Rectifier Linear Unit (ReLU), tanh is used as it takes the value between -1 and 1 can be written as Eq. (2).

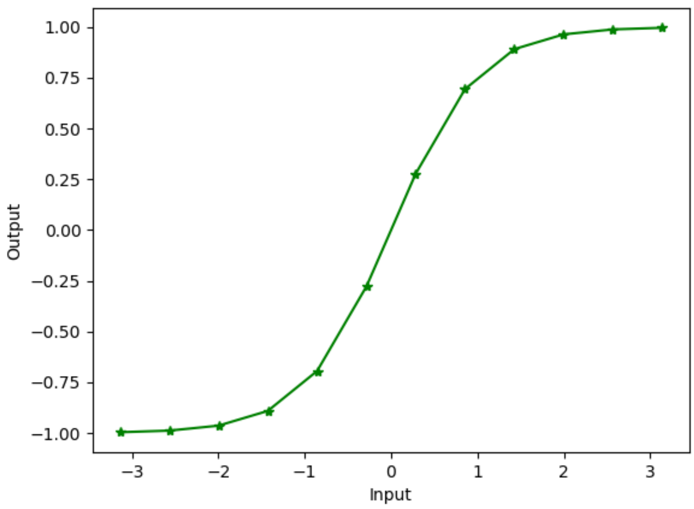

Fig. 1 Graph of tanh

$$f(x) = \frac{2}{1 + \exp(-2x)} - 1 \tag{2}$$

The weights and biases are updated to reflect the propagation errors. Weights are updated by the following equation, written as Eq. (3) and Eq. (4).

$$\Delta w_{ij} = (l) E_j O_j \tag{3}$$

$$w_{ij} = w_{ij} + \Delta w_{ij} \tag{4}$$

Where the variable *l* is the learning rate which basically means how fast a neural network will converge. Small learning rate will mean the system will learn at a slow pace which means it would converge gradually, however, it will increase the computational requirements. A large value of learning rate would mean the system would not converge as weight changes would be huge that optimizer will make the loss worse.

Fig. 2 shows an example of an architecture of the ANN with one input layer, two hidden layers, and one output layer. First input layer contains four neurons, the first hidden layer contains three neurons, the second hidden layer contains two neurons, and the output layer contains one neuron.

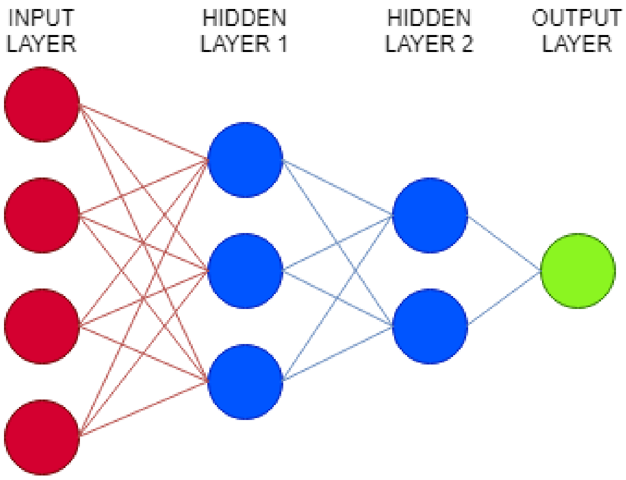

Fig. 2 ANN Architecture

B. Long Short Term Memory

Sepp Hochreiter and Juergen Schmidhuber proposed a variation of the Recurrent Neural Network (RNN) by the name of Long Short Term Memory (LSTM) in order to resolve the problem of vanishing gradient. LSTM is basically incorporation of memory that preserves the error. Back-propagation of this error can be done through layers and time. They allow RNN to keep on learning over many steps by maintaining a more constant error which allows causes and effects to be remotely linked together.

In LSTM, information is stored in a gated cell outside the flow of the RNN. These gates, implemented using sigmoid multiplication, can open and close which enable cells to make decisions and allow certain processes to take place. Decisions are made by these cells via gates on what to store, when to read, and initiate erasures. These cells behave just like a computer's memory but unlike digital computer memory, these cells are analog and are suitable for neural networks because of the fact that they have a great advantage of being differentiable.

Similar to the nodes in the neural network, these cells receive signals. Based on signals they receive, they allow blocking or passing of information on the basis of its importance and strength. This is done by filtering that information using their own set of weights. These weights just like in neural networks are adjusted through the learning process of RNN. The cells learn through the iterative process of making estimations when the data should be allowed to enter, leave, or to be deleted. This is done by adjustment of weights using gradient descent.

By looking at Fig. 3, the information from the network flows into the cells. The combination of present input and past cell state is fed to the cell itself and each of its gates which will decide how to handle the information.

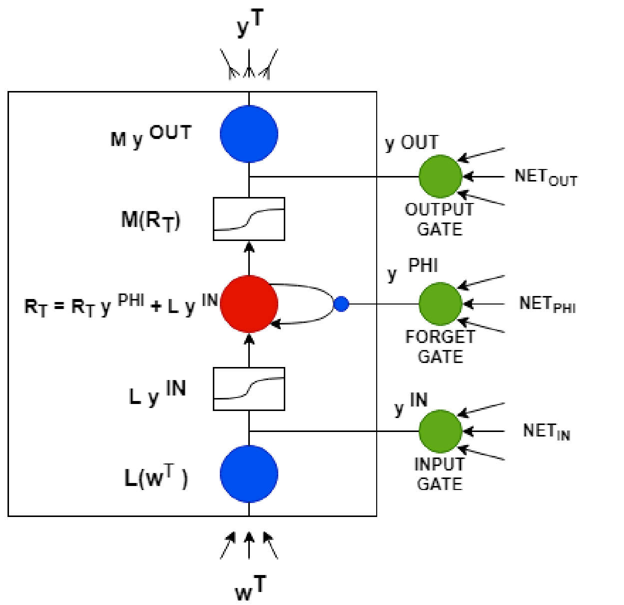

Fig. 3 Inputs to the gated cell

These gates determine if to let the present information enter, erase the present state of the cell, or allow that state to affect the output of the network at the present time step. In Fig. 3, R_T is the present state of the cell and Ly is the current input to the cell. Each gate in a cell can be open or shut down and recombination of the gate state will take place at every step. The arrows in Fig. 3 show that these memory cells can either forget their state or retain their state, either to be written to or to be read from at each time step.

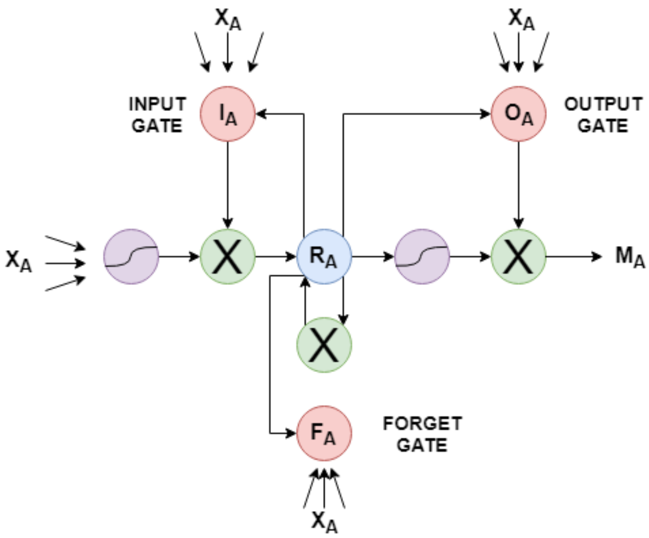

Fig. 4 Gates in LSTM cell

Fig. 4 shows an LSTM cell in which x_t is the inputs at a time step and M_A is the output of the cell to the hidden layer. Different jobs are assigned by cells in LSTM to addition and multiplication in the input transformation. Subsequent cell state can be determined by the simple multiplication of the cell's current state with the new input. This secret of LSTM cell helps preserve the constant error.

Each cell has three gates that are for input, output, and forgetting and each of these gates are weighted using a set of weights. The forget gate is a linear identity function where if the gate is open, the current state of the cell gets multiplied by one in order to propagate one-time step forward.

Fig. 5 shows gates in LSTM cell in action where green triangles represent closed gates, and dark red circles represent open gates. The forget gates are represented by the green triangles and dark red circles running horizontally downwards the hidden layer.

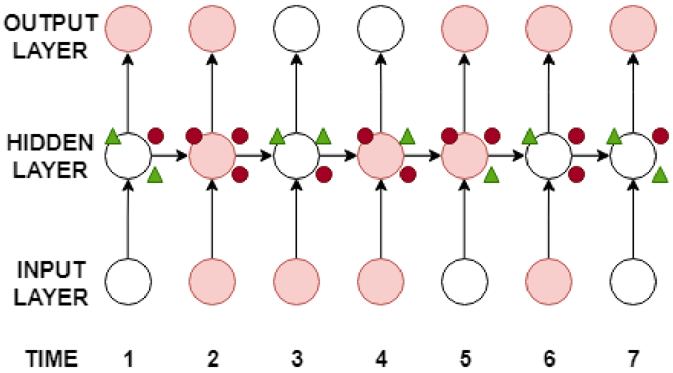

Fig. 5 LSTM cells in action

A noticeable difference between feedforward neural networks and RNN is that of the mapping of input to the output. The feedforward neural networks map one input to one output (one image, one label), while RNN can map one-many (one image, many labels), many-one (voice classification), or many-many (translations).

C. Natural Language Processing

A sub-field of machine learning is the Natural Language Processing (NLP) which focuses on making computers understand and process languages spoken by humans in order to get computers close to the level of humans in understanding languages. Computers do not have an intuitive understanding of how languages work and hence it cannot read between the lines. But the latest advancement in machine learning such as deep learning and computational linguistics are in pursuit to bridge the gap between communication done by humans and computer understanding it.

NLP works using rule-based algorithms which enables it to do tokenization, parsing, part-of-speech tagging, detection and identification of semantic relationships, and lemmatization. This is basically what humans have learned in grade school. NLP breaks down language into simpler and shorter pieces and then tries to figure out the relationship between them and explores how those pieces would create meaning when joined.

NLP algorithms can perform a vast variety of high-level tasks which includes content categorization, contextual extraction, document summarization, machine translation, topic discovery and modeling, speech-to-text and text-to-speech conversions, and sentiment analysis

IV. Proposed Methodology and Architecture

ANN proves to be the best when it comes to machine learning on complex data, the proposed model is an ANN with six layers with one layer as input, four LSTM hidden layers, and one layer for output. Four hidden layers have five hundred and twelve neurons each. A better version of the logistic function, tanh, is used because of its range of values from -1 to 1. The advantage of using tanh as the activation function is that the zero inputs would be mapped to near zero and negative inputs to more negative values. Table 1 shows all the parameters used in the model. Ensemble learning approach is also used with a second model consisting of three hidden layers with one thousand neurons while other parameters remain the same as the first model. The results of using the first model alone and using the ensemble approach are compared in the result section.

Table 1 Proposed Parameters of the Model

Parameter	***Values***
Input and Output Layer	1
Hidden Layers	4
Neurons per hidden layer	512
Dropout	10 percent
Epoch	100
Batch size	32
Activation function	tanh
Output neurons	1
L1 regularization	0.00001
Loss function	MSE
Optimizer	RMSprop

To clearly understand how this algorithm basically learns, assume if X number of input columns are given, it learns to shift the column that is intended to be predicted from 0 to length(X)-1 and only predict one value which exists in the last index. This training approach is basically known as 'one right shift concept'.

A major part of the algorithm is explained by this concept approach but what makes a difference is by using daily news information with the prediction algorithm. NLP is used to analyze the daily news titles calculating the amount of negativity, neutrality, positivity, and compound (a kind of net sentiment). Logic says that current news and stock data cannot directly affect the next day's price as stocks move in the random direction. But the current day's news directly affects the current day's stock index. Though it can be said that current information will have some effect on the next day. If this current stock information and next day's news information is used, it now seems logical to predict what stock price would be at the end of the next day.

This real-time news information is generally available for all the users at any time from online resources. Extraction of this daily news can be made using scrapping tools available on Python. In order to extract daily news related to market, a scrapper written in Python using two libraries: Requests (for fetching data), and Requests-HTML (in order to make scrapping as simple and intuitive as possible). After scrapping is done, required news related to business, politics, economic policies, and the market is separated from all the unwanted information. Another Python API, Natural Language Tool Kit (NLTK), is used for natural language processing and sentiment analysis is carried out using Valence Aware Dictionary and Sentiment Reasoner (VADER) which is sensitive to both intensity and polarity of emotions as suggested by Hutto, C.J. [25].

This API processes the news information and polarizes into four different categories: compound, positive, negative, and neutral. If the news has a positive polarity it means that news has a positive effect on stock markets. While if the news tends to have a negative polarity, it means it has a negative effect on stock markets. Table 2 shows a complete summary of the first model created using the summary method in Python.

Table 2 Summary of the Model

Layer (type)	Output Shape	Param #
lstm_1 (LSTM)	(None, 7, 25)	2700
dropout_1 (Dropout)	(None, 7, 25)	0
lstm_2 (LSTM)	(None, 7, 25)	5100
dropout_2 (Dropout)	(None, 7, 25)	0
lstm_3 (LSTM)	(None, 7, 25)	5100
dropout_3 (Dropout)	(None, 7, 25)	0
lstm_4 (LSTM)	(None, 25)	5100
dropout_4 (Dropout)	(None, 25)	0
dense_1 (Dense)	(None, 1)	26
activation_1 (Activation)	(None, 1)	0

Total params: 18,026
Trainable params: 18,026
Non-trainable params: 0

V. DATA ACQUISITION

Real-time news data that reflects the flow of the stock market and have a variety was required for this experimental study. Past eight years data of various stock companies was collected which was from 3rd January 2011 to 7th November 2018. Because of this wide time range, an automated system was necessary to extract useful information. A news scrapper was used to scrap that necessary information and simple Python data processing tools to process the stock data.

PyQuery and LXML were used as the backend for fetching and preparing data from the DAWN news website [26] for news information and KSE Stock Index [27] for stock data of various companies. Daily stock data contains five values as Open, High, Low, Close, and Volume.

Due to the weekends and public holidays, the dynamics of stock market price data is not understood completely as it is closed during these days and do not function. After the extraction of data, news information and stock data of the same day were merged and gaps of missing data were ignored. For pre-processing of news data, discernment of news on the basis of tags or token is performed. As the news is split in different tags, only those tags were selected or considered that contains a class of theme--Pakistan from the official news site of DAWN and after that, those news titles that have a tag of the article were read.

Fig. 6 to Fig. 9 shows the acquired stock data in graphs which contain an average of 2567 instances with each instance having 5 values. Table 3 shows those five values of one of the stock companies, Table 4 shows merged data, and Table 5 shows merged data with calculated news sentiments.

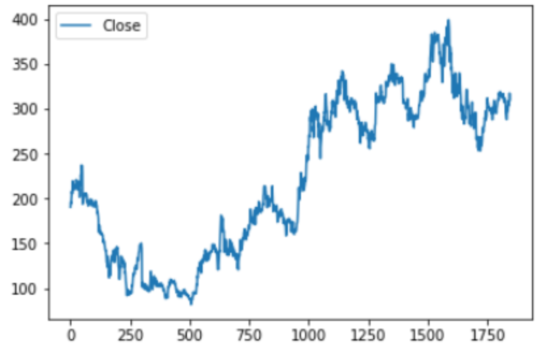

Fig. 6 ENGRO stock closing price

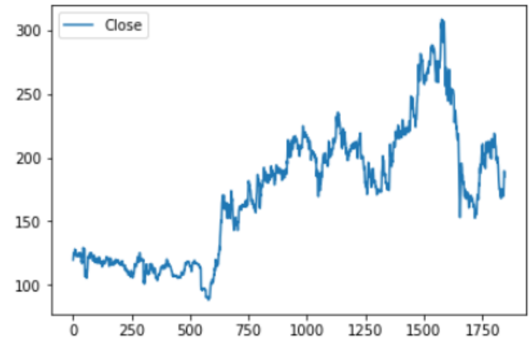

Fig. 7 HBL stock closing price

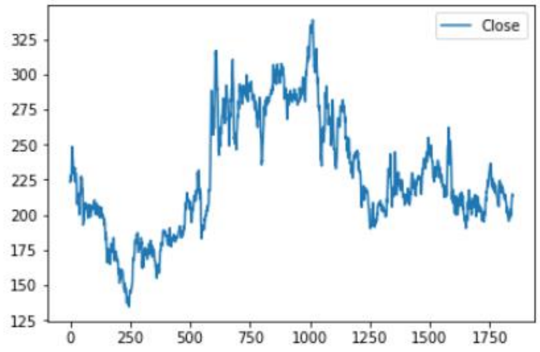

Fig. 8 MCB stock closing price

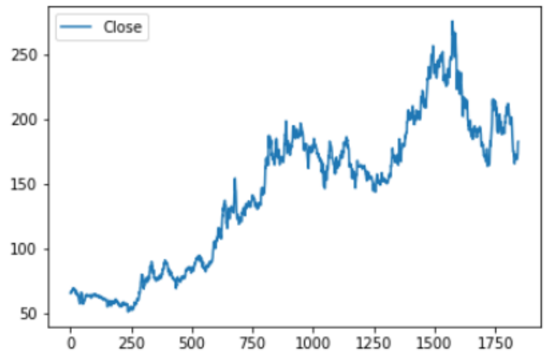

Fig. 9 UBL stock closing price

Table 3 Stock Data of MCB

	Symbol	Date	Open	High	Low	Close	Volume
0	MCB	04-Jan-2010	219.60	220.40	215.20	215.95	2478180
1	MCB	05-Jan-2010	216.89	220.75	216.20	219.25	3720502
2	MCB	06-Jan-2010	220.05	225.00	220.05	223.40	5189731
3	MCB	07-Jan-2010	223.40	224.89	221.50	222.44	2032048
4	MCB	08-Jan-2010	223.00	223.25	219.50	220.42	1752809

Table 4 Merged Stock Data with News

	Symbol	Date	Open	High	Low	Close	Volume	headlines
0	MCB	03-Jan-2011	226.00	227.00	221.50	223.32	646442	Sui cantonment turned into military college.De...
1	MCB	04-Jan-2011	223.32	229.50	223.32	227.97	1601540	Big loss to progressive forces, says Sherry.Bl...
2	MCB	05-Jan-2011	227.00	232.00	225.60	226.27	2181684	Naming new Punjab governor a test for Zardari....
3	MCB	06-Jan-2011	227.00	227.65	224.20	226.16	1034935	Violence in Pakistan declines: US report.Power...
4	MCB	07-Jan-2011	227.50	227.99	223.80	225.99	1247251	President vows to expose Taseerâ€™s killers.Be...

Table 5 Merged Data with Sentiments Calculated

	Date	compound	neg	neu	pos	Open	High	Low	Close	Volume
0	03-Jan-2011	-0.8934	0.172	0.746	0.082	226.00	227.00	221.50	223.32	646442
1	04-Jan-2011	-0.6996	0.133	0.806	0.06	223.32	229.50	223.32	227.97	1601540
2	05-Jan-2011	-0.9442	0.319	0.638	0.044	227.00	232.00	225.60	226.27	2181684
3	06-Jan-2011	-0.7906	0.179	0.726	0.094	227.00	227.65	224.20	226.16	1034935
4	07-Jan-2011	-0.5859	0.14	0.766	0.094	227.50	227.99	223.80	225.99	1247251

For the sake of comparison with the benchmark and state-of-the-art models, this proposed LSTM model was applied to NASDAQ and S&P500 data as well which was obtained from online resources.

VI. Results

Predicting the next day stock is extremely difficult but the integration of the daily news sentiments produced extremely good results as shown in Table 6 and Table 7. The proposed ANN LSTM model and ANN LSTM model with ensemble were able to achieve an accuracy of more than 94% and with an ensemble approach to get as high accuracy as 97%. This novel approach with an optimized set of parameters was able to achieve 97% accuracy on NASDAQ data as compared to Xiaotan's model who was able to achieve 72% prediction accuracy and Amin's model who was able to achieve 94% and 96% with TANSIG and LOGSIG as activation functions respectively on NASDAQ data.

Xiao Ding who proposed event-driven deep learning model was able to achieve 66% accuracy on S&P500 data vs our model which was able to achieve 94% accuracy on the same data.

This proposed model with ANN LSTM on KSE stock was applied and was tested on the unseen data producing the results as shown in Table 6.

Table 6 ANN LSTM without Ensemble on KSE Data

S. No.	***Stock Symbol***	***Accuracy***
1	ENGRO	92.34%
2	HBL	94.21%
3	MCB	90.22%
4	UBL	84.10%

This proposed model with the ensemble approach was applied to the same data and was tested on the same unseen data producing the results as shown in Table 7.

Table 7 ANN LSTM with Ensemble on KSE Data

S. No.	***Stock Symbol***	***Accuracy***
1	ENGRO	96.54%
2	HBL	97.33%
3	MCB	94.45%
4	UBL	92.87%

Fig. 10 to Fig. 13 shows graphs with a comparison of predicted trend and the actual trend of different stocks produced by ANN LSTM with the ensemble learning using optimized parameters of Table 1.

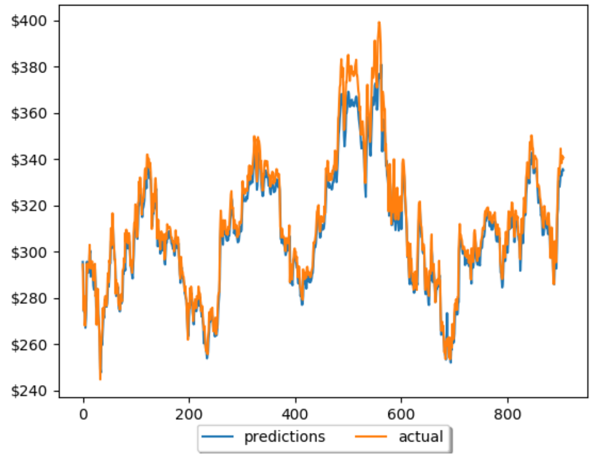

Fig. 10 Test on ENGRO unseen data. Legend (Blue: predictions, Orange: actual)

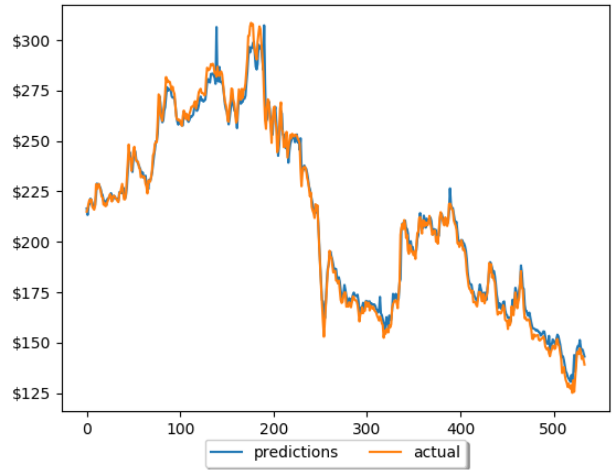

Fig. 11 Test on HBL unseen data. Legend (Blue: predictions, Orange: actual)

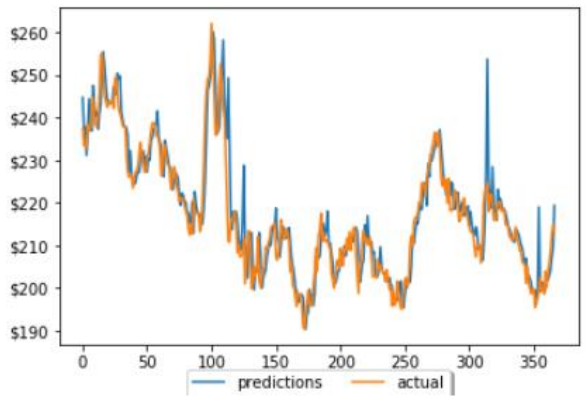

Fig. 12 Test on MCB unseen data. Legend (Blue: predictions, Orange: actual)

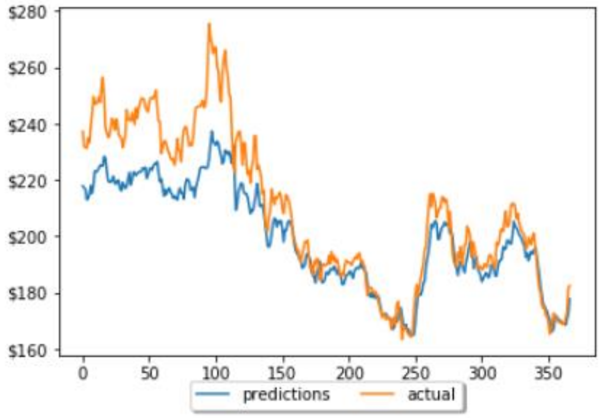

Fig. 13 Test on UBL unseen data. Legend (Blue: predictions, Orange: actual)

VII. Conclusion

ANN LSTM with the ensemble is an excellent choice when it comes to any time-series problem as it comprises the prediction power of the neural network and memory of LSTM architecture. This experimental study concludes that through the proposed model daily prediction of stock can be done. Usage of Python modules such as Pandas, Scikit-learn, and Numpy are brilliant for data processing, Matplotlib for data visualization, Keras for network modeling, and Jupyter Notebook as iPython IDE. Experimental accuracy and error of this model as shown in Table 6 and Table 7 clearly indicate that this proposed model can work brilliantly with any stock available in any exchange of the world.

The code is available on the GitHub link [28].

VIII. Acknowledgment

This research was partially supported by Dr. Adeel Yusuf. I thank my colleagues from NUST who provided insight and expertise that greatly assisted the research. I thank Dr. Adeel Yusuf for assistance with this research and for comments that greatly improved the manuscript.

IX. References

[1] Gholamiangonabadi D, Mohseni Taheri SD, Mohammadi A, Menhaj MB, editors. Investigating the performance of technical indicators in electrical industry in Tehran's Stock Exchange using hybrid methods of SRA, PCA and Neural Networks. Therm Power Plants IEEE 2014;2014:75–82.

[2] Leung MT, Daouk H, Chen A. Forecasting stock indices: a comparison of classification and level estimation models. Int J Forecast. 2000;16(2):173–190.

[3] Mostafa MM. Forecasting stock exchange movements using neural networks: Empirical evidence from Kuwait. Expert Syst Appl. 2010;37(9):6302–6309.

[4] https://www.investopedia.com/terms/t/technicalanalysis.asp.

[5] http://stockcharts.com/school/doku.php?id=chart_school:overview:fundamental_analysis.

[6] Şenol D, Özturan M. Stock price direction prediction using artificial neural network approach: the case of Turkey. J Artif Intell. 2008; 1(2):70–77.

[7] Huang W, Nakamori Y, Wang S. Forecasting stock market movement direction with support vector machine. Comput Oper Res. 2005; 32(10):13–22.

[8] Kara Y, Boyacioglu MA, Baykan ÖK. Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange. Expert Syst Appl. 2011; 38(5):11–19.

[9] Vellido A, Lisboa PJ, Vaughan J. Neural networks in business: a survey of applications (1992–1998). Expert Syst Appl. 1999; 17(1):51–70.

[10] Zhang G, Patuwo BE, Hu MY. Forecasting with artificial neural networks: The state of the art. Int J Forecast. 1998; 14(1):35–62.

[11] Fernando FR, Christian GM, Simon SR. On the profitability of technical trading rules based on artificial neural networks: Evidence from the Madrid stock market. Econ Lett. 2000; 69(1):89–94.

[12] Lu C. Integrating independent component analysis-based denoising scheme with neural network for stock price prediction. Expert Syst Appl. 2010; 37(10):56–64.

[13] Versace M, Bhatt R, Hinds O, Shiffer M. Predicting the exchange traded fund DIA with a combination of genetic algorithms and neural networks. Expert Syst Appl. 2004; 27(3):17–25.

[14] Specht DF. A general regression neural network. IEEE Trans Neural Netw. 1991; 2(6):68–76.

[15] Wang Z, Wang L, Szolnoki A, Perc M. Evolutionary games on multilayer networks: a colloquium. Eur Phys J B. 2015; 88(5):1–15.

[16] Wang Z, Kokubo S, Jusup M, Tanimoto J. Universal scaling for the dilemma strength in evolutionary games. Phys Life Rev. 2015; 14:1–30. doi: 10.1016/j.plrev.2015.04.033 PMID: 25979121

[17] Wang Z, Zhao DW, Wang L, Sun GQ, Jin Z. Immunity of multiplex networks via acquaintance vaccination. Europhys Lett. 2015; 112(4):48002–48007.

[18] R. Tsaih, Y. Hsu and C. Lai, "Forecasting S&P 500 stock index futures with a hybrid AI system", 1998.

[19] X. ZHU, H. WANG, L. XU and H. LI, "Predicting stock index increments by neural networks: The role of trading volume under different horizons", 2008.

[20] A. Moghaddam, M. Moghaddam and M. Esfandyari, "Stock market index prediction using artificial neural network", 2016.

[21] E. Guresen, G. Kayakutlu and T. Daim, "Using artificial neural network models in stock market index prediction", 2011.

[22] P. Oncharoen and P. Vateekul, "Deep Learning for Stock Market Prediction Using Event Embedding and Technical Indicators," *2018 5th International Conference on Advanced Informatics: Concept Theory and Applications (ICAICTA)*, 2018.

[23] M. Dixon, D. Klabjan and J. Bang, "Classification-based financial markets prediction using deep neural networks", 2017.

[24] Gallant SI. Neural network learning and expert systems. MIT Press, Cambridge, 1993.

[25] Hutto, C.J. & Gilbert, E.E. (2014). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. Eighth International Conference on Weblogs and Social Media (ICWSM-14). Ann Arbor, MI, June 2014

[26] https://www.dawn.com/archive/

[27] http://www.ksestocks.com/QuotationsData

[28] https://github.com/ZainUlMustafa/Stock-Prediction-using-News-Info-Sentiment