PREDICTION OF STOCK PRICE WITH ENHANCED DEEP LEARNING MODEL

M.MOOHAMBIKAI,
ASSISTANT PROFESSOR/CSE,
RVS COLLEGE OF ENGINEERING,
DINDIGUL.

B.BHAKIRATHI,
PG SCHOLAR,
RVS COLLEGE OF ENGINEERING,
DINDIGUL.

***Abstract*— Prediction of the stock price and in specific the trend the stock price will follow is an important task in various perspectives. Recurrent neural networks are used for this task. Experimental analyses were made with different time stepss and different number of Long Short Term Memory (LSTM) layers. The dataset that has been used is the Google stock price of the past five years. It is a time series data which contains stock details of 1274 days. The performance parameters that are used for evaluating the model are explained variance score, r^2score and the pearson coefficient . The results show the comparative performance of the recurrent neural network with different number of time steps and different number of LSTM layers. It has been observed that the model with 60 Time steps and 4 LSTM Layer performs better.**

Keywords: Stock Prediction, Time Series Analysis, Recurrent Neural Networks, Long Short Term Memory.

I.Introduction

Time series stock data refers to the prices of a particular share over a time period. It contains various intense data required for predicting the future stock price of the particular share. Various models are in use for time series prediction of the share prices, this includes the conventional models based on statistics [1-3], with the introduction of machine learning, and many such models have also been used for stock price prediction [4]. There are also works that employs artificial neural networks for stock price prediction [5-7].

The objective of the work is to use recurrent neural networks for predicting the stock price and the trend of the stock price. The rest of the paper is organized as follows, the next section explains the observations made with the various state of art models, the third section explains the recurrent neural networks and the last section gives the results obtained.

II. Study of state of art models.

This section explains the various works that have been used for predicting the stock movement.

Text content from the financial reports is used by [8] to predict the stock price. Two methods are used model based and visualization, the former is used for predicting the stock prices and the later is used for assessing the prediction. Six models are employed on the 8k financial dataset. The best result is achieved with the supervised PV and the accuracy achieved is 68%.

[9] proposed multi source multiple instance learning model for predicting the stock price. The multiple sources include quantitative data of stock, result of sentiment analysis made with social media data and events extracted from the web news. The authors believe that usage of different sources would yield a better prediction accuracy rather than relying on a single source. Multi instance learning algorithm is employed, which labels individual instances. Data used ranges from 1st January 2015 to 31st December 2016. The framework is tested with support vector Machine, tensor based learning approach , nested multi instance learning model, open IE multi instance learning model, without RBM multi instance learning model and without multi hinge lose multiple instance learning model. It is observed from the results that,

news events contribute much in predicting the stock price followed by the stock's quantitative data and then the data extracted from social media. The quantitative data considered are average price, market index, turnover rate. The proposed model improves the accuracy by 6.7% for 2015 data and 9.4% for 2016 data when compared to other models, but the achieved accuracy is only 59% and 60% for the years 2015 and 2016 respectively.

[10] used deep learning model to predict the trend in the stock price. The factors considered for the prediction are data retrieved from the financial news combined with the sentiment dictionary. A better recurrent neural network model is proposed to analyze the text. In addition to this, a new sentiment word embedding model based on financial dataset and sentiment dictionary is created. The network used is TGRU .

Raw daily S&P 500 Index stock prices downloaded from Yahoo Finance was in CSV format is used as the dataset. The training part of the dataset has 439,304 articles, validation set contains 2,305 articles and the test set contains 98,468 articles. The accuracy achieved by the authors is 66.32 which is comparatively higher than the other models.

[11] considered 10-Q form as their source for stock prediction. The 10-Q form is a quarterly report that contains security changes, analysis and discussion by the management, risk of the market etc.

The goal of the authors is to test the usability of text contents in 10-Q reports on stock predictions. Therefore, a simple experiment is designed at the cost of prediction rates for the purpose of preliminary testing of words in 10-Q reports.

The model is restricted to features generated from the corpus. the corpus consists of 17,939 10-Q documents published in between March 2004 to October 2017. Training dataset consists of reports that were published 30, 60, 90- days prior to August 8th, 2016 and test data set consists of the 2,000 most recent 10 documents within the corpus.

A feed forward multi layer neural network is designed with the features extracted from the 10-Q forms. The accuracy of the prediction over different day ranges lays only around 54%.

[12] included financial news as an input for predicting the stock price. The author has proposed a method called Attention-based LSTM(At-LSTM). The goal of the authors is to leverage public released financial news and train a model named Attention- based LSTM (At-LSTM) to make prediction on directional changes of the stock prices of the different companies.

The model consists a Recurrent Neural network(RNN) to encode the news text and capture the context information,self attention mechanism is applied to distribute attention on most relative words, news and days . The model also take advantages from the rapid development of deep neural networks.The optimization algorithm used in the model is Adadelta optimization algorithm and the loss function used is cross entropy loss.

The number of news used for training , validation and testing are 445262, 55658, 55658 respectively.various base line models are considered for comparison and the At-LSTM model produces the highest maximum accuracy of 65.53%.

[13] designed a prediction model based on Bidirectional Gated Recurrent Unit(BGRU). The model considers both historical stock prices and online financial news.

The experiments examine the influence of the news on predicting the polarity of the stock change in each time interval. In addition, the model had an improvement in the accuracy on both the Standard & Poor's 500 (S&P 500) stock index prediction and the individual stock prediction.

The Keras 1.1.0 deep learning library in Python is used to implement this experimentation. The first layer is the embedded layer that used 32 length vectors to represent each word. The total number of words of interest is limited to the 20,000 most frequent words, and the rest are removed. The sequence length (number of words) in each review varies, so each review is constrained to 2,000 words, truncating long reviews and pad the shorter reviews with zero values.The next layer is the LSTM layer or BGRU layer with 128 memory units (smart neurons). Keras provides this capability with parameters on the LSTM layer and GRU layer the dropout-W for configuring 20 percent the input dropout and dropout-U for

configuring 20 percen the recurrent dropout.

Finally, because this is a classification problem .Softmax output layer is used with a single neuron and a sigmoid activation function to make 0 or 1 predictions for the two classes (up and down) in the problem

The authors predict movements over a period of more than a day. However, the influence of the news is being reduced over time. The BGRU method has obtained the best performance with an accuracy of 59.98%.
The model achieves a accuracy of 65% in predicting the individual stock price.

[14] proposed a neural network model called hierarchical complementary attention network. When the previous models takes into account only the news titles as a factor for the prediction of the stock price, the authors takes into consideration the content of the news also.

The authors proposed new neural network architecture, namely Hierarchical Complementary Attention Network (HCAN). In particular, HCAN adopts a two-level attention mechanism, which includes a word-level attention and a sentence-level attention, to quantify the importances of the words and sentences in given news. To ensure that the information captured in the title and content is complementary, a new measurement called score- inverse similarity (S-IS) is introduced for calculating the attention weights. Evaluation is made with new datasets. Experimental results show that HCAN outperforms the state-ofthe- art techniques. the financial news dataset obtained from Reuters during January 2007 to December 2012, which contains 47487 news after removing duplicates and empties is used as the dataset for financial news and the stock price data from Yahoo- Finance is also used.
The features extracted from the news content and the stock data are taken as inputs for the neural network. An accuracy of 61.38% is achieved with this model.

[15] proposed a model for stock price prediction, the model considers the psychological and behavioural features retrieved from the social media sites to predict the stock price. The dataset contains 18.39 million tweets from 2780 stocks. Only a part of this is considered for research which contains
6.48 million tweets from 284 thousand users. The reason for choosing this is due to its stability in the particular period and activeness of the share market. Support vector machine and the multilayer perceptron algorithms are used. The maximum accuracy achieved only with stock specific features is 57%.

[16] addresses the problem of handling the joint impact that arises when multiple sources such as the events related to products or stocks, stock data are considered for predicting the stock price. A heterogeneous data which contains the web news and stock quantitative data is used as the dataset. An extended coupled hidden markov model is proposed and compared with the base line models such as Support vector machine, tensor based learning approach, coupled matrix and tensor model etc. Various parameters such as the accuracy, F1score and Mathews correlation coefficiency is calculated and the extended coupled hidden markov model proves to produce better result in predicting the movement of the stock price.The accuracy achieved is 62.70%

[17] tries to analyse the mail transactions that take place within the company between the persons holding key responsibilities and their relation with the changes in the stock price. It is also observed that there exists such a relationship and the stock price movements can be predicted with the content of these mail communications. Discretization of communication network and stock price is made first and the prediction of stock price with communication network is considered as a pattern discovery problem. Then prediction rules are formulated to find the stock price from the communication network. The Enron e- mail dataset is used for this purpose and itcontains
12.6 million words. The average accuracy achieved is 80%.

[18] proposed a multi-task stock prediction model that takes multi-source data fusion in addition to the stock correlations. the model first utilizes tensor to integrate the multi-sourced data, including financial Web news, investors' sentiments extracted from the social network and some quantitative data on stocks. An improved sub-mode coordinate algorithm is proposed by the authors. It also employs the Long

Short-Term Memory neural network to predict the stock fluctuation trends. The multi-source data used in the experimentation are Quantitative data, web news data and Social media data. The quantitative data of stocks of the two datasets are both collected from *Wind* which contains various financial data. The parameters considered are stock turnovers, P/E ratios, P/B ratios and PCF ratios that are commonly used indices for stock trading and evaluation as quantitative indexes. In addition, to that industry index is also collected to calculate the stock correlation matrix. For web news data, the data collected are 90,361 news articles for A-share stocks and HK stocks respectively from Wind, including the titles and the publication time. Each article is assigned to the corresponding stock. For Social media data, 1179,926 postings from Jan.1, 2015 to Dec. 31, 2016 for the stocks is collected from Guba. 6 base line models are used for comparison. Though this model (Sub-Mode coordinate algorithm +LSTM) produces higher accuracy than the other models, the accuracy produced is only 62.97% and 63.31% with two different datasets.

[19] proposed a Numerical based Attention method for dual source stock market prediction. The model selects the stock data , based on the numerical data retrieved from the news data.The authors have proposed an attention-based method to effectively exploit the complementarily between news and numerical data in predicting stock prices. The stock trend that has been reflected in the news is converted in to numerical data. The method can effectively filter the noise and make full use of the trend information in news.
Three datasets are built by the authors from two sources, China Security Index 300 (CSI300) and Standard & Poor's 500 (S&P500). The parameters considered for evaluation are Mean square error and accuracy. The maximum accuracy achieved is 60.8%.

[20] proposed a model for predicting the stock price, which combines the concept of sentiment analysis with support vector machine, day- of - week is considered and the result shows an accuracy of 89%. Other concepts such as opinion mining are also used for stock price prediction [21] but shows only lower accuracy. Similar low accurate result is also noted in [22] where CNN Model is employed.

Various models and state of art mechanisms that are used for predicting the stock trend is discussed and most of the models seem to provide only less accurate results. The next section explains the recurrent neural network model.

III Experimentation

The dataset contains stock data of 1274 days from 02.01.2015 to 24.12.20219 leaving the holidays. The parameter that is considered for the prediction is the opening price of the day. The data is split into train data and the test data. The data of the last 20 days are considered as the test data. First the model is trained with the training data and the stock price of the last 20 days are predicted which is compared with the actual values.

The data that are fed to the model is normalized with the scalar function. The normalized value of the individual values is calculated with the following formula.

$$x_{new} = (x - x_{min})/(x_{max} - x_{min})$$

where,
x is the value which has to be normalized. X_{min} is the minimum value in the considered data and X_{max} is the maximum value in the considered data.

The recurrent neural network models that were build and compared consists of various time steps and different number of LSTM Layers. The core idea is to use LSTM layers as a method of increasing the efficiency of the model by addressing the gradient descent problem that commonly occurs in recurrent neural networks. Number of epochs used in all the models is 100 and the number of units in each LSTM layer is 50. The different combinations considered are 20 time steps with 1 LSTM Layer ,20 time steps with 4 LSTM Layers, 60 time steps with 1 LSTM Layer and 60 time steps with 4LSTM Layers .the output layer consists of a single unit. Time steps here refers to the number of days for which the error value is

calculated and back propagated to solve the gradient descent problem and hence reduce the error value. Usage of LSTM layers helps in considerably reducing the error rate. The performance parameters that are considered are explained variance score, r2 score and pearson coefficient.

Explained variance score is usually used for predicting the explained variance regression score.

$$\text{Explained variance score} = 1 - \frac{variance\{y - y'\}}{variance\{y\}}$$

Where,

y= target output y'=estimated target output

R2 score is defined as follows

$$\text{R2 score} = 1 - \frac{\sum_{i=1}^{n} (y_i - y'_i)^2}{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}$$

Where,

y_i = Actual value

y'_i = Predicted Value

$$\hat{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

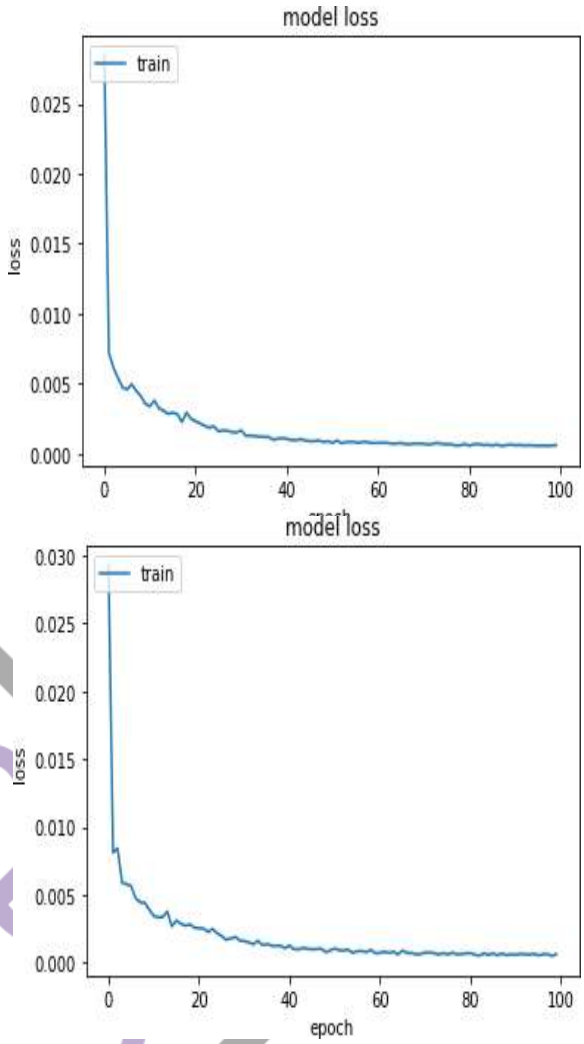
Both these values are good measures of regression models. The next performance parameter considered is the pearson coefficient. Pearson coefficient is normally used for finding the relation between two set of values. But pearson coefficient is calculated here between the actual stock price and the predicted stock price in order to analyse how similar the trend lines are.

The values obtained for the above specified parameters are discussed in the next section.

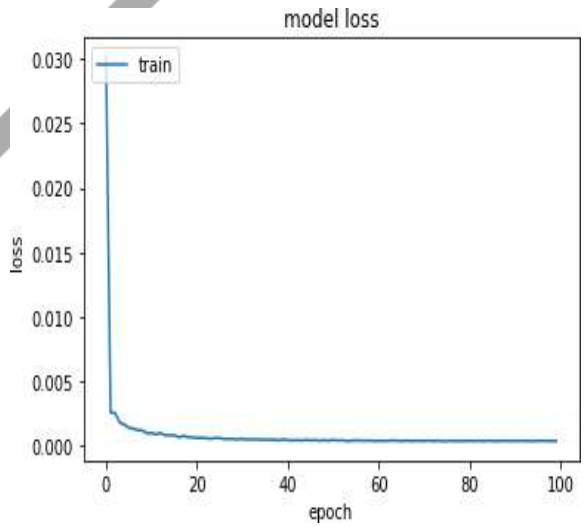
IV Experimental Results

The experimental results obtained with the recurrent neural networks for different number of time stepss and different numbers of LSTM Layers are given below.

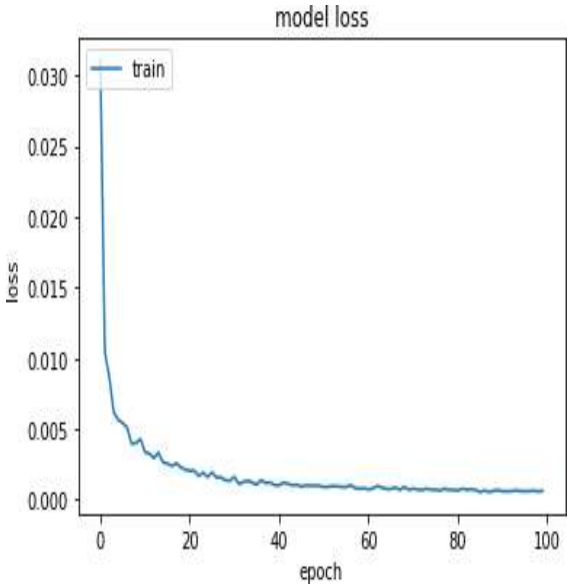
The following figure represents the Loss occurred in the model



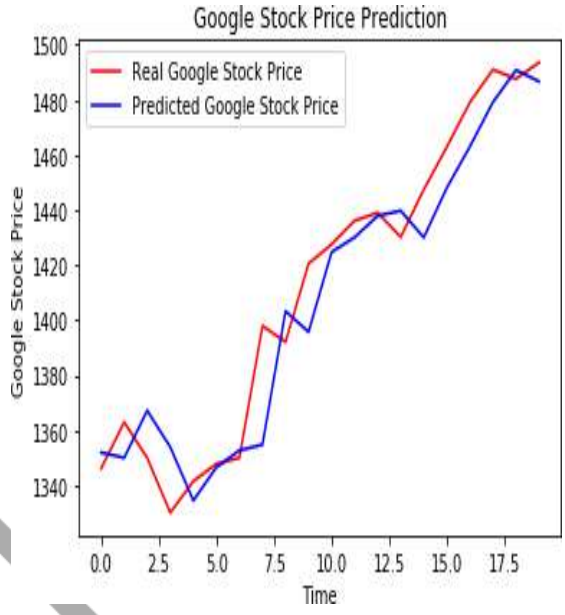
b)20 time steps 4 LSTM Layers



c)60 time steps 1 LSTM Layer

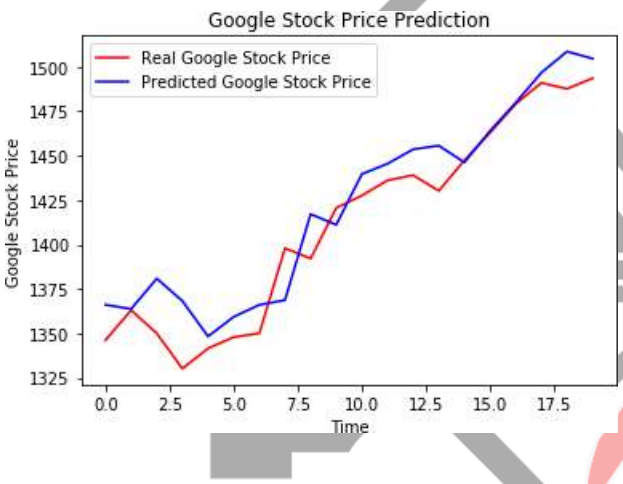


d)60 time steps 4 LSTM Layers

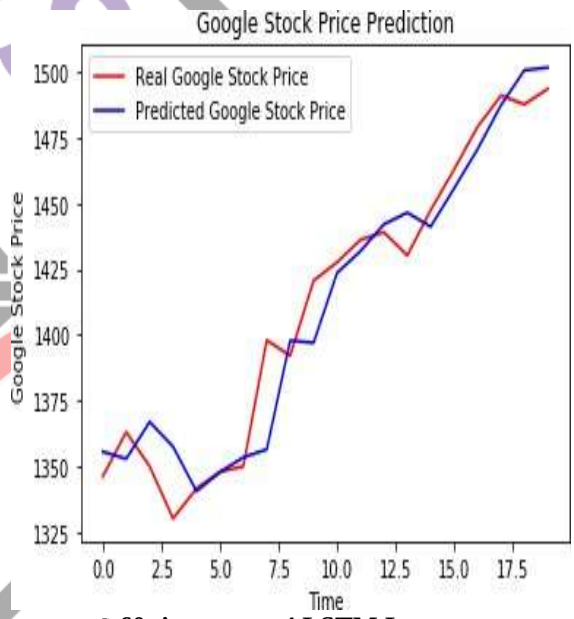


c) 60 time stamps 1 LSTM Layer

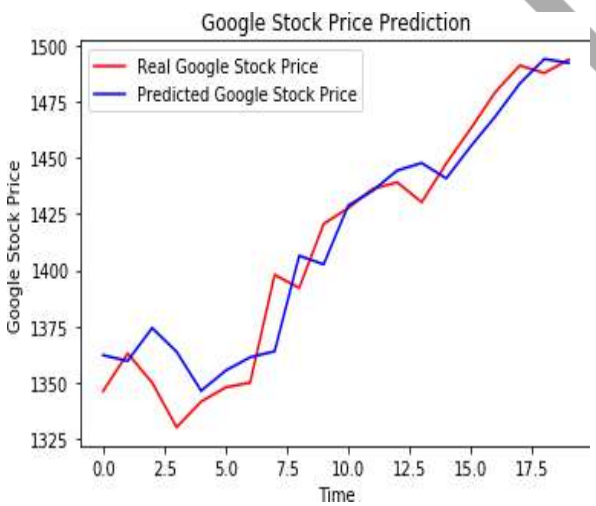
Figure 1 Model Loss of RNN for different time stamps and different Number of LSTM Layers



a) 20 time stamps 1 LSTM Layer



d) 60 time stepss 4 LSTM Layers



b) 20 time Stamps 4 LSTM Layers

Figure 2 stock price prediction with RNN for different time stamps and different Number of LSTM Layers

The performances of the models are made with explained variance score and r2score.

The following figure represents the explained variance score of the four models.

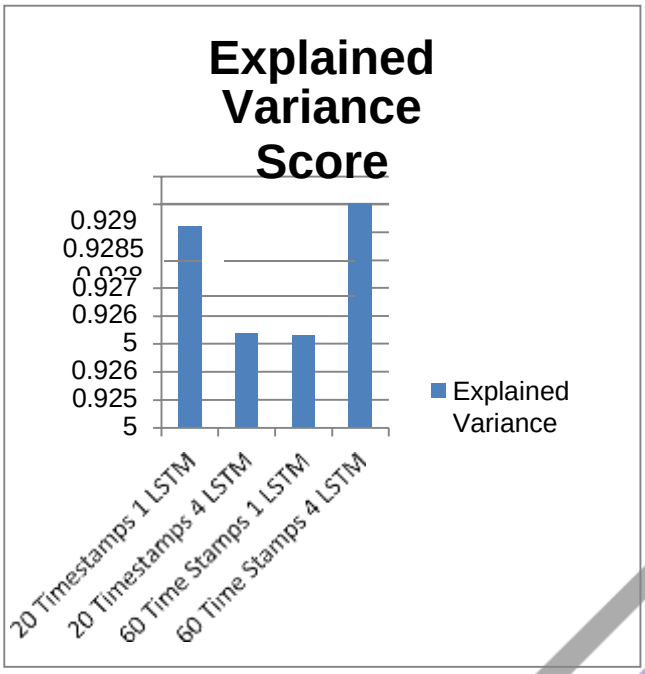


Figure 3 Explained Variance score for different models

The following figure represents the r2score of the four models.

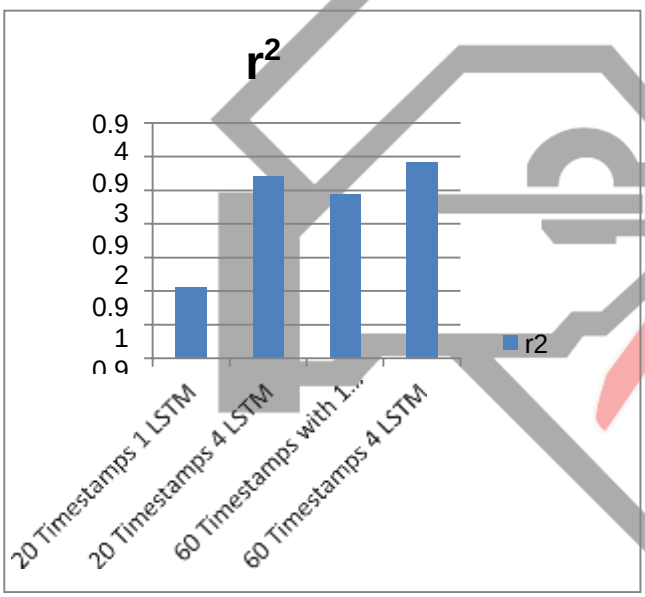


Figure 4 R^2score of different models

The pearson coefficient is also calculated for the four models in order to get an idea of the similarity in trend between both the actual and the predicted values. Obtained pearson coefficients are given below.

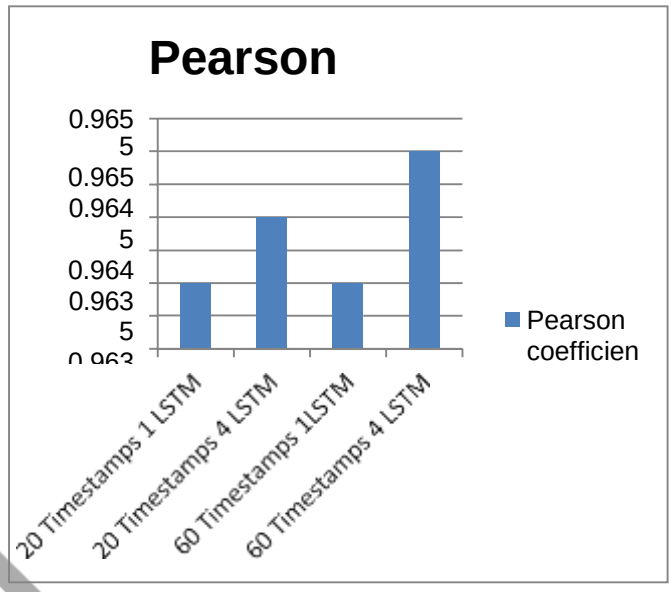


Figure 5 pearson coefficient of different models

Though the pearson coefficient is used to find the relation between two sets of values, it is believed that it can also be used as a mechanism in comparing the trend of the actual and the predicted stock prices. It is observed from the results the model with 60 Time stepss and 4 LSTM Layers perform well than the other models.

V conclusion

Recurrent neural networks with different time steps and different number of LSTM layers are employed and tested with their performance in predicting the stock price and in specific the trend of the stock price movement. It has been observed that all the models performs considerably well and comparatively the model with 60 time steps and 4 LSTM Layer performs well than the other models.

Reference

[1] C. Stolojescu, I. R. Cristina, S. Moga, P. Lenca, and A. Isar, ``A wavelet
based prediction method for time series," in *Proc. Int. Conf. Stochastic*
Modeling Techn. Data Anal., Chania, Greece, 2010.

[2] C. Stolojescu, A. Cusnir, S. Moga, and A. Isar, ``Forecasting WiMAX
BS traf_c by statistical processing in the wavelet domain," in *Proc. Int.*

Symp. Signals, Circuits Syst. (ISSCS), Jul. 2009, pp. 1_4.

[3] A. J. Conejo, M. A. Plazas, R. Espinola, and A. B. Molina, ``Day-ahead electricity price forecasting using the wavelet transform and ARIMA models," *IEEE Trans. Power Syst.*, vol. 20, no. 2, pp. 1035_1042, May 2005.

[4] N. Nava, T. Di Matteo, and T. Aste, ``Financial time series forecasting using empirical mode decomposition and support vector regression," *Risks*, vol. 6, no. 1, p. 7, 2018.

[5] W. Bao, J. Yue, and Y. Rao, ``A deep learning framework for financial time series using stacked autoencoders and long-short term memory," *PLoS ONE*, vol. 12, no. 7, 2017, Art. no. e0180944.

[6] F. Ye, Z. Liming, Z. Defu, F. Hamido, and G. Zhiguo, ``A novel forecasting method based on multi-order fuzzy time series and technical analysis," *Inf. Sci.*, vols. 367_368, pp. 41_57, Nov. 2016.

[7] A. K. Rout, ``Forecasting _nancial time series using a low complexity recurrent neural network and evolutionary learning approach," *J. King Saud Univ.-Comput. Inf. Sci.*, vol. 29, no. 4, pp. 536_552,2017.

[8] Kim, Misuk et al. "Stock price prediction through sentiment analysis of corporate disclosures using distributed representation." *Intell. Data Anal.* 22 (2018): 1395-1413.

[9] Zhang, Xi & Qu, Siyu & Huang, Jieyun & Fang, Binxing & Yu, Philip. (2018). Stock Market Prediction via Multi-Source Multiple Instance Learning. IEEE Access. PP. 1-1. 10.1109/ACCESS.2018.2869735.

[10] Dang, L. Minh & Sadeghi-Niaraki, Abolghasem & Huynh, Huy & Min, Kyungbok & Moon, Hyeonjoon. (2018). Deep Learning Approach for Short-Term Stock Trends Prediction based on Two-stream Gated Recurrent Unit Network. IEEE Access. PP. 1-1. 10.1109/ACCESS.2018.2868970.

[11] Lee, Wonho & Suh, Bongwon. (2018). Modeling Stock Prices with Text Contents in 10-Q Reports. 224-229. 10.1109/SNPD.2018.8441051.

[12] Liu, Huicheng. (2018). Leveraging Financial News for Stock Trend Prediction with Attention- Based Recurrent Neural Network. arXiv:1811.06173

[13] Huy D. Huynh, L. Minh Dang, and Duc Duong. 2017. A New Model for Stock Price Movements Prediction Using Deep Neural Network. In *Proceedings of the Eighth International Symposium on Information and Communication Technology* (SoICT 2017). ACM, New York, NY, USA, 57-62. DOI: https://doi.org/10.1145/3155133.315520

[14] Liu, Qikai & Cheng, Xiang & Su, Sen & Zhu, Shuguang. (2018). Hierarchical Complementary Attention Network for Predicting Stock Price Movements with News. 1603-1606. 10.1145/3269206.3269286.

[15] Zhang, Xi & Shi, Jiawei & Wang, Di & Fang, Binxing. (2017). Exploiting Investors Social Network for Stock Prediction in China's Market. Journal of Computational Science. 10.1016/j.jocs.2017.10.013. [16]Zhang, Xi et al.(2018) "Enhancing stock market prediction with extended coupled hidden Markov model over multi-sourced data." *Knowledge and Information Systems* pp: 1-20.

[17] P. Zhou, K. C. C. Chan and C. X. Ou, "Corporate Communication Network and Stock Price Movements: Insights From Data Mining," in *IEEE Transactions on Computational Social Systems*, vol. 5, no. 2, pp. 391- 402, June 2018.

[18] J. Huang, Y. Zhang, J. Zhang and X. Zhang, "A Tensor-Based Sub-Mode Coordinate Algorithm for Stock Prediction," *2018 IEEE Third International Conference on Data Science in Cyberspace (DSC)*, Guangzhou, 2018, pp. 716-721.

[19] Liu, Guang & Wang, Xiaojie. (2018). A Numerical-based Attention Method for Stock Market Prediction with Dual Information. IEEE Access. PP. 1-1. 10.1109/ACCESS.2018.2886367.

[20] R. Ren, D. D. Wu and T. Liu, "Forecasting Stock Market Movement Direction Using Sentiment Analysis and Support Vector Machine," in *IEEE Systems Journal*, vol. 13, no. 1, pp. 760-770, March 2019.

[21] Ali Derakhshan, Hamid Beigy,Sentiment analysis on stock social media for stock price movement

prediction,Engineering Applications of Artificial Intelligence,Volume 85,2019,Pages 569-578.

[22] M. Wen, P. Li, L. Zhang and Y. Chen, "Stock Market Trend Prediction Using High-Order Information of Time Series," in *IEEE Access*, vol. 7, pp. 28299-28308, 2019.