Predict Stock Price using Deep Learning

Hira Farooqi
Department of Computer science
Ripha International University
Lahore, Pakistan
hirafarooqi14@gmail.com

Warda Fiaz
Department of Computer science
Riphah international university
Lahore, Pakistan
Wardafiaz4@gmail.com

Sidra Rani
Department of Computer science
Riphah international university
Lahore, Pakistan
17mcs1556@gmail.com

Khadija Tahir
Department of Computer science
Riphah international university
Lahore, Pakistan
khadijatahir65@gmail.com

Abstract—the way of prognosis the stock prices have been a tough labor or effort for a lot of analyzers and investigation. Actually inspectors got maximum engrossment in the exploration area of stock price prognosis. On behalf of the quality and victorious outlay, a lot of researchers and analyzers. The dataset was preprocessed accommodate for real analysis so in this paper I will also concentrated on data preprocessing of the fresh dataset subsequence to preprocessing data. I will evaluate the use of RNN and LSTM The methodology of predicting stock price is explained in numerous ways. For example moving averages, linear regression, k-nearest neighbors and ARIMA and Prophet. These are the methods that one can examine by own self and differentiate their labor with the KERAS RNN and LSTM.

Keywords—Stock price prediction, Data preprocessing, Python, recurrent neural network (RNN), Long short term memory (LSTM), mean square error (MSE)

I. INTRODUCTION

The capital retail is fundamentally an accumulation of various purchaser and trader. Stock in common constitute or possession. Assert or declare on business by a specific separation of people. The aim to control the upcoming value of commodities market as known as stock market divination. The prophecy is look forward to be strong precise and logical and the structure must work according to the actual scenario plan and should be effectively.

Data preprocessing is the part of information of data extract and convert data into a logical arrangement sensitive data is ordinal or generally deficient and usually accommodate or carry many issues. The data preprocessing require examine or investigate values working for unconditional break in to training or coaching and test set and finally a characteristics or attributes scaling to control .

In the money affairs world stock traffic or market is one of the supreme project or enterprise. Board of trade projection is rent or take action of testing to control or govern the coming value or specific numbers of a property is different or additional or extra economic tool operate on an economic traffic. This paper describes the projection of a capital work on deep investigation or research. The scientific or technological and basic or the time string or train examination is cast off or second hand by nearly the bondholder while manufacturing the price prediction. The computer language is pre owned to project the capital retrial employee of deep research or studying the python.

Throughout this paper I also use methodology of deep learning as well as the implementation is on the RNN and LSTM. Play is a deep neural network with loops which can hold knowledge throughout the node.In short, recurrent neural To warn the future activities, the Channel utilizes its knowledge about past encounters. Recurrent frameworks are worthwhile because they can loop variables that start the API to more complex operations.
LSTM is the forward category of recurrent neural network in which the details are associated with old condition preserve. These are dissimilar from RNN as and RNN's proceed to discover the connection between the previous and ongoing information this specifies the interring of detail which is the proximately compact then that to LSTM. The leading motive at

the back of using this model market prognosis is that a prognosis on the basis of huge File Number and usually basis on the long term history of the market so LSTM manage and adjust error by providing the help to the RNN in to holding data for previous step madding prognosis more preside since stock market includes working of large information the incline concerning to the array.

Fig 1: Framwork
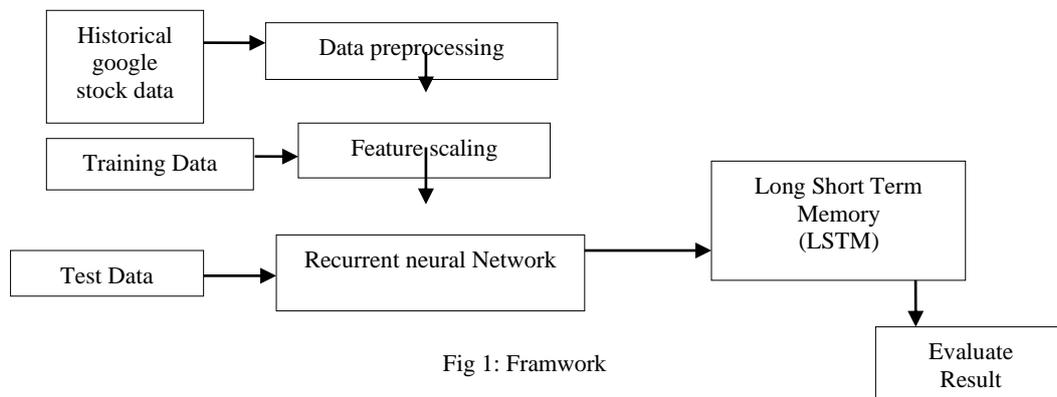
II. LITERATURE REVIEW

Our literature survey was largely about discovering similar programming algorithm, to see if they can be adapted to our usage case ,i.e operating on market price in real time info. Online Stock Learning, and Collection of Online Features.However, as we couldn't consider any plausible adaptation i agreed to look at these for stock price forecast Analyze the major drawbacks of the existing systems Only the same, and see how we can build on them. I had also zeroed upon this similarity of simulation models (with in sense of inventory data Dynamic, temporary inter stock rates) as that of the basic problem that i wanted to resolve.

A short notice the hunt for common answers to the above problem has brought us to RNNs are and LSTMs. Usage of an LSTM after you have agreed I consulted a neural network to predict stocks Number of papers studying gradient descent concept And the separate forms. I concluded our survey in literature View how you're using the ancestry slope to ease the loads of the LSTM network and to improve the process.

III. PROBLEM STATEMENT

The prediction of the stock market is described as an attempt identify and offer a robust understanding of stock prices

Citizens should learn the economy and asset values and predict them. The latest earning figure is usually described using the dataset. Use the data collection. Therefore, you may not focus on a single dataset be appropriate for the forecast and can provide an outcome imprecise. So, we are looking at the analysis deep learning and convergence of various datasets predict the patterns in the inventory and the business.The issue of supply measurement price stays a commodity problem when there isn't a better algorithm for stock market prediction suggested. Forecasting the success of the stock market it's pretty rough. The bond process normally takes place the thoughts of the thousands of creditors have been calculated. The analysis of the financial markets includes capacity to forecast the impact on creditors of recent developments. There may be occurrences a declaration by the government official, a slice of business campaigns reports concerning fraud etc. It can also be a global event just as fast currency and asset fluctuations and so on. All such incidents impact company profits, which in effect impacts investors' emotions are influenced. This is not within the reach of nearly all analysts will forecast accurately and reliably this fully virtualized. Both these considerations lead to market price it's really impossible to estimate. When the appropriate details has been obtained, could then be used to train and construct a computer predictive outcome.

Common stock market research methods and stock price forecast provide qualitative assessments that refer at a stock's historical history and overall company's reputation as whole and, or mathematical studies that deal exclusively with crushing statistics and assessing trends of stock price fluctuations. Later on is also done by genetic algorithms (GA) or artificial neuronal models (ANNs), yet they struggle to catch a link between market values in the context of long-term temporal based items Another major problem is leveraging RNNs for stock prediction. These are usually induced by 2 variables: the weight is initialized by opportunity and the amounts nearer to a network edge.Reducing equity market research is an effective way to the input data dimensionality and the function added selection of shortlist algorithms for a core set of functions with the maximum effects on asset rates or the exchange rate through currencies Mercados. This approach does not, however, find long-term investing approaches, because it does not take the entire background of consideration of trends; in fact there is no allowance for Exterior detection.

Research Questions:

- When can validation accuracy be greater than training accuracy for Deep leanring models?
- How to preprocess time series data?

IV. METHODOLOGY

The application of various factors, such as network topology, planning method, etc., will create specific types of algorithms.. We have the trial for considered Persistent Neural and Long-short term Network memory. This segment addresses our program approach. We have many phases in our system, which are as follows:

Step 1: Raw Data

Raw data usually applies to data tables In which each column contains a sighting, so each column represents a matrix reflecting the characteristics of each sighting. Data are often made reference to in this layout as clean, flat, main, nuclear, and unit record data.

Step 2: Pre-processing of data

The pre-processing period includes

A) Experience flexibility: Part of data reduction along with especially essential for numerical knowledge

B) Data transformation: Normalisation.

C) Data cleaning: Full the incomplete details.

D) Software Management: Information File Application.

Upon transformation of the data collection into a new server, server Knowledge is separated into preparation and research sets such that assess. The teaching qualities here are considered as the most modified principles. Research data were held at 5-10 per cent complete dataset.

Step 3: Feature scaling

Just the apps to be inserted into this layer Neural Network is picked. We must pick the feature from the date, free , high, small, close and distance.

Step 4: Neural Network Preparation

Information are now loaded into the neural network and informed applying arbitrary preferences and weights to prediction. Our LSTM layout consists of a sequential input stratum next 2 LSTM layers and a thick ReLU layer activation, and eventually a thick, flat performance layer Item enabled. Keras applies the application of the neural network further:

```
regressor = Sequential()
regressor.add(LSTM(units     =     50,return_sequences     =
True,input_shape = (x_train.shape[1],1)))
regressor.add(LSTM(units = 50,return_sequences = True))
regressor.add(Dropout(0.2))
regressor.add(LSTM(units = 50))
regressor.add(Dropout(0.2))
```

Code: NN implementation in keras

Step 5: TrainingGeneration

The value produced by the output layer in this layer the goal amount of the RNN is contrasted. The mistake difference between goal and output achieved and the back propagation algorithm reduces the interest that changes channel weight values.

Analysis:

In determining the mean square error (MSE), I can use method. Inaccuracy or move from goal to target the interest received the value of Standard deviation are reduced. The real number of MSE in all the error's mean / average rectangle. The framework MSE is very popular and allows it very nice metric for specific purposes of number projections defect. MSE is an outright error in contrast to the same expect amplifies and punishes severe errors.

$$\mathbf{MSE} = \frac{1}{n} \sum_{i=1}^{n} (y_i - \tilde{y}_i)^2$$
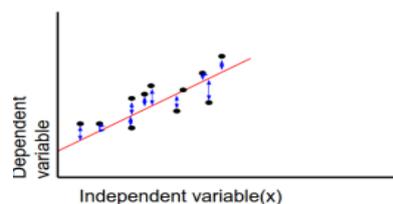

Fig 1: MSE value

Experimental Work:

In this paper I am using the recurrent neural network and long short term memory (LSTM) for storying the scalable and trained data. LSTMs are particularly effective as they can store past knowledge in sequential-prediction issues.

```
                    ┌──────────┐
                    │  INPUT   │
                    └──────────┘
                         │
                         ▼
                    ┌──────────┐
                    │  LSTM    │
                    └──────────┘
  Dropout: 0.2          │              Output: 185
                        ▼
                    ┌──────────┐
                    │  LSTM    │
                    └──────────┘
  Dropout: 0.2          │              Output: 185
                        ▼
                   ┌───────────┐
                   │  DENSE    │
                   └───────────┘
                        │              Output: 32
                        ▼
                  ┌────────────┐
                  │  DENSE     │
                  └────────────┘
                        │                   1
                        ▼
                  ┌────────────┐
                  │  OUTPUT    │
                  └────────────┘
```
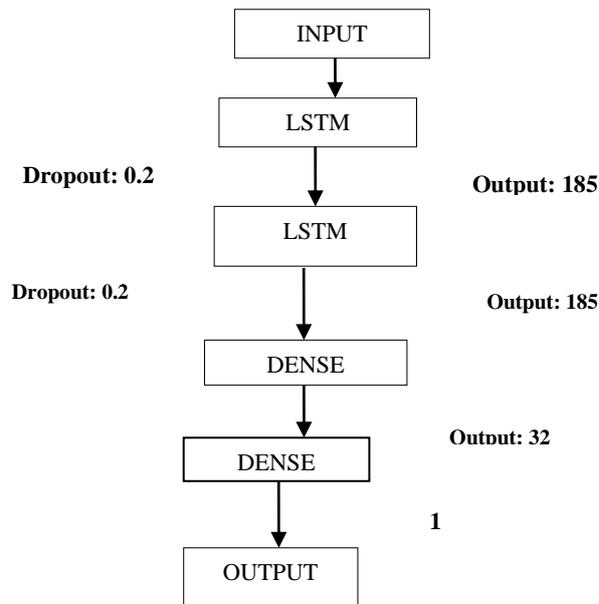
Fig 2: LSTM Layer

LSTM is the specialized Recurrent-NeuralNetworks (RNN) variant for which information can be accessed the prior condition is still in nature. They differ from RNNs because long-term addictions and RNNs are involved works to identify the connection between existing and traditional the information currently available. It displays the time details to LSTM is comparatively lower. The key goal behind such a concept market forecasting is based on forecasts huge volume of data and are usually subject to market long-standing history[6]. LSTM, then regulates fault by supplying RNN assistance retention of knowledge for older phases Forecast. It is critical in our situation, because an inventory 's previous price is essential to its potential demand forecast. In order to help load and modify google data sets, I will continue with importing NumPy for mathematical estimation, Mattplotlib for drawing diagrams and Pandas. The next move is to install the Free and Large Sectors I am using in our simulation for the training dataset. I check the dataset head and its checking for that which type of data i used. I learned from history practice with profound learning models that data must be scale for optimal efficiency. We can use MinMaxScaler for Scikit- Learn and scaled our dataset to null or one. LSTMs assume that our data would typically be in a 3D array in a specified system. I start generating and transforming data into an array with NumPy in 60 time stages. Then, using X train tests, 60 timestamps or each function or so per, i can turn that information into a 3D component series.

Layer	Type	Shape	Param #
LSTM_13	LSTM	60, 50	10400
Dropout_7	Dropout	60, 50	0
LSTM_14	LSTM	60, 50	20200
Dropout_8	Dropout	60, 50	0
LSTM_15	LSTM	50	20200
Dropout_9	Dropout	50	0
Dense_3	Dense	1	51

Table 1: Regressor Summary

In this paper I am the long short term memory which using to shape the trained data that have 10400 parameters and the 60 is the timestamps (it is tells that how much time required for 50 units diamension). In the table 1 have LSTM 15 have 50 unites that have not show the time stamps for predicting the price it reliable for the forecasting. DENSE is shows that the density of the regressor that apply prediction its have the shape 1 and it parameters are included in 51 otherwise total parameter are 50,851 and also train parameter have same 50,851. It means that the training all the parameters.The LSTM layer is added and some dropout layers are added to prevent overlap. The following reasons are used to incorporate the LSTM layer:
1- 50 units that are the production room dimensionality
2- return sequences = True that defines if the last item will be returned in the item series or the entire series
3- As the form of our training set, input shape.

If the drop-out thresholds are specified, we state 0.2, which implies that 20% of the layers are removed. Then, the Dense layer defining the unit output is applied. We then use the common adame optimizer to compile our model and describe the failure as a mean squarred error. The mean of the squared errors is calculated. I am going to fi next. Then I suit the model with a batch size of 32, operating over 100 epochs. Note that it can take several minutes to finish working based on the configuration of your machine.

I must do a few items after load in the test set in order to predict potential market prices:

1- Combine the instruction package and the 0 axis evaluation package.
2- Set period to 60 (as seen above)
3- To convert the latest dataset, using MinMaxScaler
4- Restructure the sample as before

I must use the inverse transform to bring market prices back in a functional layout after creating the predictions. Garaphical prediction of the price that have show in the graph are as follow:
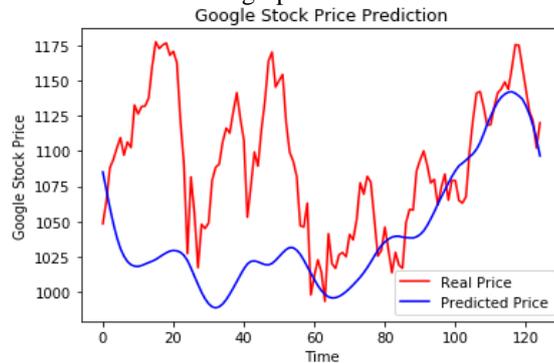

Fig 2: Google stock price prediction

In that graph easily understand the stock prices by RNN and LSTM. So two graph lines are mentioned in that graph as shown in fig. Red line graph shows the actual price of stock in market which can trade on actual price onanother hand , blue line graph shows the predictive prices which are owned by ourselfs. In this, procedure stock price of product will be reduce to actual price and gave the discounts on products when prices is low in market then the cutomer purchase the product easily and product sale will be up.when the sale of product and demand of product is up then retrieve to its original price. Then market will be the same price and the bussniss will grow and grow.

Conclusion

In this paper I am using deep learning technique that have two algorithms one Neural network is recurrent (it is in the network of using LSTM) for the storage of predicted data and it can predict the stock price that have prediction in the unites and time stamps. So many stock pricing strategies, such as shifting averages, linear regression and K-Nearest neighbours, are also available. Such procedures can be tested and connected individually to the Keras LSTM. A variety of many other stock Price prediction strategies are usable, including changing average, linear regressiong and K-Nearrest neighbours. These methods may be independently checked and linked to the Keras LSTM.

References

[1] RakhiMahant, TrilokNathPandey, Alok Kumar Jagadev, and SatchidanandaDehuri ―Optimized Radial Basis Functional Neural Network for Stock Index Prediction,‖ International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT) - 2016.

[2] Kai Chen, Yi Zhou and FangyanDai ―A LSTM-based method for stock returns prediction: A case study of China stock market,‖ IEEE International Conference on Big Data,2015.

[3] A.U.S.S Pradeep, SorenGoyal, J. A. Bloom, I. J. Cox, and M. Miller, ―Detection of statistical arbitrage using machine learning techniques in Indian Stock market,‖ IIT Kanpur, April 15, 2013.

[4] Prashant S. Chavan, Prof. Dr. Shrishail. T. Patil ―Parameters for Stock Market Prediction,‖ Prashant S Chavan et al, Int.J.Computer Technology & Applications, Vol 4 (2),337-340.

[5] NeelimaBudhani, Dr. C. K. Jha, Sandeep K. Budhani―Prediction of Stock Market Using Artificial Neural Network,‖ International Conference on Soft Computing Techniques for Engineering and Technology (ICSCTET)- 2014.

[6] SharvilKatariya, Saurabh Jain―Stock Price Trend Forecasting using Supervised Learning Methods.

[7] DhirajMundada, GauravChhaparwal, SachinChaudhari, and TruptiBhamare― Stock Value Prediction System,‖ International Journal on Recent and Innovation Trends in Computing and Communication, April 2015

[8] Hochreiter, Sepp, and Jürgen Schmidhuber― Long short-term memory,‖ Neural computation 9.8 (1997): 1735-1780