International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 6 June 2022, pp: 711-724 www.ijaem.net ISSN: 2395-5252

Estimation of Company Stock Prices Using MachineLearningalgorithm

N.Udaya Kumar[1], L.Surya Roshini[2], K. SaiCharan Teja[3] , K. Tarun[4], K.Satya Sai Chandu[5]

Department of CSE, Raghu Institute Of Technology,Visakhapatnam,Andhra Pradesh,India
Email:
Student[2,3,4,5]B.Tech(COMPUTERSCIENCEENGINEERING)
Raghu Institute of Technology,Visakhapatnam,AndhraPradesh,India

Submitted: 01-06-2022 Revised: 10-06-2022 Accepted: 15-06-2022

ABSTRACT
Machine learning is a mechanism of data analysis that automates analytical model building. It is a type ofAI based on the idea that systems can learn from data, identify patterns and make decisions with leasthuman intervention. The process of learning begins with observation of data, such as examples, directexperience, or instruction, in order to see for patterns in data and make better decisions in the future basedon the examples that we supply. The primary goal is to allow the computers learn automatically withouthuman intervention and adjust actions accordingly. Machine learning has important applications in theStockprice prediction.
The art of forecasting the stock prices has been a difficult task for many of the researchers, investors andanalysts. In fact, investors are highly attentive in the research area of stock price prediction. For a goodandsuccessfulinvestment,manyinvestorsare keen inknowingthefuturestate ofthe stock market.
In this way, we present a recurrent neural network (RNN) and Long Short-Term Memory (LSTM)approachto predictstockmarketindices.

I. INTRODUCTION

1.1 PURPOSE

The demand of stock market trading is growing rapidly, which is encouraging researchers to find out newmethodsforthepredictionusingnewtechniques.Theforecastingtechniqueisnotonlyhelpstheresearchers but it also helps investors and any person dealing with the stock market. In order to helppredictthe stock price, aforecastingmodelwith goodaccuracyisrequired.

1.2 SCOPE

The stock market refers to common markets that exist for issuing, buying, and selling stocks that trade ona stock exchange or over-the-counter. Stocks, also known as equities and F&O's represent fractionalownership in a company, and the it is a place where investors can buy and sell ownership of suchinvestibleassets.Anefficientlyfunctioningstock marketiscontemplatecriticaltoeconomicdevelopment,asitgivescompaniestheabilityto quickly access capital fromthepublic.

1.3 MOTIVATION

Accuracy plays an important role to predict the stock market. Although many algorithms are available forthis purpose, electing the most accurate one continues to be the fundamental task in getting the bestresults.In orderto reachthose result,inthiswe have used LSTMalgorithm.

1.4 WHATISSTOCKMARKET?

The stock market refers to common markets that exist for issuing, buying and selling stocks that trade onastock exchange orover-the-counter.

Stocks,alsoknownasequitiesandF&O'Srepresentfractional ownershipinacompany,andtheit isaplacewhere investors canbuy and sell ownership of such investible assets. Anefficientlyfunctioningsharemarketisclassifydifficulttoeconomicgrowth,asit givescompaniestheabilityto quickly access capitalfromthe public.

A share market is the collection of buyers and sellers of stocks (also called shares), which representownership claims on businesses; these may include securities listed on a common stock exchange, as wellas stock that is only traded privately, such as shares of private companies which are sold to investorsthroughequity rushfunding ways.

Stocks can be classify by the country where the company is located. For example, Nestle and Novartis arelocated in Switzerland and traded osn the SIX Swiss Exchange, so they may be considered as part of theSwiss share market.

II. LITERATURESURVEY

Nonlinearity and high volatility of financial time series have made the stock price predict is critical.However, thanks to recent growth in deep learning and methods such as long short-term memory (LSTM)and convolutional neural network (CNN) models, significant improvements have been obtained in theanalysis of this type of data. Further, empirical mode decomposition (EMD) and full ensemble empiricalmode decomposition (CEEMD) algorithms decomposing time series to different frequency spectra areamong the types that could be effective in analyzing financial time sequence. Based on these theoreticalframeworks, we create novel hybrid algorithms, i.e., CEEMD-CNN-LSTM and EMDCNN-LSTM, whichcould extract deep features and time sequences, which are finally applied to one-step-ahead prediction.The way it suggested algorithm is that when fixing these models, some collaboration is establishedbetween them that could enhance the analytical power of the model. The practical findings accept thisclaim and indicate that CNN along, side LSTM and CEEMD or EMD could enhance the predictionaccuracyand outperformothercounterparts.

Predicting Stock Prices Using Genetic Algorithms (GA) or Artificial Neural Networks (ANN's) areimplemented earlier and these algorithms can be functionated with the low accuracy and low predictions.so,we needto predictthecompany stock priceswith highaccuracyand highpredictions.

III. PROPOSEDALGORITHMS

Recurrent neural networks (RNN) are the state-of-the-art algorithm for sequential data and are used byApple's Siri and Google's voice search. It is the first algorithm that remembers its input, due to an internalmemory,whichmakesitperfectlysuitedformachinelearningproblemsthat involvesequentialdata.

IV. METHODOLOGY

Longshort-termmemory(LSTM)networksareanextensionforrecurrentneuralnetworks,whichbasically extends the memory. Therefore, it is well suited to learn from important experiences that haveverylong-time lagsin between.

LSTMs enable RNNs to remember inputs over a long period of time. This is because LSTMs containinformation in a memory, much like the memory of a computer. The LSTM can read, write and deleteinformationfromits memory.

Traditionalneuralnetworkscan’tdothis,anditseemslikeamajorshortcoming.Forexample,imagine youwant to classify what kind of event is happening at every point in a movie. It’s unclear how a traditionalneuralnetworkcouldusetsreasoningaboutpreviouseventsinthe filmto informlaterones.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 6 June 2022, pp: 711-724 www.ijaem.net ISSN: 2395-5252

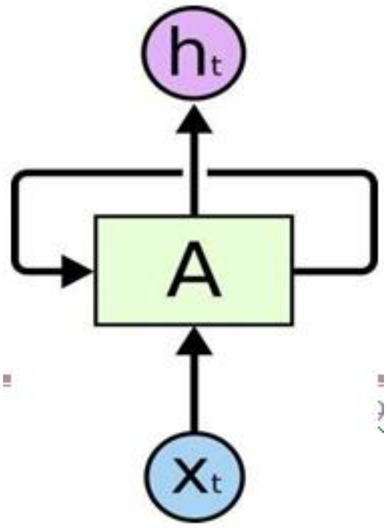


FIGURE 4.1 Recurrent Neural Networks have loops. In the above diagram, a chunk of neural network

Loopallowsinformationtobepassedfromone stepofthe network tothe next.

These loops make RNN seem kind of mysterious. However, if you think a bit more, it turns out that theyaren't all that different than a normal neural network. A recurrent neural network can be thought of asmultiple copies of the same network, each passing a message to a successor. Consider what happens if weunrolltheloop

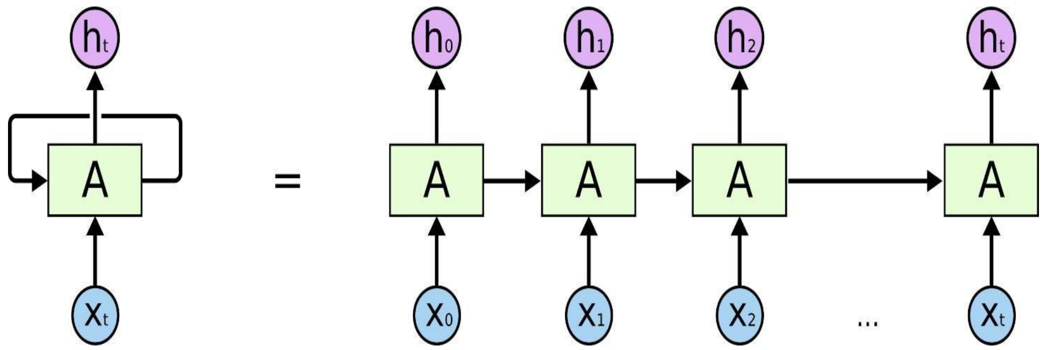


FIGURE4.2 Anunrolledrecurrentneuralnetwork.

This chain - like nature reveals that RNN are intimately related to sequences and lists. They're the naturalarchitectureof neuralnetwork to usefor suchfiles.

Oneofthe appealsof RNNis theideathattheymightbe able toconnectpreviousinformation tothe

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 6 June 2022, pp: 711-724 www.ijaem.net ISSN: 2395-5252

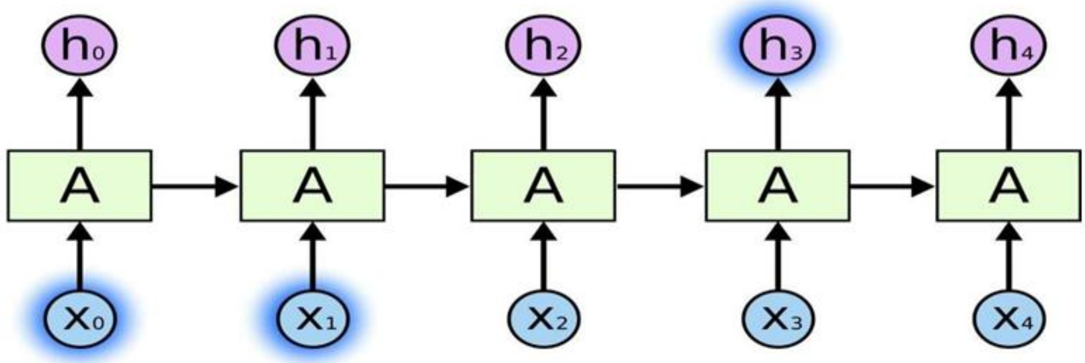
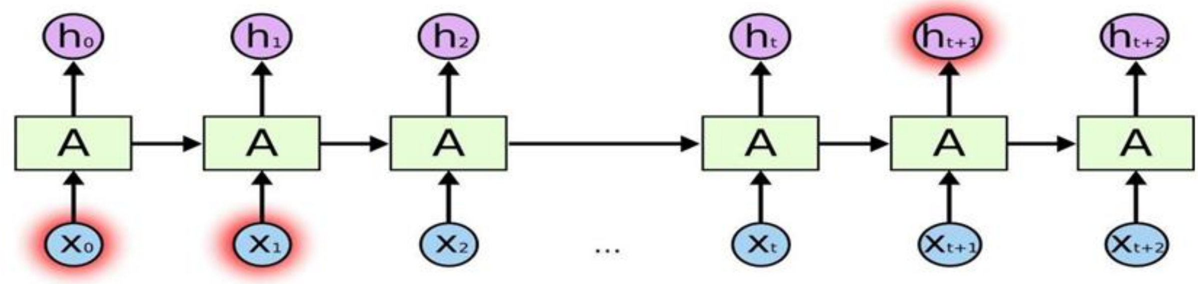


FIGURE4.3

present task, such as using previous back video frames might inform the understanding of the presentstructure.If RNNcoulddothis, they'dbe extremelyuseful.Butcanthey? It depends.
Sometimes, we only need to look at recent information to perform the present task. In such cases, wherethegapbetweentherelevantinformationandtheplacethatit'sneededissmall,RNNscanlearntousethepastinformation

But there areal so cases where we need more condition. Consider trying to predict the last word in the text"IgrewupinFrance…IspeakfluentFrench."Recentinformationsuggeststhatthenextwordislikelythe name of a language, but if we want to narrow down which language, we need the condition of France,fromfurtherback.It'sentirelypossibleforthegapbetweentherelevantinformationandthepointwhereit is neededtobecomeverylarge.
Unfortunately,asthat gapgrows,RNNsbecomeunabletoreadtoconnecttheinformation.Intheory,



RNNs are absolutely capable of handling such "long-term dependencies." A human could carefully pickparameters for them to solve toy problems of this form. Sadly, in practice, RNN don't seem to be able toread them. The problem was explored in depth by Hochreiter (1991)[German]and Bengio, et al. (1994),whofoundsome prettyfundamentalreasonswhyit mightbedifficult.

V. LSTM NETWORKS

Long Short Term Memory networks–usually just called "LSTMs"–area special kind of RNN, capable oflearning long-term dependencies. They were introduced by Hochreiter&Schmidhuber (1997).1They worktremendouslywellon alarge variety ofproblems, andarenow widely used.
LSTMsareexplicitlydesignedtoavoidthelong-termdependencyproblem.Rememberinginformation forlongperiodsoftimeispracticallytheirdefaultbehavior.

Longshort-termmemorynetworksarean extensionforrecurrentneuralnetworks,whichbasicallyextendsthe memory. Therefore it is well suited to learn from important event that have very long time lags inbetween.

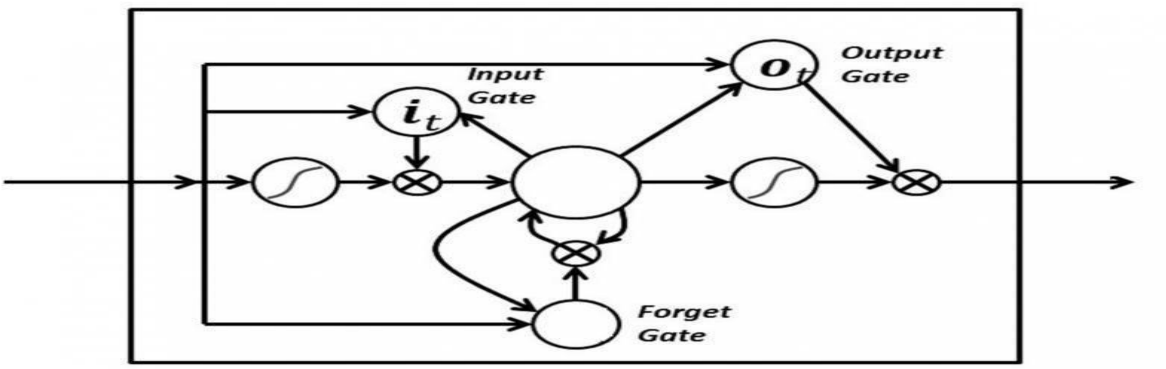


FIGURE5.1

The units of an LSTMs are used as structure units for the layers of a RNN, often called an LSTMsnetwork. This is because LSTMs contain info in a memory, much like the memory of a computer. TheLSTMscan read, write anddeleteinformationfromitsmemory.

The gates in an LSTM areana log in the form of sigmoids, meaning they range from zero to one. The factthat they are analog enablesthemtodo back propagation.

VI. PERFORMANCE ANALYSIS IN BETWEEN OF PROPSED SYSTEMS AND EXISISTING SYSTEMS

Theperformanceofourproposedstockpredictionsystem,whichusesanLSTMmodel,wasrelated with a simple Artificial Neural Network model on five different stocks of varying sizes of data. Threeclassification of stock were chosen depending upon the size of dataset. Small data set is a stock for whichonlyabout10yearsofdataishandye.g.,DixonHughes.Amediumsizeddatasetisastockforwhichdata up to 25 years is handy, with examples including Cooper Tire & Rubber and PNC Financial.Similarly, a large data set is one for which more than 25 years of stock data is available; Citi groupand American Airlines are ideal examples of the same. Variables such as the training split, dropout,numberof layers, numberof neurons,andactivationfunction remainedthesamefor alldatasetsforbothLSTMandANN

Data Size	Stock Name	**LSTM (RMSE)**	**ANN (RMSE)**
Small	Dixon Hughes	**0.04**	**0.17**
Medium	Cooper Tire & Rubber	**0.25**	**0.35**
Medium	PNC Financial	**0.2**	**0.28**
Large	CitiGroup	**0.02**	**0.04**
Large	Alcoa Corp	**0.02**	**0.04**

TABLE6.1RESULTS

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 6 June 2022, pp: 711-724 www.ijaem.net ISSN: 2395-5252

VII. EXISISTING SYSTEMS

1.5 Vanishinggradientdescentproblem:

Inmachinelearning,thevanishinggradientproblemise ncounteredwhentrainingartificialneuralnetworkswit hgradient-basedlearningmethodsand backpropagation.Insuch methods,eachoftheneuralnetwork'sweightsreceivesa nupdateproportionaltothepartialderivativeoftheerror functionwithrespecttothecurrentweightineachiterati onofinstruction.

1.6 LinearRegression:

Linear regression was less reactive to normalization techniques as opposed to the polynomial regressiontechniques. Some reasonable outcomes were appearing prior in the study even when a small number offeatures were used without normalization, while this lead to the polynomial regression models to overflow.

1.7 StochasticGradientDescent(SGD):

At first, it appeared that Stochastic Gradient Descent would be an exact fit to a problem of this type forlong term price prediction. As the dataset that was used only covered the time period of 2005-2013 thetraining data could only provide a maximum of (365 ∗ 8) = 2920 training samples to be used. But, thestockexchangeisnotopeneveryday ofthe year,thereforethis numberwould be significantly lower.PROPOSEDSYSTEMS

Accuracy plays lead role in stock market prediction .Although many algorithms are available for thispurpose,electingthemostaccurateonecontinuesto be thefundamentaltask ingettingthebest results.
In order to pull off this, we used LSTM algorithm. This involves training the algorithms, executing them,getting the results, comparing various performance parameters of the algorithm and finally obtaining themost accurate outcome.

1.8 FeasibilityStudy:

Preliminary investigation examinesproject feasibility thelikelihood thesystemwillbehandy to theorganization.Themainpurposeofthefeasibilitystu dyistotesttheTechnical,OperationalandEconomical feasibility for adding new modules and debugging old running system. All systems areachievable if they are given unlimited resources and limitless time. There are aspects in the feasibilitystudypartof the preliminaryinvestigation.

1.8.1 TECHNICALFEASIBILITY:

To determine whether the suggested system are technically feasible, we should take into consideration thetechnicalissuesinvolvedbehindthesituation.Tech nicalfeasibilitycenterontheexistingcomputersystem and to what scale it can support the proposed addition. Python and it's libraries are technologysoftwarewhichare helpfulin developingDataAnalytics.
So,thereisnoneedforadditionalpurchase.

1.8.2 OPERATIONALFEASIBILITY:

Proposed projects are beneficial only if they can be turned out into information system that will meet theuser’s operating requirements. Operational feasibility features of the project are to be taken as animportant part of the application implementation. This system is operational feasible since the users areknown with the technologies and hence there is no need to gear up the personnel to use the system. Alsothesystemisvery friendlyand easy to use.

8.1.3. ECONOMICFEASIBILITY:

Todecidewhetheraprojectiseconomicallyfeasible,we havetolookintovariouscomponentsas:

- Maintenancecosts
- Long-termreturns
- Costbenefit analysis

The proposed system is computer based. It requires average computing capabilities which is very primaryrequirementand can beaffordedby an organization.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 6 June 2022, pp: 711-724 www.ijaem.net ISSN: 2395-5252

VIII. SYSTEMARCHITECTURE

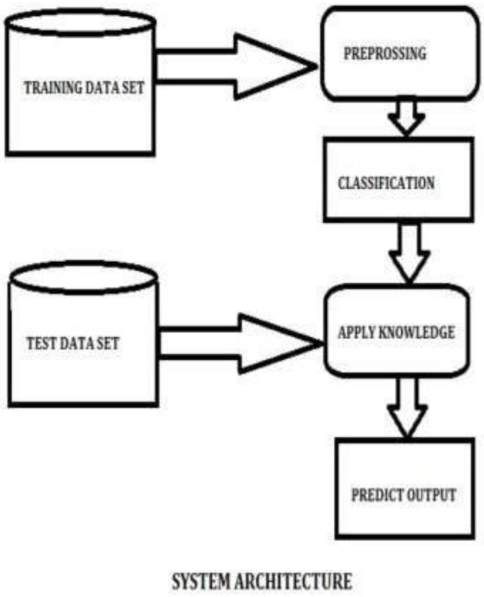


FIGURE9.1SYSTEMARCHITECTURE

IX. DATASET

we have taken dataset from Kaggle website of tesla stock prices and predicted the future stock price fortesla the dataset consists of stock data from 2010 to 2020 and it consists of 2416 rows and 7 columns ofdata and the prices for the previous year stock and by using the recurrent neural network with LSTMmodel we have implemented sequential data and to predict the future stock price of Tesla .where recurrentneuralnetworkwithLSTMmodelholdlargea mountofdatafora longperiodof timeandanalyze thedata in sequential way if there are any gaps in stock data also it analysis the data and gives the output thatisfuture stock price.

1.9 TRAININGDATA:

Training Data is nothing but enriched or labeled data you need to train your models. You might just needto collect more of it to improve your model accuracy. But, the possibility of using your data is pretty lowbecause,asyoubuild agreatmodelyouneedgreattrainingdata atscale.

1.10 TESTDATASET:

The test set is a set of observations used to assess the performance of the model using some performancemetric.Itisimportantthatno observationsfromthetrainingsetareinvolved the testset.

1.11 PREPROCESSING:

pre-processing is main step in Machine Learning as the quality of data and the useful information that canbeobtained fromitdirectlyaffectsthe ability ofourmodeltolearn.

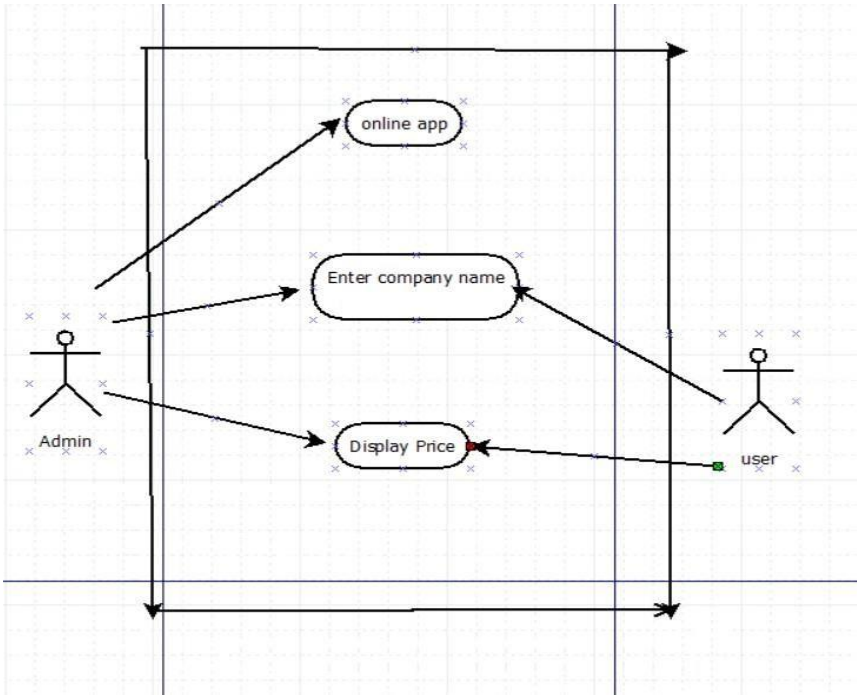


FIGURE10.3.1USECASEDIAGRAM

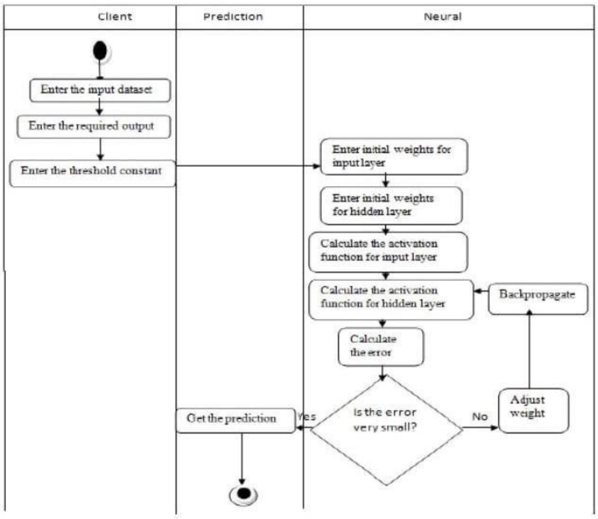


FIGURE 10.3.2ACTIVITYDIAGRAM

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 6 June 2022, pp: 711-724 www.ijaem.net ISSN: 2395-5252

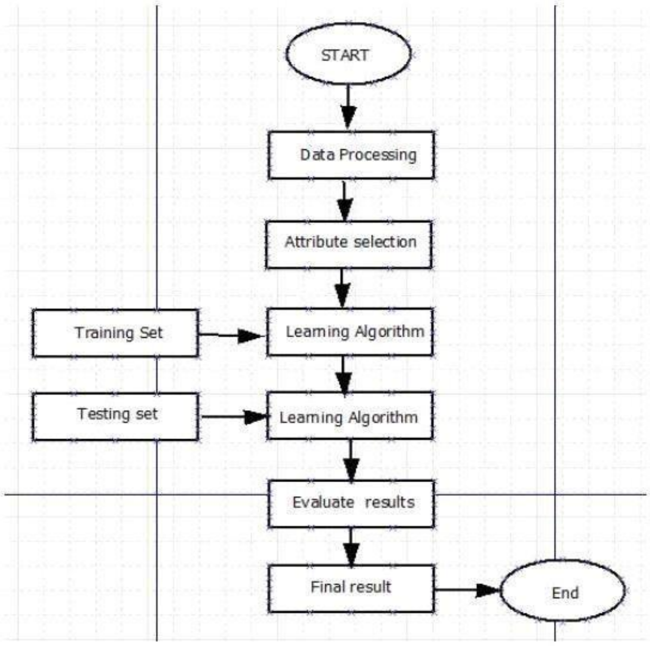


FIGURE10.3.3FLOWCHART

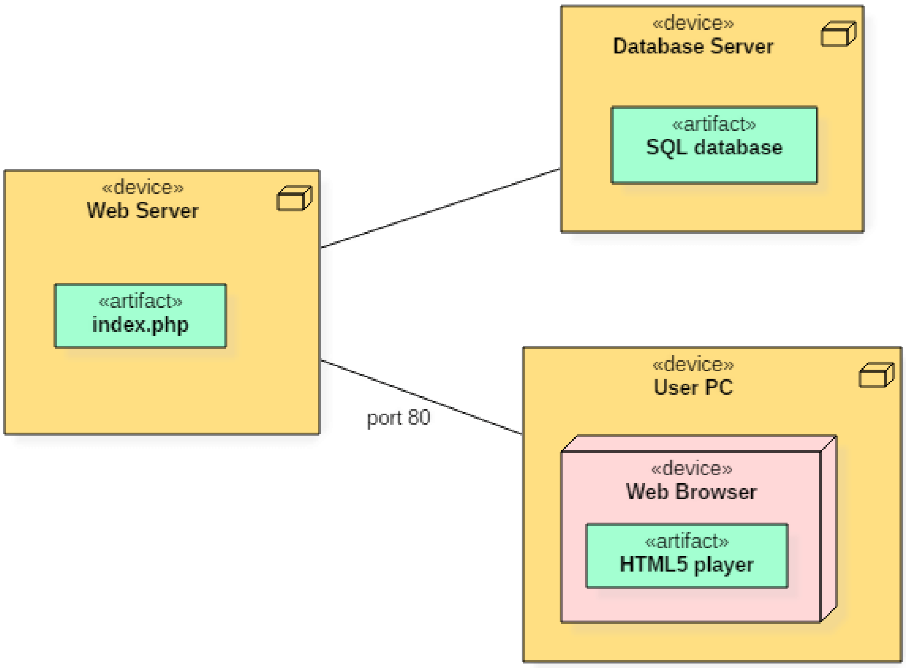


FIGURE10.3.4COMPONENTDIAGRAM

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 6 June 2022, pp: 711-724 www.ijaem.net ISSN: 2395-5252

X. LIBRARIES

MathPandasNumPySkLearnKerasMatplotLibDenseSequential

1.12 Example

from pandas_datareader import data# Only get theadjustedclose.
aapl = data.DataReader("AAPL", start='2015-1-1', end='2015-12-31',
data_source='yahoo')aapl.plot(title='AAPLAdj. ClosingPrice')

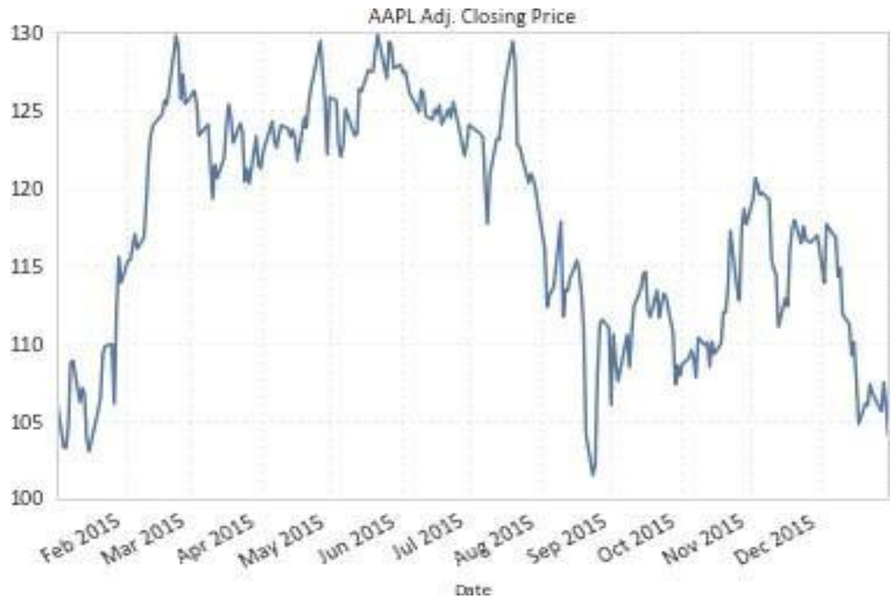


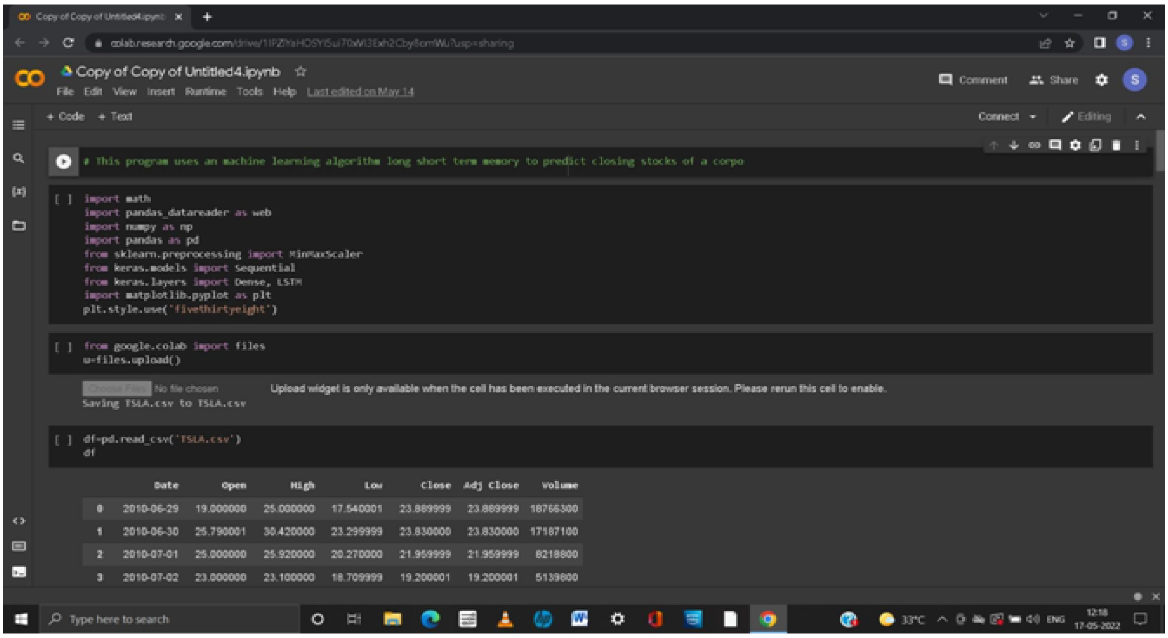
FIGURE11.1 EXAMPLEFOR11

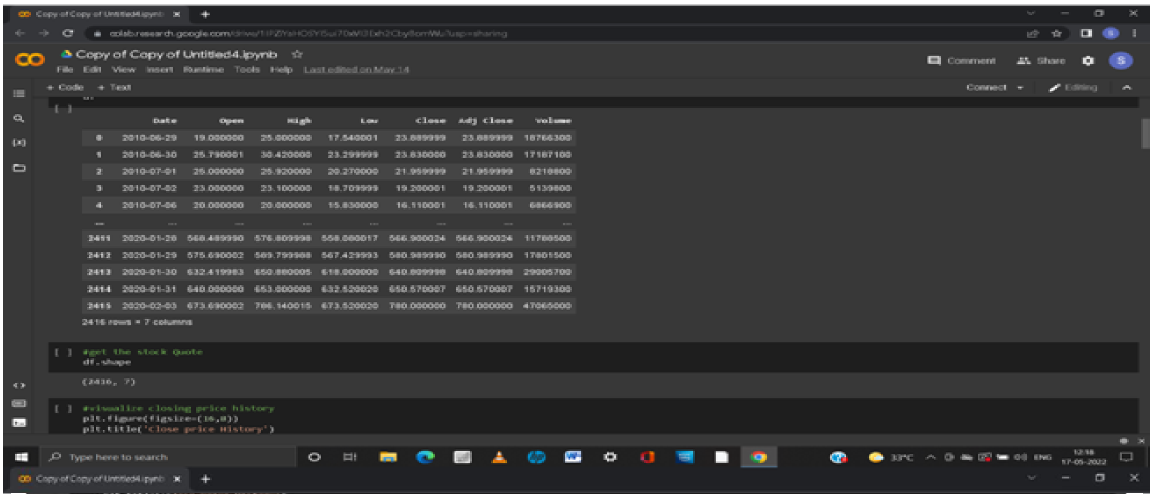
IMPLENTATION:-
#build the LSTM Modelmodel=Sequential()
model.add(LSTM(50,return_sequences=True,input_shape
=
(x_train.shape[1],1)))model.add(LSTM(50,return_sequences=False))
model.add(Dense(25))model.add(Dense(1))#compilethemodel
model.compile(optimizer = 'adam', loss = 'mean_squared_error')#Trainthemodel
model.fit(x_train, y_train, batch_size=1, epochs=1)#createtheexistingdataset
#create a new array Containing scaled values from index 1543 to 2003test_data =scaled_data[training_data_len-60: ,:]x_test=[]
y_test = dataset[training_data_len: , :]for i in range(60, len(test_data)):x_test.append(test_data[i-60:i,0])

XI. RESULTS AND DISCUSSION:-

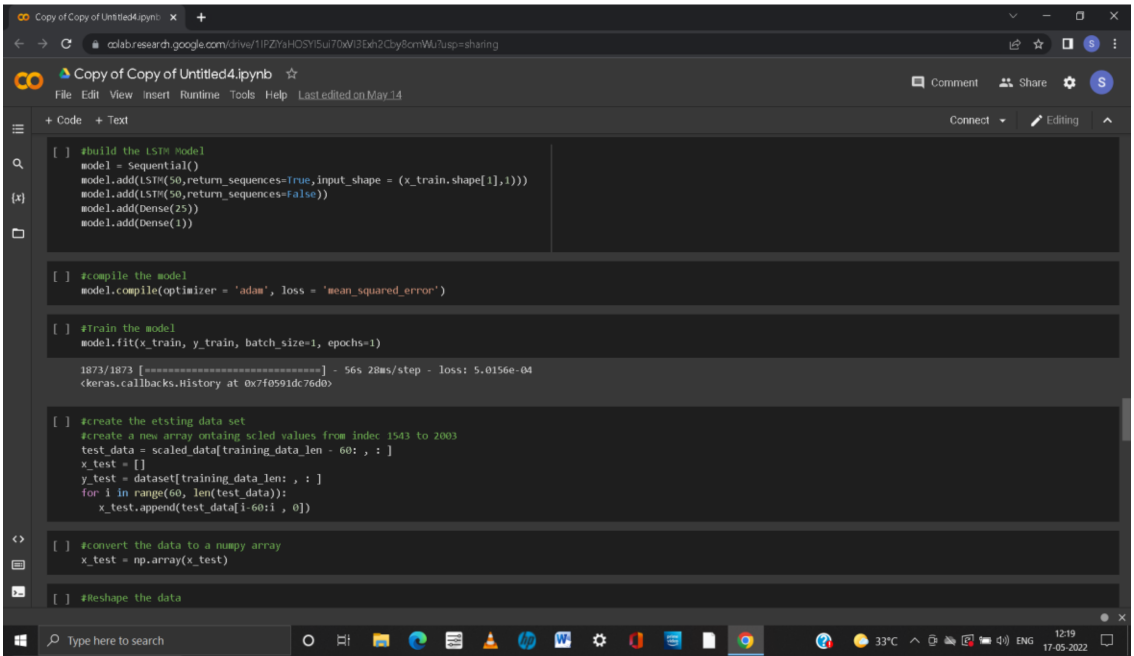
WehaveimplementedRecurrentneuralnetworkwithLSTMmodelforbeststockpricesprediction.In this model it takes all the information about the previous years stock prices that is when they started the stockmarket and then it analysis the data in a sequenctional way and it predict the price for the future. By this it helpsthe investors and traders to put their returns for the future profit. So,we have taken the dataset of Tesla from theyear 2010 to 2020 that is of 2416 rows of data with their previous year prices anddetails to analyse and predictthe stock price for the future and implemented LSTM model hold large amount of data for a long period of timeand analyse the data in sequenctional way if there are any gaps in stockdata also it analysisthe data and givesthe output that is future stock price. We have taken the dataset from the kaggle that is Tesla stock details andimplementedLSTMmodeland predictedthe future price forthe Teslastock.

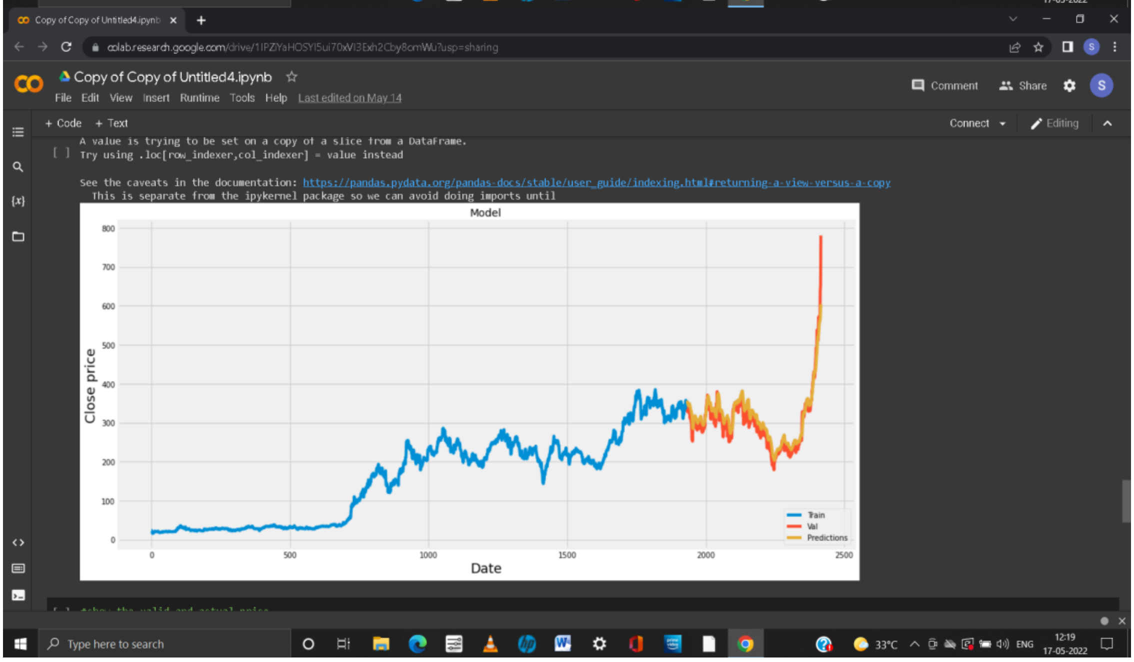
International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 6 June 2022, pp: 711-724 www.ijaem.net ISSN: 2395-5252

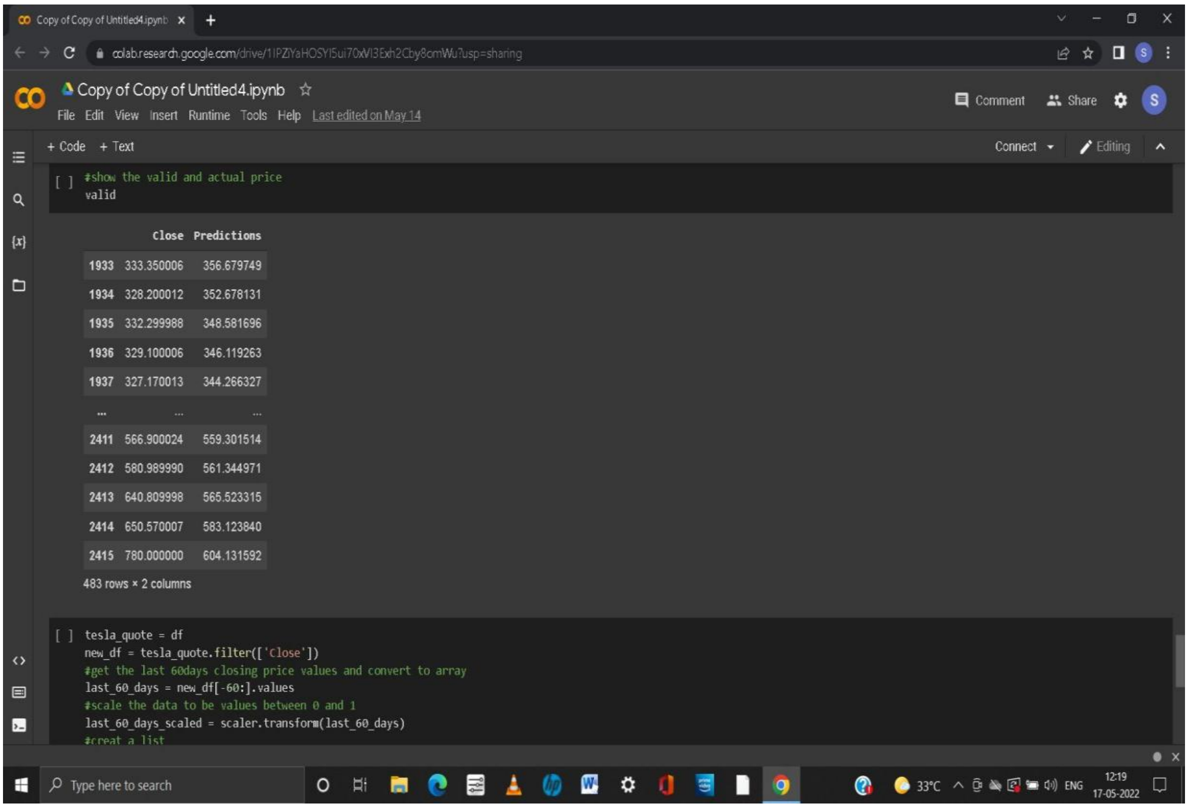


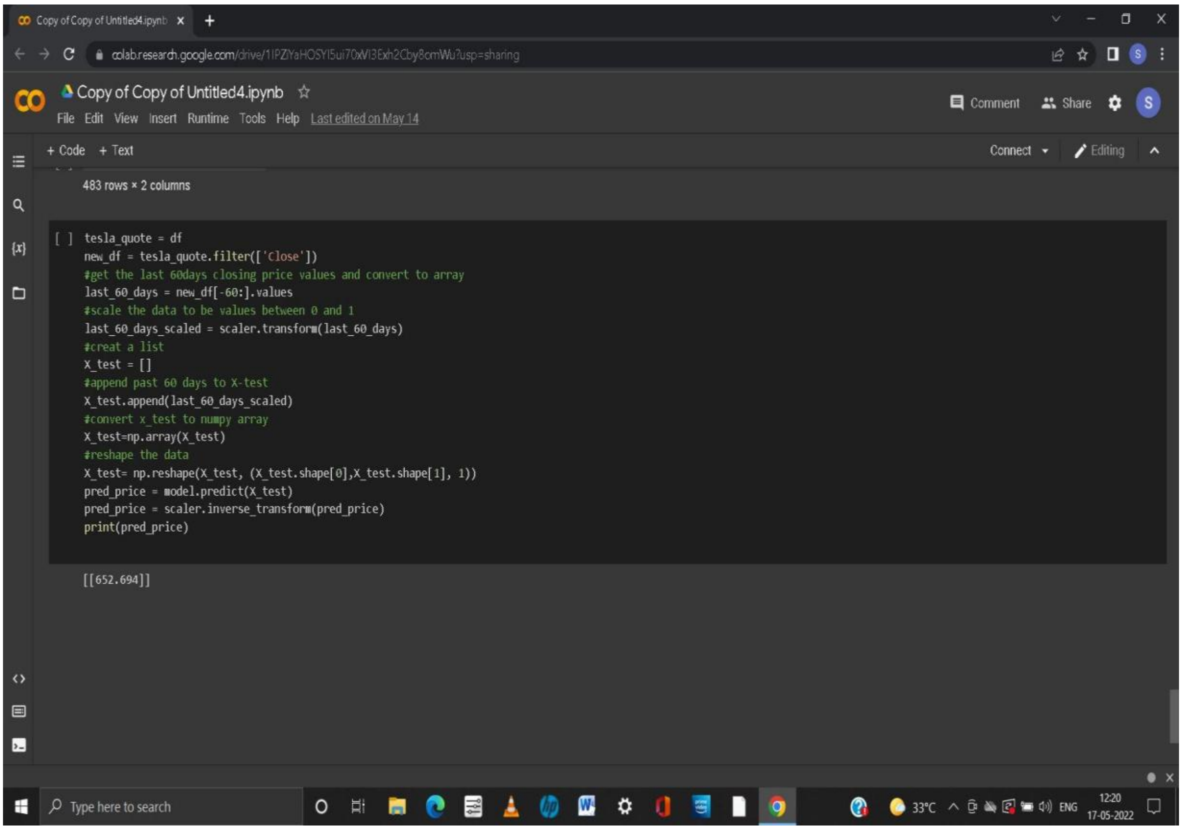


International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 6 June 2022, pp: 711-724 www.ijaem.net ISSN: 2395-5252






Copy of Copy of Untitled4.ipynb
colab.research.google.com/drive/1IPZiYaHOSYI5ui70xVI3Exh2Cby8omWu?usp=sharing
Copy of Copy of Untitled4.ipynb
File Edit View Insert Runtime Tools Help Last edited on May 14
Comment
Share
+ Code + Text
Connect
Editing
#show the valid and actual price
valid
Close Predictions
1933 333.350006 356.679749
1934 328.200012 352.678131
1935 332.299988 348.581696
1936 329.100006 346.119263
1937 327.170013 344.266327
2411 566.900024 559.301514
2412 580.989990 561.344971
2413 640.809998 565.523315
2414 650.570007 583.123840
2415 780.000000 604.131592
483 rows × 2 columns
tesla_quote = df
new_df = tesla_quote.filter(['Close'])
#get the last 60days closing price values and convert to array
last_60_days = new_df[-60:].values
#scale the data to be values between 0 and 1
last_60_days_scaled = scaler.transform(last_60_days)
#creat a list
Type here to search
33°C
ENG
12:19
17-05-2022


Copy of Copy of Untitled4.ipynb
colab.research.google.com/drive/1IPZiYaHOSYI5ui70xVI3Exh2Cby8omWu?usp=sharing
Copy of Copy of Untitled4.ipynb
File Edit View Insert Runtime Tools Help Last edited on May 14
Comment
Share
+ Code + Text
Connect
Editing
483 rows × 2 columns
tesla_quote = df
new_df = tesla_quote.filter(['Close'])
#get the last 60days closing price values and convert to array
last_60_days = new_df[-60:].values
#scale the data to be values between 0 and 1
last_60_days_scaled = scaler.transform(last_60_days)
#creat a list
X_test = []
#append past 60 days to X-test
X_test.append(last_60_days_scaled)
#convert x_test to numpy array
X_test=np.array(X_test)
#reshape the data
X_test= np.reshape(X_test, (X_test.shape[0],X_test.shape[1], 1))
pred_price = model.predict(X_test)
pred_price = scaler.inverse_transform(pred_price)
print(pred_price)
[[652.694]]
Type here to search
33°C
ENG
12:20
17-05-2022

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 6 June 2022, pp: 711-724 www.ijaem.net ISSN: 2395-5252

XII. CONCLUSION

TheLSTMmodelcan betunedforvarious parameterssuchas changingthe numberofLSTMlayers,addingdropoutvalueandincreasingthenumberofepochs.ButarethepredictionsfromLSTMsufficienttoidentifywhetherthe stockprice willincreaseordecrease?Certainlynot! stockpriceisaffectedbythenews aboutthecompany andotherfactorslikedemonetizationormerger/demergerofthecompanies.Thereiscertainintangiblepartaswellwhichcanoftenbeimpossibletopredictbeforehand.

Timeseriesforecastingisaveryfascinatefieldtoworkwith.Thereisa insightinthegroupthatit'sacomplexfield,andwhilethereisa grainoftruthinthere,it'snotsohardonceyougetthehangofthebasictechnique

REFERENCES

[1] LSTM BASED STOCK PRICE PREDICTION 1Prof. PritamAhire, 2Hanikumar Lad, 3SmitParekh, 4Saurabh Kabrawala 1Professor, 2Student, 3Student, 4Student 1Computer Engineering,1DYPatilInstituteof Engineeringand Technology, Pune,India2021

[2] UsingNeural NetworkstoForecastStockMarket Prices, RamonLawrence.This paper is a survey on the application of neural networks in forecasting stock market prices.With their ability to discover patterns in nonlinear and chaotic systems, neural networks offer theabilitytopredictmarket directionsmore accuratelythancurrenttechniques.2019

[3] StockMarketPredictionUsingHybridApproach,VivekRajput,SarikaBobde. The objective of this paper is to construct a model to predict stock value movement usingtheopinionminingand clusteringmethod topredictNational Stock Exchange(NSE).

[4] Appliedattention-basedLSTMneuralnetworksinstockprediction.,Cheng,Li-Chen,Yu-HsiangHuang, and MuEn Wu. Prediction of stocks is complex due to dynamic, complex, and chaoticenvironment of the stock market. several studies predict that stock value movements are usingdeeplearningmodels.2018

[5] Usmani, Mehak, Syed Hasan Adil, Kamran Raza, and Syed SaadAzhar Ali. "Stock marketprediction using machine learning techniques." In 2016 3rd international conference on computerandinformation sciences (ICCOINS), pp. 322-327. IEEE,2016.

[6] RakhiMahant,TrilokNathPandey,AlokKumarJagadev,andSatchidanandaDehuriOptimized Radial Basis Functional Neural Network for Stock Index Prediction,InternationalConferenceonElectrical,Electronics, andOptimizationTechniques (ICEEOT)-2016.