

# Predict Stock Market's Fluctuating Behaviour : Role of Investor's Sentiments on Stock Market performance

Uma Gurav <sup>#1</sup>, Prof.Dr.S.Kotrappa <sup>\*2</sup>

<sup>#</sup> Research Scholar, Dept. of Computer Science and engineering,  
K.L.E's Dr.M.S.S CET  
Belagavi,India

<sup>\*</sup>Dept. of Computer Science and engineering,  
K.L.E's Dr.M.S.S CET  
Belagavi,India

<sup>1</sup>umabgurav@gmail.com

<sup>2</sup>kotrappa06@gmail.com

**Abstract** - Long term historical records of the stock markets are widely used in technical research to define, understand and analyze stock market's time series trends and patterns which can be used to generate huge profits during trading sessions. Even though, technical analysis using different technical measures have been shown to be helpful in forecasting market patterns, formulating specific trading rule is a challenging task. In this research paper, we have tried to analyze investor's sentiments considering US presidential elections and effects of Covid 19 as an explicit fluctuating factor affecting stock market performance. In addition to this, in this research work, we have tried to identify correct and better trading rules and trading points, technical indicators to be considered using mathematical formulations, to determine when to buy or sell stocks. Thus, given dynamically varying stock market behaviour in high frequency trading environment, it is important to integrate market sentiments into forecasting operations. This paper combines sentiments into stock forecasting model using the log bilinear (LBL) model for short term stock market's sentiment pattern learning and recurrent neural (RNN) for long term sentiments pattern learning which achieves better performance than deep learning based stock price forecasting existing methodologies.

**Keywords** — Sentiment Analysis, machine learning, social networking platform, stock price forecasting, Time series Analysis

## I. INTRODUCTION

The movement of stock prices is related to different aspects like company's profile, including government policies, inflation rates and economic conditions. A popular methodology used for stock market forecasts in financial

engineering[1] is a technical analysis that uses mathematical formulas and graphs known as technical indicators to evaluate stock price fluctuations, with the purpose of identifying patterns that are used for profitable trade gains. The success of stock trading depends largely on getting the exact time to sell or purchase stocks. It is very difficult to anticipate, despite developments in technical analysis of stock prices and trading points. According to the Efficient Market Hypothesis, Fundamental analysis or technical analysis does not lead, to higher Returns on Investments(ROI)[2-5], which are consistently above average. Computer science studies are concerned with the examination of non-conventional metrics or external variables that assist in the analysis of the stock market[6]. Research focused primarily on exploring native parameters, real-life social networking sites[7], investor sentiments[8,38], news articles search engines [9-11]. Considering Covid 19 as an explicit factor for stock market performance fluctuations, Initially Chinese stock market affected a lot during the month of December 2019. However the Republic of China has recovered from this epidemic and its economy has substantially benefited from what it was in December 2019. In[13], small-medium-sized companies have been shown to be significantly affected by the pandemic as the Chinese stock index has almost fallen by 8 percent on 3 February[37]. Whereas, RIL -Reliance stock prices increased with Facebook and Google have obtained certain percentage of it's stake during COVID19 pandemic. Now, US presidential elections is also playing major role as investor's sentiments varies over time. This shows capturing market sentiment is the key in building efficient stock forecasting model. Literature survey shows the conventional machine learning algorithms as, support vector machine(SVM), neural network(NN), random forest(RF) based stock predictive models are not efficient when



behavior is dynamic for a single session or a day. In addition to this, dynamic nature of stock market can be captured using deep learning methods such as RNN with gated recurrent and LSTM[36-38], convolution NN, deep NN[39-40], general adversary network. Deep learning models play very important role in modeling long term sentiment behavior, but cannot capture short term market sentiment behavior. As per our knowledge, no prior work has considered modeling both long and short term sentiment context (LSSC) together in stock price forecasting.

## II. LITEARATURE SURVEY

This section presents survey of existing stocks price prediction methodologies using existing machine learning (ML) and deep learning techniques. Recently machine learning (ML) techniques have played an important role in most of the areas especially in some of the complicated applications, it includes various methodologies. In general, optimization is the key in building efficient forecasting models. NN[18], reinforcement learning(RL) [20], Neural Networks are recently widely utilized ML methodologies. ML modern methods in various domains, for example speech recognition, image classification, various predictive models and Natural language processing(NLP) are designed using either RNN or Convolution Neural Networks (CNN), modified back propagation neural networks (MBNN) [24]. Some research showed LSTM (Long Short Term Memory) networks [21,24,28] outclassed random forests (RF) and Deep Neural Networks. In [22] explained an RL methodologies using recurrent neural network architecture for addressing gradient descent problems during training process. Correspondingly, [23] designed mechanism for the forecasting error loss and direction forecasting loss. This architecture is known as GAN-FD for loss of path prediction and loss of prediction error. For better stock price estimation, the GAN-FD uses CNN and LSTM [24-27]. Nevertheless the simulation of short-term and long-term business behavioural dynamics does not achieve successful learning tradeoffs. Therefore, GAN-FD fails considerably when the market's behaviors are highly volatile in nature. Further, market's short-term and long term sentiment pattern is not considered. For building efficient stock price forecasting model it is important to consider dynamically varying market sentiments.

### A. Findings, Gaps & Limitations

1. Existing stock Forecasting models fails to consider long and short term market variance context in design and mathematical modeling of the stock predictive models.
2. Most of existing stock price forecasting models under highly volatile nature of stock market are designed using implicit data and very limited work is carried out considering explicit data.

3. Existing stock forecasting models do not consider session variance between feature-sets that plays major role in improving prediction accuracies.

4. Explicit data plays very important role in predicting investor's mood in highly volatile, uncertain stock market environment. As investor's sentiments varies over time, market performance also varies. This issue needs to be addressed.

## III. RESEARCH MOTIVATION

This motivated the research work to consider investor's sentiment as an external feature for improving stock market price prediction and to indicate when to buy or sell stocks. This paper presents sentiment conscious stock forecasting model(SCSPF) combining both short and long term sentiment features. First, the work carried out sentiment analysis on twitter dataset and marked it as "0" shows negative sentiment and "1" shows positive sentiment, technical indicators considered based on polarity, subjectivity of the sentiments. Then, this sentiment forecasted data is added to the stock information dataset and is trained using SCSPF. The SCSPF is built by combining both RNN and LBL which captures long-term market features and short-term market features respectively. Thus, proposed Sentiment Conscious SC-RNN-LBL model can be used for predicting stock market patterns based on current market sentiments. Our main focus is on volatility modeling, time series in finance [14] and one main vital advantage of volatility modeling in TSF is, it's relevant for finance market practitioners and academic researchers.

### A. Research Work Contribution is as follows :

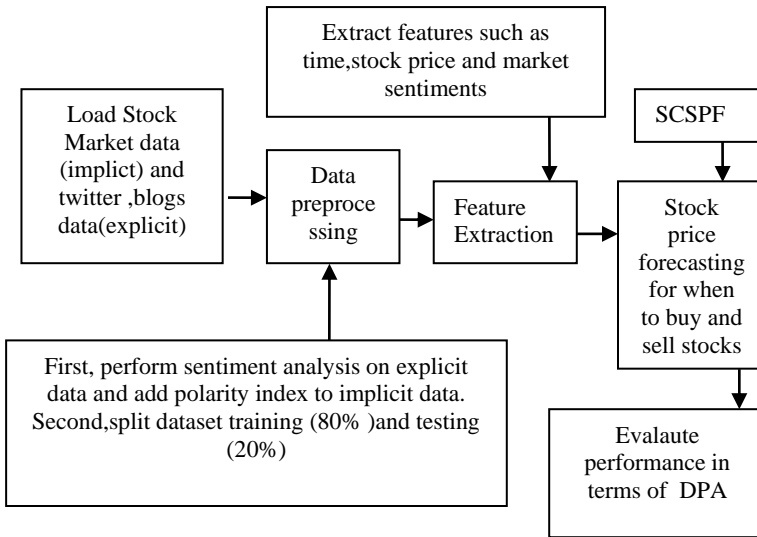
1. Sentiment conscious RNN-LBL (SC-RNN-LBL) learning model can capture Long and short term sentiment context (LSSC) features more efficiently, which obtains investor's sentiment behaviour within short-term context for current trading sessions. Thus, predicts stock peak value & time duration more efficiently and can be used for forecasting prices of short term sells.
2. Improved Bayes classification model used to classify tweets. Existing Bayes theorem multiplies entire feature probabilities together, the zero probabilities of any class will turn the overall probabilities to Zero. For addressing this issue, Laplace add-one smoothing methodology is introduced
3. SC-RNN-LBL model can effectively indicate when to buy and sell stocks. As compared to long-term market sentiments results, short-term bin sentiment variance plays a very important role in potential stock's buying or selling
4. Experiment outcome shows the proposed sentiment conscious RNN-LBL based stock predictive model

achieves better DPA performance compared to other existing deep learning based stock price predictive models.

5. Considered US presidential elections related twitter tweets for sentiment analysis and combined this indexed /labelled sentiments with US stock Market indices..

#### IV. Sentiment conscious stock forecasting system model(SCSPF):

This section presents sentiment conscious RNN-LBL model for forecasting when to buy or sell stock as shown in Fig. 1. Firstly, system architecture of sentiment conscious stock market forecasting is presented. Second, discusses about methods used for carrying out sentiment analysis and create sentiment indexed stock dataset. Then, this work discusses the algorithm details of sentiment conscious stock forecasting model using both LSSC and stock pattern contextual features using RNN-LBL.



**Fig. 1. Block diagram of Sentiment Conscious RNN-LBL model for forecasting stock price.**

Stock price forecasting and indication of when to buy or sell stock is a challenging task under highly fluctuating and volatile environment. Recently, forecasting when to buy or sell stocks with one-step considering a period of one minute time interval are the major issues for stock market forecasting algorithm designer or traders. Especially, under highly volatile market conditions sentiment analysis during pandemic, any uncertainties such as war and natural disasters ,elections plays important role. In this paper, the stock price volatility of respective stocks and market index are considered one minute prior, utilizing markets sentiments from tweets/news and using past long-term and short-term

historical information from stock markets. The forecasting problem is mathematically represented in this section below. Let us assume  $\mathcal{U}_t$  as elementary feature that includes market sentiments features and  $Y$  represent the stock closing price of respective stocks with the interval of one minute at time  $t$  is described as follows:

$$t = 1, 2, \dots, T \quad (1)$$

where  $T$  represent the maximum time interval considered for forecasting. For respective stock historical records combined long and short-term history elementary stock indicator feature  $\mathcal{U}$  is represented as follows:

$$\mathcal{U} = \{\mathcal{U}_1, \mathcal{U}_2, \dots, \mathcal{U}_T\} \quad (2)$$

and the historical stock market closing price  $\mathcal{V}$  with different market sentiments are represented as follows :

$$\mathcal{V} = \{\mathcal{V}_1, \mathcal{V}_2, \dots, \mathcal{V}_T\} \quad (3)$$

The objective of this paper is to forecast stock price  $\mathcal{V}_{T+1}$  for a period of every one minute time interval considering highly volatile stock market environment.

#### A. Sentiment conscious stock forecasting mathematical model(SCSPF)::

This section briefs about the sentiment analysis tool used for analysis stock market news/tweets information. This work uses open source VADER (Valence Aware Dictionary for Sentiment Reasoning) sentiment analysis tool [34], [35] for analyzing market sentiments. We design the model just considering two polarities only such as positive ("1") and negative ("0") where score varies from -1 to 1 and scores greater than 0 is considered to be positive sentiment and anything equal or below 0 is considered as negative sentiments. For analyzing market sentiments VEDAR provides different machine learning algorithms linear regression , logistic regression and Naïve Bayes classifier[35]. This work uses naïve Bayes classifier as it is very efficient when used for real life data. Further, they are efficient in avoiding over fitting problems and can more efficiently deal with dimensionality. The conditional probabilities of naïve Bayes classifier is defined using following equation:

$$P(a|b) = \frac{P(b|a)P(a)}{P(b)} \quad (4)$$

The classification models establish the computed class  $d$  among set of classes  $d \in D$  for respective blog/document  $e$ . Thus the computed classes are defined as :

$$d = \operatorname{argmax}_{d \in D} P(d|e) \quad (5)$$

After applying conditional probabilities from Eq. (4) and (5) we will have,

$$d = \operatorname{argmax}_{d \in D} P(d|e) = \operatorname{argmax}_{d \in D} \frac{P(e|d)P(d)}{P(e)} \quad (6)$$

Solving above equation and utilizing the probabilities of  $P(e|d)$ , we can obtain

$$b = \operatorname{argmax}_{d \in D} P(b_1, b_2, \dots, b_o|d)P(d) \quad (7)$$

where  $b_1, b_2, \dots, b_o$  describes the feature sets of blog  $e$ . Nonetheless, this work assumes that the position of word doesn't impact classification outcomes and probability  $P(b_j|d)$  are independent in nature considering respective class  $d$ . Thus, we update the equation as follows:

$$P(b_1, b_2, \dots, b_o|d) = P(b_1|d) \cdot P(b_2|d) \cdot \dots \cdot P(b_o) \quad (8)$$

Thus using Eq. (7) and (8) the improved naïve Bayes classification model can be constructed as follows:

$$D_{NB} = \operatorname{argmax}_{d \in D} P(d) \sum_{b_j \in B} P(b_j|d) \quad (9)$$

For applying the above classification model for stock market news/tweets here we consider the word index position  $x_j$  within the documents by replacing  $b_j$  with  $x_j$ . Then

$$D_{NB} = \operatorname{argmax}_{d \in D} \log P(d) + \sum_{j \in \text{positions}} \log P(x_j|d) \quad (10)$$

considering feature sets in log space the Eq. (7) is updated as follows:

The Bayes classification model is trained as follows. In Eq. (8) this work first computes  $P(d)$  and  $P(x_j|d)$ . For that, let us consider  $O_d$  representing the total number of

tweets/blogs considered in training that belongs to class  $d$  and  $O_{doc}$  representing the total number of tweets/blogs. Then,

$$P(d) = \frac{O_d}{O_{doc}} \quad (11)$$

The probability of word  $x_j$  belong to class  $C$  is established using following equation

$$P(x_j|d) = \frac{\text{count}(x_j, d)}{\sum_{x \in W} \text{count}(x, d)} \quad (12)$$

where  $\text{count}(x_j, d)$  represents number of times  $x_j$  occurs in class  $d$ ,  $V$  defines overall word vocabulary. As seen from Eq. (7) the Bayes multiplies entire feature probabilities together, the zero probabilities of any class will turn the overall probabilities to Zero. For addressing this , applies Laplace add-one smoothing methodology as described below:

$$P(x_j|d) = \frac{\text{count}(x_j, d) + 1}{\sum_{x \in W} \text{count}(x, d) + 1} = \frac{\text{count}(x_j, d) + 1}{\sum_{x \in W} \text{count}(x, d) + |W|} \quad (13)$$

Using above equation we compute the market sentiment score and update the polarity index as shown in Figure 2.

	Date	Label	Trump Tweets	Joe Biden Tweets	Open	High	Low	Close	Adj Close	Volume
998	10/20/2020	1.0	Any deal completed before the fiscal curb must...	"We have a responsibility not only to vote... ..	28245.84961	28575.02930	28243.03906	28308.78906	28308.78906	350220000
999	10/21/2020	1.0	Watch as I humiliate a dais full of "talent" ...	"We're on the precipice of an American century...	28270.66016	28450.64063	28196.39063	28210.82031	28210.82031	299420000
1000	10/22/2020	1.0	@mstang_13 I hope so.	"Every time there's a problem in Congress, I g...	28197.41992	28421.31055	28040.17969	28363.66016	28363.66016	330320000
1001	10/23/2020	0.0	@HyperspaceIT Thanks--@HuffingtonPost is a to...	"@timkaine and I believe, we really believe, t...	28409.65039	28436.51953	28149.82031	28335.57031	28335.57031	347300000
1002	10/26/2020	0.0	@pokerswayne @VanityFair Very wrong dopey!	"We are endowed by our creator with certain in...	28185.82031	28185.82031	27370.16016	27685.38086	27685.38086	404150000
1003	10/27/2020	0.0	Worst ever issue of @VanityFair magazine—bad f...	"If we come together, there's nothing, nothing...	27651.17969	27707.68945	27457.96094	27463.18945	27463.18945	366400000
1004	10/28/2020	0.0	Our new @MissUniverse Olivia Culpo is not only...	"We're America! We don't scare easily... We ne...	27102.14063	27102.14063	26497.39063	26519.94922	26519.94922	496150000
1005	10/29/2020	1.0	Back by popular demand. @GiulianaRancic and @B...	"Let's make history tomorrow. Let's send a sig...	26480.98047	26891.89063	26290.77930	26659.10938	26659.10938	419770000
1006	10/30/2020	1.0	Miss Universe contestants are amazing—the most...	1. Make sure you know how to vote today: https...	26572.26953	26639.17969	26143.76953	26501.59961	26501.59961	504820000

Fig. 2. Twitter Sentiment Analysis of Tweets by Presidential election candidates

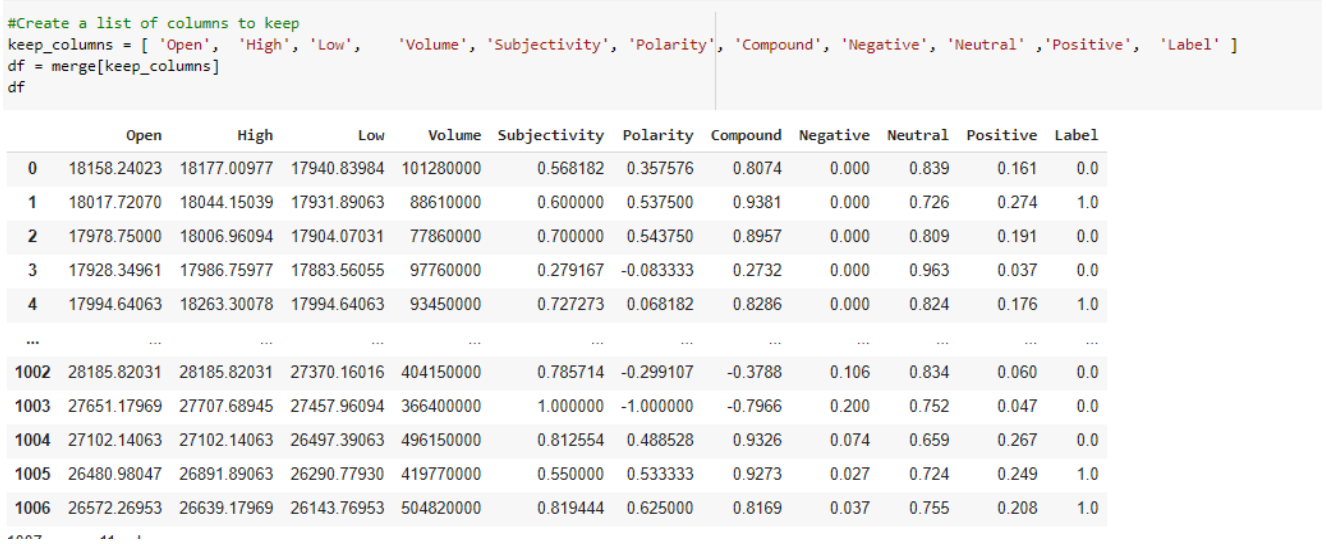


Fig. 3. Polarity , Subjectivity ,Compound, Negative, Neutral and Positive score of sentiments

This work presents SC stock market pattern analysis model by learning both short term behaviour and long-term behaviour using LBL and RNN, respectively similar to LSTM model. First let us define the RNN model architecture, which consists of several hidden layers with input layers and output layers. Hidden Layer's activation parameters are obtained as follows :

$$i_{\ell}^{\nu} = f(\mathcal{X}i_{\ell}^{\nu} + \mathcal{D}s_{w_{\ell}^{\nu}}), \quad (14)$$

where,  $i_{\ell}^{\nu} \in \mathbb{S}^e$  describes hidden details of stock  $\nu$  at time instance, position  $\ell$  in time series,  $s_{w_{\ell}^{\nu}} \in \mathbb{S}^e$  describes the details of the  $\ell^{th}$  input stock of a particular stock market  $\nu$ . The activation function (AF) is defined by  $f(i)$  and the present stock's transition matrix (TM) is given as:

$$\mathcal{D} \in \mathbb{S}^e \quad (15)$$

previous status is given as:

$$\mathcal{W} \in \mathbb{S}^e. \quad (16)$$

$\mathcal{D}$  obtains volatile shifting behaviour from stocks and  $\mathcal{X}$  propagates signals from the time series. The Eq. (14) is iteratively executed to calculate status at each time instance.

The RNN architecture for learning long-term stock market characteristics is as shown in Figure 4. With regard to the study of the stock market time series, for the repeated patterns, the hidden layer context is more complicated in nature. RNN faces difficulties in the time series sequence in learning short-term context. This study presents the bilinear log (LBL) model, a deterministic model, with a single linear layer for addressing, as shown in Figure 4. A linear forecast that can be illustrated using the following equation:

$$i_{\ell}^{\nu} = \sum_{j=0}^{\sigma-1} \mathcal{D}_j s_{w_{\ell-j}^{\nu}}, \quad (17)$$

Where  $\mathcal{D}_j \in \mathbb{S}^{e * e}$  is a transition matrix,  $\sigma$  the total number of elements modelled in time series is the total

number of elements modelled in time series. In LBL, in the proposed technique, a precise transition matrix [TM] is modelled for each position in the time series SC matrices to obtain feature sets of different aspects of feelings. For the respective stock market  $\nu$ , the hidden description is then computed as follows.

$$i_{\ell}^{\nu} = \mathcal{X}i_{\ell}^{\nu} + \sum_{j=0}^{\sigma-1} \mathcal{D}_j \mathcal{N}_{c_{\ell-j}^{\nu} r_{\ell-j}^{\nu}}, \quad (18)$$

This research work strengthens our previous RNNLBL[19] model by combining sentiment-based variance information and the SC RNN-LBL model.

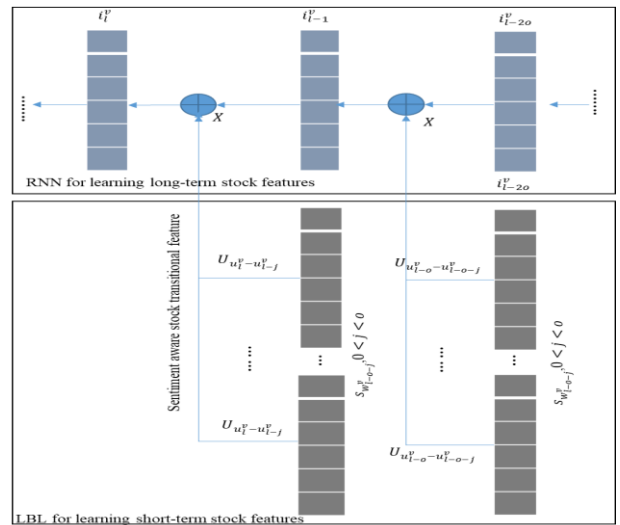


Fig 4. The architecture of sentiment conscious RNN-LBL model.

From Figure 4, considering respective stock market  $\nu$ , the position  $l$  are established using following equation:

$$i_l^v = Xi_{l-o}^v + \sum_{j=0}^{o-1} U_{u_l^v-u_{l-1}^v} S_{w_{l-j}^v}, \quad (19)$$

where  $u_l^v$  depicts current time,  $u_{l-1}^v$  depict time information of each stock of each layer of sentiment conscious RNN-LBL model, and  $U_{u_l^v-u_{l-1}^v} \rightarrow$  time-specific transitional matrix of sentiment session variation  $u_{l-1}^v - u_l^v$  between  $u_{l-1}^v$  and  $u_l^v$ . The time-specific transitional matrices aids in collecting time specific sentiment pattern features with respect to recent tweet information. Considering these the Eq. (19) updated as described below :

$$i_l^v = Xi_0^v + \sum_{j=0}^{l-1} U_{u_l^v-u_{l-1}^v} S_{w_{l-j}^v}, \quad (20)$$

where  $i_l^v = v_0$  with respect to initial status of respective stock within stock market environment. For obtaining varying sentiment nature of stock with respect to different time intervals, This work establishes sentiment conscious transitional matrices in modeling sentiment conscious RNN-LBL model as described using following equation :

$$i_l^v = Xi_{l-o}^v + \sum_{j=0}^{l-1} U_{u_l^v-u_{l-1}^v} N_{c_{l-m}^v} S_{w_{l-j}^v}. \quad (21)$$

Lastly, predict whether a stock  $v$  will exhibit certain patterns or sentiments  $C$  with respect to stock  $W$  at successive position  $l + 1$  are computed as follows

$$z_{v,l+1,c,w} = (t_l^v)^U N_{cS_w} = (i_l^v + v_v)^U N_{cS_w}. \quad (22)$$

In order to learn, time variance related information continuously, to create larger sizes of time-specific transitional matrices(TM), this approach suffers from over fitting problems. By splitting time variations into equivalent windows this issue can be resolved.

#### IV . Results and Discussion

This section presents performance comparative analysis and evaluation of proposed SCSPF model with existing deep learning based stock prediction methodology [23]. This work uses GAN-FD model for comparison because it achieves much better result than existing LSTM based stock forecasting methodology [25-27]. The proposed SCSPF model is implemented using python and C# programming language. The Tweets dataset and finance news for evaluating sentiment is obtained from following sites which are publically available [30-33]. This work evaluated performance using same data used by GAN-FD methodology[23] which is base research paper. The data is composed of 244 trading points. Further, experiments are carried out on datasets from January 01,2012 to October 31<sup>st</sup>, 2020. More detailed description of dataset used can be obtained from [23] and downloaded [33]. Direction Prediction Accuracy (DPA) used for performance evaluation. 80% of training data set and for testing 20% of dataset is considered. Experiments carried out for different stocks at time  $t$ , a prediction is done for next time  $t + 1$  using training models. Considering,  $T_0$  is the total number time intervals for forecasting. Actual stock value depicted as  $Y_t$  and forecasted value is described as  $\hat{Y}_t$ .

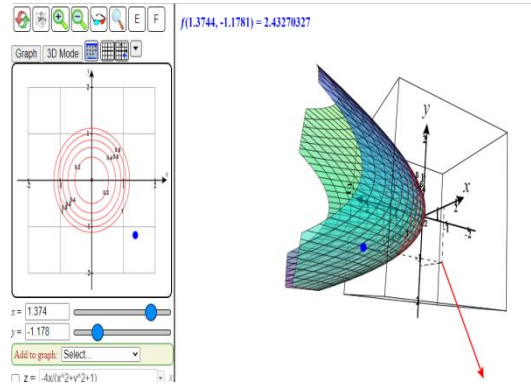


Fig. 5. Gradient Descent curve analysis for fluctuating parameters calculation

As shown in Figure 5 . As model senses fluctuations in the stock market conditions, steepest gradient and cost function varies over time.

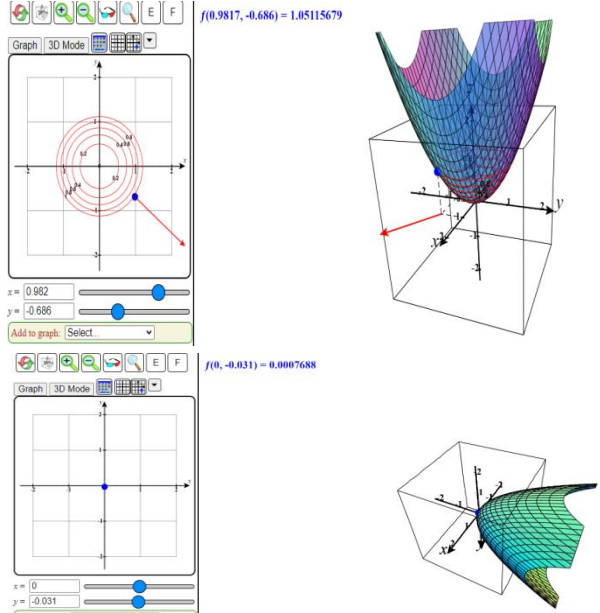


Fig. 6. Gradient Descent parameters calculations - global Minima

From Figure. 6, shown how steepest gradient affects cost function. Moving away along the gradient, going for global minima , cost function also decreases considerably.

#### A. DPA Performance Evaluation

Further ,to evaluate performance of SCSPF model under highly dynamic and volatile environment. Corresponding tweet dataset obtained from Yahoo Finance dataset [31] and Yahoo news [32] dataset, respectively. The stock price forecasting performance is evaluated between January 01 ,2016 to June 30<sup>th</sup> 2020 period.This section evaluated the SCSPF model's DPA performance. The metric DPA

determines a metric of accuracy that is measured as follows:

$$DPA = \frac{100}{T_0} \sum_{t=1}^{T_0} I_t, \tag{24}$$

where

$$I_t = \begin{cases} 1 & \text{if } (Y_{t+1} - Y_t)(\hat{Y}_{t+1} - Y_t) > 0 \\ 0 & \text{otherwise} \end{cases} \tag{25}$$

Comparative Analysis of The DPA outcome of SCSPF model and existing deep learning methods is shown in TABLE I. The graphical representation of DPA performance achieved by SCSPF model and existing deep learning based stock price forecasting methodology is shown in Figure 5. From results it is seen an average DPA performance achieved by LSTM-FD, GAN-FG, RNN-LBL, and SCSPF model is 0.6423, 0.68585, 0.7822, and 0.80805, respectively. The RNN-LBL model improves DPA performance by 17.88% and 12.32% over LSTM-FD and GAN-FD, respectively. Similarly, the SCSPF model improves DPA performance by 20.51%, 15.123% and 3.2% over LSTM-FD GAN-FD and RNN-LBL model, respectively. From results it can be seen SCSPF model achieves much superior DPA performance than existing deep learning methodologies. For evaluating such scenarios this work considered evaluation using Apple(Inc.) stock and corresponding tweet dataset (in Covid 19 conditions, US elections Tweets [32])stock market related dataset obtained from Yahoo Finance dataset [31] and Yahoo news dataset [32], respectively. The DPA performance achieved considering sentiment index is shown in Fig.5. From result it is seen the RNN-LBL and SCSPF model achieves a DPA of 0.0085 and 0.0059, respectively. From result it is seen the SCSPF improves forecasting performance by 18.26% over RNN-LBL considering stock market sentiments. Thus, SCSPF can really work really well with better profit for the investor. Experiments are carried out to test performance of the SCSPF model, DPA values are shown in Table 1.

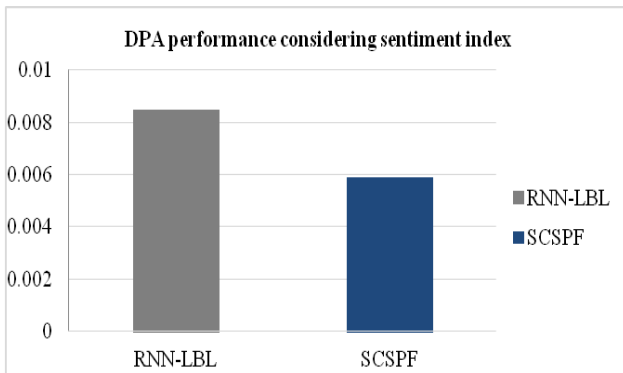


Fig 7. DPA performance of SCSPF model and RNN-LBL methodologies considering sentiment index.

Date	Tweet	Length	ID	Likes	Retweets
2020-09-20 13:37:58	MAKE AMERICA GREAT AGAIN!	25	1307675370322145282	361071	59997

Fig 8. Sample Tweets likes,retweet (US Presidential Election)

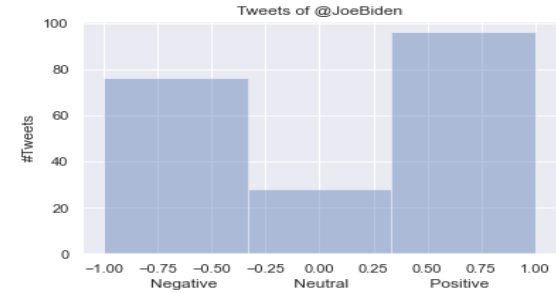


Fig 9. Tweets Sentiment Analysis.



Fig 10 Prediction of Stock market index for the period 2012-2021.

Figure 8 & 9 shows sample tweet and sentiment analysis of the tweets. It is clear from Figure 10 stock market index for the Apple (Inc.) decreased considerably during Covid -19 pandemic period from end of the 2019 and predicts improvement in stock market performance from end 2020 to 2021 period, amid US election polls.

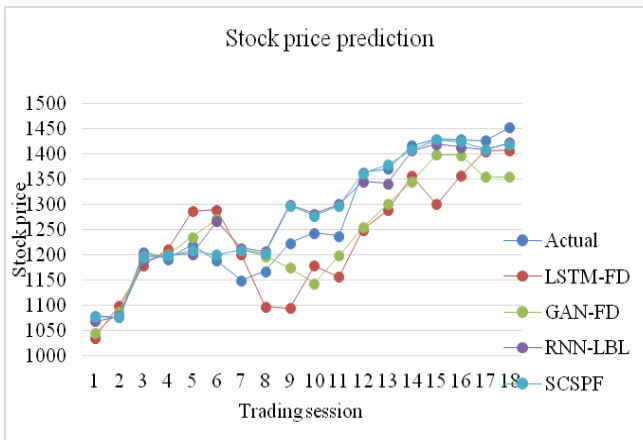
TABLE I

Methods	DPA
LSTM-FD	0.6423
GAN-FD	0.68585
RNN-LBL	0.7822
Proposed SCSPF model	0.80305

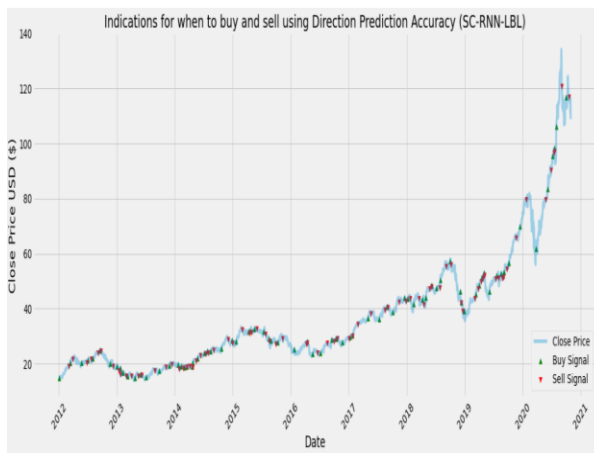
Comparitive Performance Analysis in terms of DPA

**B. Indications for trading :When to buy and sell stocks**

This section presents stock price prediction given by proposed RNN-LBL, SCSPF over existing stock forecasting model such as LSTM-FD and GAN-FD. The stock price forecasted by various existing models is shown in Figure 6. From Figure it is seen the RNN-LBL model highly correlate with actual stock price value when compared with LSTM-FD and GAN-FD. Further, incorporating sentiment factor in stock market analysis aid in improving forecasting performance which is experimentally shown in Figure 7 as SCSPF model achieves much better correlation with respect to actual stock price value when compared with RNN-LBL model. In Figure 12 when to buy and sell stock for achieving better profits is shown. For achieving better profit or loss it is important to identify the peak of both high and low. From figure it can be seen can establish the peak more efficiently. Thus, using SCSPF for stock recommendation of buying and selling will aid in achieving profit.



**Fig. 11 Stock price prediction by SCSPF over existing stock price forecasting model.**



**Fig 12. Indications for when to buy and sell using Direction Prediction Accuracy**

Here in Figure 12 shows when to buy stocks of a particular company using Green colored signal and when to sell the stocks shown using red colored signal. As stock prices starts increasing or decreasing gradient descent in SCSPF model senses and captures those fluctuations in the market and gives indications to the user.

**V. CONCLUSION**

The stock market is inherently unpredictable in nature. Stock market forecasting is a difficult job to take into consideration external fluctuating variables such as investor sentiments that differ over time according to the current unpredictable situations. It is therefore important to catch the feelings of the market. In this research paper, a literature survey on current stock price forecasting on various deep learning models is performed, but considering LSSC, very less work has been done. This research implemented the Sentiment Conscious SCSPF stock forecasting model to solve the problem of reducing forecasting errors and producing higher returns during the US election polls and the Covid-19 pandemic for short span of sales. The SCSPF divides the current session into several bins and compares these session bins to create stock behavioural patterns. It can be seen from the findings that a Direction Prediction Accuracy of 0.6423, 0.68585, 0.7822 and 0.80305 is achieved by the LSTM-FD, GAN-FD, RNN-LBL and SCSPF models respectively. The SCSPF increases the performance of forecasting by 18.26 percent over the current RNNLBL model. Therefore, SCSPF performs very well with higher returns on investments in a high frequency trading environment that is highly competitive and volatile in nature. In addition, it also provides investors with information about whether to purchase or sell stocks in order to achieve greater profitability.

**REFERENCES**

- [1] Q. Huang, J. Yang, X. Feng, A. W. Liew and X. Li, Automated Trading Point Forecasting Based on Bicluster Mining and Fuzzy Inference, in IEEE Transactions on Fuzzy Systems, 28(2) (2020) 259-272, doi: 10.1109/TFUZZ.2019.2904920.
- [2] J. Lee, R. Kim, Y. Koh and J. Kang, Global Stock Market Prediction Based on Stock Chart Images Using Deep Q-Network, in IEEE Access, 7, 167260-167277, (2019) doi: 10.1109/ACCESS.2019.2953542,.
- [3] J. H. Kim, A. Shamsuddin, and K. P. Lim, Stock return predictability and the adaptive markets hypothesis: Evidence from century-long U.S. data, Journal of empirical finance, 18(5) (2011) 868–879.
- [4] G. Schwert, Anomalies and market efficiency, Handbook of the Economics of Finance, 1 (2003) 939–974.
- [5] H. R. Stoll and R. E. Whaley, Transaction costs and the small firm effect, The Journal of Financial Economics. 12(1) (1983) 57–79.
- [6] C. H. Park and S. H. Irwin, What do we know about the profitability of technical analysis? Journal of Economic Surveys, 21(4) (2007) 786– 826.
- [7] J. Bollen, H. Mao, and Z. X., Twitter mood predicts the stock market, Journal of computational science, 2(1) (2011) 1–8.
- [8] D. Huang, F. Jiang, J. Tu, and G. Zhou, Investor Sentiment Aligned: A Powerful Predictor of Stock Returns, The Review of Financial Studies, 28(3) (2014) 791–837.
- [9] R. P. Schumaker and H. Chen, Textual analysis of stock market prediction using breaking financial news: The azfin text system,



- ACM Transactions on Information Systems, 27(2) (2019).
- [10] T. Preis, H. S. Moat, and S. H.E., Quantifying trading behavior in financial markets using google trends, Scientific reports, 3 (2013) 1684.
- [11] World Health Organisation. 2020a, Public Health Emergency of International Concern Declared.30. [https://www.who.int/docs/default-source/coronaviruse/transcripts/ihremergency-committee-for-pneumonia-due-to-the-novel-coronavirus-2019-ncov-press-briefing-transcript30012020.pdf?sfvrsn=c9463ac1\\_2](https://www.who.int/docs/default-source/coronaviruse/transcripts/ihremergency-committee-for-pneumonia-due-to-the-novel-coronavirus-2019-ncov-press-briefing-transcript30012020.pdf?sfvrsn=c9463ac1_2)
- [12] World Health Organisation. 2020b , WHO Characterizes COVID-19 as a Pandemic. 11 March. [https://www.who.int/docs/default-source/coronaviruse/transcripts/who-audio-emergencies-coronavirus-press-conference-full-and-final-11mar2020.pdf?sfvrsn=cb432bb3\\_2](https://www.who.int/docs/default-source/coronaviruse/transcripts/who-audio-emergencies-coronavirus-press-conference-full-and-final-11mar2020.pdf?sfvrsn=cb432bb3_2)
- [13] Duan, Hongbo, Shouyang Wang, and Cuihong Yang , Coronavirus: Limit Short-term Economic Damage. Nature 578 (7796): 515 (2020).
- [14] J. Han, X. Zhang and F. Wang, Gaussian Process Regression Stochastic Volatility Model for Financial Time Series, in IEEE Journal of Selected Topics in Signal Processing, 10(6) (2016) 1015-1028, doi: 10.1109/JSTSP.2016.2570738.
- [15] S.-G. Deng, and T.-H. Yeh, Using least squares support vector machines to the product cost estimation, Journal of Chung Cheng Institute of Technology, 40(1) (2011) 1-16.
- [16] F. Kaytez, M. C. Taplamacioglu, E. Cam, and F. Hardalac, Forecasting electricity consumption: A comparison of regression analysis, neural networks and least squares support vector machines, International Journal of Electrical Power and Energy Systems, 67 (2015) 431-438.
- [17] J. Chou and T. Nguyen, Forward Forecast of Stock Price Using Sliding-Window Metaheuristic-Optimized Machine-Learning Regression, in IEEE Transactions on Industrial Informatics. 14(7) (2018) 3132-3142, doi: 10.1109/TII.2018.2794389.
- [18] Gurav U., Sidnal N. Predict Stock Market Behavior: Role of Machine Learning Algorithms. In: Bhalla S., Bhateja V., Chandavale A., Hiwale A., Satapathy S. (eds) Intelligent Computing and Information and Communication. Advances in Intelligent Systems and Computing, 673 (2018). Springer, Singapore. [https://doi.org/10.1007/978-981-10-7245-1\\_38](https://doi.org/10.1007/978-981-10-7245-1_38)
- [19] Uma Prashant Gurav, S. Kotrappa, RNNLBL : A Recurrent Neural Network and Log Bilinear based Efficient Stock Forecasting Model ", International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, 9(4) (2020).
- [20] T. Fischer and C. Krauss, Deep learning with long short-term memory networks for financial market predictions, European Journal of Operational Research, 270(2) (2018) 654–669.
- [21] S. Hochreiter, and J. Schmidhuber, Long short-term memory, Neural computation, (9) (1997) 1735–1780,.
- [22] Y. Deng, F. Bao, Y. Kong, Z. Ren, and Q. Dai, Deep direct reinforcement learning for financial signal representation and trading, IEEE Transactions on Neural Networks and Learning Systems, 28(3) (2016) 653–664.
- [23] Zhou, Xingyu & Pan, Zhisong & Hu, Guyu & Tang, Siqi & Zhao, Cheng. Stock Market Prediction on High-Frequency Data Using Generative Adversarial Nets. Mathematical Problems in Engineering. (2018) 1-11. 10.1155/2018/4907423.
- [24] GURAV UMA and Kotrappa Dr. S., Lbl - Lstm : Log Bilinear And Long Short Term Memory Based Efficient Stock Forecasting Model Considering External Fluctuating Factor (2020). International Journal of Engineering and Advanced Technology (IJEAT), ISSN: 2249 – 8958, 9(4) (2020), Available at SSRN: <https://ssrn.com/abstract=3697046>.
- [25] Ding, G., Qin, L. Study on the prediction of stock price based on the associated network model of LSTM. Int. J. Mach. Learn. & Cyber. **11**, 1307–1317. <https://doi.org/10.1007/s13042-019-01041-1>, 2020.
- [26] Li, Hui & Hua, Jinjin & Li, Jinqiu & Li, Geng. (2020). Stock Forecasting Model FS-LSTM Based on the 5G Internet of Things. Wireless Communications and Mobile Computing. (2020) 1-7. 10.1155/2020/7681209,.
- [27] X.Teng, T. Wang, X. Zhang, L. Lan and Z. Luo. Enhancing Stock Price Trend Prediction via a Time-Sensitive Data Augmentation Method, Hindawi Complexity, Volume 2020, Article ID 6737951, (8) (2020) <https://doi.org/10.1155/2020/6737951>,.
- [28] R. Ren and D. Wu, An Innovative Sentiment Analysis to Measure Herd Behavior, in IEEE Transactions on Systems, Man, and Cybernetics: Systems, (2018) doi: 10.1109/TSMC.2018.2864942,.
- [29] Y. He, C. Wang and C. Jiang, Correlated Matrix Factorization for Recommendation with Implicit Feedback, in IEEE Transactions on Knowledge and Data Engineering. (2017) doi: 10.1109/TKDE.2018.2840993.
- [30] Covid dataset available at: <https://ieee-dataport.org/open-access/coronavirus-covid-19-tweets-dataset>, last accessed (30)
- [31] Twitter dataset available at: <https://twitter.com/ppollingnumbers?lang=en>.
- [32] Yahoo News available at <https://in.finance.yahoo.com/topic/latestnews>, last accessed July 30<sup>th</sup>.
- [33] Chinese stock dataset can be downloaded using Wind Financial Terminal Available at: <http://www.wind.com.cn>, last accessed (30).
- [34] A. L. Iao and A. C. Fajardo, Sentiment Analysis of Tweet Messages using Hybrid Approach Algorithm, 2019 17th International Conference on ICT and Knowledge Engineering (ICT&KE), Bangkok, Thailand, (2019) 1-6, doi: 10.1109/ICTKE47035.2019.8966887.
- [35] Hutto, C.J. & Gilbert, Eric. (2015). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. Proceedings of the 8th International Conference on Weblogs and Social Media, ICWSM 2014.J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., Oxford: Clarendon, (2) (1892) 68-73.
- [36] U. Gurav and N. Sidnal, Adaptive Stock Forecasting Model using Modified Backpropagation Neural Network (MBNN), 2018 International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS), Belgaum, India, (2018). 380-385, doi: 10.1109/CTEMS.2018.8769290.
- [37] Gurav Uma and Kotrappa, Dr. S., Impact of COVID-19 on Stock Market Performance Using Efficient and Predictive LBL-LSTM Based Mathematical Model (August 10, 2020). International Journal on Emerging Technologies, 11(4): 108-115(2020), DOI: 10.13140/RG.2.2.19198.41283. SSRN: <https://ssrn.com/abstract=3670565>,
- [38] Uma Gurav, Dr. Nandini Sidnal, Opinion mining for reputation evaluation on unstructured big data, International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), 4(4) (2015).
- [39] Dr. Ferdinand Kartriku, Robert Sowah, Charles Saah, Deep Neural Network: An Efficient and Optimized Machine Learning Paradigm for Reducing Genome Sequencing Error, International Journal of Engineering Trends and Technology 68(9) (2020):27-30.
- [40] Narendra Mohan, Recognition of Skin Diseases using Deep Neural Network Optimized by Group Teaching Algorithm, International Journal of Engineering Trends and Technology 68(9) (2020):109-120.