

Sentiment Aware Stock Price Forecasting using the SA-RNN-LBL Learning Model

Uma Prashant Gurav

Dept. of Computer Science and Engineering
KIT's College of Engineering
Kolhapur, India
umabgurav@gmail.com

S. Kotrappa

Dept. of Computer Science and Engineering
KLE Dr. MSSCET
Belgaum, India
kotrappa06@gmail.com

Abstract—Stock market historical information is often utilized in technical analyses for identifying and evaluating patterns that could be utilized to achieve profits in trading. Although technical analysis utilizing various measures has been proven to be helpful for forecasting and predicting price trends, its utilization in formulating trading orders and rules in an automated system is complex due to the indeterminate nature of the rules. Moreover, it is hard to define a specific combination of technical measures that identify better trading rules and points, since stocks might be affected by different external factors. Thus, it is important to incorporate investors' sentiments in forecasting operations, considering dynamically the varying stock behavior. This paper presents a sentiment aware stock forecasting model using a Log BiLinear (LBL) model for learning short term stock market sentiment patterns, and a Recurrent Neural Network (RNN) for learning long-term stock market sentiment patterns. The Sentiment Aware Stock Price Forecasting (SASPF) model achieves a much superior performance compared to standard deep learning based stock price forecasting models.

Keywords—sentiment analysis; data mining; machine learning; social networking platform; stock price forecasting; time series

I. INTRODUCTION

Analyzing and forecasting future stock prices has attracted much research interest. Success in stock trading totally relies on determining the right time to sell or buy stocks. However, predicting accurate stock prices is a very challenging task due to the congenitally noisy environment and the high volatility of price variations [1]. Stock price movement is related to a number of elements including government policies, organization's performance, inflation rates and economic environment. Technical analysis examines the variations of a stock through mathematical formulas and charts, called technical indicators, aiming to discover patterns that could be utilized for profitable gains in trading, and it is a usual approach in financial engineering for stock market prediction [1]. Despite the advances in technical analysis, it is difficult to accurately predict stock prices and trading points in financial markets, mostly because of the volatile investors' sentiments during uncertainties. Authors in [2] proclaimed that the stock market is very efficient and trading prices totally reflect the available data, which is referred as the Efficient Market Hypothesis (EMH), while fundamental or technical analysis

would not lead to profit that is consistently over-average for investors. Nonetheless, not all researchers agree with EMH [3], as some conventional methods in financial economics proclaim that stock markets have anomalies [4-5]. Most studies use technical analysis, which includes older stock volumes and prices, in order to demonstrate trading profits and predict future stock market prices [6]. Research studies focused on discovering native parameters, mostly used web data such as Social Networking Services (SNS) [7], investor's sentiments [8], news articles [9], or search engines [10], and explained that search query frequencies and investors' sentiments from SNS platforms provide vital information for predicting upcoming stock prices, especially during global uncertainties such as the recent CoViD-19 pandemic. Initially, during December 2019 to March 2020 CoViD-19 cases were significantly high in the Republic of China affecting its stock market. Due to the rigid expenditure on rents, wages, and bank interests the stability of the banking system was affected and small-medium sized organizations were significantly impacted as the Chinese stock index nearly dropped by 8% on February 3rd [11-13]. On the other side, Reliance Industries Limited (RIL) stock prices were continuously growing as Facebook and Google obtained certain percentage of its stake during the CoViD-19 pandemic.

During the previous years, researchers mostly focused on volatility modeling, evaluated as the standard deviation of an asset's return, showing the variability in Finance Time Series (FTS) data [14], which play a crucial role in effective stock forecasting. Extensive surveys were carried out showing limitations of the existing Machine Learning (ML) based stock forecasting models, based on random forest, neural networks, and support vector machines, as they proved inefficient in highly volatile market behaviors. Furthermore, in order to confront the dynamic nature of stock markets, a number of deep learning models such as RNNs with gated recurrent and long-short term memories, Convolutional Neural Networks (CNN), Deep Neural Networks (DNN), and General Adversary Networks (GAN) [15] have been presented. Neural networks and Reinforcement Learning (RL) methods are the most widely utilized ML methods [16, 17]. Existing deep learning models have limitations, as although they are efficient in modeling long-term market sentiment behaviors they do not detect short-term market sentiment behaviors. Modern ML methods have been adapted in various domains such as speech recognition,

Corresponding author: Uma Prashant Gurav

image classification, and natural language processing, and they are designed using either RNNs or CNNs [18]. Long Short Term Memory (LSTM) [19-22] networks outperform random forests and DNNs. In [23], an RL method was presented using RNN architecture for addressing gradient descent problems during the training process. Correspondingly, in [15] a forecasting error loss and direction prediction accuracy method was proposed, noting that Generative Adversarial (GA) training can be used to integrate the loss parameter for acquiring the long-term outcomes. The GAN-FD used CNN and LSTM [24-26] for predicting more accurately stock prices. Nonetheless, it failed to get a good learning tradeoff performance of modeling both short and long-term market behavioral contexts. Therefore, the GAN-FD suffers considerably when market's behaviors are high volatile in nature, while market's short-term and long term sentiment pattern is not considered [15].

II. THE SENTIMENT AWARE STOCK PRICE FORECASTING MODEL (SASPF)

The highly dynamic stock forecasting environment during the pandemic requires the modeling of both short and long-term sentiment behaviors of the market. To the best of our knowledge, no prior work has considered capturing market sentiment in dynamic manner for forecasting stock price. This paper presents a Sentiment Aware Stock Price Forecasting (SASPF) model combining short and long term sentiment features. At first, a sentiment analysis was carried out on review datasets, marked it as “-1” for negative, “0” for neutral and “1” for positive sentiment. This sentiment forecast data were added to the stock information datasets and trained using a sentiment aware stock forecasting model. The model was built utilizing both an RNN and an LBL. The RNN was used for detecting long-term market features, while the LBL was used for detecting short-term features. The main contributions of this paper are:

- A sentiment aware RNN-LBL (SA-RNN-LBL) learning model is presented for forecasting stock prices considering dynamically varying stock sentiment.
- The SA-RNN-LBL model can capture both short-term and long-term market sentiment features more efficiently.
- The SA-RNN-LBL model can obtain multiple market sentiment behaviors within short-term context for current trading sessions. Thus, it can predict the stock value fluctuation time more efficiently and can be used for forecasting the price of short sells.
- The experimental outcome shows the proposed sentiment aware RNN-LBL based stock forecasting model attain better RMSRE and DPA performance than existing [15] deep learning based stock price forecasting model.

High-quality one-step forecasting provides significant information for risk assessment and management in trading environments, especially during pandemics. This work aims to forecast the price fluctuation of individual stocks or the market index one step ahead, using historical data and market sentiments derived from tweets.

A. Mathematical Formulation

This problem can be mathematically formalized as follows: X_t represents a set of basic indicators including sentiment parameters and Y denotes the closing price of a stock for 1-minute intervals at time t , described as:

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

Given the historical basic indicators, X is described as:

$$X = \{X_1, X_2, \dots, X_T\} \quad (2)$$

The closing price Y is described as:

$$Y = \{Y_1, Y_2, \dots, Y_T\} \quad (3)$$

The goal is to predict the closing price Y_{T+1} for the next minute considering the market sentiment factor. Lets consider a set of stock market and set of stock within stock market as follows:

$$U\{u_1, u_2, \dots\}, \quad (4)$$

$$V = \{v_1, v_2, \dots\} \quad (5)$$

This work considers stock market environment, which is composed of following stock volatility behavior such as:

$$C = \{c_1, c_2, c_3, c_4\} \quad (6)$$

Similarly, in the stock performance pattern there exist:

$$C = \{c_1, c_2, c_3\} \quad (7)$$

behaviors. The final task is to predict the future price of a particular stock in a stock market using the sentiment aware stock price forecasting methodology.

B. Sentiment Analysis Tool for Predicting Market Sentiments from Tweet Data

The sentiment analysis model detects and marks market sentiments having positive (1), neutral (0), and negative (-1) polarity. However, using a state-of-art sentiment analyzer tool is bulky in nature and induces high computation overhead. For addressing this problem, a sentiment analysis method by employing Partial Text Entailment (PTE) was utilized. The PTE is used for grouping similar tweets by measuring the semantic similarities among different tweets. This aided in improving the speed of market sentiment prediction, meeting the real-time requirements of stock forecasting models.

C. Model's Description

This section presents the sentiment aware stock market pattern analysis model by learning both short and long-term behaviors using LBL and RNN respectively similar to the LSTM model. The RNN model is composed of a set of input layers, multiple hidden layers, and an output layer. The activation parameters of the hidden layers are obtained as:

$$i_\ell^v = f(\mathcal{X}i_\ell^v + \mathcal{D}s_{w_\ell^v}) \quad (8)$$

where, $i_\ell^v \in \mathbb{S}^e$ depicts the hidden illustration of stock v at time instance (i.e., position) ℓ in a series, and $s_{w_\ell^v} \in \mathbb{S}^e$ depicts the illustration of the ℓ^{th} input stock of a particular stock market v . The Activation Function (AF) is represented by $f(i)$

and the Transition Matrix (TM) of the present stock is represented as:

$$\mathcal{D} \in \mathbb{S}^e \quad (9)$$

while the previous status is represented as:

$$\mathcal{W} \in \mathbb{S}^e. \quad (10)$$

where \mathcal{D} can obtain stock's present volatility (i.e. behavior pattern) and \mathcal{X} can propagate time series signals. Equation (8) is executed iteratively to obtain or compute the status of each time instance in a time series sequence.

The RNN is composed of multiple hidden layers. Hidden layer information is dynamic in nature with respect to stock market behavior sequence where the pattern is repetitive. Thus, the RNN faces problems in learning short-term patterns in a stock market behavior sequence. For addressing this limitation, an LBL was appended with a single linear hidden layer, and it can be considered as a deterministic model. Using LBL for stock market behavior forecasting, the absolute forecasting representation of time sequence is constructed based on stock market score/price input and TMs at each time instance. The next time instance is a linear forecasting, described as:

$$i_\ell^\sigma = \sum_{j=0}^{\sigma-1} \mathcal{D}_j s_{w_\ell^\sigma-j}, \quad (11)$$

where $\mathcal{D}_j \in \mathbb{S}^{e \times e}$ depicts the TM for the respective time instance in a stock market score behavior sequence and σ is the number of elements modeled in a time sequence. Each position in the time sequence is modeled with a precise TM. In general, LBL finds difficulties in efficiently learning long-term contexts in stock market behavior sequence. Furthermore, the market behaves dynamically depending on market sentiments (i.e. the moods of customers and investors). Thus, it is important to capture such information into stock price forecasting. Thus the proposed method considers sentiment aware matrices for obtaining feature sets of various aspects of sentiments. Thus, for the respective stock market σ at ℓ (i.e. hidden description) is computed as:

$$i_\ell^\sigma = \mathcal{X} i_\ell^\sigma + \sum_{j=0}^{\sigma-1} \mathcal{D}_j \mathcal{N}_{c_{\ell-j}^\sigma} s_{w_\ell^\sigma-j}, \quad (12)$$

where $\mathcal{N}_{c_{\ell-j}^\sigma} \in \mathbb{S}^{e \times e}$ are sentiment aware TMs with respect to the j^{th} stock of a stock market σ . The cold start problem is overcome by assuming $i_0^\sigma = v_0$. It should be noted that if any change in sentiments is detected, the proposed method can obtain underlying feature sets of different sentiments exhibited in past sessions. However, the existing stock forecasting model in general doesn't consider session variance between features, playing a significant role in improving prediction accuracies. This happens because short-term market sentiment variance plays a significant part in future buying or selling of stocks compared to long-term market sentiment information. Thus, this work improves the previous RNN-LBL model by incorporating time based sentiment feature variance information, modeling a sentiment aware RNN-LBL model. The sentiment aware RNN-LBL stock market forecasting model is shown in Figure 1.

Considering the respective stock market v , the position l is established using:

$$i_l^v = \mathcal{X} i_{l-1}^v + \sum_{j=0}^{\sigma-1} U_{u_l^v-u_{l-1}^v} s_{w_{l-1}^v-j}, \quad (13)$$

where u_l^v represents current time, u_{l-1}^v represents time information of each stock of each layer of the model, and $U_{u_l^v-u_{l-1}^v}$ is the time-specific TM of sentiment session variation $u_{l-1}^v - u_l^v$. The time-specific TMs aid in collecting time specific sentiment pattern features with respect to recent tweet information. Considering these, (13) is updated as:

$$i_l^v = \mathcal{X} i_0^v + \sum_{j=0}^{l-1} U_{u_l^v-u_{l-1}^v} s_{w_{l-1}^v-j}, \quad (14)$$

where $i_l^v = v_0$ with respect to the initial status of the respective stock within the stock market environment. For obtaining the varying sentiment nature of stock market with respect to different time intervals, this work established sentiment aware TMs in the model, described by:

$$i_l^v = \mathcal{X} i_{l-1}^v + \sum_{j=0}^{l-1} U_{u_l^v-u_{l-1}^v} N_{c_{l-m}^v} s_{w_{l-1}^v-j}. \quad (15)$$

Finally, the prediction of whether a stock v will exhibit certain patterns or sentiments c with respect to stock w at successive position $l + 1$ is calculated by:

$$z_{v,l+1,c,w} = (t_\ell^\sigma)^U N_{c s_w} = (i_\ell^v + v_v)^U N_{c s_w}. \quad (16)$$

where t_ℓ^σ is the current sentiment of the stock market σ with respect to session instance ℓ , with static hidden and dynamic feature illustration being $v_\sigma \in \mathbb{S}^e$ and i_ℓ^σ , respectively.

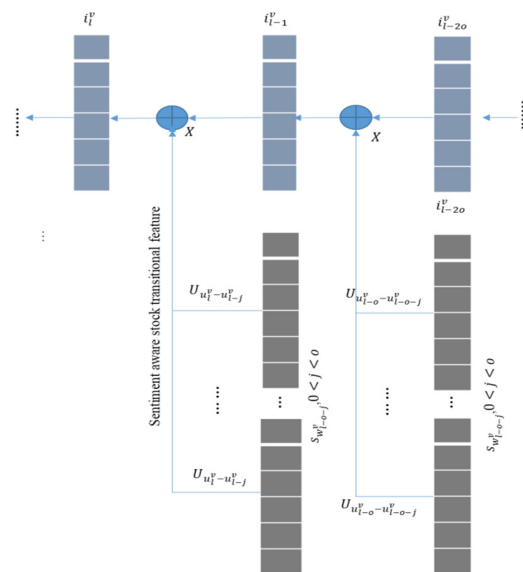


Fig. 1. The architecture of the sentiment aware RNN-LBL model.

In order to learn exclusive possible time variances in a continuous manner, it is important to establish large sized time-specific TMs, but this introduces over-fitting issues. This problem can be addressed by dividing time variance into equal windows. This model estimates transitional matrices of lower and upper limits of time windows/bins, utilizing linear interpolation of entire time variance in time windows, in order to establish the corresponding TMs. Experiments were conducted in order to evaluate the outcome of the proposed SASPF model and are described below.

III. RESULTS AND DISCUSSION

This section presents the performance evaluation of the proposed SASPF model over other prediction models. The GAN-FD model [15] was chosen for comparison as it achieved much better results than existing LSTM based stock forecasting methods [21, 24-28], as investors’ sentiments are considered as a major contributing parameter.

```

iCreate a list of columns to keep
keep_columns = [ 'Open', 'High', 'Low', 'Volume', 'Subjectivity', 'Polarity', 'Compound', 'Negative', 'Neutral', 'Positive', 'Label' ]
df = merge(keep_columns)
df
    
```

	Open	High	Low	Volume	Subjectivity	Polarity	Compound	Negative	Neutral	Positive	Label
1988	11432.080844	11759.959961	11388.040039	212830000	0.267549	-0.048568	-0.9983	0.212	0.729	0.059	0
1987	11729.669922	11867.110352	11675.530273	183190000	0.374896	0.121956	-0.9977	0.202	0.738	0.061	1
1986	11781.700195	11782.349609	11681.519531	173580000	0.536234	-0.044302	-0.9975	0.225	0.684	0.091	0
1985	11632.809570	11633.780273	11453.339644	102550000	0.364021	0.011398	-0.9571	0.132	0.767	0.102	0
1984	11532.070312	11718.280273	11450.889648	159790000	0.375099	-0.040677	-0.9644	0.148	0.758	0.094	1
...
4	17355.210938	17355.210938	17063.080078	138740000	0.352311	-0.014015	-0.9882	0.189	0.717	0.094	0
3	17190.509766	17489.720703	17190.509766	112190000	0.352649	0.046560	-0.9809	0.146	0.788	0.066	1
2	17456.019531	17704.509766	17456.019531	106380000	0.389617	0.052622	-0.9715	0.128	0.816	0.056	1
1	17712.759766	17930.609375	17711.800781	133030000	0.382566	0.011243	-0.9858	0.191	0.721	0.089	1
0	17524.240234	18002.380859	17916.910156	82160000	0.320261	-0.035458	-0.9982	0.235	0.724	0.041	1

1989 rows x 11 columns

Fig. 2. Results obtained for sentiment index (positive, negative, neutral and compound).

The proposed SASPF model was implemented using python. The dataset for evaluating sentiments was obtained from public available sources [29-31]. This work evaluated performance using the same datasets used by GAN-FD [15]. Data were composed of 244 trading points ranging from January 1 to December 31st 2016. More detailed description of the dataset used can be obtained from [15], while it can be downloaded from [32]. Direction Prediction Accuracy (DPA) and Root Mean Squared Relative Error (RMSRE) were used for performance evaluation. The 80% of the data was used for training, while the remaining 20% was used for testing. Experiments were conducted for different stocks considering time t , a forecasting was performed for the next time interval $t + 1$ using the proposed learning methods. The total time interval for the forecasting was T_0 . The actual stock value is Y_t and the forecasted value is \hat{Y}_t .

A. RMSRE Performance Evaluation

This section evaluated the RMSRE’s performance achieved by SASPF model over existing deep learning stock price forecasting methods. The RMSRE metric was computed as:

$$RMSRE = \sqrt{\frac{1}{T_0} \sum_{t=1}^{T_0} \left(\frac{Y_{t+1} - \hat{Y}_{t+1}}{Y_{t+1}} \right)^2} \quad (17)$$

A low RMSRE value indicates that forecasting is good (i.e. it defines it as closer to the actual stock price). The RMSRE outcomes of the SASPF model and existing deep learning based stock price forecasting methods are shown in Table I. The graphical representations of RMSRE performance achieved by SASPF and existing deep learning based stock price forecasting methods are shown in Figure 3.

TABLE I. RMSRE PERFORMANCE COMPARISON

Method used	RMSRE
LSTM-FD	0.0197
GAN-FD	0.00885
RNN-LBL	0.00675
SASPF	0.005465

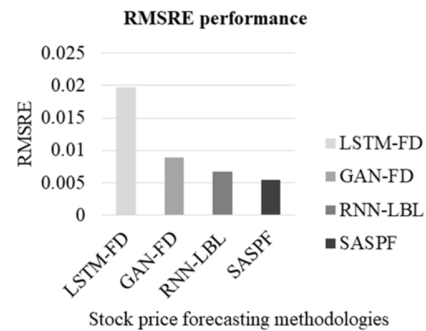


Fig. 3. RMSRE performance of SASPF model and existing deep learning based stock price forecasting methods.

The results showed that the average RMSRE performance achieved by LSTM-FD, GAN-FD, RNN-LBL, and SASPF was 0.0197, 0.00885, 0.00675, and 0.005465, respectively. The RNN-LBL model had lower RMSRE by 65.736% and 23.73% over LSTM-FD and GAN-FD, respectively. Similarly, the SASPF model had lower RMSRE by 72.41%, 38.59% and 19.48% over LSTM-FD, GAN-FD and RNN-LBL models, respectively. The results show that SASPF model achieves much superior RMSRE performance than the existing deep learning based stock price forecasting models.

In order to further evaluate the performance of the SASPF model, a performance evaluation was performed under highly dynamic and volatile environments. For evaluating such scenarios, this work considered the evaluation using the Reliance Jio communication stock and a corresponding tweet dataset obtained from Yahoo Finance [30] and Yahoo news [31] datasets, respectively. The stock price forecasting performance was evaluated between March 1st and June 30th 2020. The RMSRE performance achieved considering the sentiment index is shown in Figure 4. The results show that RNN-LBL and SASPF model achieves an RMSRE of 0.0085 and 0.0059, respectively. SASPF had improved forecasting performance by 30.59% over RNN-LBL considering stock market sentiments. Thus, SASPF can work really well under highly dynamic and volatile environments such as the CoViD-19 pandemic scenario.

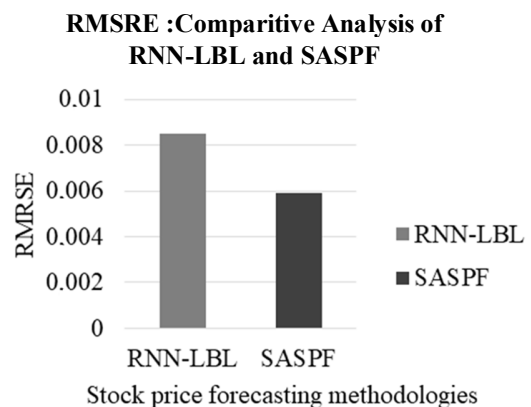


Fig. 4. Comparitive Analysis of SASPF and RNN-LBL models in terms of RMSRE

B. DPA Performance Evaluation

This section evaluated the DPA performance achieved by the SASPF model over existing deep learning stock price forecasting methods. The higher value of DPA describes higher revenue for the investor. The DPA metric is computed as:

$$DPA = \frac{100}{T_0} \sum_{t=1}^{T_0} I_t, \quad (18)$$

where:

$$I_t = \begin{cases} 1 & \text{if } (Y_{t+1} - Y_t)(\hat{Y}_{t+1} - Y_t) > 0 \\ 0 & \text{otherwise} \end{cases} \quad (19)$$

The DPA outcomes of the SASPF and existing deep learning based stock price forecasting methods are shown in Table II, while the graphical representation of the results is shown in Figure 5.

TABLE II. DPA PERFORMANCE

Method used	DPA
LSTM-FD	0.6423
GAN-FD	0.68585
RNN-LBL	0.7822
SASPF	0.80305

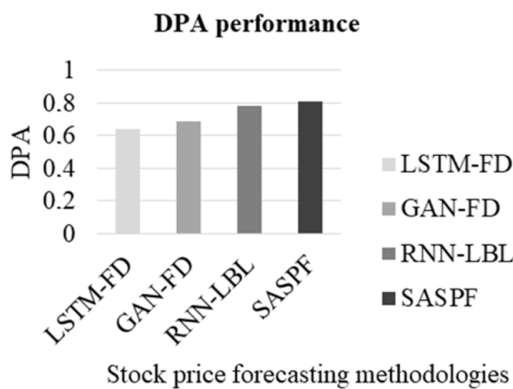


Fig. 5. DPA performance of SASPF and existing stock forecasting methods.

The results showed that the average DPA performance achieved by LSTM-FD, GAN-FG, RNN-LBL, and SASPF model was 0.6423, 0.68585, 0.7822, and 0.80805, respectively. The RNN-LBL model had better DPA performance by 17.88% and 12.32% over LSTM-FD and GAN-FD. Similarly, the SASPF model had better DPA performance by 20.51%, 15.123%, and 3.2% over LSTM-FD, GAN-FD, and RNN-LBL model, respectively. The SASPF model achieves superior DPA performance than the existing deep learning based stock price forecasting models. Furthermore, in order to evaluate the SASPF model's performance under highly dynamic and volatile environments, the model was evaluated using Reliance Jio communication stock and corresponding tweet datasets obtained from [30, 31]. Its stock price forecasting performance was evaluated with data between March 1st and June 30th of 2020. The DPA performance achieved considering the sentiment index is shown in Figure 6. The results show that the SASPF and RNN-LBL achieved DPAs of 0.0085 and 0.0059 respectively. The SASPF had better forecasting performance

by 18.26% over the RNN-LBL when considering stock market sentiments. Thus, SASPF can work really well under highly dynamic and volatile environments such as the CoViD-19 pandemic scenario, providing better profits for an investor.

DPA: Comparitive Analysis of RNN-LBL and SASPF

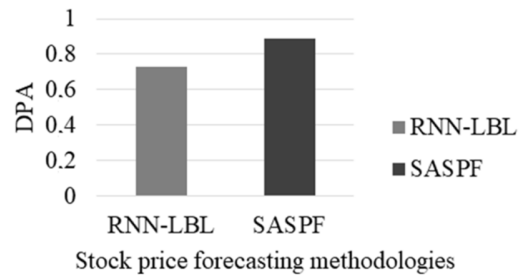


Fig. 6. Comparitive Analysis of SASPF and RNN-LBL models in terms of DPA.

IV. CONCLUSION

Predicting stock prices is a challenging task, especially when considering external fluctuating factors such as a pandemic situation, since stock markets are volatile in nature as user sentiments, behaviors, and patterns vary over time. During the pandemic situation some stocks such as oil price fell, and contrarily the stock prices of gold and mobile communication sector increased significantly. Thus, it is important to capture market sentiments in order to design a better forecasting model. Very few researches integrated market sentiments into stock prices forecasting models. In order to predict better profits, reduce forecasting errors, and achieve higher returns for short sell, this paper presented a sentiment aware stock price forecasting model. The SASPF divides the current session into multiple bins and compares them to establish stock behavior patterns. Results showed the SASPF improved forecasting performance by 18.26% over RNN-LBL when considering stock market sentiments. Thus, SASPF can work really well under highly dynamic and volatile environments such as the CoViD-19 pandemic scenario, providing better profits for investors. Experimental outcomes showed that the SASPF model achieved better RMSRE and DPA performance than existing deep learning based stock price forecasting methods. Future work should consider designing an efficient stock price forecasting framework combining both micro and macro-economic data with different machine learning and deep learning methods.

REFERENCES

- [1] Q. Huang, J. Yang, X. Feng, A. W.-C. Liew, and X. Li, "Automated Trading Point Forecasting Based on Bicluster Mining and Fuzzy Inference," *IEEE Transactions on Fuzzy Systems*, vol. 28, no. 2, pp. 259–272, Feb. 2020, 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," *IEEE Access*, vol. 7, pp. 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*, vol. 18, no. 5, pp. 868–879, Dec. 2011, doi: 10.1016/j.jempfin.2011.08.002.

- [4] G. W. Schwert, "Chapter 15 Anomalies and market efficiency," in *Handbook of the Economics of Finance*, vol. 1, Elsevier, 2003, pp. 939–974.
- [5] H. R. Stoll and R. E. Whaley, "Transaction costs and the small firm effect," *Journal of Financial Economics*, vol. 12, no. 1, pp. 57–79, Jun. 1983, doi: 10.1016/0304-405X(83)90027-2.
- [6] C. H. Park and S. H. Irwin, "What Do We Know About the Profitability of Technical Analysis?," *Journal of Economic Surveys*, vol. 21, no. 4, pp. 786–826, 2007, doi: 10.1111/j.1467-6419.2007.00519.x.
- [7] J. Bollen, H. Mao, and X. Zeng, "Twitter mood predicts the stock market," *Journal of Computational Science*, vol. 2, no. 1, pp. 1–8, Mar. 2011, doi: 10.1016/j.jocs.2010.12.007.
- [8] D. Huang, F. Jiang, J. Tu, and G. Zhou, "Investor Sentiment Aligned: A Powerful Predictor of Stock Returns," *The Review of Financial Studies*, vol. 28, no. 3, pp. 791–837, Mar. 2015, doi: 10.1093/rfs/hhu080.
- [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*, vol. 27, no. 2, p. 12:1–12:19, Mar. 2009, doi: 10.1145/1462198.1462204.
- [10] T. Preis, H. S. Moat, and H. E. Stanley, "Quantifying Trading Behavior in Financial Markets Using Google Trends," *Scientific Reports*, vol. 3, no. 1, Art. no. 1684, Apr. 2013, doi: 10.1038/srep01684.
- [11] "WHO | Novel coronavirus (COVID-19)," WHO. http://www.who.int/bulletin/online_first/COVID-19/en/ (accessed Sep. 30, 2020).
- [12] World Health Organisation, "Virtual press conference on COVID-19," Mar. 11 2020. [Online]. Available: https://www.who.int/docs/default-source/coronaviruse/transcripts/who-audio-emergencies-coronavirus-press-conference-full-and-final-11mar2020.pdf?sfvrsn=cb432bb3_2
- [13] H. Duan, S. Wang, and C. Yang, "Coronavirus: limit short-term economic damage," *Nature*, vol. 578, no. 7796, pp. 515–515, Feb. 2020, doi: 10.1038/d41586-020-00522-6.
- [14] J. Han, X.-P. Zhang, and F. Wang, "Gaussian Process Regression Stochastic Volatility Model for Financial Time Series," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 6, pp. 1015–1028, Sep. 2016, doi: 10.1109/JSTSP.2016.2570738.
- [15] X. Zhou, Z. Pan, G. Hu, S. Tang, and C. Zhao, "Stock Market Prediction on High-Frequency Data Using Generative Adversarial Nets," *Mathematical Problems in Engineering*, Art. no. 4907423, Apr. 2018, doi: 10.1155/2018/4907423.
- [16] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.
- [17] U. Gurav and N. Sidal, "Predict Stock Market Behavior: Role of Machine Learning Algorithms," in *Intelligent Computing and Information and Communication*, Singapore, 2018, pp. 383–394, doi: 10.1007/978-981-10-7245-1_38.
- [18] P. Chakraborty and C. Tharini, "Pneumonia and Eye Disease Detection using Convolutional Neural Networks," *Engineering, Technology & Applied Science Research*, vol. 10, no. 3, pp. 5769–5774, Jun. 2020.
- [19] U. Gurav and S. Kotrappa, "Lbl - Lstm: Log Bilinear And Long Short Term Memory Based Efficient Stock Forecasting Model Considering External Fluctuating Factor," *International Journal of Engineering and Advanced Technology*, vol. 9, no. 4, pp. 2057–2063, Apr. 2020, doi: 10.35940/ijeat.D8680.049420.
- [20] Uma Gurav, Dr. Nandini Sidal, "Opinion mining for reputation evaluation on unstructured big data", *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, Volume 4 Issue 4, April 2015.
- [21] U. Gurav and D. S. Kotrappa, "Impact of COVID-19 on Stock Market Performance Using Efficient and Predictive LBL-LSTM Based Mathematical Model," *International Journal on Emerging Technologies*, vol. 11, no. 4, pp. 108–115, Aug. 2020.
- [22] S. L. Sanga, D. Machuve, and K. Jomanga, "Mobile-based Deep Learning Models for Banana Disease Detection," *Engineering, Technology & Applied Science Research*, vol. 10, no. 3, pp. 5674–5677, Jun. 2020.
- [23] 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*, vol. 28, no. 3, pp. 653–664, Mar. 2017, doi: 10.1109/TNNLS.2016.2522401.
- [24] G. Ding and L. Qin, "Study on the prediction of stock price based on the associated network model of LSTM," *International Journal of Machine Learning and Cybernetics*, vol. 11, no. 6, pp. 1307–1317, Jun. 2020, doi: 10.1007/s13042-019-01041-1.
- [25] H. Li, J. Hua, J. Li, and G. Li, "Stock Forecasting Model FS-LSTM Based on the 5G Internet of Things," *Wireless Communications and Mobile Computing*, Art. no. 7681209, Jun. 2020, doi: 10.1155/2020/7681209.
- [26] X. Teng, T. Wang, X. Zhang, L. Lan, and Z. Luo, "Enhancing Stock Price Trend Prediction via a Time-Sensitive Data Augmentation Method," *Complexity*, Art. no. 6737951, Feb. 2020, doi: 10.1155/2020/6737951.
- [27] U. Gurav and S. Kotrappa, "RNNLBL: A Recurrent Neural Network and Log Bilinearbased Efficient StockForecastingModel," *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, vol. 9, no. 4, pp. 1676–1682, Feb. 2020, doi: 10.35940/ijitee.D1555.029420.
- [28] U. Gurav and N. Sidal, "Adaptive Stock Forecasting Model using Modified Backpropagation Neural Network (MBNN)," in *2018 International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS)*, Dec. 2018, pp. 380–385, doi: 10.1109/CTEMS.2018.8769290.
- [29] *Coronavirus (COVID-19) Tweets Dataset*, IEEE DataPort. [Online]. Available: <https://iee-dataport.org/open-access/coronavirus-covid-19-tweets-dataset>. Accessed: Jul. 30, 2020.
- [30] "Yahoo Finance." <https://in.finance.yahoo.com/> (accessed Jul. 30, 2020).
- [31] "Yahoo Finance." <https://in.finance.yahoo.com/topic/latestnews/> (accessed Jul. 30, 2020).
- [32] "Wind Financial Terminal – Chinese Stock Dataset." <https://www.wind.com.cn> (accessed Jul. 30, 2020).