Stacked Bidirectional Long Short-Term Memory for Stock Market Analysis

Jing Yee Lim
Faculty of Information Science & Technology
Multimedia University
Jalan Ayer Keroh Lama, 75450, Melaka, Malaysia.
1171303306@student.mmu.edu.my

Kian Ming Lim
Faculty of Information Science & Technology
Multimedia University
Jalan Ayer Keroh Lama, 75450, Melaka, Malaysia.
kmlim@mmu.edu.my

Chin Poo Lee
Faculty of Information Science & Technology
Multimedia University
Jalan Ayer Keroh Lama, 75450, Melaka, Malaysia.
cplee@mmu.edu.my

2021 IEEE International Conference on Artificial Intelligence in Engineering and Technology (IICAIET) | 978-1-6654-2899-6/21/$31.00 ©2021 IEEE | DOI: 10.1109/IICAIET51634.2021.9573812

***Abstract*— Stock market prediction is a difficult task as it is extremely complex and volatile. Researchers are exploring methods to obtain good performance in stock market prediction. In this paper, we propose a Stacked Bidirectional Long Short-Term Memory (SBLSTM) network for stock market prediction. The proposed SBLSTM stacks three bidirectional LSTM networks to form a deep neural network model that can gain better prediction performance in the stock price forecasting. Unlike LSTM-based methods, the proposed SBLSTM uses bidirectional LSTM layers to obtain the temporal information in both forward and backward directions. In this way, the long-term dependencies from the past and future stock market values are encapsulated. The performance of the proposed SBLSTM is evaluated on six datasets collected from Yahoo Finance. Additionally, the proposed SBLSTM is compared with the state-of-the-art methods using root mean square error. The empirical studies on six datasets demonstrates that the proposed SBLSTM outperforms the state-of-the-art methods.**

Keywords*— *Stock market prediction, Stacked Bidirectional Long Short-Term Memory, Long Short-Term Memory

I. Introduction

Stock market analysis is one of the most popular topics in the finance field. However, the price of the stock market is complex and changing over time unexpectedly and abruptly. This has gained the attention of the researchers to try to propose a variety of methods to obtain best performance for the stock market prediction. To date, the stock market analysis methods developed in both conventional statistical methods and machine learning techniques. Statistical methods such as Simple Moving Average (SMA) [1], Autoregressive Integrated Moving Average (ARIMA) [2], and machine learning techniques such as Artificial Neural Networks (ANN) [2], Deep Neural Networks (DNN) [4], have been widely applied in stock market prediction. Unfortunately, these methods are unable to learn the temporal details in stock market price time-series data. Therefore, these methods did not achieve promising performance in stock market prediction.

In view of this, this paper proposes a deep neural network stock market prediction model named Stacked Bidirectional Long Short-Term Memory (SBLSTM) which adopts a bidirectional structure that not only capture the temporal information of stock market price in forward direction, but also in the backward direction. Moreover, the Bidirectional Long Short-Term Memory (BiLSTM) layers are stacked in both forward and backward direction. This can further enhance the performance in the stock market prediction task.

This research is supported by the Fundamental Research Grant Scheme of the Ministry of Higher Education under the award number FRGS/1/2019/ICT02/MMU/03/7.

978-1-6654-2899-6/21/$31.00 ©2021 IEEE

II. Related Works

Generally, the approaches used to do prediction on stock market price can be classified into three main categories, which are statistical methods, machine learning methods, and also deep learning methods.

In an early work [1], SMA was proposed to predict the currency rates of Forex. In their work, SMA averaged the input values over a period of time. They then evaluated the performance of their proposed SMA using Mean Square Error (MSE). Another work by Ariyo et. al. [3] proposed ARIMA for this time series prediction task. In statistical methods, the complexity of the data needs to be linear. Stock market data, on the other hand, is a non-linear time-series data. As a result, these statistical approaches are ineffective for stock market forecasting. This has prompted the researchers to investigate machine learning approaches, for example, ANN, and Long Short-Term Memory (LSTM), given the nonlinear mapping and generalization capabilities.

In [2], ARIMA model was compared with the ANN model for stock market prediction. ARIMA and ANN models had achieved the Forecast Error (FE) of 0.439 and 0.604 respectively. This proved that the ANN model can outperform the statistical methods in the prediction of the stock market price values. Another work by Vijh et. al. [5] also proposed ANN to forecast the stock market price and compared with the performance of the Random Forest model. The prediction results proved that ANN performs better than Random Forest in the prediction of the stock market price. Later, Bommareddy et. al. proposed to use Linear Regression to perform prediction of the values of the stock price. Several pre-processing steps were carried out on the stock data to extract the compact features before feed into the Linear Regression. Some researchers hybrid the ANN with computational intelligence methods to aim for higher performance in stock market prediction. In [6], Genetic Algorithm (GA) was proposed to combine with ANN to optimize the performance of a single ANN model. The proposed GA eliminated the local convergence problem of the ANN model which was trained by the backpropagation algorithm. Another work [7] proposed a prediction method by combining ANN model with a Decision Tree and achieved higher performance in stock market prediction. These researches summarized that ANN has the strong capability in analyzing the complex pattern of the time series data.

Inspired by the successes of deep learning in many vision [8, 9, 10, 11] and audio [12, 13] tasks, researchers began to explore deep learning methods in stock market prediction. In

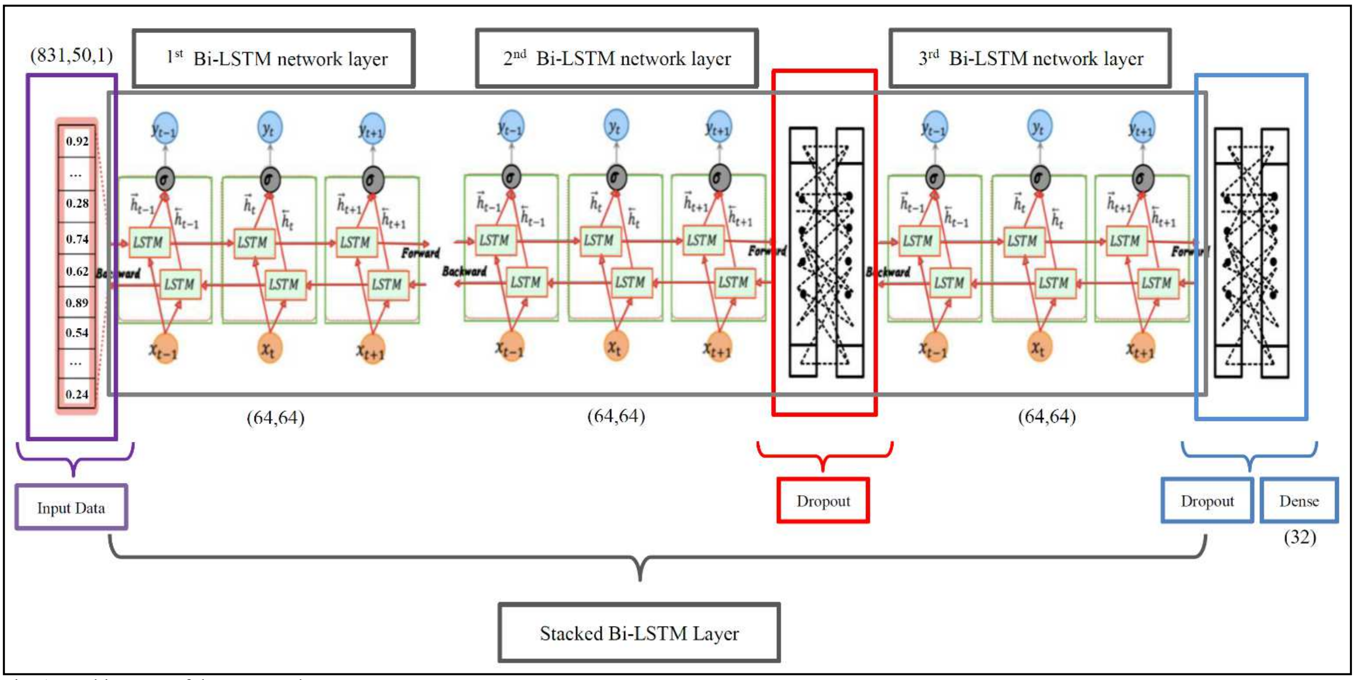

Fig. 1. Architecture of the proposed SBLSTM.

[2], Convolutional Neural Network (CNN) was proposed to forecast the price of the stock market. The proposed CNN consists of input layer, convolution layer, pooling layer, and also output layer. In order to capture the temporal information in the stock market data, researchers tend to use Recurrent Neural Network (RNN)-based neural networks. In [14], RNN was proposed to perform the stock market prediction. LSTM is a variant of RNN where it manages to capture long-term dependencies in the time series data. In [15], different network architectures of LSTM were proposed to perform prediction on stock price and achieved good performance. Later, Althelaya et. al. [16] further improved the performance of stock market prediction by stacking multiple LSTM layers. LSTM networks only capture the temporal information in unidirectional. In recent works, researchers started to examine the bidirectional computation for the time series data prediction [16]. Similarly, Jia et. al. [6] analyzed the stock market price using Bidirectional Long Short-Term Memory (BLSTM) which captures the temporal information from both forward and backward directions. In [4], a combination of CNN and BLSTM was proposed for stock market prediction. In general, deep learning-based approaches manage to achieve higher performance in stock market analysis as compared to statistical and ANN methods.

III. Method

This section describes the components and the architecture of the proposed Stacked Bidirectional LSTM (SBLSTM). The stock price prediction aims at predicting the future stock price based on the historical stock price.

Before building the prediction model, normalization is applied on the data. The data is normalized into the range of [0, 1] while preserving the shape of the original distribution of the data. The normalization ensures the data with different units and ranges to be modeled more realistically.

The proposed SBLSTM model consists of two layers of bidirectional LSTM network, a dropout layer, followed by another layer of bidirectional LSTM network, a dropout layer, a dense layer, and a prediction layer. The stacked bidirectional LSTM architecture is able to better encode the long-term temporal dependencies over the stock price data, thus gaining better performance in the stock price prediction. Fig. 1 illustrates the architecture of the proposed SBLSTM model. Based on Fig. 1, the input stock market data will be fed into the proposed SBLSTM. It then goes into a stack of three Bi-LSTM layers to learn the representation of the stock data before the network predicts the price of the stock. Dropout is adopted to reduce the overfitting during the training.

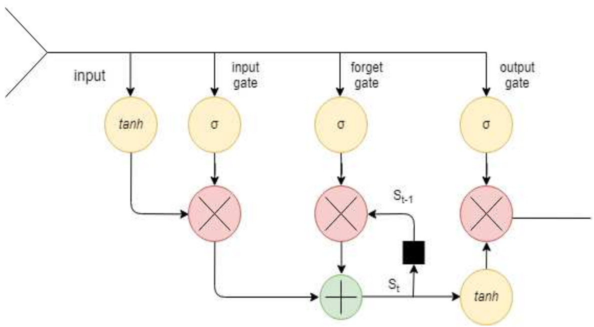

Fig. 2 The architecture of a single LSTM cell.

There are two LSTM cells in every Bi-LSTM cell, where each LSTM cell stores the information in both forward and backward directions. A LSTM cell is shown in Fig. 2. The temporal information can be retained longer with LSTM cell. The computation in a LSTM network is continuous and differentiable. Therefore, this memorization can be changed adaptively using optimization methods with the temporal information.

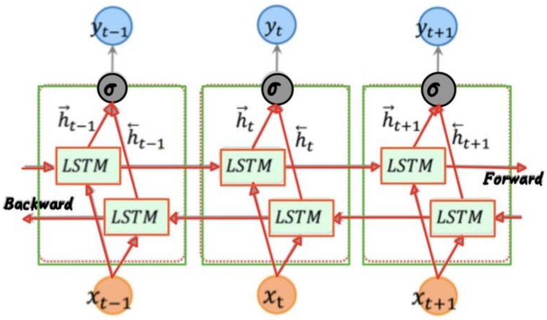

Fig. 3 Close diagram of a Bi-LSTM cell.

Fig. 3 shows the close diagram of a Bi-LSTM cell. The upper layer is the LSTM network layer in forward direction and the bottom layer is the LSTM network layer that is used to process the sequence data in backward direction. The input data flowing in a forward direction starting from time $T-n$ until $T-1$ is used to compute $\vec{h}$ iteratively in the upper layer of the LSTM network layer. On the other hand, $h^{\leftarrow}$ is computed based on the input data flows in the backward direction from time $T-n$ to $T-1$. The equation to compute the output vector, y_t, in the Bi-LSTM layer is defined as:

$$y_t = \sigma\,(\overrightarrow{h_t}\,, h_t^{\leftarrow}) \tag{1}$$

where σ is the sigmoid function used to combine the two sequences of the output. In the last bidirectional LSTM, the final output vector, which is the last element, y_{T-1}, will be used to predict the next-day stock price in the next iteration.

IV. Experimental and Results

There are six datasets used in this project to conduct the experiments which are: Alphabet Inc Class C (GOOG), Apple Inc. (APPL), Tesla Inc. (TSLA), Starbucks Corporation (SBUX), Alibaba Group Holding Ltd. (BABA), and Walmart Inc. (WMT). These datasets are obtained from Yahoo Finance. For all six datasets, we use the 5 years stock market data starting from 1st January 2015 to 1st January 2020.

In this work, we split the data into 70 % for training and 30% for testing. From the training set, we then take 30% of the training samples as the validation set. To evaluate the performance of the model, Root Mean Square Error (RMSE) is selected as the evaluation metric:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (y_i - \hat{y}_i)^2} \tag{2}$$

Table I presents the hyperparameter settings of the proposed model.

TABLE I. Hyperparameter setting of the SBLSTM model.

Hyperparameter	**Setting**
Bi-LSTM	64, 64
Bi-LSTM	64, 64
Dropout	0.1
Bi-LSTM	64, 64
Dropout	0.1
Dense	32
Dense	1

Table II shows the stock market prediction performance comparison between LSTM [17], Simple Moving Average [1], Linear Regression [18], XGBoost [19], and the proposed method SBLSTM in RMSE. The proposed SBLSTM outperforms the state-of-the-art methods in all six datasets. From the results, the proposed SBLSTM achieves 0.0006, 0.0028, 0.0020, 0.0010, 0.0022, and 0.0005 for AAPL, GOOG, TSLA, WMT, BABA, and SBUX, respectively.

TABLE II. Performance evaluation comparison of the proposed SBLSTM with state-of-the-art methods

Methods	**Root Mean Square Error (RMSE)**					
	Datasets					
	AAPL	**GOOG**	**TSLA**	**WMT**	**BABA**	**SBUX**
LSTM [17]	0.968	21.202	1.839	1.036	4.190	1.011
Simple Moving Average [1]	0.884	19.631	1.974	1.009	3.727	1.085
Linear Regression [18]	1.149	25.300	2.315	1.420	4.398	1.454
XGBoost [19]	0.803	19.316	1.782	0.950	3.251	1.002
Proposed SBLSTM	**0.0006**	**0.0028**	**0.0020**	**0.0010**	**0.0022**	**0.0005**

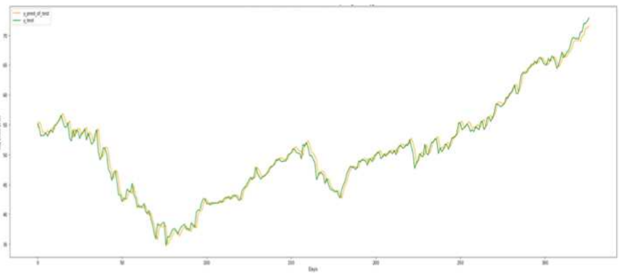

Fig. 4 Prediction of the SBLSTM on AAPL

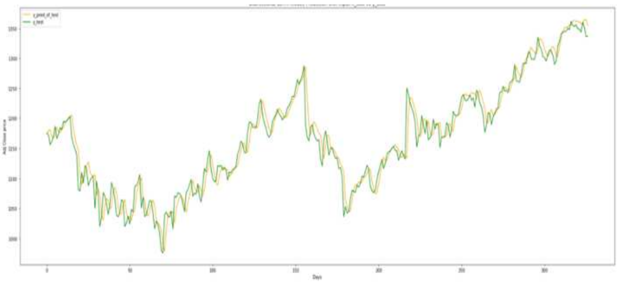

Fig. 5 Prediction of the SBLSTM on GOOG

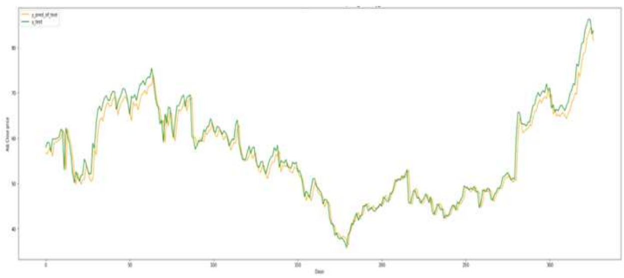

Fig. 6 Prediction of the SBLSTM on TSLA

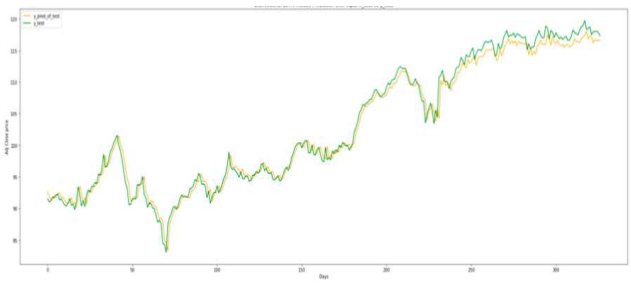

Fig. 7 Prediction of the SBLSTM on WMT

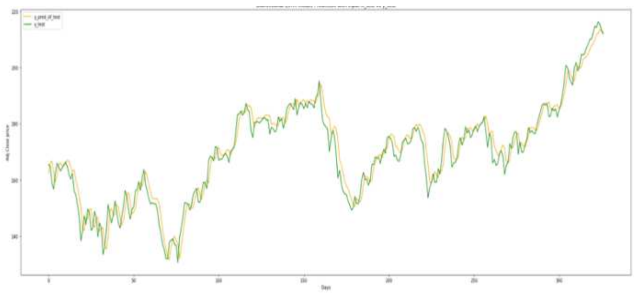

Fig. 8 Prediction of the SBLSTM on BABA

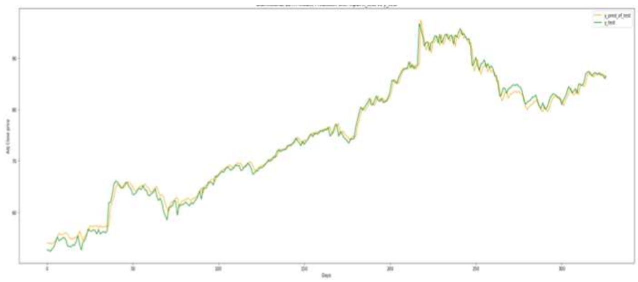

Fig. 9 Prediction of the SBLSTM on SBUX

For qualitative evaluation, the performance of the proposed SBLSTM is visualized in Fig. 4 – Fig. 9. These figures plot the testing data and the predicted stock market values using the proposed SBLSTM on the six datasets. In the figures, the green line represents the original stock price of the testing data. On the other hand, the yellow line is the predicted price from the proposed SBLSTM. Noticeably, the proposed SBLSTM predicts the stock price closely to the actual stock price in testing data.

V. Conclusion

In this paper, we propose SBLSTM for stock market price prediction. The proposed SBLSTM stacks three bidirectional LSTM networks. In contrast to LSTM-based methods, the proposed bidirectional LSTM captures the temporal information in both forward and backward directions. The long-term dependencies from the past and future stock market values are obtained. The proposed SBLSTM outperformed all five existing prediction methods for all six datasets in the experiment.

References

[1] O. Chantarakasemchit, S. Nuchitprasitchai, and Y. Nilsiam, "Forex Rates Prediction on EUR/USD with Simple Moving Average Technique and Financial Factors," In 2020 17th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), pp. 771-774, 2020.

[2] A.A. Adebiyi, A.O. Adewumi, and C.K.Ayo, "Comparison of ARIMA and artificial neural networks models for stock price prediction," Journal of Applied Mathematics, 2014.

[3] A.A. Ariyo, A.O. Adewumi, and C.K. Ayo, "Stock price prediction using the ARIMA model," UKSim-AMSS 16th International Conference on Computer Modelling and Simulation, pp. 106-112, IEEE, 2014.

[4] L.S. Chong, K.M. Lim, and C.P. Lee, "Stock Market Prediction using Ensemble of Deep Neural Networks," In 2020 IEEE 2nd International Conference on Artificial Intelligence in Engineering and Technology (IICAIET), pp. 1-5, 2020.

[5] M. Vijh, D. Chandola, V.A. Tikkiwal, and A. Kumar, "Stock closing price prediction using machine learning techniques," Procedia Computer Science, vol. 167, pp. 599-606, 2020.

[6] M. Qiu, and Y. Song, "Predicting the direction of stock market index movement using an optimized artificial neural network model," PloS one, vol. 11(5), pp. e0155133, 2016.

[7] C.F. Tsai, and S.P. Wang, "Stock price forecasting by hybrid machine learning techniques," In Proceedings of the international multiconference of engineers and computer scientists, vol. 1, no. 755, pp. 60, 2009.

[8] Y.S. Tan, K.M. Lim, and C.P. Lee, "Hand gesture recognition via enhanced densely connected convolutional neural network," Expert Systems with Applications, vol. 175, pp. 114797, 2021.

[9] P.S. Tan, K.M. Lim, and C.P. Lee, "Human Action Recognition with Sparse Autoencoder and Histogram of Oriented Gradients," In 2020 IEEE 2nd International Conference on Artificial Intelligence in Engineering and Technology (IICAIET), pp. 1-5, 2020.

[10] Y.S. Tan, K.M. Lim, C. Tee, C.P. Lee, and C.Y. Low, "Convolutional neural network with spatial pyramid pooling for hand gesture recognition," Neural Computing and Applications, pp.1-13, 2020.

[11] K.M. Lim, A.W. Tan, and S.C. Tan, "A four dukkha state-space model for hand tracking", Neurocomputing, vol. 267, pp. 311-319, 2017.

[12] P.S. Tan, K.M. Lim, C.P. Lee, and C.H. Tan, "Acoustic Event Detection with MobileNet and 1D-Convolutional Neural Network," In 2020 IEEE 2nd International Conference on Artificial Intelligence in Engineering and Technology (IICAIET), pp. 1-6, 2020.

[13] X. Xia, R. Togneri, F. Sohel, Y. Zhao, and D. Huang, "A survey: neural network-based deep learning for acoustic event detection," Circuits, Systems, and Signal Processing, vol. 38(8), pp. 3433-3453, 2019.

[14] G. Petneházi, "Recurrent neural networks for time series forecasting," arXiv preprint arXiv:1901.00069, 2019.

[15] M. Jia, J. Huang, L. Pang, and Q. Zhao, "Analysis and research on stock price of LSTM and bidirectional LSTM neural network," In 3rd International Conference on Computer Engineering, Information Science & Application Technology (ICCIA 2019), pp. 467-473, 2019.

[16] K.A. Althelaya, E.S.M. El-Alfy, and S. Mohammed, "Stock market forecast using multivariate analysis with bidirectional and stacked (LSTM, GRU)," In 2018 21st Saudi Computer Society National Computer Conference (NCC), pp. 1-7, 2018.

[17] H. Jia, "Investigation into the effectiveness of long short term memory networks for stock price prediction," arXiv preprint arXiv:1603.07893, 2016.

[18] S.R. Bommareddy, K.S.S. Reddy, P. Kaushik, K.V. Vinay Kumar, and V.R. Hulipalled, "Predicting the stock price using linear regression," Int J Adv Res Comput Sci, vol. 9(3), pp. 81-85, 2018.

[19] S. Dey, Y. Kumar, S. Saha, and S. Basak, "Forecasting to classification: Predicting the direction of stock market price using xtreme gradient boosting," PESIT South Campus, 2016.