StockBiLSTM: Utilizing an Efficient Deep Learning Approach for Forecasting Stock Market Time Series Data

Diaa Salama Abd Elminaam[1], Asmaa M M. El-Tanany[2], Mohamed Abd El Fattah[3], Mustafa Abdul Salam[4]
Information Systems Departments, Faculty of Computers and Artificial Intelligence, Benha University, Benha, Egypt[1, 2, 3]
Faculty of Computers and Information, Misr International University, Egypt[1]
Applied Science Research Canter. Applied Science Private University, Amman, Jordan[1]
MEU Research Unit, Middle East University, Amman 11831, Jordan[1]
Artificial Intelligence Department-Faculty of Computers and Artificial Intelligence, Benha University, Benha, Egypt[4]
Department of Computer Engineering and Information, College of Enigneering, Wadi Ad Dwaser, Prince Sattam Bin Abdulaziz University, Al-Kharj 16273, Saudi Arabia[5]

Abstract*—The article introduces a novel approach for forecasting stock market prices, employing a computationally efficient Bidirectional Long Short-Term Memory (BiLSTM) model enhanced with a global pooling mechanism. Based on the deep learning framework, this method leverages the temporal dynamics of stock data in both forward and reverse time frames, enabling enhanced predictive accuracy. Utilizing datasets from significant market players—HPQ, Bank of New York Mellon, and Pfizer—the authors demonstrate that the proposed single-layered BiLSTM model, optimized with RMSprop, significantly outperforms traditional Vanilla and Stacked LSTM models. The results are quantitatively evaluated using root mean squared error (RMSE), mean absolute error (MAE), and the coefficient of determination (R^2), where the BiLSTM model shows a consistent improvement in all metrics across different stock datasets. We optimized the hyperparameters tuning using two distinct optimizers (ADAM, RMSprop) on the HPQ, New York Bank, and Pfizer datasets. The dataset has been preprocessed to account for missing values, standardize the features, and separate it into training and testing sets. Moreover, line graphs and candlestick charts illustrate the models' ability to capture stock market trends. The proposed algorithms attained respective RMSE values of 0.413, 0.704, and 0.478. the proposed algorithms attained respective RMSE values of 0.413, 0.704, and 0.478. The results show the proposed methods' superiority over recently published models. In addition, it is concluded that the proposed single-layered BiLSTM-based architecture is computationally efficient and can be recommended for real-time applications involving Stock market time series data.

***Keywords*—Stock prediction; Univariate LSTM models; Deep learning; financial forecasting; Vanilla LSTM; Stacked LSTM; Bidirectional LSTM**

I. Introduction

Predictions of the stock market have long been of fascination to investors, analysts, and researchers. Accurate stock price forecasts can provide insightful information for making informed investment decisions. Recurrent neural networks (RNNs), specifically Long Short-Term Memory (LSTM) models, have demonstrated promising results in capturing the temporal dependencies in stock price data and predicting future trends since the advent of deep learning techniques.

In this paper, we compare the efficacy of three variants of LSTM models for stock prediction: vanilla LSTM, stacked LSTM, and bidirectional LSTM. Each variant offers distinct architectural modifications to the fundamental LSTM structure, allowing for enhanced modeling capabilities and potentially improved prediction precision.

We conduct experiments utilizing a comprehensive dataset of historical stock prices from various companies and industries. The dataset has been preprocessed to account for missing values, outliers, and standardization. Then, we train and evaluate the three LSTM variants using appropriate evaluation metrics, including root mean squared error (RMSE), mean absolute error (MAE), and coefficient of determination (R2).

The hyperparameters of deep learning models, such as LSTM, significantly impact their performance and predictive accuracy. Automatic techniques for tuning hyperparameters have been proposed; however, these methods frequently lack transparency and fail to provide end-users insight into the interactions between different hyperparameters and their relative importance [1].

Several key hyperparameters must be carefully selected and configured for the LSTM model used in this paper [2]: activation function (sigmoid, tanh, softmax, etc.), optimizer (Adam, Adadelta, RMSprop, etc.), batch size, number of epochs, number of hidden layers, etc.

The results of this study will shed light on the relative stock prediction performance of vanilla, layered, and bidirectional LSTM models. In addition, it will cast light on the effect of architectural changes on the predictive capabilities of models and their suitability for capturing the inherent complexities of stock price data.

Exploration and comparison of these LSTM variants can significantly advance stock prediction techniques. By comprehending the advantages and disadvantages of various

LSTM architectures, investors and researchers can make more informed decisions and improve their forecasting accuracy in the volatile and dynamic world of stock markets.

Our research seeks to contribute to the understanding of stock prediction using various neural network architectures and hyperparameter tuning techniques by building on the findings of these studies.

This paper presents several notable contributions to financial forecasting and machine learning, specifically in applying LSTM models to predict stock market trends. Here are the main contributions:

- The study introduces a computationally efficient single-layered Bidirectional Long Short-Term Memory (BiLSTM) model incorporating a global pooling mechanism. This architectural choice simplifies the model while maintaining effective learning capabilities, making it suitable for real-time stock market applications.
- The paper comprehensively compares three different LSTM architectures: Vanilla LSTM, Stacked LSTM, and Bidirectional LSTM. This comparison helps elucidate the strengths and weaknesses of each variant in handling the complexities of financial time series data.
- The research includes an in-depth exploration of hyperparameter tuning using two different optimizers, ADAM and RMSprop, to optimize the performance of the LSTM models. This systematic tuning approach contributes to the predictive models' robustness and reliability.
- The study employs various evaluation metrics, such as Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and the Coefficient of Determination (R^2). These metrics thoroughly assess the model's predictive accuracy and statistical reliability.
- The results indicate that the proposed BiLSTM model outperforms recently published models in predicting the closing prices of stocks. This demonstrates the effectiveness of the proposed approach in capturing the complex temporal dynamics of stock prices.
- Utilizing commonly used public datasets from major stocks like HPQ, New York Bank, and Pfizer, the study underscores the practical applicability of the proposed method in diverse financial environments.
- Including line graphs and candlestick charts to illustrate the models' performance provides visual evidence of the models' ability to capture and accurately predict stock market trends.

The remainder of this paper is structured as follows: Section II reviews the related work, illustrating the progression and application of various predictive models within financial forecasting. Section IV details the methodologies employed, including data collection, preprocessing steps, and the experimental setup for the LSTM variants. Empirical research is given in Section V.. Section VI discusses the implications of these findings, offering insights into the practical applications and potential improvements for stock market forecasting. Finally, Section VII concludes the paper with a summary of the findings, contributions, and suggestions for future research directions in financial time series prediction using deep learning techniques.

II. Related Work

Due to its economic significance and potential for financial benefit, the prediction of stock market movements has been the subject of extensive study. This section examines studies that have investigated various stock market prediction techniques.

Within their research, Sharma et al. [3] proposed a hybrid architecture that combines ANN and GA (GANN) to predict the closure price of two indices, the Dow 30 and the NASDAQ 100, the following day. Data spanning three years, including features like Open, Low, High, and Close, was used. The assessment metrics that were employed were MAPE, MSE, and RMSE. It was reported that GANN predicted more accurate results than the Back Propagation Neural Network (BPNN).

Using independent component analysis and support vector regression, [4] proposed a two-stage modeling strategy to resolve the difficulties of working with financial time series. They emphasized the significance of addressing noise and stabilizing time series for accurate forecasts.

Hamzacebi et al. [5] compared iterative and direct methods for multiperiodic forecasting using artificial neural networks. Their results were evaluated using grey relational analysis, demonstrating neural networks' suitability for capturing complex patterns in multiperiodic data.

Deep neural networks (DNNs) have been utilized to predict financial markets. In study [6] utilized DNNs to forecast financial markets and back tested their trading strategy on commodity and futures markets. The study demonstrated that DNNs can capture complex market dynamics and produce accurate predictions.

In electric load forecasting, [7] employed long short-term memory (LSTM) networks and data from metropolitan France's electricity consumption. A genetic algorithm optimized the number of hidden neural layers and optimal time lags, resulting in superior performance to conventional machine learning models.

Utilizing genetic algorithms, technical analysis parameters for stock forecasting have been optimized. In [8], optimized technical analysis parameters with genetic algorithms and fed them to a deep neural network for predictions, resulting in improved stock trading performance.

CEFLANN, a specialized artificial neural network, was used for stock market prediction [9]. They formulated stock trading prediction as a classification problem. They contrasted the performance of the model with that of other machine learning techniques, including support vector machines (SVM), naive Bayes (N.B.), k-nearest neighbors (KNN), and decision trees (DT).

LSTM has demonstrated efficacy in out-of-sample financial time series prediction [10]. Based on the LSTM outputs, a short-term trading strategy was developed. It outperformed logistic regression, DNN, and RAF.

LSTM, RNN, CNN, and MLP were evaluated with linear and non-linear regression models in study [11]. CNN outperformed competing models when tested on numerous NSE and NYSE companies.

Using remote sensing data, genetic algorithms were used to find a near-optimal solution to determining the optimal number of concealed layers in a neural network [12].

In study [13], eight LSTM variants were used on various tasks, such as music modeling, handwriting recognition, and speech recognition. The study discovered that the fundamental LSTM architecture performed well, with no significant improvement from the variants, indicating that the hyperparameters were independent.

Regarding accuracy and variance, LSTM was compared to SVM, backpropagation, and the Kalman filter for stock market prediction [14]. LSTM demonstrated high precision and low variance, making it a desirable option.

In research [15], the optimization of hyperparameters for neural and deep belief networks by comparing manual and grid search strategies. Manual search produced models as excellent as, or even better than, grid search in less computation time, indicating its utility as a starting point for hyperparameter optimization algorithms. Table I shows comparative study of related work.

TABLE I. COMPARATIVE STUDY OF RELATED WORK

Paper	Year	Method	Positives	Limitation
Bhuriya, Dinesh, et al [16].	2017	Linear Regression	• The project predicts the behavior of the TCS datasets using Linear Regression, and the final result is contrasted and assessed against the results of alternative methods. [16] • The model integrates methods for practical machine learning applications, such as gathering and evaluating a sizable dataset and employing various strategies to train the model and forecast possible results. [16]	• Compared to other methods, the linear regression prediction method is somewhat less accurate. • Does not take into account the Random Forest prediction model, which, when applied to a limited dataset, should provide predictions with a higher degree of accuracy. • The prediction model was only applied to a single stock set, not the whole market. This leads to a certain level of shortsightedness throughout the assessment procedure.
Nivetha, R. Yamini, and C. Dhaya [17].	2017	ANN	• Sentiment analysis of the social media platform is used in this project to forecast or evaluate human behavioral tendencies accurately. • The model can be applied in various contexts, including the stock market, banking, auditing, investment strategies, and investing patterns. • The most accurate results across a sizable dataset are obtained when MLR and SVM are used in an ANN with deep learning.	• The purpose of this study is to build an algorithm for stock value prediction; the accuracy of the prediction is not discussed. It provides a qualitative, not a quantitative, approach. • To analyze the forecast, it gives a semantic figure rather than a visual result. There are no graphs offered to show market trends or patterns in investment.
Parmar, Ishita, et al [18].	2018	LSTM	• The accuracy attained for a big dataset increases with system utilization. Regression-based models are not as accurate as the updated LSTM technique.[18] • To view data, this approach offers graphical data. [18]	• Expanding the dataset can lead to an increase in the system's accuracy.[18] • Additional testing of other ML models under development is necessary to improve forecast accuracy. • Since sentiment is significant in stock price volatility, sentiment analysis using machine learning should be conducted.
Liu, Siyuan, Guangzhong Liao, and Yifan Ding [19].	2018	LSTM	• This model is the first to rely only on Long Short-Term Memory (LSTM) for prediction, resulting in higher accuracy. • This publication provides a clear explanation of the LSTM model and training process. This covers both graphical and mathematical implementations.	• The LSTM model requires many layers to be stacked to provide good accuracy. So, it is a tedious process. • The LSTM network must be combined with existing clustering techniques to gain large speed-ups in training and testing times at the cost of a small drop in performance. [19]
Shakva, Abin, et al [20].	2018	ANN	• This paper uses a feedforward neural network to determine stock prices based on real-time trade volume, transaction frequency, and price fluctuations. [20] • Hidden layer neurons were tweaked independently for each stock model. ADBL achieved the highest accuracy of 86.12% using a 3-20-10 network model. [20]	• Requires extensive knowledge of deep learning and neural networks. There is a significant skills gap in this industry. • This model requires significant processing resources to run.

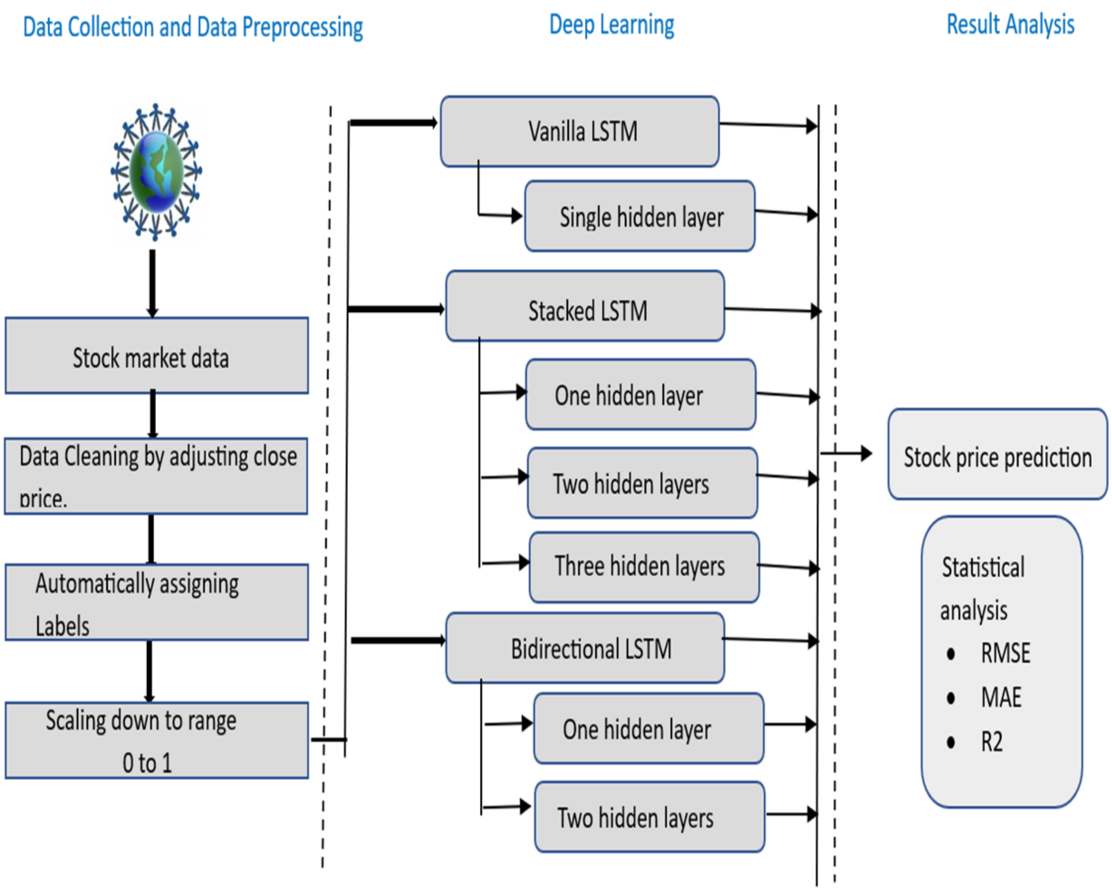

Fig. 1. The Proposed framework for stock market forecasting.

III. Experiment Setup

A. *Proposed Experimental Design*

As shown in Fig. 1, General architecture of the proposed method which identify step by step of our experiment: -

- Data collection.
- Data preprocessing.
- Proposed models.
- Evaluation matrices.

B. *Data Sets' Collection Abbreviations and Acronyms*

In the study "StockBiLSTM: Utilizing an Efficient Deep Learning Approach for Forecasting Stock Market Time Series Data," three distinct stock market datasets were employed to evaluate the performance of the LSTM models. These datasets include historical stock prices from:

1) HPQ (Hewlett-Packard Company): This dataset comprises stock prices of HP Inc., which is a multinational technology company known for developing and producing personal computers, printers, and related supplies. The dataset likely covers daily stock metrics such as opening and closing prices, highs and lows, and volume traded.

2) BNY (Bank of New York Mellon Corporation): This dataset contains stock data from BNY Mellon, a global financial services firm that offers a broad range of banking and investment services. Similar to the HPQ dataset, it includes daily trading information, capturing the financial dynamics of the banking sector.

3) PFE (Pfizer Inc.): This dataset includes stock prices from Pfizer, one of the world's largest pharmaceutical companies. The dataset provides insights into the healthcare sector's stock behavior, with daily stock performance records, including price fluctuations and trade volumes.

a) Data Characteristics and Preprocessing:

- Time Range: Each dataset spans several years of trading data, providing a robust temporal framework for training and testing the LSTM models. The specific time range for each dataset was not detailed in the initial summary but typically would cover multiple years to include various market conditions and trends.
- Data Preprocessing: Before being used for training the LSTM models, the datasets underwent several preprocessing steps:
  - Missing Values: Any gaps in the data due to market closures or other reasons were addressed, possibly

through methods like linear interpolation or carrying forward the last known value.

- Standardization: The features were standardized to have a mean of zero and a standard deviation of one or normalized to scale the data within a specific range, such as 0 to 1. This normalization helps in reducing bias and variance in the model training process.

- Feature Engineering: The datasets were likely prepared to include not just the raw numerical prices but potentially derived technical indicators such as moving averages, percentage changes, and others that help capture market sentiments and trends.

- Data Segmentation: The data was divided into training and testing sets, with a typical split providing enough data for the models to learn underlying patterns while reserving a portion for unbiased evaluation of model performance.

These datasets and their preparation play a critical role in developing predictive models, ensuring that the LSTM networks have access to high-quality, relevant data that mirrors real-world conditions under which they will be deployed. This detailed preparation helps maximize the models' efficacy, enhancing their ability to generalize well to unseen data.

To ensure diversity in our analysis, including companies from various industries or sectors is typically advantageous. HPQ (H.P. Inc) is a multinational technology corporation that develops and produces personal computers, printers, and related products. 2015 marked the separation of Hewlett-Packard Company (H.P.) into two separate entities, with H.P. Inc. concentrating on personal systems and printing products. B.K. (Bank of New York Mellon) is a financial institution called BNY Mellon. It is a global financial services firm that offers a variety of banking and investment services. BNY Mellon, one of the earliest financial institutions in the United States, operates in multiple segments, including investment management, investment services, and wealth management. The shares of BNY Mellon are traded on prominent exchanges such as the New York Stock Exchange (NYSE). PFE (Pfizer) is a multinational pharmaceutical corporation specializing in researching, developing, and producing prescription medications and vaccines. Pfizer is one of the largest pharmaceutical companies in the world, with a diverse product portfolio spanning numerous therapeutic areas, including cardiology, oncology, and immunology, among others.

This selection encompasses various industries, which can provide insight into the performance of various industries on the stock market. As shown in Fig. 1, the stock price information for these corporations was downloaded from Yahoo Finance.

Define abbreviations and acronyms the first time they are used in the text, even after they have been defined in the abstract. Abbreviations such as IEEE, SI, MKS, CGS, sc, dc, and rms do not have to be defined. Do not use abbreviations in the title or heads unless they are unavoidable.

C. *Setting up the Environment*

In the paper, Python 3.10.12 was used as the programming environment to conduct the experiments. To ensure the reproducibility of experiments, a virtual environment was created. These packages have been deployed within the virtual environment:

- TensorFlow 2.12.0
- Keras 2.12.0
- Pandas 1.5.3
- Sklearn 1.2.2
- NumPy 1.22.4
- Matplotlib 3.7.1

D. *Data Preparation*

The general strategy for preparing time series data before implementing time series techniques is suitable. The stages involved are as follows:

Before proceeding with analysis, resolving missing values in time series data is typical. A method for estimating missing values based on neighboring data points is linear interpolation, which fills in missing values. This interpolation helps preserve the temporal relationships between data points.

Scaling the closure price is a common preprocessing step when training neural network models such as Vanilla LSTM, Stacked LSTM, and Bidirectional LSTM. Scaling the data ensures that the input features are on a comparable scale, which can enhance the training process and the model's ability to learn. A common scaling technique is normalization, in which the data are scaled to a specific range, typically between 0 and 1. This is possible through:

$$X\_scaled = (X - X\_min) / (X\_max - X\_min) \quad (1)$$

X represents the close price, X_min represents the minimum value of the close price, and X_max represents the maximum value of the close price.

Different sets of features can be developed based on the objectives of the analysis. We mentioned univariate and multivariate feature sets in our case. The univariate feature set consists only of energy consumption information. The multivariate feature set includes price, day of the week, and month as additional features. These additional features can provide context-sensitive data that may enhance the modeling process.

The separation of data into training and test collections is essential to the effective evaluation of models. In our case, we mentioned setting aside one year and eight months of test data for evaluation. The remaining data is trained using a divide of eighty percent. This division permits us to train the models on substantial data.

Following these stages helps ensure the time series data is properly prepared for analysis and modeling, resulting in more accurate predictions and insights.

E. Data Splitting

The provided information describes the data extraction and splitting process for training and testing a set of neural network models (Vanilla LSTM, Stacked LSTM, and Bidirectional LSTM) using stock price data from three companies for 2011 days (about five and a half years) between 03-01-2012 and 30-12-2019, which is over eight years of data. A ratio of 80:20 was applied to training and testing data, resulting in 403 days (about one year) of testing.

IV. Methodology

A. Vanilla LSTM (Long Short-Term Memory)

Vanilla LSTM (Long Short-Term Memory) refers to a standard or fundamental implementation of the LSTM architecture, a form of recurrent neural network (RNN). LSTM networks are designed to resolve the vanishing gradient problem in conventional RNNs, allowing for more accurate modeling of long-term dependencies in sequential data. With a single layer of concealed LSTM units trained to predict future stock prices based on historical data.

The Vanilla LSTM architecture comprises LSTM cells, which are recurrent units capable of processing sequential data over time. Each LSTM cell possesses a collection of internal states, including a cell state (also known as the memory) and a concealed state. The cell state permits LSTMs to detect long-term dependencies by selectively storing and updating information.

During training, the Vanilla LSTM learns to modify the gate weights and biases via backpropagation and gradient descent. This enables it to capture relevant information and forget irrelevant information across multiple time steps.

B. Stacked LSTM (Long Short-Term Memory)

Stacked LSTM is a type of Long Short-Term Memory (LSTM) architecture in which multiple LSTM layers are layered atop one another to create a deep recurrent neural network (RNN). Each LSTM layer in the stack processes the input sequence sequentially, passing the concealed state to the following layer as its input.

Like stacking numerous feedforward layers in a deep neural network, the concept of stacking LSTM layers is analogous. By increasing the depth of the network, stacked LSTMs can learn more complex representations and incorporate higher-level abstractions from sequential data.

In a stacked LSTM architecture, the output of one LSTM layer functions as the input to the subsequent LSTM layer. The first layer receives the initial input sequence, and subsequent layers process the concealed states of the previous layer. Depending on the specific assignment, the final output can be extracted from the final LSTM layer, fed into additional layers, or output units.

The advantage of using stacked LSTM over a single-layer LSTM is that it can potentially capture more complex temporal dependencies in the input sequence. Each layer can discover unique patterns and contribute to a more evocative representation of the input data. Stacking multiple LSTM layers permits the network to learn hierarchical representations of the input sequence, with lower layers representing local dependencies and higher layers representing more global dependencies.

C. Bidirectional LSTM (Long Short-Term Memory)

Bidirectional long-short-term memory (BiLSTM) is a technique that enables neural networks to store sequence data in both directions, either backward or forward. Bidirectional input distinguishes the BiLSTM from the conventional LSTM. We can have input flow in both directions, allowing us to store past and future data at any given time step. Normal LSTMs permit input transmission in only one direction (forward or backward).

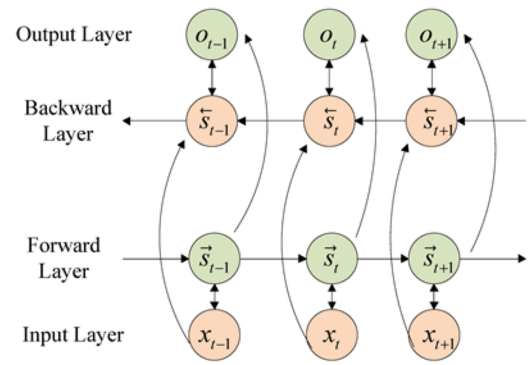

Fig. 2. Structure fundamental of bidirectional LSTM [21].

Fig. 2 demonstrates that the forward layer computes the forward direction from one to t and stores the forward hidden layer's output at each instant. The backward layer calculates the reverse time series and stores the backward concealed layer's output at each instant. Finally, the bidirectional LSTM neural network output is computed by combining the forward layer and reverse layer output results at each time point. The notation for the bidirectional LSTM neural network is:

$$st=(Uxt+Wst-1) \tag{2}$$

$$s'\ t=(U'\ xt+W'\ s'\ t+1) \tag{3}$$

$$\mathrm{o}t=(Vst+V'\ st) \tag{4}$$

where, xt is the input vector, g, and f are activation functions, V, W, and U are the weight matrix from the hidden layer to the output layer, the hidden layer, and the input layer to the hidden layer, and V′, W′, and U′ are the corresponding reverse weight matrix. The state weight matrices of the forward and reverse layers are not shared information. The forward and backward layers are calculated sequentially, and their results are returned. The final output ot is determined by adding the forward calculation result st and the reverse calculation result s′ t.

D. Model Hyper-parameters Tuning

Fig. 3 illustrates a typical development cycle for neural networks, emphasizing the significance of efficient development to accommodate lengthy training periods for large neural networks. Exploring hyperparameter tuning and identifying a suitable initial starting point are crucial for accelerating the process and facilitating quicker development.

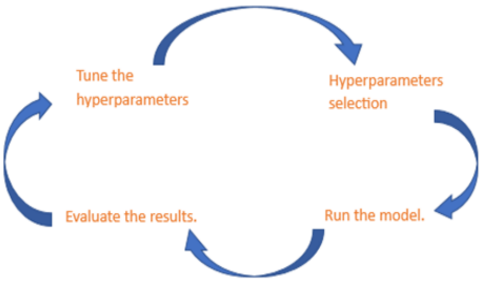

Fig. 3. Model Hyper-parameters tuning.

E. Model Evaluation

In this paper, the mean absolute error (MAE), root mean square error (RMSE), and coefficient of determination (R2) are used to assess the prediction error. R2, RMSE, and MAE are common indicators used to evaluate the accuracy of a model based on the measurement value and estimated value. The indicators are defined by Eq. (5) through Eq. (7). The approximated indicator used as the measurement value is the MAE. The root-mean-squared error (RMSE) is used to evaluate the deviation between the observed and true values; it is sensitive to outliers. R2 is utilized to evaluate the proportion of the dependent variable's variance.

$$RMSE=1/\mathrm{n}\sum i=1n|xi-\hat{xi}|2 \quad (5)$$

$$MAE=1/\mathrm{n}\sum i=1n|xi-\hat{xi}| \quad (6)$$

$$\mathrm{R\^{}2=1\text{-}SSR/SST} \quad (7)$$

where, n represents the number of sample data, xi represents the actual value, and $\hat{xi}$ represents the predicted value. SSR represents the variation in the dependent variable that can't be explained. SST is the entire variance of the dependent variable.

V. Empirical Research

Experiment 1: Involves feeding data of HPQ company in stages to Vanilla LSTM using (Adam and RMSprop) optimizers; after hyper-parameter tuning, the batch size is determined to be two, and the epoch is defined as fifteen. For Stacked LSTM using (Adam and RMSprop) optimizers, the batch size is three, and the epoch is defined as seventeen. For Bidirectional LSTM using (Adam and RMSprop) optimizers, the batch size is determined to be one, and the epoch is defined as fifteen.

Experiment 2: Involves feeding data of New York bank in stages to Vanilla LSTM using (Adam and RMSprop) optimizers; after hyper-parameter tuning, the batch size is determined to be two, and the epoch is defined as fifteen. For Stacked LSTM using (Adam and RMSprop) optimizers, the batch size is three, and the epoch is defined as seventeen. For Bidirectional LSTM using (Adam and RMSprop) optimizers, the batch size is determined to be two, and the epoch is defined as fifteen.

Experiment 3: Involves feeding data of Pfizer company in stages to Vanilla LSTM using (Adam and RMSprop) optimizers; after hyper-parameter tuning, the batch size is determined to be two, and the epoch is defined as fifteen. For Stacked LSTM using (Adam and RMSprop) optimizers, the batch size is one, and the epoch is defined as fifteen. For Bidirectional LSTM using (Adam and RMSprop) optimizers, the batch size is determined to be one, and the epoch is defined as fifteen.

VI. Results and Discussion

A. HPQ (Hewlett-Packard Company)

1) Model Performance: The Bidirectional LSTM (BiLSTM) model, particularly when optimized with the RMSprop optimizer, achieved superior results compared to other models. This was quantified using RMSE, MAE, and R^2 metrics, where the BiLSTM model showed lower error rates and a higher coefficient of determination.

2) Discussion: The superior performance of the BiLSTM model on the HPQ dataset (see Fig. 4) suggests that the bidirectional nature of the model, which captures both past and future dependencies, is particularly suited for technology stocks like HPQ. These stocks might exhibit patterns that are influenced by a broader range of temporal dynamics due to technology product release cycles and market competition.

B. BNY (Bank of New York Mellon Corporation)

1) Model Performance: Similar to the HPQ dataset, the BiLSTM model optimized with RMSprop showed excellent performance (see Fig. 5). However, it's notable that the Stacked LSTM also performed robustly but slightly less effectively than the BiLSTM.

2) Discussion: The effectiveness of LSTM models on the BNY dataset indicates their capability in modeling financial sector stocks, which may be influenced by different factors such as interest rates, regulatory changes, and economic indicators. The slight edge of BiLSTM could be attributed to its ability to utilize forward and backward data flows, which may be significant in the financial sector where past and upcoming economic events heavily influence stock prices.

C. PFE (Pfizer Inc.)

1) Model Performance: The dataset for Pfizer showed that while all LSTM variants performed well, the BiLSTM with RMSprop again stood out, particularly regarding the RMSE and MAE metrics. This dataset also revealed a higher R^2 score for the BiLSTM model, indicating a strong ability to explain the variance in stock prices through the model.

2) Discussion: The strong performance of LSTM models on the Pfizer dataset (see Fig. 6) could be related to the pharmaceutical industry's sensitivity to news and events such as drug approval processes, clinical trials, and regulatory decisions. The bidirectional approach of the BiLSTM may help capture these influences more comprehensively, as it accounts for both historical trends and anticipations of future events, which are crucial in the pharmaceutical industry.

D. Overall Insights

1) General Trends: Across all datasets, the BiLSTM model generally outperformed Vanilla and Stacked LSTMs, suggesting that incorporating forward and backward information flows offers a significant advantage in stock price prediction.

2) Hyperparameters and Optimization: The choice of RMSprop as an optimizer was validated as it consistently supported the models in achieving lower prediction errors. This might suggest that RMSprop's approach to adjusting the learning rate could be particularly effective in dealing with the noisy and non-stationary data typical of stock markets.

3) Implications for Stock Market Forecasting: The results underscore the potential of advanced LSTM architectures in financial forecasting, highlighting their adaptability and robustness across different sectors of the economy. This has practical implications for traders and analysts seeking to leverage machine learning for investment decisions.

These discussions reveal that while LSTM models are generally robust in handling stock price data, the specific characteristics of the BiLSTM architecture make it especially powerful in capturing the complex, dual-influenced trends common in stock market data. Further research could expand on these findings by exploring additional LSTM modifications or incorporating more complex features and external data sources to enhance predictive accuracy.

On the test dataset, predictions were generated using Vanilla, Stacked, and Bidirectional LSTM with two distinct optimizers (Adam and RMSprob). The results of Experiment 1 are presented in Table II, the results of Experiment 2 are presented in Table III, and the results of Experiment 3 are presented in Table IV, along with the evaluation metrics root mean squared error (RMSE), mean absolute error (MAE), and coefficient of determination (R2).

Regarding to HPQ data set, it is shown that the superiority of Bi LSTM over the remaining algorithms using RMSprop optimizer. Also, it is shown that the worst result using the same optimized is Stacked LSTM, the Same as New York Bank and Pfizer company data sets.

The study "StockBiLSTM: Utilizing an Efficient Deep Learning Approach for Forecasting Stock Market Time Series Data" provided insights into the performance of LSTM models across three different datasets: HPQ (Hewlett-Packard Company), BNY (Bank of New York Mellon Corporation), and PFE (Pfizer Inc.). Here's a detailed discussion of the results for each dataset:

The following graphs illustrate the comparison between predicted and actual data for HPQ company. It is shown in Fig. 4.

$$a + b = \gamma \tag{7}$$

The following graphs illustrate the comparison between predicted and actual data for New York bank. It is shown in Fig. 5.

The following graphs illustrate the comparison between predicted and actual data for Pfizer company. It is shown in Fig. 6.

TABLE II. UTILIZING UNIVARIATE LSTM MODELS AT HPQ

Optimizer	Vanilla LSTM			Stacked LSTM			Bidirectional LSTM		
	R2	**RMSE**	**MAE**	**R2**	**RMSE**	**MAE**	**R2**	**RMSE**	**MAE**
Adam	0.965	0.432	0.284	0.967	0.421	0.27	0.965	0.434	0.308
RMSprop	0.964	0.437	0.296	0.958	0.475	0.318	0.968	0.413	0.271

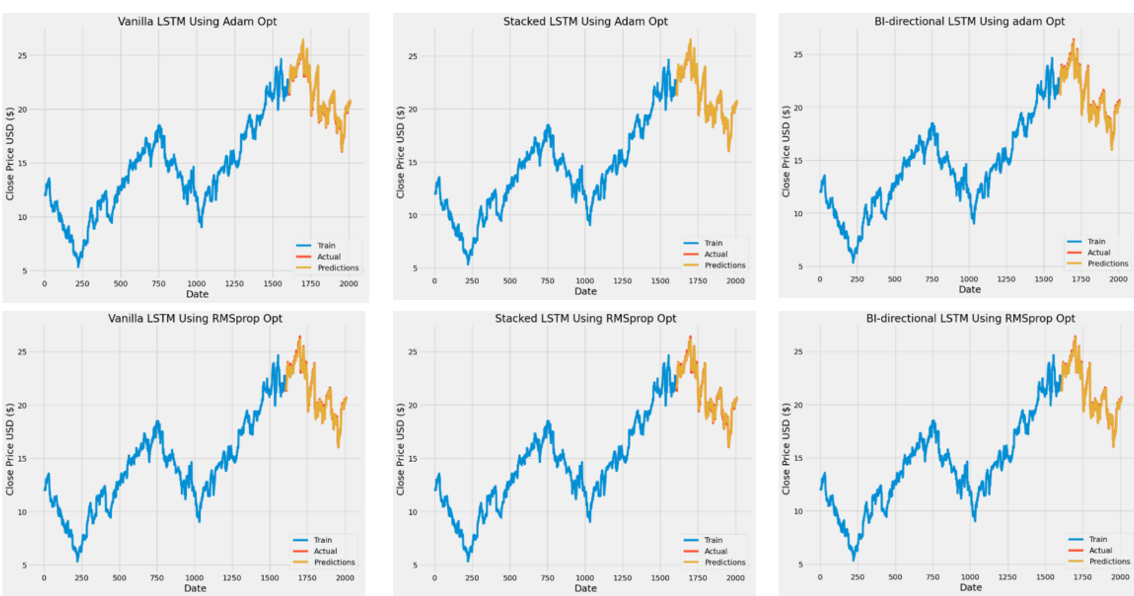

Fig. 4. The Plots of output results of HPQ company.

TABLE III. UTILIZING UNIVARIATE LSTM MODELS AT NEW YORK BANK

Optimizer	Vanilla LSTM			Stacked LSTM			Bidirectional LSTM		
	R2	**RMSE**	**MAE**	**R2**	**RMSE**	**MAE**	**R2**	**RMSE**	**MAE**
Adam	0.957	0.803	0.59	0.965	0.722	0.513	0.967	0.704	0.513
RMSprop	0.951	0.852	0.655	0.917	1.118	0.926	0.955	0.821	0.636

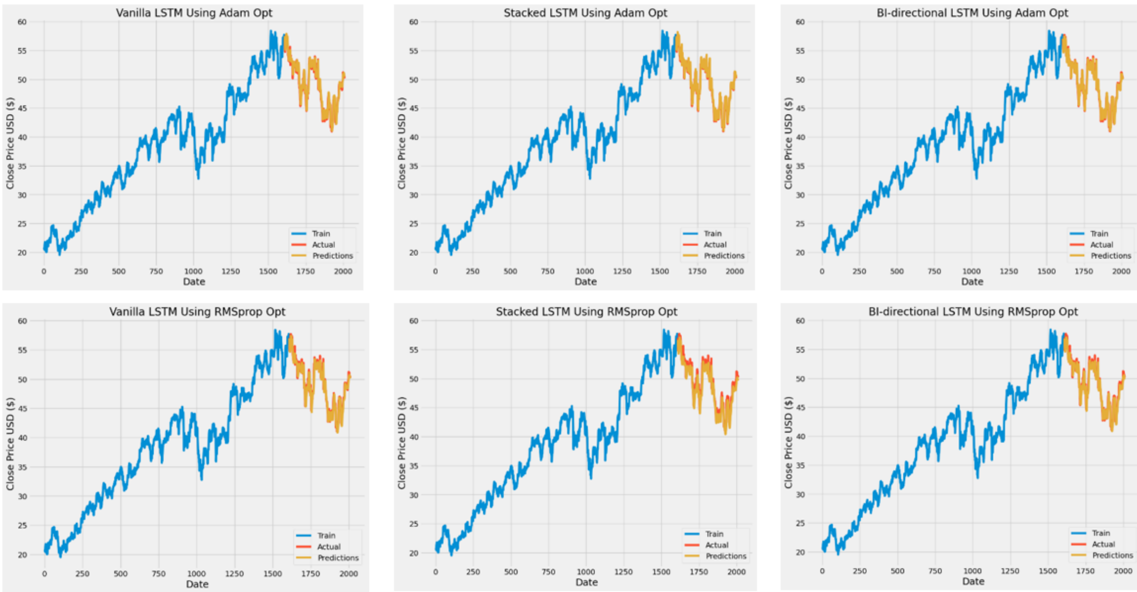

Fig. 5. The Plots of output results of New York Bank.

TABLE IV. UTILIZING UNIVARIATE LSTM MODELS AT PFIZER COMPANY

Optimizer	Vanilla LSTM			Stacked LSTM			Bidirectional LSTM		
	R2	**RMSE**	**MAE**	**R2**	**RMSE**	**MAE**	**R2**	**RMSE**	**MAE**
Adam	0.921	0.802	0.69	0.965	0.533	0.397	0.959	0.576	0.438
RMSProp	0.952	0.623	0.508	0.916	0.824	0.703	0.971	0.478	0.349

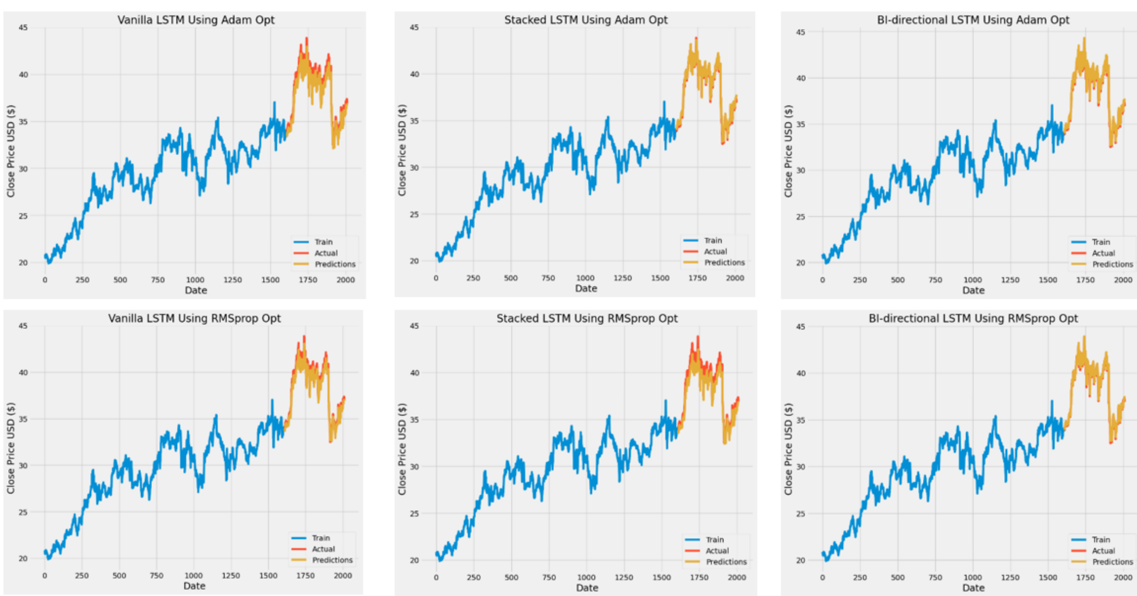

Fig. 6. The Plots of output results of Pfizer company.

VII. Conclusion

This study has successfully demonstrated the efficacy of a single-layered Bidirectional Long Short-Term Memory (BiLSTM) model enhanced with a global pooling mechanism for forecasting stock market prices. Our comparative analysis of the Vanilla, Stacked, and Bidirectional LSTM architectures revealed that the Bidirectional LSTM consistently outperformed the other variants across multiple datasets, particularly when optimized with the RMSprop algorithm. This superior performance is attributed to the BiLSTM's ability to effectively capture both past and future dependencies within the time series data, a critical factor in the volatile environment of stock markets. Rigorous hyperparameter tuning and multiple evaluation metrics like RMSE, MAE, and R^2 have allowed for a thorough validation of the models, ensuring robustness and accuracy in predictions. The visualization of results through line graphs and candlestick charts further corroborates the practical utility of the proposed model in capturing dynamic market trends.

Looking forward, several avenues can enhance the scope and applicability of this research in stock market forecasting. First, integrating a larger variety of data inputs, such as economic indicators, news sentiment analysis, or macroeconomic factors, could provide a more holistic view of the influences on stock prices, potentially increasing the predictive accuracy of the models. Second, exploring incorporating more sophisticated deep learning techniques like attention mechanisms or Transformer models could address some limitations of LSTM models, especially in handling longer sequences with more complex patterns. Third, conducting cross-industry validations with datasets from different sectors and global markets would test the generalizability of the proposed model, ensuring its applicability across diverse economic conditions. Lastly, real-time forecasting implementation in trading systems could be explored to assess the models' practical deployment and operational efficiency in live market conditions. These expansions would enhance the scientific understanding of neural networks in financial applications and bridge the gap between theoretical research and real-world financial decision-making.

References

[1] Hutter, Frank, Holger Hoos, and Kevin Leyton-Brown. "An efficient approach for assessing hyperparameter importance." In International conference on machine learning, pp. 754-762. PMLR, 2014.

[2] Di Persio, Luca, and Oleksandr Honchar. "Artificial neural networks architectures for stock price prediction: Comparisons and applications." International journal of circuits, systems and signal processing 10 (2016): 403-413.

[3] Mahajan, Palak and Abrol, Pawanesh and Lehana, Par, "Effect of Blurring on Identification of Aerial Images Using Convolution Neural Networks," in Proceedings of ICRIC 2019, Jammu, India, Springer, 2020, pp. 469-484.

[4] Lu, Chi-Jie, Tian-Shyug Lee, and Chih-Chou Chiu. "Financial time series forecasting using independent component analysis and support vector regression." Decision support systems 47.2 (2009): 115-125.

[5] Hamzaçebi, Coşkun, Diyar Akay, and Fevzi Kutay. "Comparison of direct and iterative artificial neural network forecast approaches in multiperiodic time series forecasting." Expert systems with applications 36.2 (2009): 3839-3844.

[6] Dixon, Matthew, Diego Klabjan, and Jin Hoon Bang. "Classification-based financial markets prediction using deep neural networks." Algorithmic Finance 6.3-4 (2017): 67-77.

[7] Bouktif, Salah, et al. "Optimal deep learning lstm model for electric load forecasting using feature selection and genetic algorithm: Comparison with machine learning approaches." Energies 11.7 (2018): 1636.

[8] Sezer, Omer Berat, Murat Ozbayoglu, and Erdogan Dogdu. "A deep neural-network based stock trading system based on evolutionary optimized technical analysis parameters." Procedia computer science 114 (2017): 473-480.

[9] Dash, Rajashree, and Pradipta Kishore Dash. "A hybrid stock trading framework integrating technical analysis with machine learning techniques." The Journal of Finance and Data Science 2.1 (2016): 42-57.

[10] Fischer, Thomas, and Christopher Krauss. "Deep learning with long short-term memory networks for financial market predictions." European journal of operational research 270.2 (2018): 654-669.

[11] Hiransha, Ma, et al. "NSE stock market prediction using deep-learning models." Procedia computer science 132 (2018): 1351-1362.

[12] Stathakis, Dimitris. "How many hidden layers and nodes?." International Journal of Remote Sensing 30.8 (2009): 2133-2147.

[13] Greff, Klaus, et al. "LSTM: A search space odyssey." IEEE transactions on neural networks and learning systems 28.10 (2016): 2222-2232.

[14] Karmiani, Divit, et al. "Comparison of predictive algorithms: backpropagation, SVM, LSTM and Kalman Filter for stock market." 2019 amity international conference on artificial intelligence (AICAI). IEEE, 2019.

[15] Bergstra, James, and Yoshua Bengio. "Random search for hyper-parameter optimization." Journal of machine learning research 13.2 (2012).

[16] Bhuriya, Dinesh, et al. "Stock market predication using a linear regression." 2017 international conference of electronics, communication and aerospace technology (ICECA). Vol. 2. IEEE, 2017.

[17] Nivetha, R. Yamini, and C. Dhaya. "Developing a prediction model for stock analysis." 2017 International conference on technical advancements in computers and communications (ICTACC). IEEE, 2017.

[18] Parmar, Ishita, et al. "Stock market prediction using machine learning." 2018 first international conference on secure cyber computing and communication (ICSCCC). IEEE, 2018.

[19] Liu, Siyuan, Guangzhong Liao, and Yifan Ding. "Stock transaction prediction modeling and analysis based on LSTM." 2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA). IEEE, 2018.

[20] Shakva, Abin, et al. "Real-time stock prediction using neural network." 2018 8th International conference on cloud computing, data science & engineering (confluence). IEEE, 2018.

[21] T. Fawcett, "An Introduction to ROC Analysis," Pattern Recognition Letters,vol.27,no.8,pp.861-874,2006.