electronics

Article

Predicting Close Price in Emerging Saudi Stock Exchange: Time Series Models

Abdullah H. Al-Nefaie [1,2] and Theyazn H. H. Aldhyani [1,3,*]

[1] The Saudi Investment Bank Chair for Investment Awareness Studies, The Deanship of Scientific Research, The Vice Presidency for Graduate Studies and Scientific Research, Al-Ahsa 31982, Saudi Arabia
[2] Department of Quantitative Methods, School of Business, King Faisal University, Al-Ahsa 31982, Saudi Arabia
[3] Applied College in Abqaiq, King Faisal University, Al-Ahsa 31982, Saudi Arabia
* Correspondence: taldhyani@kfu.edu.sa

Abstract: The forecasting of stock prices is an important area of research because of the benefits it provides for individuals, corporations, and governments. The purpose of this study is to investigate the application of a key of study to the prediction of the adjusted closing price of a particular firm. Estimating a stock's volatility is one of the more difficult tasks that traders must undertake. Investors are able to mitigate the risks associated with their portfolios and investments to a greater extent when stock prices can be accurately predicted. Prices of stocks do not move in a linear fashion. We propose artificial intelligence (AI) for multilayer perceptron (MLP) and long short-term memory (LSTM) models to predict fluctuations on the Saudi Stock Exchange (Tadawul). This paper focuses on the future forecasting of the stock exchange in the communication, energy, financial, and industrial sectors. The historical records from Tadawul were used as a basis for data collection for these sectors, in time periods from 2018 to 2020. For the purpose of predicting the future values of various stock market sectors, the AI algorithms were applied over a period of 60 days. They demonstrated highly effective performance when simulated using input data, which was carried out to validate the proposed model. In addition, the correlation coefficient (R) of the LSTM and MLP models for predicting the stock market in four sectors in the Saudi Stock Exchange (Tadawul) was >0.9950, which indicates that the outcomes were in good agreement with the predicted values. The outcomes of the forecasts were provided for each method based on four different measures. Among all the algorithms utilized in this work, LSTM demonstrated the most accurate findings and had the best capacity for model fitting.

Keywords: artificial intelligence; deep learning; Saudi Arabia; Saudi stock exchange; statistical analysis

check for updates

Citation: Al-Nefaie, A.H.; Aldhyani, T.H.H. Predicting Close Price in Emerging Saudi Stock Exchange: Time Series Models. *Electronics* **2022**, *11*, 3443. https://doi.org/10.3390/electronics11213443

Academic Editor: Ahmad Taher Azar

Received: 4 October 2022
Accepted: 24 October 2022
Published: 25 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Due to the inherent unpredictability of stock prices over the long term, the process of valuing stocks and making predictions about their future performance is difficult. The dated market hypothesis asserts that it is not possible to forecast stock values and that stocks act in a haphazard manner [1]. However, recent technical analyses have shown that the majority of stock values are reflected in previous records; as a result, movement trends are essential for effectively forecasting stock values [2]. In addition, the groupings and movements of the stock market are influenced by a number of economic factors, including political events, general economic conditions, the index of commodity prices, investors' expectations, the movements of other stock markets, and the psychology of investors [3]. High market capitalization is a factor that is considered when determining the value of stock groupings. When attempting to obtain statistical information from the value of stock prices, a variety of technical parameters must be considered [4]. In most cases, stock indices are derived from the values of stocks that have substantial market investment, and they

frequently provide an evaluation of the current state of economic activity in each nation. For instance, research has shown that the size of a country's stock market capitalization has a beneficial effect on the country's overall economic growth [5]. The unpredictability inherent in the movement of stock prices makes investments generally high risk for their owners. In addition, it is typically challenging for governments to identify the current position of the market. Stock prices tend to be dynamic, non-parametric, and nonlinear; as a result, they frequently contribute to the poor performance of statistical models and make it difficult to forecast precise values and movements [6,7].

Since they encourage monetary and capital gain [8], stock exchanges have developed into an essential component of the economy. A stock market is a network of economic transactions that allows for the buying and selling of shares in various companies. The equity or market share of an organization is a reflection of the ownership claims made on that organization. This could include shares that were acquired through the public stock market or through individual trading, such as the shares of private companies that were sold to investors. The process of moving money from small individual investors to large trading participants, such as banks and companies, is referred to as transactions in stock markets. Investing in the stock market is considered a profession fraught with peril because of the likelihood of unstable behavior [9].

Investors stand to gain quite a bit from using stock market prediction (SMP) strategies that are carried out accurately. The accurate forecasting of stock prices may give shareholders helpful guidance in making appropriate decisions regarding whether to buy or sell shares in their company. SMP refers to the process of attempting to forecast the value that will be placed on a stock in the future [10]. Over the years, numerous strategies for predicting the price of stocks have been presented. In general, they can be broken down into one of four categories. The first type of analysis is known as basic analysis, and it is constructed using information that is freely accessible to the public. The second type of study is known as technical analysis, which entails making recommendations based on historical data and pricing. The third approach involves utilizing machine learning and data mining on enormous amounts of data that have been compiled from a variety of different sources. The final method is called sentiment analysis, which uses previously reported news, articles, or blog posts as the basis for its forecasts [11]. The combination of the two most recent groups is more recent than the combination of the two groups that came before it, and research suggests that they may exert more effect on the decision of whether to purchase or sell stocks [12].

A stock market is a key pillar that plays an active role in the growth of both developed and emerging economies. It functions as an intermediary between borrowers and lenders in both types of economies. The term "efficient stock market" refers to the degree to which a market is able to successfully incorporate stock prices, which reflect the value of the underlying stocks. This market has the potential to hasten the pace of economic development by encouraging individuals to save more money and make more effective use of available resources. The integration of stock markets has emerged as a significant topic in the field of finance literature in recent years and has garnered widespread currency as a result of the high value it holds for the parties concerned. This integration may result in significant economic development across the economy, an improvement in the allocation of capital, a reduction in expenses associated with capital, and an increase in the efficiency of risk sharing [13].

Using artificial intelligence models known as multilayer perceptron (MLP) and long short-term memory (LSTM) to demonstrate forecasting performance in the thin emerging Saudi Stock Exchange (Tadawul) of the Gulf countries, we break new ground in the existing literature that is pertinent to estimating stock market volatility. This is because we use these models to break new ground in the existing literature that is pertinent to estimating stock market volatility [14,15].

- The Saudi market, which has established itself as one of the most important markets worldwide and in the Middle East over the past few years, is particularly noteworthy

because it is home to the Aramco company, which is widely regarded as the most valuable company on the planet in terms of market value.
- Using advanced AI models, this study contributes to the area by giving the investment community a model that can be used to evaluate and anticipate the risks associated with investing in the Saudi stock market.
- This study can be helpful to Saudi investors in that it can provide them with information that can guide them in making informed decisions regarding their investments and the diversification of their portfolios.

2. Related Work

Numerous studies on various methods of market prediction have been carried out in academic domains. To evaluate the movement of stock prices, Long et al. [16] investigated a deep neural network model using public market data and transaction records. The results of the experiments demonstrated that bidirectional LSTM was capable of predicting stock prices for the purposes of making financial judgments, and the method achieved the best performance when compared to other prediction models. Pang et al. [17] conducted research to enhance a novel neural network algorithm in the hope of producing more accurate stock market forecasts. To analyze the movement of the stock market, they suggested LSTM models equipped with an integrated layer, as well as LSTM models equipped with an automatic encoder. According to the findings, the deep LSTM model that had an embedded layer performed the best, and the accuracy of the model for the Shanghai composite index was 57.2% and 56.9%, respectively. Kelotra and Pandey [18] constructed a model that successfully analyses swings in the stock market by employing a predictor that is known as the deep convolutional LSTM model. This model was used in their model. The findings showed that the model acquired a low MSE and RMSE of 7.2487 and 2.6923, respectively, after it was trained using an optimization method based on Rider's monarch butterfly. This was shown by the fact that it achieved a low mean square error (MSE). The fact that it attained a low value for each of these criteria was evidence that this was the case.

An improved approach of emotions analysis was used by Bouktif et al. [19] in their investigation of whether or not the direction of stock market trends may be predicted. They were curious as to whether or not it was possible to forecast the path that the trend would take. The proposed technique beat earlier sentiment-based stock market prediction methods that incorporated deep learning and had an accuracy of 60% when predicting stock market movements. Zhong and Enke [20] proposed utilizing a big dataset from the SPDR S&P 500 ETF to examine return direction using 60 economic and financial variables. This analysis will be conducted using the dataset. Deep neural networks, artificial neural networks (ANNs), and principal component analysis were utilized in order to accomplish the task of projecting the daily future of stock market index returns (PCA). According to the findings, deep neural networks performed significantly better than other potential data classifiers when based on PCA representations of the data. According to the findings, deep neural networks performed significantly better than other potential data classifiers when based on PCA representations of the data. The feature optimization was achieved by Das et al. [21] by taking into consideration the social and biological components of the firefly technique. They approached the problem by incorporating the evolutionary background of objective value selection into their methodology. According to the findings, the firefly model, which utilized an evolutionary framework and made predictions using the online sequential extreme learning machine (OSELM) method, performed better than the other models tested.

Convolutional neural networks (CNNs) are a framework established by Hoseinzade and Haratizadeh [22]. This framework may be applied to various types of data collection (including different markets) to investigate features for predicting the future movement of markets. According to the findings, a significant leap forward in the performance of the prediction was accomplished compared to other recent baseline approaches. In their research, Krishna, Kumar, and Haider [23] compared the efficacy of single classifiers with that of a

multi-level classifier, which was a combination of various machine learning approaches. In their findings, they found that single classifiers were more effective than multi-level classifiers (such as decision trees, support vector machines, and logistic regression classifiers). The results of the experiments demonstrated that the multi-level classifiers performed better than the single classifiers, which resulted in a model that was more accurate and had the best predictive ability, with a growing inaccuracy of approximately 10–12%. This model was the result of the experiments that were conducted. Chung and Shin [24] employed CNN, a sort of method that belongs to the category of deep learning, to produce forecasts on the behavior of the stock market. In addition to this, a genetic algorithm (GA) was utilized in order to systematically enhance the parameters of the CNN technique. According to the findings, the approach that combines GA and CNN, which is known as GA-CNN, performed significantly better than the models that were used for comparison. The authors Sim et al. [25] stated that a fresh educational technique for learning how to predict stock prices may be to watch CNN. The use of CNNs and increasing their performance on data from the stock market were the two difficulties that this study intended to investigate and find solutions for. Wen et al. [24] employed the CNN algorithm to assess noisy temporal series based on frequent pattern occurrences. This is a unique method that has not been used before. The results showed that the approach was both more effective than traditional signal processing methods and also superior to these methods, demonstrating an accuracy improvement of between 4% and 7%. This was proved by the findings.

To anticipate the stock price for the following day based on historical stock price data, a CNN-BiLSTM-AM [24] technique was one of the options under consideration. This approach takes a number of different inputs into consideration, including opening prices, highest and lowest prices, closing prices, volume, and turnover. The ups and downs of the situation are also taken into consideration. A convolutional neural network is used to process the data entered into the system (CNN). The BiLSTM algorithm is used to learn and make predictions based on the collected feature data. By utilizing AM, it is possible to examine the effect that different time periods within a time series dataset have on prediction results. When compared to the findings of the other models, the CNN-BiLSTM-AM produced the most accurate results and had the best prediction accuracy.

Both ANN [25] and RF were utilized to conduct price comparison research and anticipate the stock's closing price for the following trading day. For projecting stock prices, ANN performed better than RF based on the values of RMSE, MAPE, and MBE. According to the findings of the experiments, ANN performed much better than the other approaches in terms of RMSE (0.42), MAPE (0.77), and MBE (0.013).

To improve the accuracy of the stock predictions, GWO-optimized ENN [26] was examined. To compute daily returns, the model that was presented made use of the stock prices that were recorded each day on the NASDAQ and the New York Stock Exchange for eight different companies. The ENN model was improved through the application of the GWO algorithm. Based on the findings, it can be concluded that GWO-ENN performs better than optimized models such as PSO-ANN and FPA-ANN. In this work, the forecasting of stock prices for the domestic stock market was performed with the help of the combination model ARI-MA-LS-SVM [27]. With the use of the PCA method, the dimensions of the input variables can be decreased, so the amount of time spent training the model was reduced. The K-NN regression model was utilized to make projections regarding stock values. To forecast the movements of stock prices in the future, the historical stock prices of a number of different firms, as well as a number of technical indicators, were analyzed. When compared to the findings of other machine learning algorithms, the experimental results performed significantly better.

3. Data and Methodology

The process of modeling and forecasting the performance of a variety of AI models is becoming increasingly important for businesses and policymakers around the world. To analyze and make predictions regarding the volatility of the Saudi Stock Exchange

(Tadawul), we employed two nonlinear time series models that were powered by AI. The MPL and LSTM nonlinear models were utilized for our applications. The framework of the Saudi Stock Exchange prediction model is presented in Figure 1.

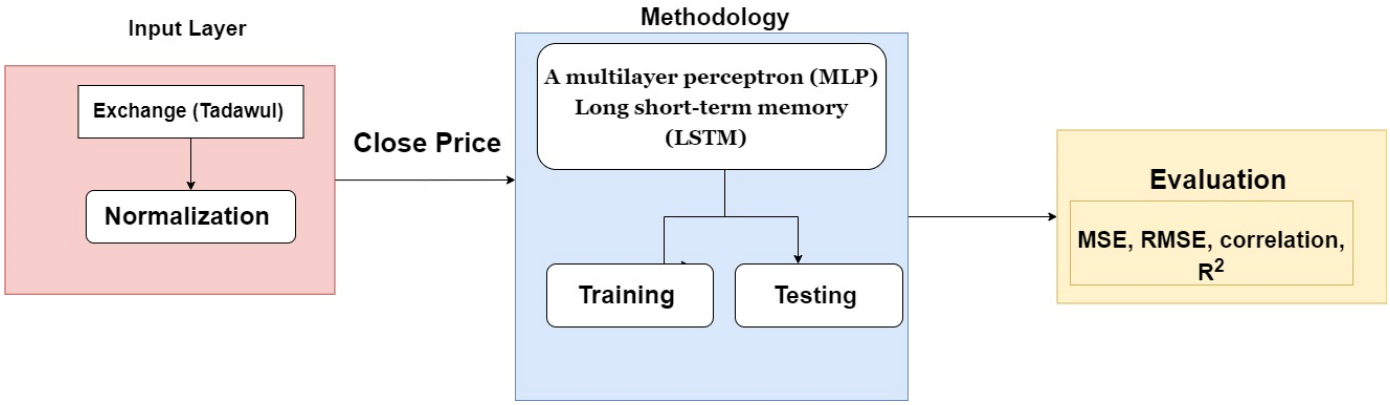


Figure 1. Framework for predicting Saudi Stock Exchange.

3.1. Dataset

The Saudi Stock Exchange (Tadawul) was transformed into a holding company, the Saudi Tadawul Group, in March 2021, and the Saudi Exchange was thereafter established as a wholly owned subsidiary of the parent business. The exchange facilitates the listing and trading of securities for local and international investors, serving as Saudi Arabia's dedicated stock exchange and the largest stock exchange in the Middle East. The Saudi Exchange is the authoritative source for the market and plays a crucial role in the Tadawul Group's long-term expansion ambitions and the provision of attractive and diverse investment possibilities for market participants. Within the Gulf Cooperation Council, the Saudi Exchange is the preeminent market and ranks ninth largest among the 67 stock exchanges that make up the World Federation of Exchanges (GCC). It is an affiliate member of the International Organization of Securities Commissions (IOSCO), the World Federation of Exchanges (WFE), and the Arab Federation of Exchanges (AFE); its stock market ranks third among its emerging market counterparts. In this research work, the dataset was collected from 2018 to 2020. The Tadawul has nine important sectors, and four were used to evaluate and test the proposed model: communication, energy, financial, and industrial. Table 1 shows the features of the dataset; the closing price was considered to predict the future values of the stock market.

3.2. Preprocessing

It is the goal of data normalization to eliminate data redundancy, which arises when numerous fields contain duplicated information. A database can be made more flexible by deleting redundant information from it. In statistical parlance, "normalization" refers to the process of rescaling the data into the interval [0, 1]. In general, standardization involves recalculating the values of the data so that they have a mean of zero and a standard deviation of one (unit variance).

$$z_n = \frac{x - x_{min}}{x_{max - x_{min}}} (New_{max_x} - New_{min_x}) + New_{min_x} \quad (1)$$

The *x_max* and *x_min* variables represent the highest and lowest possible values, respectively. The lowest number is denoted by the notation New (*min x*), and the highest number is written as New (*max x*). The normalization performance of the Saudi Stock Exchange (Tadawul) is presented in Figure 2.

Table 1. Features of the Stock Exchange (Tadawul).

Description	Feature Type	Feature
Symbol of companies	Integer	Symbol
Name of companies	String	Name of company
Trading of different of company	String	Trading name
Company operate	String	Sector
Time	Date	Date
Lowest price for the stock on that day	Decimal	Low
Highest price for the stock on that day	Decimal	High
Opening price for the stock on that day	Decimal	Open
Closing price for the stock on that day	Decimal	Close
Change of stock market on the last day	Decimal	Change
Number traded at stock market	Decimal	Value-traded
Volume traded on the day	Decimal	volume traded

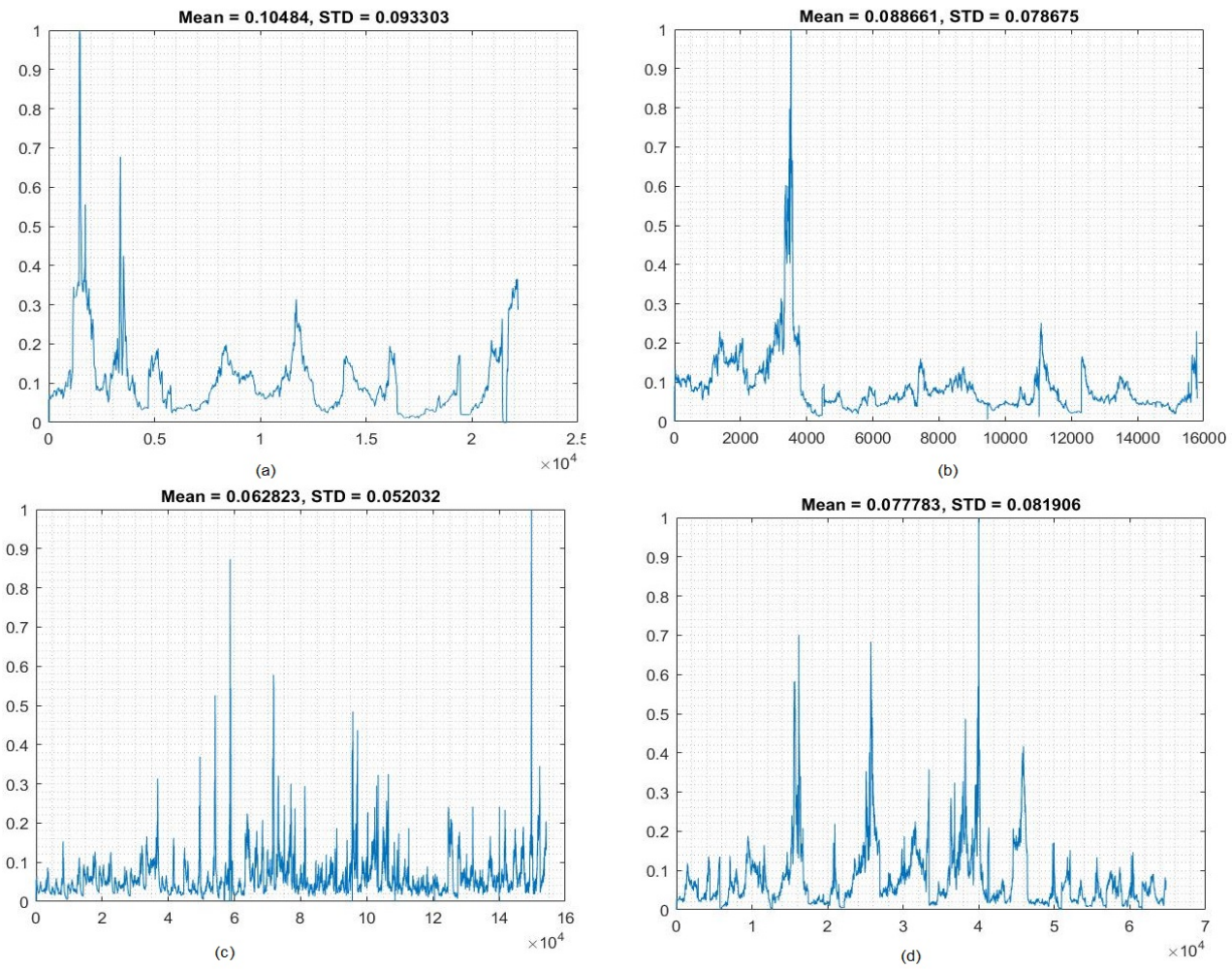


Figure 2. Normalization performance of the Saudi Stock Exchange (**a**) communication (**b**) energy, (**c**) financial and (**d**) industrial sectors.

3.3. Prediction Models

The process of modeling and forecasting the performance of several AI algorithms, such as MLP and LSTM models, is becoming increasingly important for businesses and policymakers all over the world.

3.3.1. Multilayer Perceptron (MLP)

ANNs can range from straightforward neural networks to extremely complicated neural networks with dozens of interconnected layers. Figure 3 is an abstract representation of an ANN taken from [28], showing an input layer, an output layer, and two hidden layers. Each layer's nodes are linked to one another by their common parent. More hidden layers can be added to the network to increase its depth. ANNs are a form of AI that are modeled after the way the human brain works [29]. The function of a neural network is controlled by the model of densely interconnected neurons, the network topology, and the learning method. ANNs are simulations of biological neural networks. Pattern recognition and data categorization are two specific issues that can be addressed by applying the ANN approach [30,31]. MLP is the ANN that is utilized most often, particularly in the field of environmental research. This approach can be utilized to carry out a variety of tasks, including matching features and solving pattern recognition issues. In addition, MLP can be utilized to demonstrate the classification of linear patterns that are inseparable from one another. MLPs are a representation of feedforward neural networks, which have numerous layers of units between the input and output layers. The following are some possible ways to express the output of a neutron:

$$\xi = \sum_{i=1}^{n} w_i x_i - b = w^T x - b \tag{2}$$

$$y = \sigma(\xi), \tag{3}$$

$$\sigma(\xi) = \frac{1}{1 + e^{-(\xi)}}, \tag{4}$$

where x_i is the i-th input number, w_i is the i-th input's link weight, and $w = (w_i \ldots w_n)$ is the total link weight. The sum of all the inputs is denoted by T, and each input is denoted by x_i, where x_i ranges from 1 to x_n. The letter b stands for a threshold or bias, while the number n stands for the total number of inputs. The activation function $s(x)$ is often a continuous or discontinuous function whose task is to map real numbers onto the interval. Another option is to use the sigmoidal activation function. Figure 3 shows the structure of MLP for predicting the stock market.

Pseudocode MLP

1. Initialization variable delay number, number of feed forward layers max-iteration
2. Load data set X
3. Normalization dataset (X)
4. Split the dataset 70% training and 30% as testing
5. -building and ANN architecture
6. for d = Delays
7. X = [X; x(:,Range-d)];
8. end
9. make prediction
10. calculate prediction by using evaluation metrics

3.3.2. Long Short-Term Memory (LSTM)

The LSTM model is an extension of the recurrent neural network (RNN) in its more complex form. Utilizing memory units that are able to update the previous hidden state is how this model manages to preserve long-term memory. It offers feedforward at each individual neuron [32–34]. The output of an RNN is not only reliant on the neuron input and the weight of the present neuron, but also dependent on the neuron inputs of the neurons

that came before. Due to this functionality, it is feasible to understand the temporal linkages that exist in a long-term sequence. The exploding and vanishing gradient difficulties that arise during the training of typical RNNs can be circumvented, thanks to the device's internal memory unit and gate mechanism [35–39]. Therefore, the LSTM model's internal structure is comprised of four essential parts: input gate, output gate, forget gate, and cell status. A diagram of the LSTM's internal structure is shown in Figure 4.

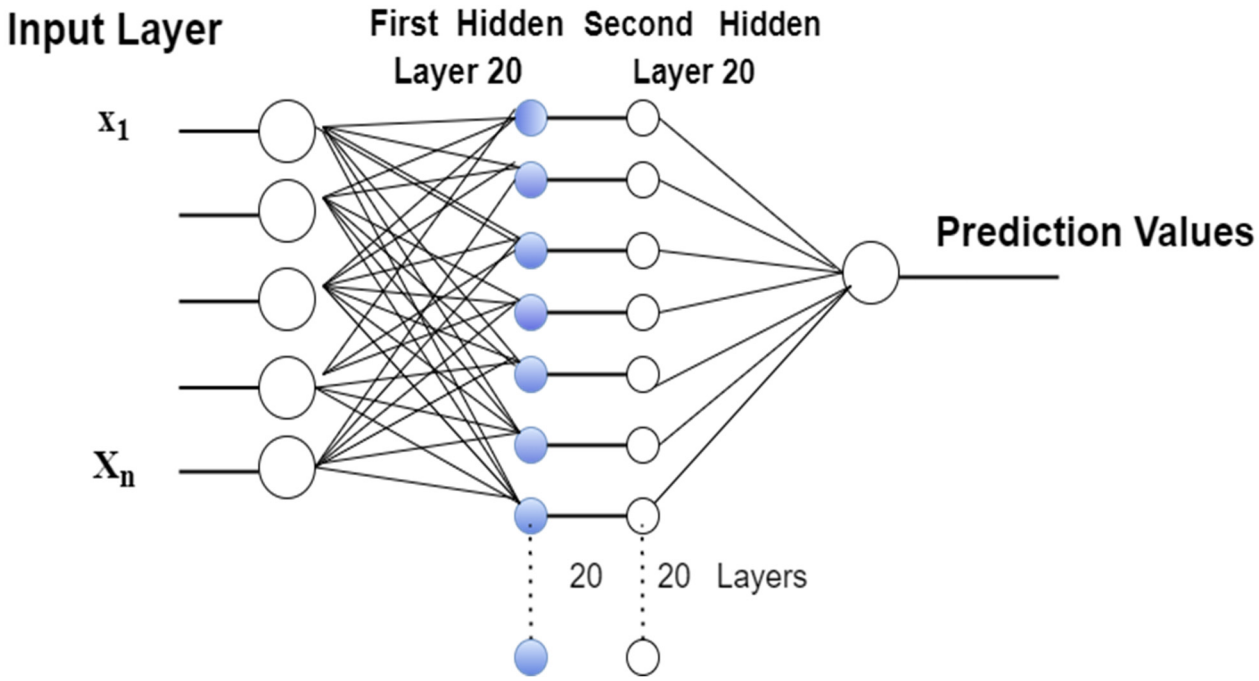


Figure 3. MLP structure to predict the stock market.

$$\text{Forget gate layer}: \; f_t = \sigma\left(W_{ef}X_t + W_{ef}h_{t-1} + W_{cf}C_{t-1} + U_f\right) \tag{5}$$

$$\text{Input gate layer}: \; i_t = \sigma\left(W_{xi}X_t + W_{hi}h_{t-1} + W_{ci}C_{t-1} + U_i\right) \tag{6}$$

$$\text{New memory cell}: \; C_t = \sigma\left(f_t c_{t-1} + i_t \tanh(W_{xc}X_t + W_{hc}\, h_{t-1} + U\right) \tag{7}$$

$$\text{Output gate layer}: \; o_t = \sigma\left(W_{xo}X_t + W_{ho}h_{t-1} + W_{co}C_{t-1} + U_o\right), \tag{8}$$

$$h_t = O_t \times \tan \mathrm{h}(C_t) \tag{9}$$

It is possible to think of an LSTM network as a gated cell. To be gated implies that the cell evaluates each piece of data and makes a decision about whether to keep it, depending on the relative relevance or weights it has been given. Input, forget, and output gates make up LSTM. Which states should be remembered, and which ones forgotten are determined by the forget gate, f_t. It is the responsibility of the input gate i_t to determine which value will be modified by the incoming signal. The state of the cell is ultimately propagated to neighboring neurons through the output gate, o_t=. It is made up of a logistic layer and a layer, with the former producing values between zero and one and the latter producing a new vector that is added to the state. Equations (4)–(8) depict LSTM equations. In a recurrent neural network, the output from the embedding layer, X_t, is fed into the hidden layer, where the weight matrix, W, is applied to produce the final output, y_t. In LSTM h_t, three gates govern the behavior of the central element, which serves as a memory cell.

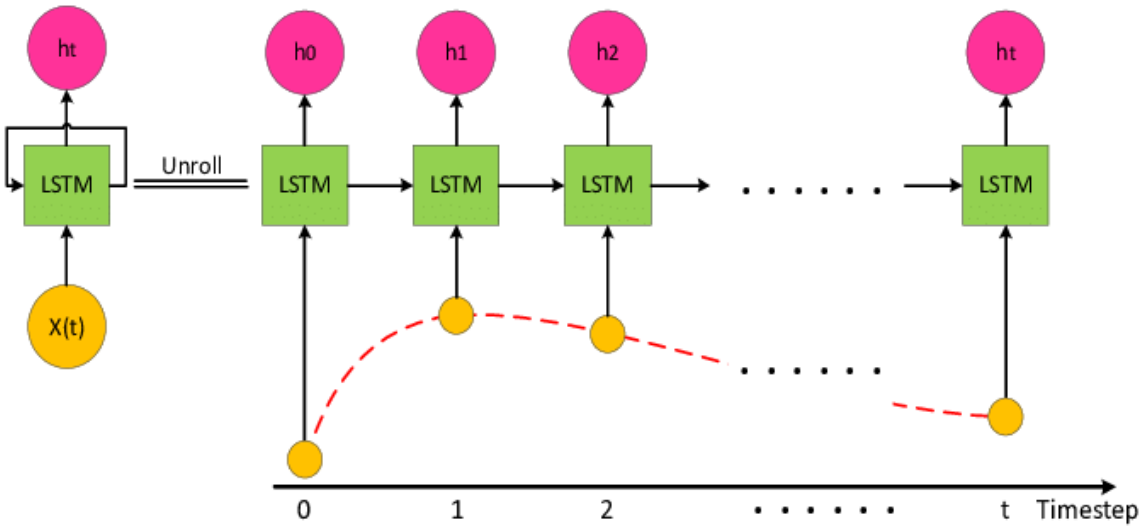


Figure 4. LSTM's internal structure.

Pseudocode LSTM

1. Import libraries
2. Load set
3. Normalize the dataset into values 0 and 1
4. Split the dataset 70% training and 30% as testing
5. Set input unites, output units, LSTM unit and optimizer
6. For epoch and batch size do
7. Train LSTM
8. End
9. Make prediction
10. calculate prediction by using evaluation metrics

3.4. Performance Measurement

Several different metrics have been used to assess the accuracy of the models' projections about price trends and movement directions. Mean squared error (MSE), Root mean squared error (RMSE), normalized root mean squared error (NRMSE), and Pearson correlation coefficient error criteria were utilized in order to assess prediction results in the experiments and estimate the performances of the models.

3.4.1. Root Mean Squared Error (RMSE)

In regression analysis, RMSE stands for the standard deviation of the errors in the predictions. Residuals not only show how the actual values are spread around a prediction model, but they also depict the distance of the actual values from the model. Residuals are another name for prediction errors. This metric depicts how tightly the data fit within the parameters of the best-fitting model. The root square mean root is average squared variance between the prediction and the actual data is the formula for the root mean square error, often known as the RMS error. The relative root mean square error (RRMSE) normalizes the total squared error in the same manner that the root mean square error (RMSE) does for the total squared error by dividing it by the total squared error of the predictor model. Both of these methods are described in more detail below. This procedure normalizes the total squared error in this particular instance. The correct formula may be found in Equation (10), which can be found.

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(y_{i,observ} - y_{i,\,estim})^2}{n}} \quad (10)$$

3.4.2. Mean Squared Error (MSE)

Mean squared error (MSE) is a measure of predictor performance that can never have a negative value (values closer to zero are better). MSE accounts for the bias and variance of a prediction model (how broadly the forecasts vary from one data sample to another) and is the second moment of error (around the origin). Similar to how the MSE measures distance from the true origin, the MSE measures distance from the mean (how close the average predicted value is to the observation). The necessary formula is given in Equation (11).

$$MSE = \frac{1}{n} \sum_{i=1}^{n} \left(y_{i,observ} - y_{i,\ pred} \right)^2 \tag{11}$$

3.4.3. R-Squared Metrics

The coefficient of determination (R^2) is a statistical metric that indicates how much of the variation in a dependent variable can be attributed to the independent variable(s) in a regression model.

$$R\% = \frac{n(\sum_{i=1}^{n} y_{i,observ} \times y_{i,\ estim}) - (\sum_{i=1}^{n} y_{i,observ})(\sum_{i=1}^{n} y_{i,\ estim})}{\sqrt{\left[n(\sum_{i=1}^{n} y_{i,observ})^2 - (\sum_{i=1}^{n} y_{i,observ})^2\right]\left[n(\sum_{i=1}^{n} y_{i,\ estim})^2 - (\sum_{i=1}^{n} y_{i,\ estim})^2\right]}} \times 100, \tag{12}$$

where $y_{i,observ}$ is the experimental value of data point i, $y_{i,\ estim}$ is the anticipated value, and n is the number of simple.

4. Results Analysis

For the purpose of the experiment, we looked at data from the Saudi Stock Exchange, which is open to the public and can be accessed at https://www.saudiexchange.sa/wps/portal/tadawul/home/ (last accessed 26 August 2022). The files contain daily stock price data for 593,821 trading days, spanning 1 January 2018 to 31 December 2020. The current analysis was designed to estimate how the communication, energy, financial, and industrial sectors would perform on the Saudi Stock Exchange. It is possible to recognize an uptrend in stock prices by employing moving average technical indicators, which are regarded as one of the stock market's technical indicators. The MSE, RMSE, R-squared, and NRMSE statistics were utilized to perform an analysis on the suggested deep learning stock prediction model. The stock marketing value is represented by variable *xi*, the forecast value is specified by variable *yi*, and the number of items in the dataset is represented by variable *n*. The variable *n* also contains information about how many elements are in the dataset. The dataset was divided into 70% training and 30% testing. Figure 5 shows a snapshot of the performance of the proposed system.

4.1. Platform for Developing the Prediction System

The proposed system went through development in a variety of contexts, including both hardware and software. MATLAB was used to develop our own algorithm for predicting the Saudi Stock Exchange (Tadawul). The application was set up to run on a computer with a CPU of Intel (R) Core (TM) i7–4770, a speed of 3.20 GHz, 8 gigabytes of memory, and 64-bit Windows 10 installed.

4.2. Results of LSTM Model

Most stock market data are nonlinear, the kind of data most typically handled by AI algorithms. In the proposed work, we considered LSTM models to predict the closing price of the Saudi Stock Exchange sectors. Table 2 shows the results of the LSTM model for four sectors in the training process in the Saudi Stock Exchange (Tadawul). It was observed that the LSTM model achieved a low prediction error MSE = 1.0836×10^{-07} for predicting the closing price of the communication sector, and the LSTM model scored low for predicting all stock market sectors.

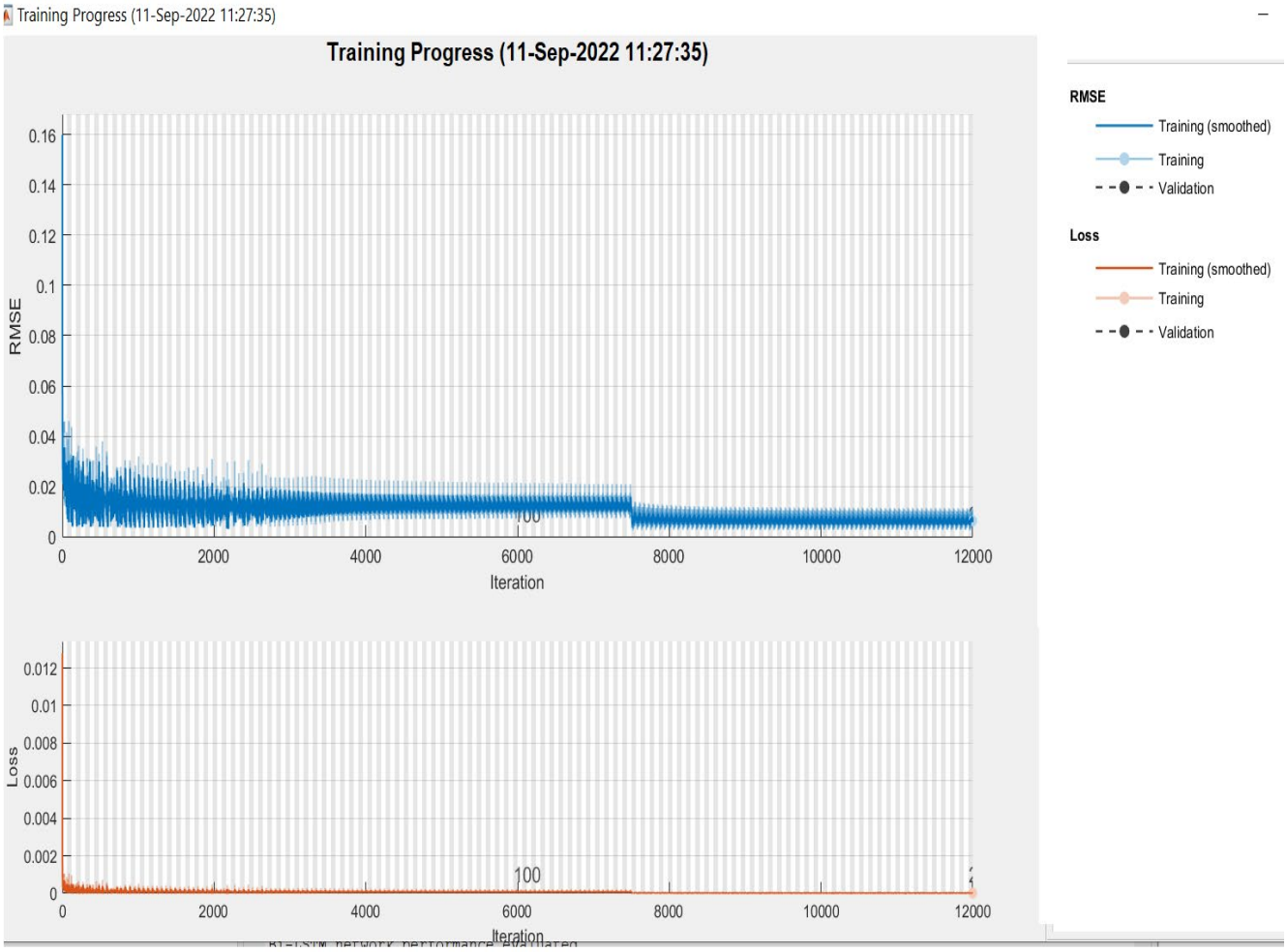


Figure 5. Snapshot of the performance model.

Table 2. Performance of the LSTM model in the training stage.

Dataset	MSE	RMSE	NRMSE	R
Communication	1.0836×10^{-07}	0.00032918	0.0056653	99.99
Energy	7.1197×10^{-05}	0.008437	0.08597	99.81
Financial	3.1061×10^{-05}	0.0055732	0.093013	99.69
Industrial	4.0276×10^{-05}	0.0063463	0.073471	99.97

Figure 6 illustrates how well the suggested LSTM model performed during the training phase. According to the findings, there is a very good correlation, as indicated by the R-value, between the values obtained via experimentation (along the X-axis) and those anticipated (along the Y-axis). The LSTM showed a high percentage of R = 99.99 in the communication, energy, financial, and industrial communication sectors.

The LSTM model was tested and validated using an unseen dataset that consisted of 30% of the total data. Figure 7 and Table 3 show the results of its performance in the testing phase. As can be seen in Figure 4, it was possible to obtain good agreement between the values that were predicted and the values that were sought for experimentation. Similar to the testing phase, high values of R (99%) were shown in all sectors. This level of agreement demonstrates that the LSTM model was accurate in its ability to forecast the amount of arsenic that would be removed from contaminated water by adsorption techniques.

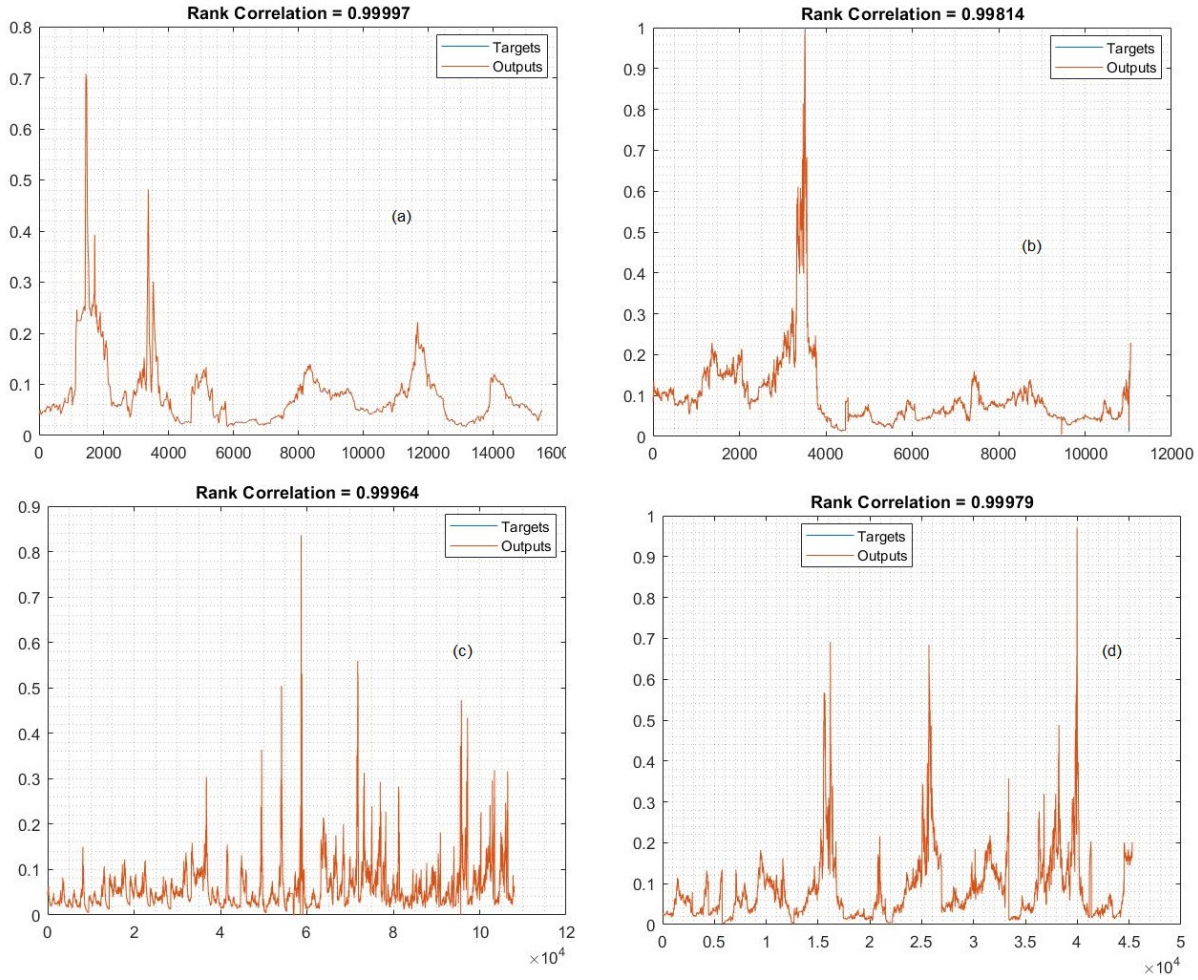


Figure 6. Prediction plot of the LSTM model in the training process (**a**) communication (**b**) energy, (**c**) financial and (**d**) industrial sectors.

Table 3. Performance of the LSTM model in the testing stage.

Dataset	MSE	RMSE	NRMSE	R
Communication	6.9207×10^{-06}	0.0026307	0.039057	99.99
Energy	2.688×10^{-05}	0.0051853	0.07919	99.75
Financial	7.4341×10^{-05}	0.008624	0.1444	99.96
Industrial	1.0516×10^{-05}	0.0032429	0.06599	99.95

Figures 8 and 9 illustrate the histogram error that occurred for the predicted values while the model was in the training and testing stages. Examining the error histogram metrics allowed for the calculation of the degree of discordance that existed between the anticipated values and the values that were achieved. These error values indicate how the predicted values deviated from the goal values, and because of this, these numbers might be negative. In addition, they show how the anticipated values differed from the goal values. The histogram errors were 0.0026, 0.00908, 8.9102×10^{-07}, and 3.4915×10^{-05} in the training process, whereas the results of LSTM in the testing phase were 1.3605×10^{-05}, 8.5609×10^{-05}, 0.00287, and 1.6726×10^{-05} in the communication, energy, financial, and industrial sectors.

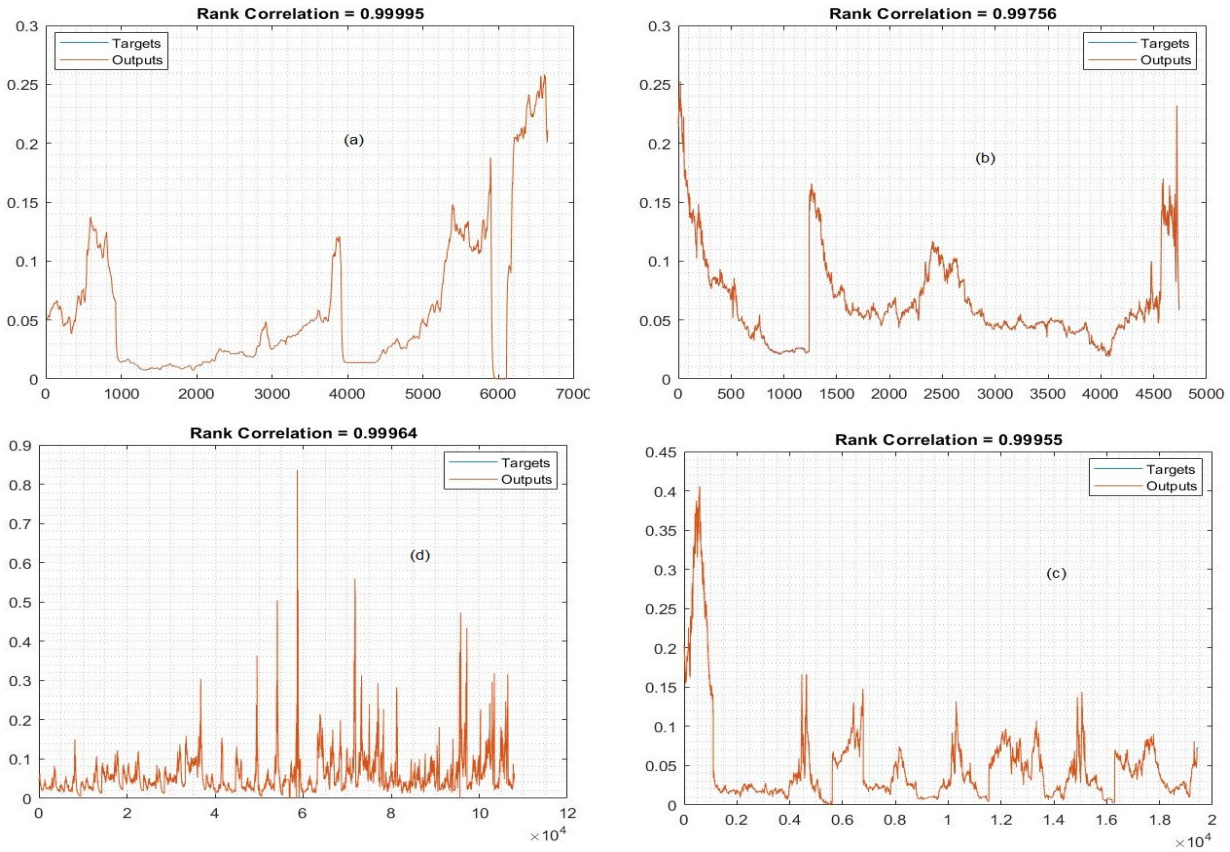


Figure 7. Prediction plot of the LSTM model in the testing process, (**a**) communication (**b**) energy, (**c**) financial and (**d**) industrial sectors.

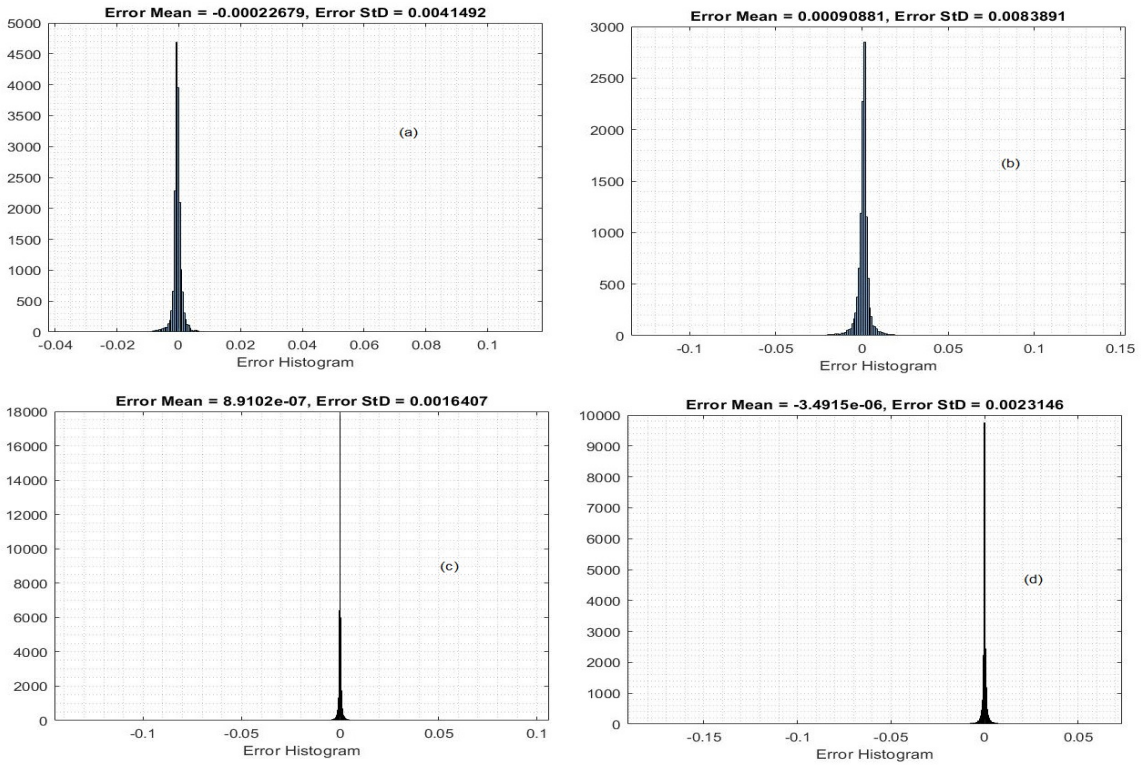


Figure 8. Histogram error of LSTM in the training phase, (**a**) communication (**b**) energy, (**c**) financial and (**d**) industrial sectors.

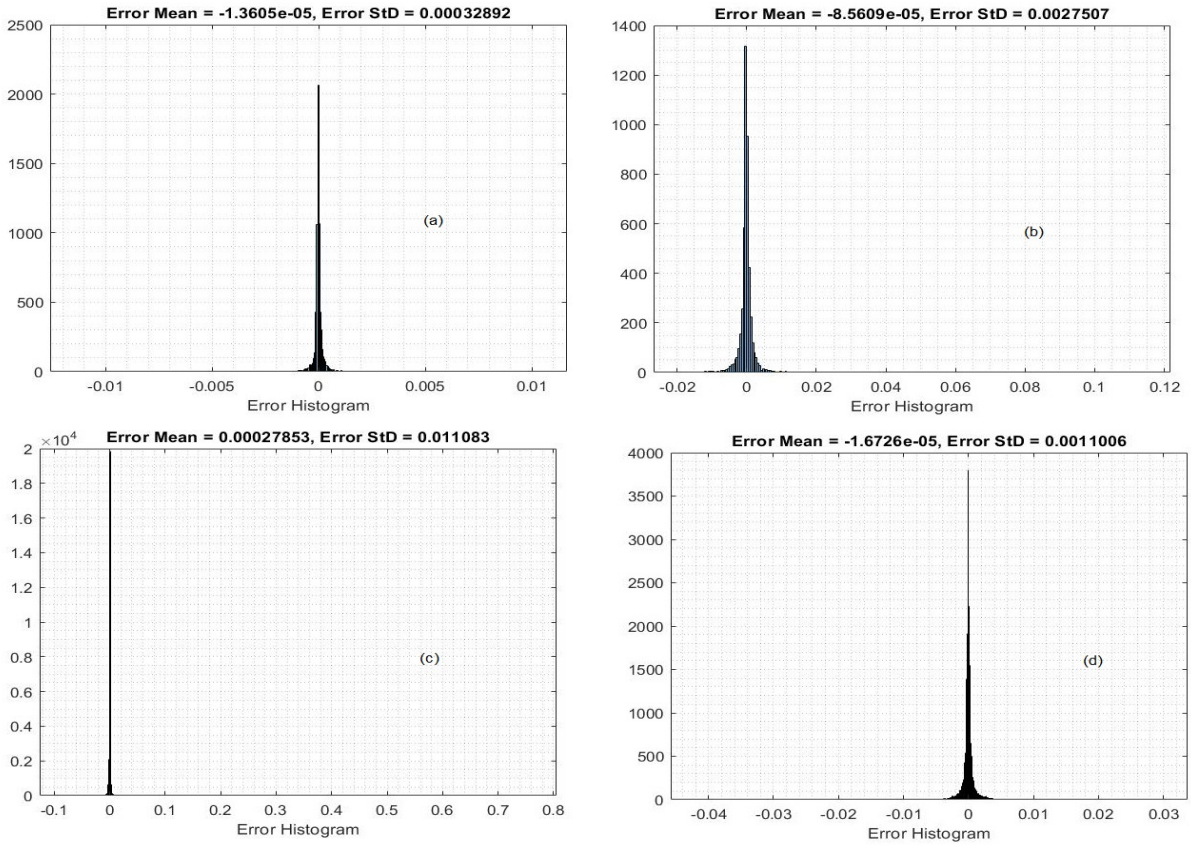


Figure 9. Histogram error of LSTM in training and testing (**a**) communication (**b**) energy, (**c**) financial and (**d**) industrial sectors.

4.3. Results of the Multilayer Perceptron (MLP) Model

Table 4 and Figure 10 present the performance of the MLP model. The mean execution time for the 100 rounds of execution of the model in the same computing environment was found to be 42.04 s. This was almost twice the time needed for the execution of the MLP model. The values of R, MSE, RMSE, and R^2 were used to conduct an analysis of the predictive ability of the model. This should be kept to a minimum for an ANN model that has a high level of efficiency. The average value of the ratio of the MSE to the mean of the actual close values yielded by the model was 1.825×10^{-07} for predictions in the communication sector. The prediction errors of the MLP model for forecasting values of the energy, financial, and industrial sectors were 2.7828×100^{-05}, 2.692×100^{-06}, and 5.3574×100^{-06}. As shown in Table 4, not only was the R-value rather high (R was found to be more than 0.9999), but the MSE, RMSE, and NRMSE values were alste quite low. The LSTM achieved lower prediction values in the communication sector with MSE = 1.825×10^{-07}. These numbers suggest that the proposed model performs very well in terms of its ability to make accurate predictions about output parameters.

Table 4. Performance of the MLP model in the training stage.

Dataset	MSE	RMSE	NRMSE	R
Communication	1.825×10^{-07}	0.0042721	0.005313	99.95
Energy	2.7828×100^{-05}	0.0052751	0.05373	99.71
Financial	2.692×100^{-06}	0.0016407	0.027383	99.71
Industrial	5.3574×100^{-06}	0.00023146	0.026796	99.85

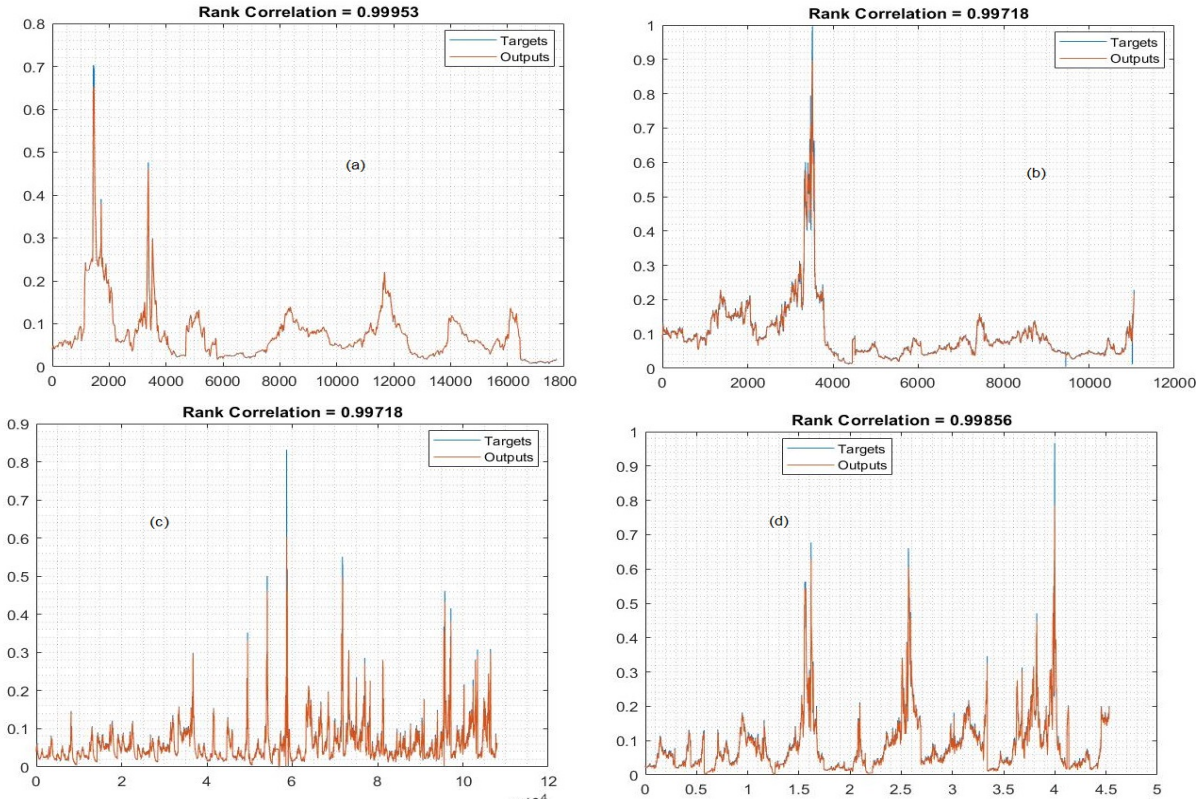


Figure 10. Prediction plot of the MLP model in the training process, (**a**) communication (**b**) energy, (**c**) financial and (**d**) industrial sectors.

Table 5 shows the results of the MLP model at testing. The observation dataset comprising 30% of the whole dataset was used to test and verify the current iteration of the LSTM model. Figure 11 demonstrates that it was possible to arrive at a satisfactory level of agreement between the anticipated values and the desired values via the process of testing. The optimal result of the MLP model to predict stock markets was MSE = 1.0836×10^{-07} in the communication sector. Figure 11 shows the agreement between the predicted and observed values, and LSTM had the highest percentage of R = 99 in the testing phase.

Table 5. Performance of the MLP model in the testing stage.

Dataset	MSE	RMSE	NRMSE	R%
Communication	1.0836×10^{-07}	0.0032918	0.005665	99.82
Energy	7.572×100^{-06}	0.0027517	0.042025	99.59
Financial	0.00012291	0.011087	0.18566	99.60
Industrial	1.211×10^{-06}	0.0011007	0.022401	99.79

In addition, the error histogram of the MLP at the testing phase of the dataset is shown in Figure 12, and Figure 13 is drawn over the zero error threshold. The histogram errors obtained by the MLP model for predicting stock markets in the training phase were 0.000226, 0.000906, 0.000959, and 0.002156, while the histogram errors of the MLP model in the testing phase were 0.00368, 0.000894, 0.0001183, and 0.001407 for the communication, energy, financial, and industrial sectors, respectively. This is further evidence of the excellent performance of the suggested ANN model.

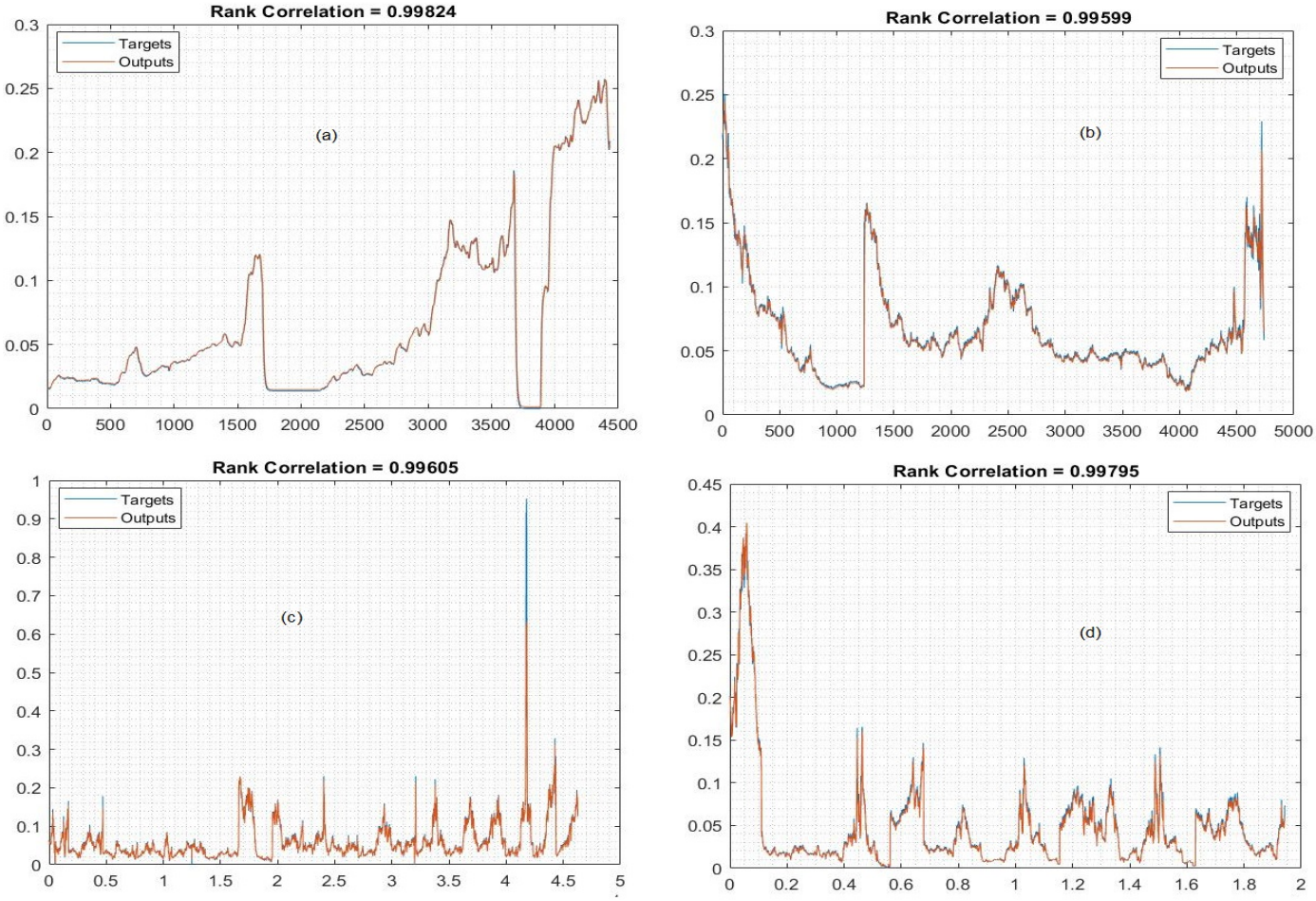


Figure 11. Prediction plot of the MLP model in the testing process, (**a**) communication (**b**) energy, (**c**) financial and (**d**) industrial sectors.

The artificial intelligence models LSTM and MLP were applied to forecast the future values of the Saudi stock exchange market. A forecast of the close prices of the stock market (Tadawul) for the four sectors in a 60-day period is presented in Figure 14. The forecasting values of the MLP model are presented in Figure 15.

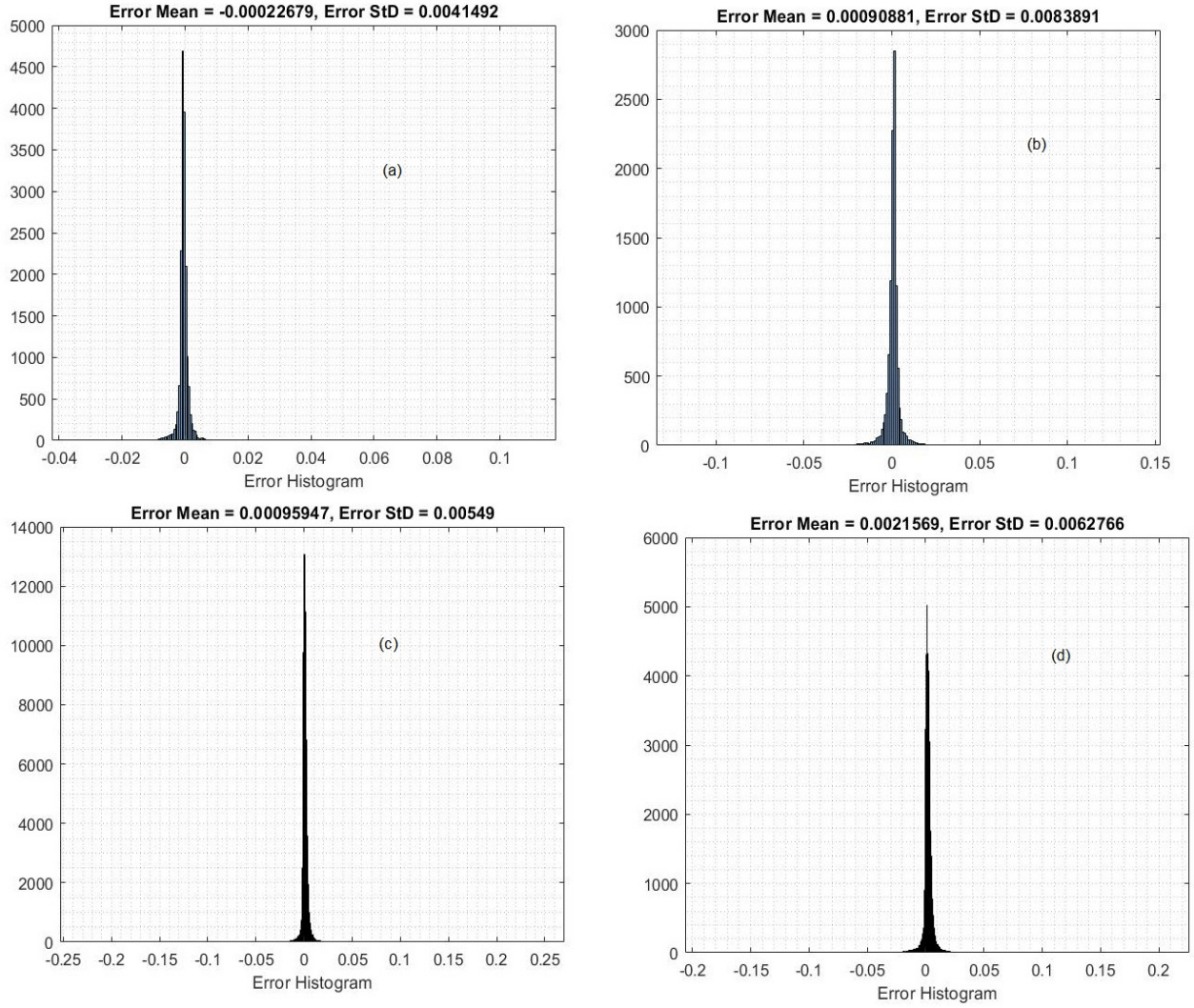


Figure 12. Histogram error of MLP in training testing (**a**) communication (**b**) energy, (**c**) financial and (**d**) industrial sectors.

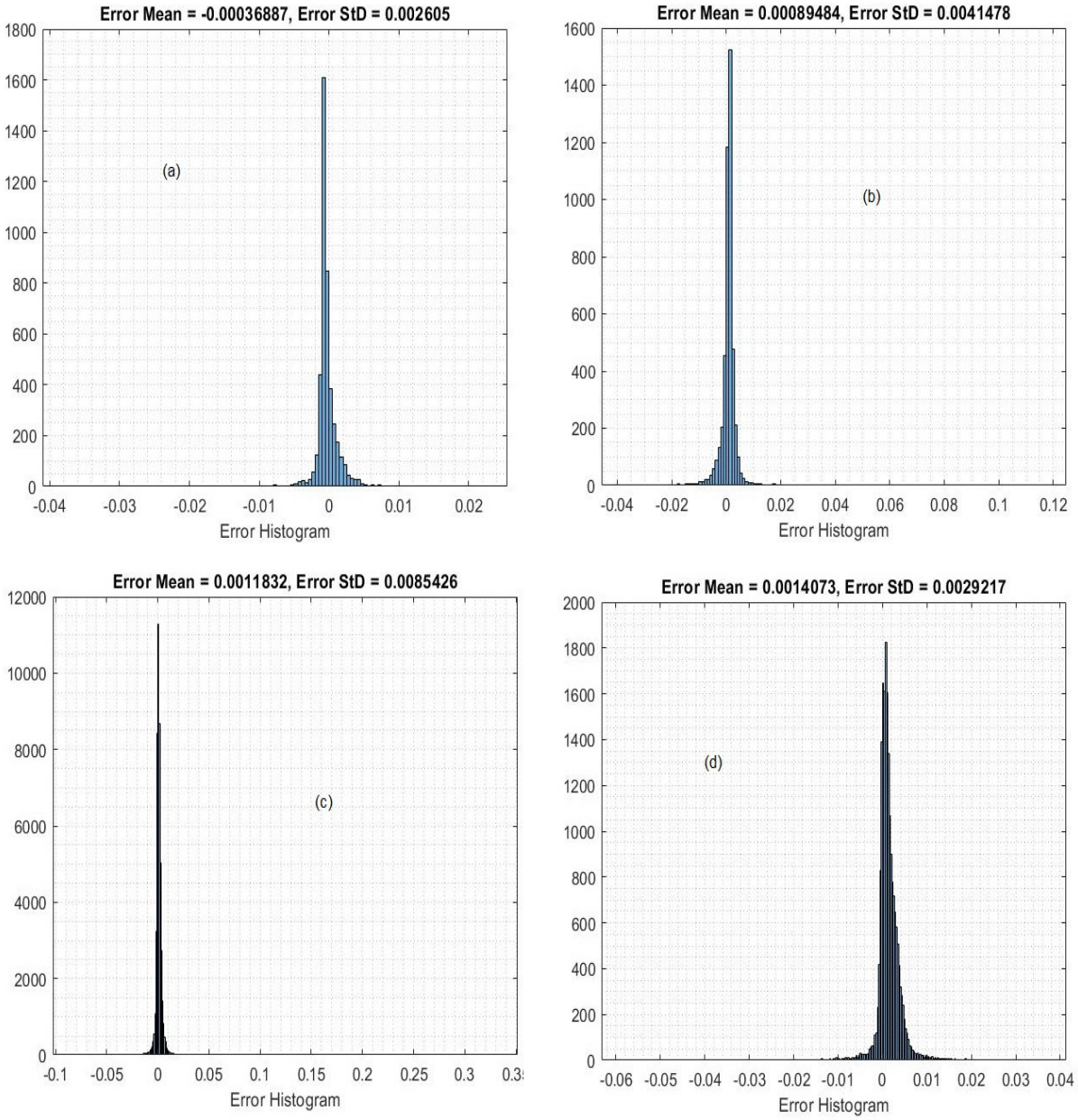


Figure 13. Histogram error of MLP in training testing, (**a**) communication (**b**) energy, (**c**) financial and (**d**) industrial sectors.

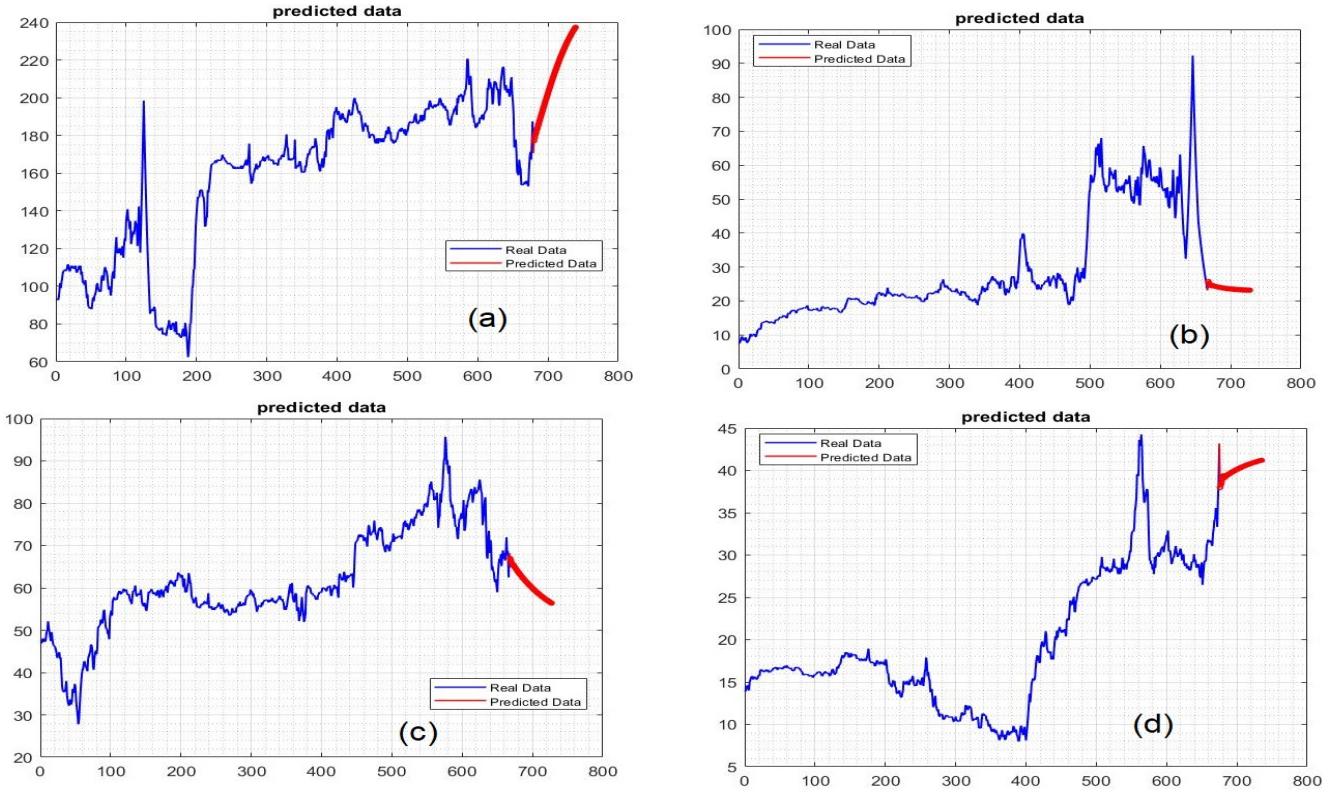


Figure 14. Forecasting values of LSTM model, (**a**) communication (**b**) energy, (**c**) financial and (**d**) industrial sectors.

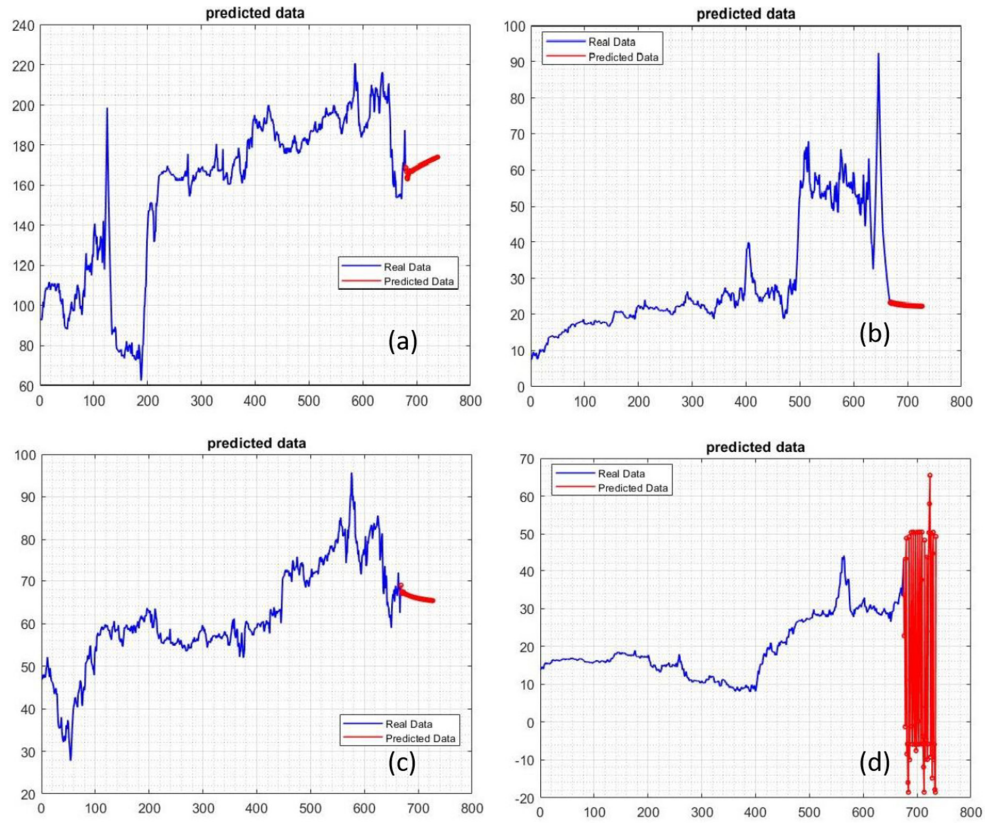


Figure 15. Forecasting values of MLP model, (**a**) communication (**b**) energy, (**c**) financial and (**d**) industrial sectors.

5. Results and Discussion

The vast majority of research that has been conducted on stock prediction with live testing indicates that previously offered approaches may be utilized in real time. These strategies could be successful under carefully monitored conditions. Nevertheless, the live testing of the forecast is a significant obstacle to overcome. Live testing presents a number of problems, such as varying pricing, noise, and unanticipated occurrences. The degree to which the price of an investment swings on the market is referred to as the market's volatility. Uncertainty and inflation are the two primary factors contributing to volatility, and the level of risk increases whenever the market is volatile. There is never a break in the effect that volatility has on our feelings. When the market is turbulent, it is difficult to make accurate predictions about stock values. The use of algorithmic trading is one factor that contributes to the volatile market. It is difficult to evaluate the appropriateness and precision of the many AI algorithms now flooding the market at a steady and rapid speed. The paradoxical character of this field of study is one of the things that makes it so interesting. In simple terms, sharing techniques that create large profits with market rivals will render the procedures obsolete. The trading of top-tier algorithms on the market is hampered and kept private as a result of this mechanism. The method or strategy that lies behind these algorithms is never revealed in written form.

During the course of the previous three decades, developing markets have been responsible for a significant increase in earnings. This is because these economies provide a diverse range of prospects for financial sector investments. The stock market in Saudi Arabia is one of the most robust in the Gulf Council nations, the Middle East, and North Africa. This is not surprising given that Saudi Arabia is often regarded as the country with the highest level of global oil production. The Saudi Stock Market, also known as Tadawul, was given its official start in the year 1984 and has since grown to become one of the most successful rising markets in the Arab world. As a direct response to the growing amount of stock that was traded in Saudi Arabia, the Tadawul All Share Index (TASI), the country's new official stock market index, was formed in the year 2001. (Tadawul 2019). According to the statistics data provided by Tadawul (2019), the number of firms that are listed on the stock exchange has climbed from 163 in 2013 to 199 at the end of 2019, with a total market value of $2,406.78 billion in the United States of America.

The LSTM and MLP artificial intelligence models achieved high accuracy. Figures 16 and 17 show the relationship between the observed values of the stock market and the prediction (output) values from the proposed models. The LSTM has a high R-squared of 100, whereas the LSTM model achieved >99% in the training and testing process. The great performance of the constructed model is indicated by the fact that the experimental values and the forecasting values. In addition, the building model was able to successfully predict the Saudi Stock Market (Tadawul) by obtaining an R^2-value of one and extremely low values. These results further support the robustness of the model.

The return and volatility behavior of the Saudi stock market, which has not been investigated in any thorough way in prior studies, was the focus of the current study, the objective of which was to explore the behavior of the Saudi stock market. We calculated the out-of-sample prediction of this volatility and assessed the effectiveness of AI models in terms of their capacity to capture the peculiarities of the Saudi stock market. Figures 18 and 19 show regression plots of the MLP model for predicting stock marketing values. The MLP has a score of more than R^2 (%) = 99.

Table 6 shows the comparison results of MLP and LSTM model for predicting Saudi Arabia Stock Exchange (Tadawul) with different existing prediction systems. It is observed that the proposed system has achieved superior performance.

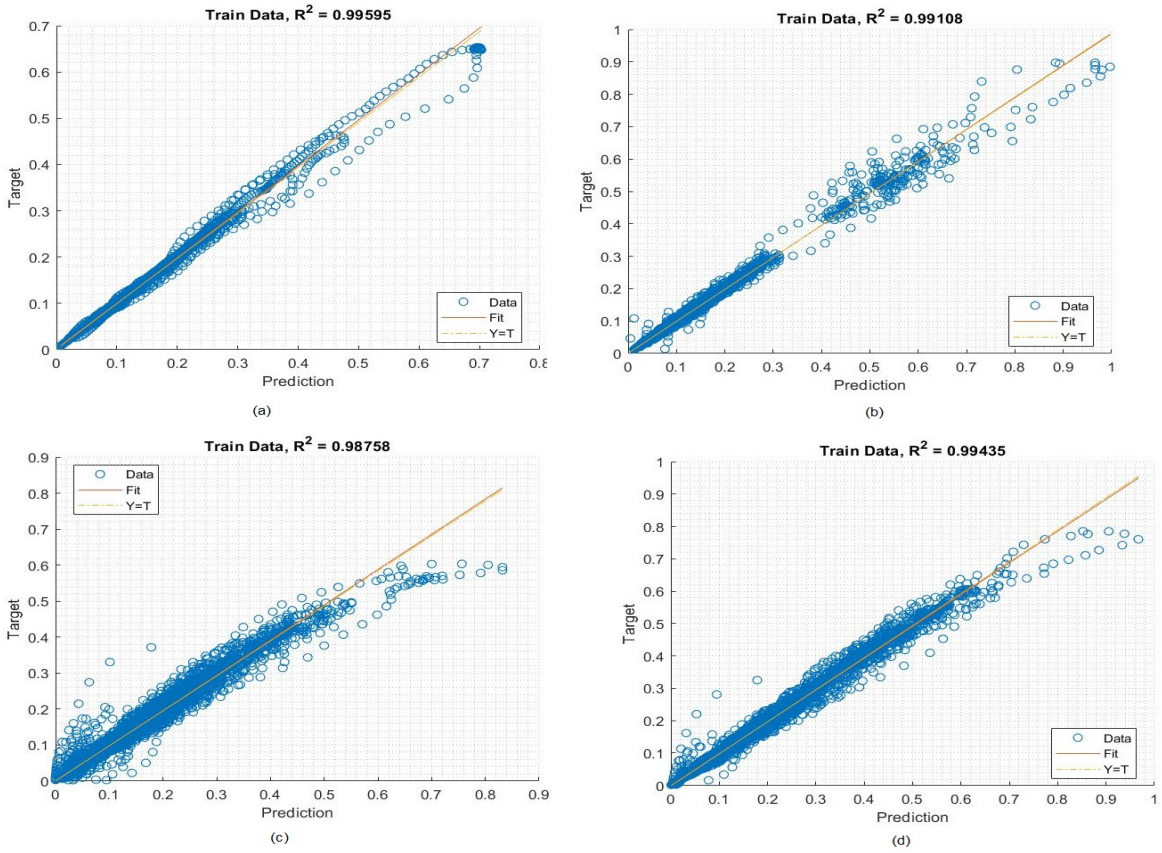


Figure 16. Regression plot of LSTM for predicting stock markets in Saudi Arabia in the training phase, (**a**) communication (**b**) energy, (**c**) financial and (**d**) industrial sectors.

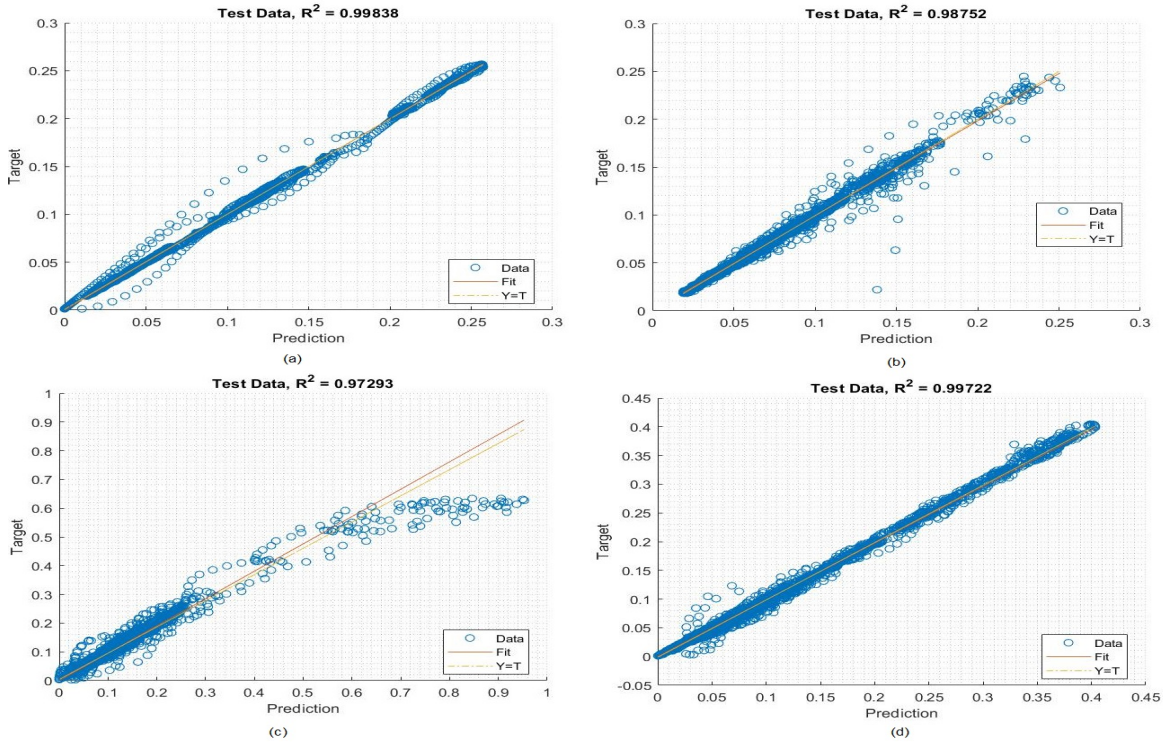


Figure 17. Regression plot of LSTM for predicting stock markets in Saudi Arabia in the testing phase, (**a**) communication (**b**) energy, (**c**) financial and (**d**) industrial sectors.

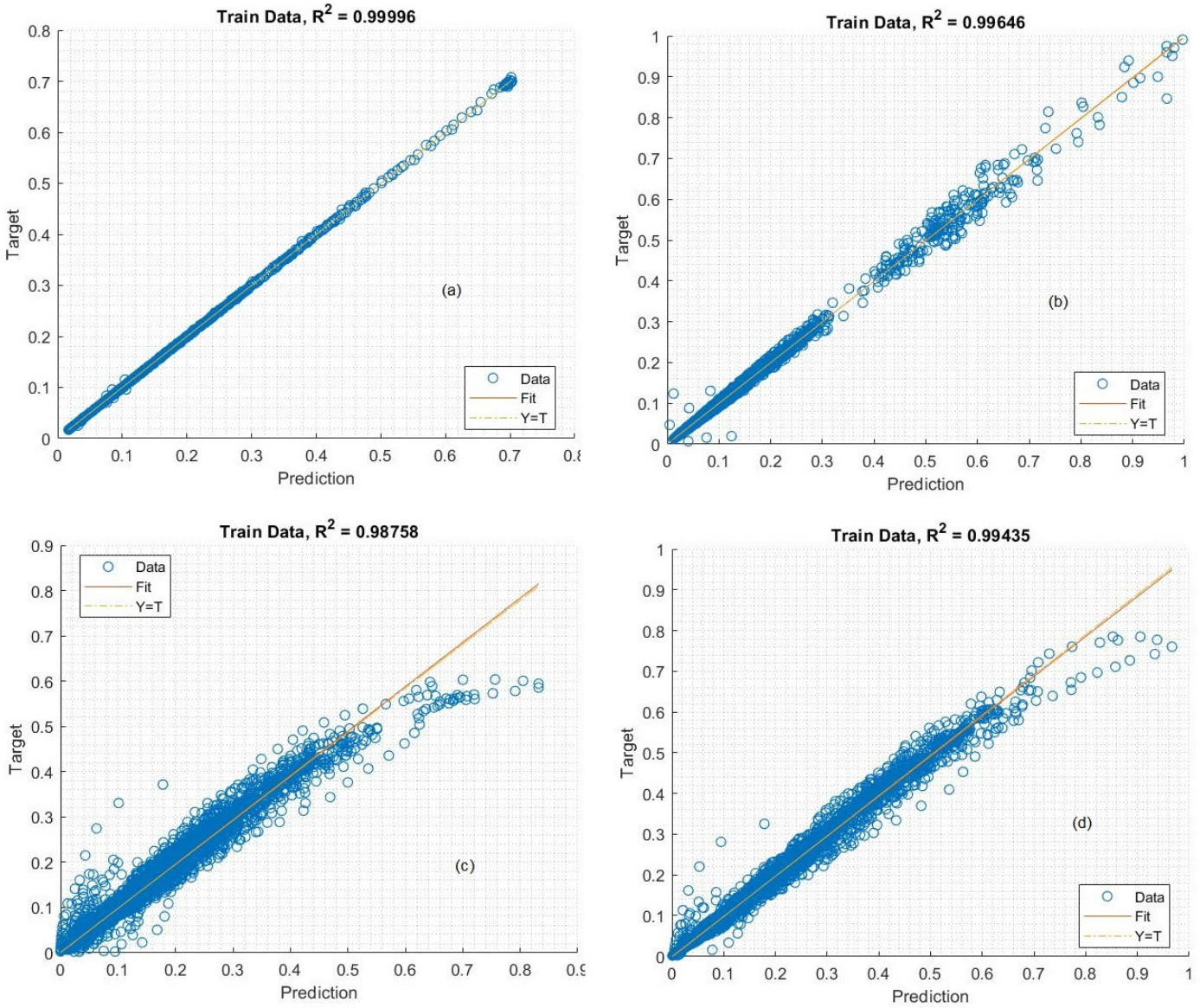


Figure 18. Regression plot of MLP for predicting stock markets in Saudi Arabia in the training phase, (**a**) communication (**b**) energy, (**c**) financial and (**d**) industrial sectors.

Table 6. Results of proposed artificial intelligence model against existing system for predicting Saudi Stock Exchange (Tadawul).

References	Model	Dataset	MSE
Ref. [40]	CR-tree	Tadawul	104.59
Ref. [14]	GARCH (1,1)	Tadawul	0.02018
Proposed model	MLP	Tadawul	0.00012291
Proposed model	LSTM	Tadawul	2.692×100^{-06}

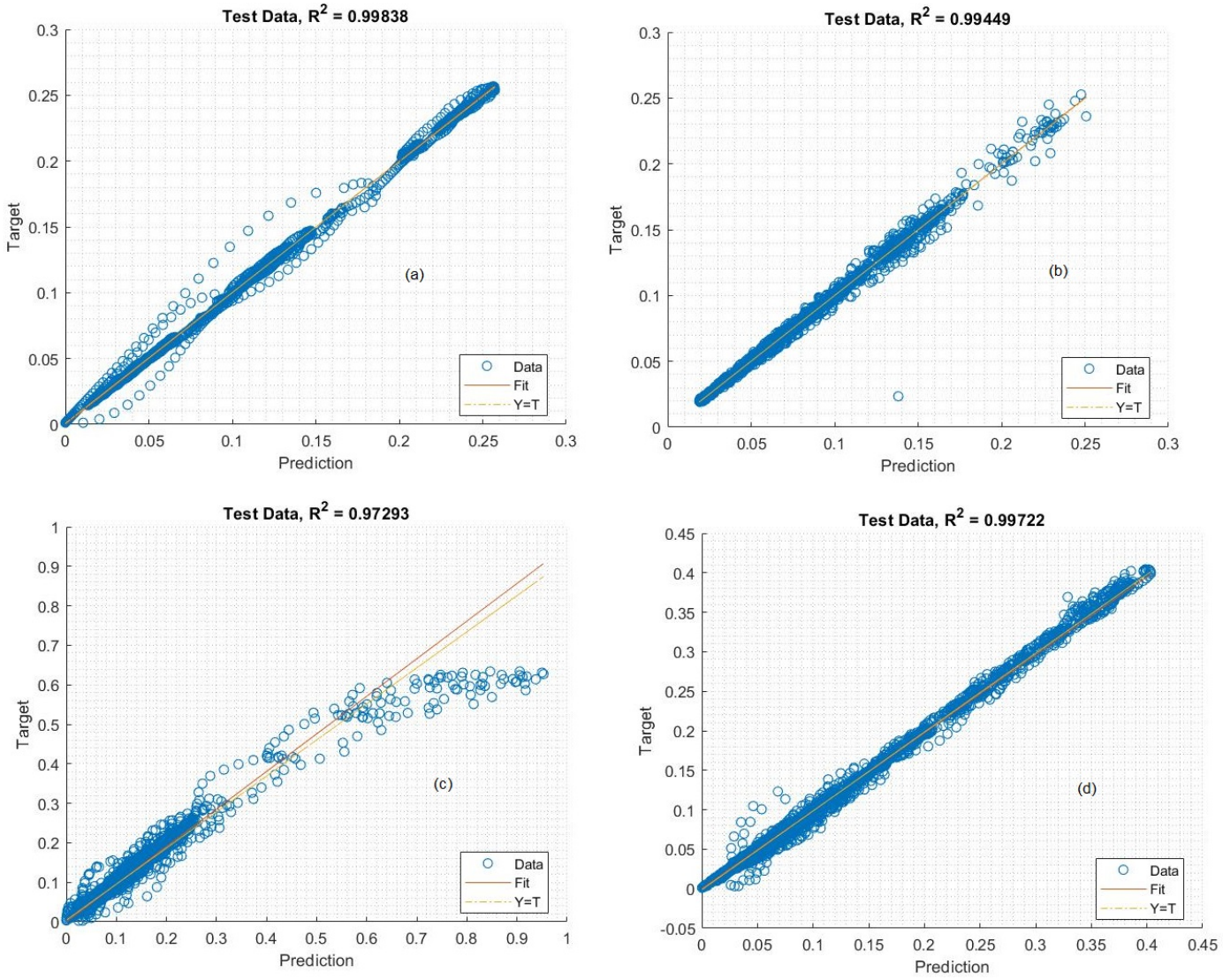


Figure 19. Regression plot of MLP for predicting stock markets in Saudi Arabia in the training phase, (**a**) communication (**b**) energy, (**c**) financial and (**d**) industrial sectors.

6. Conclusions

Forecasting stock prices is difficult owing to the noisy, dynamic, and nonlinear data present in the stock market. Investors are able to increase their profits in the financial market with the help of accurate price predictions for stocks. Identifying trends in the industry is a challenging endeavor. We proposed MLP and LSTM artificial intelligence models to predict the Saudi stock market in four important sectors: communication, energy, financial, and industrial. The proposed models were evaluated using different performance measurement metrics. Our findings have important practical implications for Saudi stock market traders, who may wish to consider these models to better understand risk in the aforementioned sectors as well as the overall riskiness of the Saudi stock market. These findings suggest that AI models may be superior to other models in the areas of risk management strategy implementation and stock pricing model development. Our findings can help policymakers estimate the riskiness of the two indices analyzed. Activities such as modeling and forecasting the performance of a variety of AI models are becoming increasingly crucial for corporations and regulators all around the world. Due to the constraints of the research, we did not make use of the sentiment information that was gleaned through our investigation into the financial markets. The significance of this restriction cannot be overstated in relation to the study. Future research will use either a

more advanced form of deep learning or a hybrid model that takes into account sentiment news analysis in addition to stock price indices. This will allow for the performance of the system that has been suggested to be improved even further.

Author Contributions: T.H.H.A. and A.H.A.-N. resources, T.H.H.A. data curation, A.H.A.-N. and A.H.A.-N.; writing—original draft preparation, T.H.H.A. and A.H.A.-N. writing—review and editing, A.H.A.-N.; visualization, T.H.H.A. and A.H.A.-N. supervision, T.H.H.A.; project administration, T.H.H.A. and A.H.A.-N.; funding acquisition, T.H.H.A. and A.H.A.-N. All authors have read and agreed to the published version of the manuscript.

Funding: This research and the APC were funded by the Saudi Investment Bank Chair for Investment Awareness Studies, the Deanship of Scientific Research, The vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Al Ahsa, Saudi Arabia [Grant No. 143].

Informed Consent Statement: Not applicable.

Data Availability Statement: https://www.saudiexchange.sa/wps/portal/tadawul/home/ (last accessed 26 August 2022).

Conflicts of Interest: The authors declare that they have no conflict of interest.

References

1. Alkhatib, K.; Khazaleh, H.; Alkhazaleh, H.A.; Alsoud, A.R.; Abualigah, L. A New Stock Price Forecasting Method Using Active Deep Learning Approach. *J. Open Innov. Technol. Mark. Complex.* **2022**, *8*, 96. [CrossRef]
2. Koukaras, P.; Nousi, C.; Tjortjis, C. Stock Market Prediction Using Microblogging Sentiment Analysis and Machine Learning. *Telecom* **2022**, *3*, 19. [CrossRef]
3. Kim, J.-M.; Han, H.H.; Kim, S. Forecasting Crude Oil Prices with Major S&P 500 Stock Prices: Deep Learning, Gaussian Process, and Vine Copula. *Axioms* **2022**, *11*, 375. [CrossRef]
4. Lehoczky, J.; Schervish, M. Overview and History of Statistics for Equity Markets. *Annu. Rev. Stat. Its Appl.* **2018**, *5*, 265–288. [CrossRef]
5. Aali-Bujari, A.; Venegas-Martínez, F.; Pérez-Lechuga, G. Impact of the stock market capitalization and the banking spread in growth and development in Latin American: A panel data estimation with System GMM. *Contaduría Y Adm.* **2017**, *62*, 1427–1441. [CrossRef]
6. Rouf, N.; Malik, M.B.; Arif, T.; Sharma, S.; Singh, S.; Aich, S.; Kim, H.-C. Stock Market Prediction Using Machine Learning Techniques: A Decade Survey on Methodologies, Recent Developments, and Future Directions. *Electronics* **2021**, *10*, 2717. [CrossRef]
7. Qian, B.; Rasheed, K. Stock market prediction with multiple classifiers. *Appl. Intell.* **2007**, *26*, 25–33. [CrossRef]
8. Khedr, A.E.; Yaseen, N. Predicting stock market behavior using data mining technique and news sentiment analysis. *Int. J. Intell. Syst. Appl.* **2017**, *9*, 22. [CrossRef]
9. Rousidis, D.; Koukaras, P.; Tjortjis, C. Social media prediction: A literature review. *Multimed. Tools Appl.* **2020**, *79*, 6279–6311. [CrossRef]
10. Gurjar, M.; Naik, P.; Mujumdar, G.; Vaidya, T. Stock market prediction using ANN. *Int. Res. J. Eng. Technol.* **2018**, *5*, 2758–2761.
11. Huang, Y.; Capretz, L.F.; Ho, D. Machine learning for stock prediction based on fundamental analysis. In Proceedings of the 2021 IEEE Symposium Series on Computational Intelligence (SSCI), Orlando, FL, USA, 5–7 December 2021; pp. 1–10.
12. Smailović, J.; Grčar, M.; Lavrač, N.; Žnidaršič, M. Predictive sentiment analysis of tweets: A stock market application. In *International Workshop on Human-Computer Interaction and Knowledge Discovery in Complex, Unstructured, Big Data*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 77–88.
13. Gkikas, D.C.; Theodoridis, P.K.; Beligiannis, G.N. Enhanced Marketing Decision Making for Consumer Behaviour Classification Using Binary Decision Trees and a Genetic Algorithm Wrapper. *Informatics* **2022**, *9*, 45. [CrossRef]
14. Al Rahahleh, N.; Kao, R. Forecasting Volatility: Evidence from the Saudi Stock Market. *J. Risk Financ. Manag.* **2018**, *11*, 84. [CrossRef]
15. Altahtamouni, F.; Masfer, H.; Alyousef, S. Dynamic Linkages among Saudi Market Sectors Indices. *Economies* **2022**, *10*, 16. [CrossRef]
16. Long, J.; Chen, Z.; He, W.; Wu, T.; Ren, J. An integrated framework of deep learning and knowledge graph for prediction of stock price trend: An application in Chinese stock exchange market. *Appl. Soft Comput.* **2020**, *91*, 106205. [CrossRef]
17. Pang, X.; Zhou, Y.; Wang, P.; Lin, W.; Chang, V. An innovative neural network approach for stock market prediction. *J. Supercomput.* **2018**, *11*, 2098–2118. [CrossRef]
18. Kelotra, A.; Pandey, P. Stock Market Prediction Using Optimized Deep-ConvLSTM Model. *Big Data* **2020**, *8*, 5–24. [CrossRef]
19. Bouktif, S.; Fiaz, A.; Awad, M. Augmented Textual Features-Based Stock Market Prediction. *IEEE Access* **2020**, *8*, 40269–40282. [CrossRef]

20. Zhong, X.; Enke, D. Predicting the daily return direction of the stock market using hybrid machine learning algorithms. *Financ. Innov.* **2019**, *5*, 4. [CrossRef]
21. Das, S.R.; Mishra, D.; Rout, M. Stock market prediction using Firefly algorithm with evolutionary framework optimized feature reduction for OSELM method. *Expert Syst. Appl. X* **2019**, *4*, 100016. [CrossRef]
22. Hoseinzade, E.; Haratizadeh, S. CNNpred: CNN-based stock market prediction using a diverse set of variables. *Expert Syst. Appl.* **2019**, *129*, 273–285. [CrossRef]
23. Kumar, K.; Haider, M.T.U. Blended computation of machine learning with the recurrent neural network for intra-day stock market movement prediction using a multi-level classifier. *Int. J. Comput. Appl.* **2019**, *43*, 733–749. [CrossRef]
24. Hung, H.; Shin, K.-S. Genetic algorithm-optimized multi-channel convolutional neural network for stock market prediction. *Neural Comput. Appl.* **2020**, *2*, 7897–7914.
25. Sim, H.S.; Kim, H.I.; Ahn, J.J. Is deep learning for image recognition applicable to stock market prediction? *Complexity* **2019**, *2019*, 4324878. [CrossRef]
26. Wen, M.; Li, P.; Zhang, L.; Chen, Y. Stock Market Trend Prediction Using High-Order Information of Time Series. *IEEE Access* **2019**, *7*, 28299–28308. [CrossRef]
27. Prachyachuwong, K.; Vateekul, P. Stock Trend Prediction Using Deep Learning Approach on Technical Indicator and Industrial Specific Information. *Information* **2021**, *12*, 250. [CrossRef]
28. Srivinay; Manujakshi, B.C.; Kabadi, M.G.; Naik, N. A Hybrid Stock Price Prediction Model Based on PRE and Deep Neural Network. *Data* **2022**, *7*, 51. [CrossRef]
29. Ratchagit, M.; Xu, H. A Two-Delay Combination Model for Stock Price Prediction. *Mathematics* **2022**, *10*, 3447. [CrossRef]
30. Aldhyani, T.H.H.; Alzahrani, A. Framework for Predicting and Modeling Stock Market Prices Based on Deep Learning Algorithms. *Electronics* **2022**, *11*, 3149. [CrossRef]
31. Ammer, M.A.; Aldhyani, T.H.H. Deep Learning Algorithm to Predict Cryptocurrency Fluctuation Prices: Increasing Investment Awareness. *Electronics* **2022**, *11*, 2349. [CrossRef]
32. Orrù, P.F.; Zoccheddu, A.; Sassu, L.; Mattia, C.; Cozza, R.; Arena, S. Machine Learning Approach Using MLP and SVM Algorithms for the Fault Prediction of a Centrifugal Pump in the Oil and Gas Industry. *Sustainability* **2020**, *12*, 4776. [CrossRef]
33. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning*; Springer: Berlin, Germany, 2017.
34. Aldhyani, T.H.H.; Alkahtani, H. A Bidirectional Long Short-Term Memory Model Algorithm for Predicting COVID-19 in Gulf Countries. *Life* **2021**, *11*, 1118. [CrossRef]
35. Iyyappan, M.; Ahmad, S.; Jha, S.; Alam, A.; Yaseen, M.; Abdeljaber, H.A.M. A Novel AI-Based Stock Market Prediction Using Machine Learning Algorithm. *Sci. Program.* **2022**, *2022*, 4808088. [CrossRef]
36. Ghosh, P.; Neufeld, A.; Sahoo, J.K. Forecasting directional movements of stock prices for intraday trading using LSTM and random forests. *Financ. Res. Lett.* **2022**, *46*, 102280. [CrossRef]
37. Maiti, M.; Vyklyuk, Y.; Vuković, D. Cryptocurrencies Co-movement Forecasting with Neural Networks. *Internet Technol. Lett.* **2020**, *3*, e157. [CrossRef]
38. Maiti, M.; Vukovic, D.; Vyklyuk, Y.; Grubisic, Z. BRICS Capital Markets Co-Movement Analysis and Forecasting. *Risks* **2022**, *10*, 88. [CrossRef]
39. Sahoo, J.P.; Tripathy, A.K.; Mohanty, M.; Li, K.C.; Nayak, A.K. *Advances in Distributed Computing and Machine Learning*; Lecture Notes in Networks and Systems; Springer: Singapore, 2022; p. 302. [CrossRef]
40. AL-Najjar, D.; AL-Najjar, H.; Al-Rousan, N.; Assous, H.F. Developing Machine Learning Techniques to Investigate the Impact of Air Quality Indices on Tadawul Exchange Index. *Complexity* **2022**, *2022*, 4079524. [CrossRef]