An Efficient Algorithm for Stock Market Prediction Using Attention Mechanism

Zena Kreem Minsoor[1,*], Ali Yakoob Al-Sultan[1]
[1]Department of Computer, College of Science for Women, University of Babylon, Babylon, Iraq
Emails: scw734.zynh.kareem@student.uobablyon.edu.iq; ali.alsultan@uobabylon.edu.iq

Abstract

Forecasting the stock market is a significant challenge in the financial industry due to its time series' complicated, noisy, chaotic, dynamic, volatile, and non-parametric nature. Nevertheless, due to computer advancements, an intelligent model can assist investors and expert analysts mitigate the risk associated with their investments. In recent years, substantial research has been conducted on deep learning models. Many studies have investigated using these techniques to anticipate stock values by analyzing historical data and technical indications. However, since the goal is to create predictions for the financial market, validating the model using profitability indicators and model performance is crucial. This article incorporates the attention mechanism model, incorporating attention from both feature and time perspectives. Utilize artificial neural networks. This approach addresses issues in time series prediction. The issue is the varying degrees of influence that many input features have on the target sequence. To tackle this, the method utilizes a feature attention mechanism to obtain the weights of distinct input features. An enhanced feature association relationship is achieved, whereas the data before and following the sequence exhibit a significant time correlation. An attention technique is employed to address this issue, allowing for the acquisition of weights at various time intervals to enhance robustness and temporal dependence. The system is applied to the three global SMs (TESLA, S&P500, and NASDAQ) datasets, the best enhancement results are 99% in Acc, and the better results improvement to minimize error in MSE, MAPE, and RMSE are 0.004, 0.004 and 0.01 respectively.

Keywords: Stock market Prediction; Deep learning; Neural networks; Sentiment analysis; Reinforcement learning; Financial markets; Trading strategies

1. Introduction

Financial evaluation and investment decisions heavily rely on analyzing stock market trends. Stock market trend analysis involves time series regression. The complexity relies on various non-numeric factors such as socio-economic situations, political circumstances, financial crises, trade conflicts, global economic downturns, and public perceptions of a corporation [1] .As a result, stock trends can show inconsistent patterns compared to historical data [2]. Economic conditions can significantly influence stock trends. Stock values typically decrease during a recession due to investors' increased risk aversion.

During periods of economic expansion, the value of stocks generally rises as investors become more optimistic about the future[3]. The performance of individual companies can also impact stock patterns. Positive events, such as higher profits and the introduction of new products, can lead to a significant increase in stock prices. Conversely, negative events, such as a decrease in sales or the recall of products, can have the opposite effect on stock prices[4].. Stock trends can be influenced by global events such as natural disasters, geopolitical tensions, or pandemics[5]. These occurrences can result in market ambiguity, prompting investors to adopt a more cautious approach towards taking risks. It is crucial to acknowledge that stock trends inherently exhibit volatility. In addition, it can be influenced by a variety of factors[6]. The task is significant, such that even a minor enhancement in stock market forecasting can result in substantial profits. The stock market can go in two directions: upwards (when stock prices rise) or downwards (when stock prices decrease) [5]. Typically, four methods exist for analyzing the stock market's direction [7]. Fundamental analysis which is the most wide-spread and popular way to evaluate the stock market

DOI: https://doi.org/10.54216/JCIM.140226
Received: January 30, 2024 Revised: April 10, 2024 Accepted: July 13, 2024

consists of analysis of macroeconomic situation, reports and plans of company. The second of the methods used most is technical analysis [5]. This strategy is by looking at the stock market chart and comparing them with the previous prices so as to predict the direction where the price would be going. [8]. Another highly advanced technology level is based on Machine Learning (ML), which can often analyze the market without too much human intervention [9]. Machine learning models work with historical data to find the trends apart from different things that might measure future market price movements. The Sentiment-based approach is a specific strategy that employs other persons' sentiments expressed in social media activity or financial news reports to assess stock market prices [8,10]. Different global researchers showed an intense interest in stock market prediction. Associated with this are papers published lately demonstrating the high potential of ML models to predict stock prices. While the models have several designs, the Network Artificial Neural (ANN) stands out in this category of [11], Decision Tree (DT) [12], Support Vector Machine (SVM) [13], Random Forest (RF) [14] ,and Long Short-Term Memory networks[15] Different types of machine learning algorithms present with diverse strengths in analyzing the same historical data. Performance depends on the nature of the data itself and the time duration of the term that the historical data is at hand. An increasingly developing area of financial and statistical studies which use different machine learning algorithms to reconstruct financial time series data for performing stock market price forecasts and performance evaluation [13].

Most machine learning-based systems have shown a primary weakness in their empirical outcomes. The ML models' performance was solely assessed based on their classification capability. While a valuable measure for evaluating machine learning models, more is needed to assess their effectiveness in stock market prediction. The following part explains the suggested method regarding feature engineering, data filtering, and model training.

2. Related Work

Several academics have researched the stock market to develop ideas for its functioning. The stock market known for its unpredictable nature and complex workings, for this problem, deep learning and machine learning techniques have been used. Most of them are described:

Patil et al. in 2020 employed deep learning models, including ARIMA, RNN, and Facebook Prophet. The researchers aimed to improve prediction accuracy by collecting substantial time series data and examining its correlation with relevant news articles. The dataset they compiled encompassed daily stock prices for S&P500 companies. RNN outperformed other models, achieving a MAPE value of 2.03 [1] . Khan et al. in 2023 models of machine learning such as RANDOM FOREST, XG Boost, ADA Boost, and ANN are used to predict the futurity of TESLA Inc. They implemented a simulation model of financial market to evaluate the risk, maximum drawdown, and profits associated with each model of machine learning. The proposed approach enhanced the profitability of the stock market and mitigated the risks linked to each machine-learning model, leading to an accuracy rate of 91.27% [9]. Chung and Shin in 2020 they assessed several models, such as the MLP neural network, the DAN2 structure, and the hybrid network that combines or uses the autoregressive conditional heteroscedasticity (GARCH) to bring forth new inputs. The evaluation of these models was performed using (MSE) and (MAD), using real daily exchange rate values of the NASDAQ Stock Exchange index. The findings indicate that the traditional artificial neural network (ANN) model, particularly the multilayer perceptron (MLP), outperforms the other models in relation to mean squared error (MSE), attaining a value of 1472.278 [11]. Chandrika and Srinivasan in 2021 Evaluated multiple models, such as the multilayer perceptron (MLP), and Utilized the Artificial Neural Network algorithm of deep learning to forecast stock indices that align with the direction of price fluctuations. The dataset utilized for numerous stock indexes, such as NIFTY 50, S&P 500, New York Stock Index, Korean Stock Index, Dow Jones Index, and Shanghai Stock Index. Many measurement measures were employed to assess the model's performance, including accuracy, precision, recall, and F1 score [16]. Zulqarnain et al. in 2020 employed combine Convolutional Neural Networks (CNN) with recurrent neural networks to elevate the signal prediction process. In particular, the model focuses on the direction of such signals based on the past time series data. Three different data sets were used to determine the prototype: HSI (the Stock Indexes of Hang Seng), DAX (the European Index), and S&P (the stock index of S&P). Consequently, the best ERR results: 56.2% for the HSI dataset, 56.1% for the DAX dataset, and 56.3% for the S&P 500. This is the result of using the proposed GRU-CNN model [17]. Shams and Shams in 2020They investigated the comparative performance of seven machine learning algorithms on four different stock returns indices, which are NASDAQ index, NYSE index, NIKKEI index, and FTSE. The main aim of their study was to assess the effectiveness of these algorithms in reducing investment risks. The results indicated that using leaked datasets, the Random Forest and Bagging algorithms performed exceptionally well [18]. Li in 2024 employed artificial intelligence tools for stock market trends forecasting by comprehensively analyzing. The main goal of the study is to assess and compare the predictive capabilities of four different methods: linear regression, support vector regression, random forest regression, and LSTM. These approaches are especially employed to predict Tesla's stock values. Linear regression

DOI: https://doi.org/10.54216/JCIM.140226
Received: January 30, 2024 Revised: April 10, 2024 Accepted: July 13, 2024

stands out as the most outstanding performer among the four approaches, reaching a notable R-squared value of 0.85 [19]. Bathla in 2020 Applied Long Short-Term Memory (LSTM) to predict fluctuations in stock values. In the experiment, the two methods LSTM and SVR were compared using different stocks index data including S&P 500, NYSE, NSE, BSE, NASDAQ and DJIA. The result of the test demonstrates that LSTM, in fact, is able to reach a significant level of accuracy [20]. Focused on optimizing the parameters of the Convolutional Neural Network (CNN) model and incorporating well-known technical indicators as effective input variables for stock price forecasting. The study utilized the S&P 500 index dataset. The experimental findings emphasized the suitability of the CNN model for constructing precise stock prediction models [21]. Gurav and Sidnal in 2018 present a new modified backpropagation neural network (MBNN) as an adaptive stock forecasting model for worldwide markets. The study encompasses trials conducted on diverse marketplaces, including the S&P 500 index, NASDAQ, and NSE. The results suggest that the MBNN model exhibits favorable performance in all of the analyzed markets [22]. Suryani and Buani in 2020 is employed a neural network model to forecast the value of stock shares using the historical TESLA dataset. They utilized a multilayer feed-forward network to address the problem. The results indicated that, in terms of predicting changes in stock value direction, none of the methods performed better than the backpropagation algorithm, achieving an accuracy of 94% [23].

3. Attention mechanism

The attention mechanism was initially employed in machine translation and has since become a fundamental technique in neural networks. The attention mechanism is an essential component of artificial intelligence, performing a critical function in various domains, including natural language processing, time series prediction, speech recognition, and computer science [24].

3.1. General form

When seeing a scene in our everyday lives, our attention is naturally drawn to the distinctive areas, which we then process rapidly. The aforementioned procedure can be expressed as the equation (1) represents the relationship between the variable "Attention" and the functions "g(x)" and "x". Here, the function g(x) can be used to denote the generation of attention, which refers to the act of focusing on the discriminative regions. The expression f(g(x),x) denotes the operation of applying attention g(x) on input x, where attention g(x) focuses on important regions and extracts relevant information. Based on the given definition, it can be observed that nearly all attention mechanisms now in existence may be expressed using the aforementioned language. Illustrations. The expressions g(x) and f(g(x),x) can be formulated for the purpose of self-attention [25].

$$Q,K,V=Linear(x) \tag{1}$$

$$g(x)=Softmax(QK) \tag{2}$$

$$f(g(x),x)=g(x)V \tag{3}$$

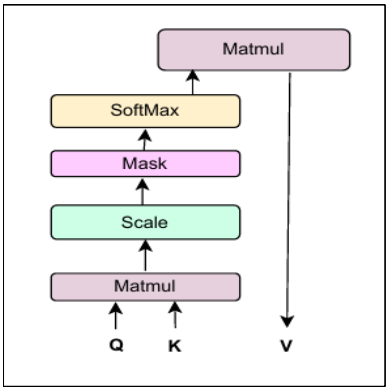


Figure 1. General form for self-attention

DOI: https://doi.org/10.54216/JCIM.140226
Received: January 30, 2024 Revised: April 10, 2024 Accepted: July 13, 2024

4. Methodology

The proposed system includes main stages the first stage is preprocessing dataset that includes transform nominal to numeric , drop duplicate value , and interpolation. The second stage involves TIS generation. The third stage is building prediction model, which is implemented by ATTENTION Model, final stage is the evaluation of result based on test data.

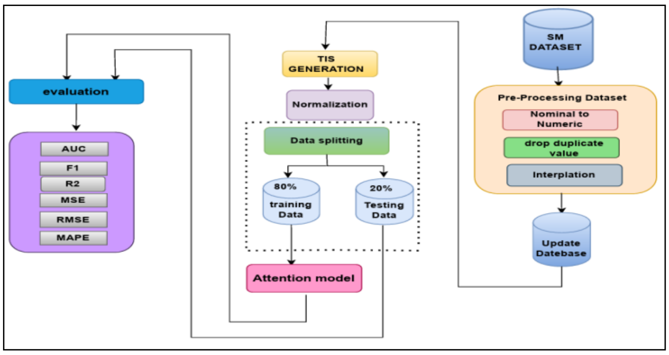


Figure 2. The proposed system architecture

4.1. Description of stock market dataset

The dataset for our study is provided from Yahoo finance's stock historical data. The study based on experimental tests with historical data set from three global stock markets such as (S&P500, Tesla, and NASDAQ). The period of S&P500 is selected from 18/11/1999 to 12/07/2022 , 5697 instances and NASDAQ from 18/11/1999 to 31/03/2020 , 5124 instances as shown visually in figure (4.1) and figure (4.2) respectively, while the period of Tesla is selected from 29/06/2010 to 03/02/2020 , 2416 instances. The historical data is collected in daily basis (Open, High, low, Close, Volume, Adj Close).

•Open: price at the opening of the particular day.

•High: resulted in an all-time high level recorded for the day

•Low: closed session of lowest traded stock of the day.

•Close: the closing price of the company's stock at the end of the day that is being talke about day

•Adjacent close: It denotes the value remaining after stock dividends distribution

•Volume: represents a combination of share and contract exchanges between the buyer and seller and sellers.

After downloading data set, datasets have saved in three tables and they have added to local database.

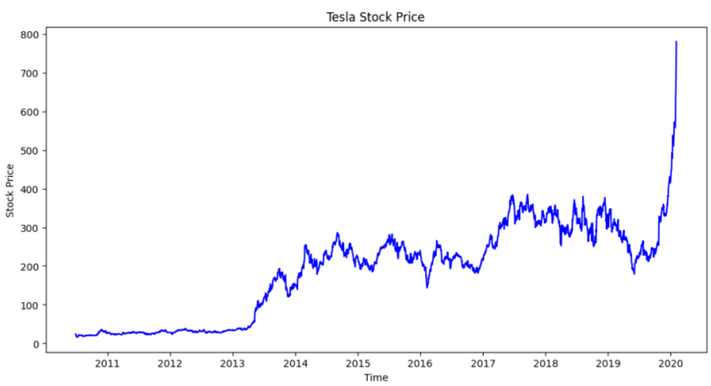


Figure 3. The period time TESLA stock market

DOI: https://doi.org/10.54216/JCIM.140226
Received: January 30, 2024 Revised: April 10, 2024 Accepted: July 13, 2024

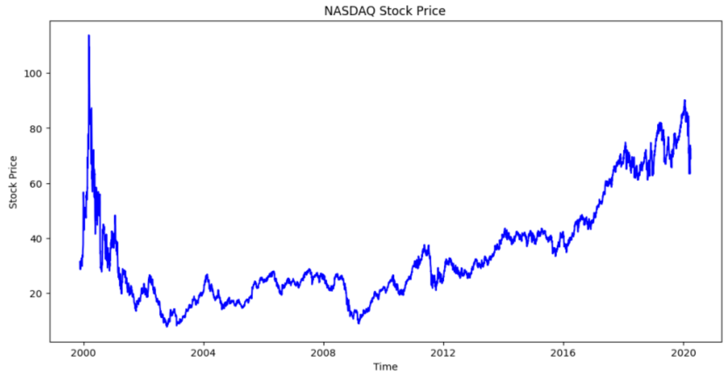


Figure 4. The period time of NASDAQ stock market

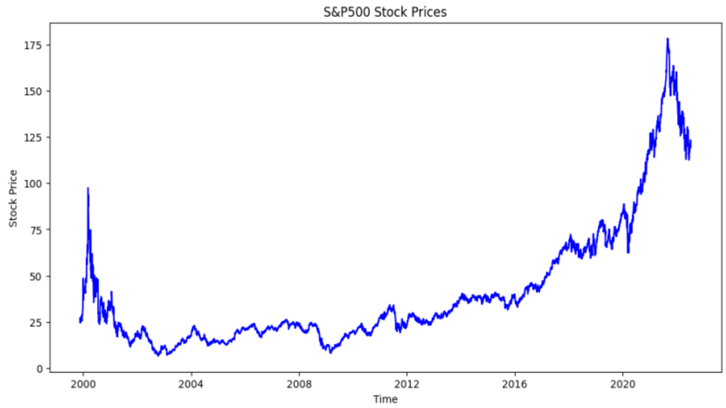


Figure 5. The period time of S&P500 stock market

Table 1: Summarization of dataset

DATASET	Period time	NO. of Features	No. of Instances
Tesla	**29/06/2010 to 03/02/2020**	**7**	**2416**
S&P500	**18/11/1999 to 12/07/2022**	**9**	**5697**
Nasdaq	**18/11/1999 to 31/03/2020**	**7**	**5124**

4.2. Preprocessing stage

The data is processed and prepared into prediction stage because actual datasets like SM dataset might have some unsuitable structure. This stage include:

4.2.1. Nominal to numeric

In this step, transform attributes to numeric form, which is performed on categorical attributes (Date). This process converts the data type from date into text after that to numeric by merging the date (day, month, and year) and deleting (/) separator digits. Finally, the text data converts into numeric.

DOI: https://doi.org/10.54216/JCIM.140226
Received: January 30, 2024 Revised: April 10, 2024 Accepted: July 13, 2024

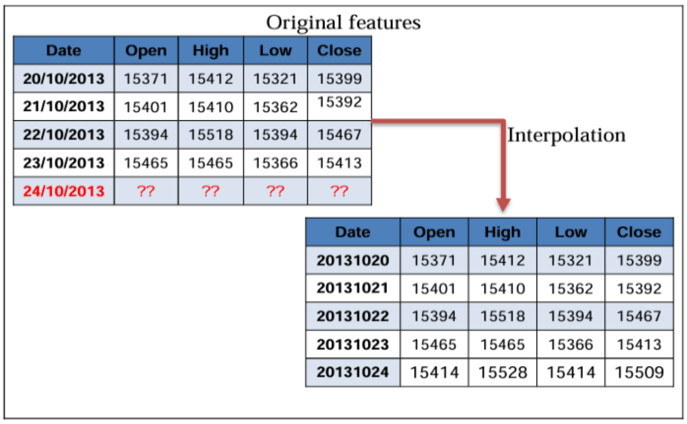
4.2.2 Duplicate rows

Are found in the data, they are all removed, with the exception of the first instance of each duplicate row. The updated information can be saved in a new variable or substituted in the existing one.

4.2.3 Interpolation

It is often observed in real-world scenarios to have missing values in a dataset. Imputation using the median entails replacing the missing data with a computed value [26] .

$$\text{MED} = \frac{P_H(t)+P_L(t)}{2} \qquad (4)$$



Date	Open	High	Low	Close
20/10/2013	15371	15412	15321	15399
21/10/2013	15401	15410	15362	15392
22/10/2013	15394	15518	15394	15467
23/10/2013	15465	15465	15366	15413
24/10/2013	??	??	??	??

Date	Open	High	Low	Close
20131020	15371	15412	15321	15399
20131021	15401	15410	15362	15392
20131022	15394	15518	15394	15467
20131023	15465	15465	15366	15413
20131024	15414	15528	15414	15509

Figure 6. The process interpolation

4.3. Technical indicators generation

TIs are used as inputs to the model to improve the accuracy [13] [27].

- Day of the Week: Some stocks may exhibit specific behaviors on certain days.
- Month: Certain months might show seasonal trends.
- Lagged Prices: The prices of the stock over previous days.
- Standard deviation (SD) is a mathematical term that quantifies the amount of variation or dispersion around an average value. Standard deviation (SD) is a statistical metric used to quantify volatility. The standard deviation is employed to quantify anticipated risk and ascertain the significance of specific price fluctuations. The calculation of SD is calculated using a particular formula:

$$\text{Standard Deviation } SD = \sqrt{\frac{(Pc(t)-SMA_t(n))^2+(Pc(t-1)-SMA_{t-1}(n))^2+\cdots(Pc(t-n+1)-SMA_{t-n+1}(n))^2}{n}} \qquad (5)$$

Relative Strength Index (RSI): Quantifies the velocity and magnitude of price fluctuations [13].

$$\text{RSI} = 100 - \frac{100}{1+(\sum_{i=0}^{n-1}\text{UP}_{i-1}/n)/(\sum_{i=0}^{n-1}\text{DW}_{i-1}/n)} \qquad (6)$$

- Moving Average Convergence Divergence (MACD): Compares short-term and long-term trends to indicate potential buying or selling opportunities. [28]

MACD Line (MACD) = EMA (Shorter Time Period) – EMA (Longer Time Period) (7)

Signal Line (Signal) = EMA (MACD Line, Signal Time Period)

- Bollinger Bands: Measures volatility and identifies overbought or oversold conditions.
- Standard Deviation: Quantifies the amount of historical volatility in a stock's price.

DOI: https://doi.org/10.54216/JCIM.140226
Received: January 30, 2024 Revised: April 10, 2024 Accepted: July 13, 2024

- Price Rate of Change (ROC): Measures the percentage difference in the current closing price and n number of days in the past.[28][13].
- ROC is computed by taking the difference between current closing price from n days closing prices ago.

$$\text{ROC}=\frac{P_c(t)-P_c(t-n)}{p_c(t-n)} \quad (7)$$

Where: Pc (t): Current Price is current price of the asset and Pc(t-n): Price n means the past price n periods ago.

Exponential moving average (EMA): is a mathematical method that assigns greater importance to recent data points, resulting in a higher level of significance in the overall computation. This technical indicator is widely popular. The EMA is employed to evaluate the trajectory of a financial asset's movement. To accomplish this goal, EMA utilizes a method known as smoothing, which entails removing the unpredictable variations in price by computing the average price over a designated time period, represented as m. The Exponential Moving Average (EMA) is a lagging indicator that relies on historical price data. The EMA lacks the ability to predict upcoming trends, but it is capable of confirming the direction of an already established trend. The Exponential Moving Average (EMA) is a calculation that gives more weight to recent values. It is determined using the following formula [29]:

$$\text{EMA (t)} =(\text{Close(t)} - \text{EMA(t-1)}) * \frac{2}{(\text{n}+1)+\text{EMA}(\text{t}-1)} \quad (8)$$

where EMA[t] is the Exponential Moving Average at time(t), Close[t] is the closing price at time [t], EMA[t-1] is the Exponential Moving Average at time[t-1], and n is the number of periods

- % R (WILLIAMS %R) the current price relationship with the high and low prices over the preceding is indicated using the Williams %R indicator. Williams %R is calculated as follows [30]: (10

$$\%R = \frac{P_H(n)-P_c(t)}{P_H(n)-P_L(n)} \quad (9)$$

- CCI (Commodity Channel Index) is a technical indicator used in technical analysis to assess the strength of a trend and identify overbought or oversold levels in the market [11][27].

$$\text{CCI} = \frac{P_C\,(t)-SMAP_C\,(n)}{(0.015)*Mean\ Deviation} \quad (10)$$

Where: The Typical Price is calculated as the mean of the high, low and close prices which is exactly for a specific time period .The Simple Moving Average (SMA) is the average of the Typical Prices over a specified number of periods. The Mean Deviation is the average of the absolute differences between each Typical Price and the corresponding SMA over the specified number of periods.

- Simple Moving Average (SMA)

The Simple Moving Average (SMA) can be determined by computing the mean price over a specified time period. This operation is performed repeatedly on a daily basis in order to create its own time series. The formula obtained is [31]:

$$SMA = \frac{p_c(t)+p_c(t-1)+\cdots+p_c(t-n-1)}{n} \quad (11)$$

4.4. MIN-MAX normalization

The input data is transformed via linear scaling to a range of values between 0 and 1, resulting in a normalised multi-dimensional time-series. The act of scaling this data enhances the efficiency of the training method for the neural network model. The value of x0 can be stated using linear scaling.[32]:

$$\upsilon{*}= \frac{(\upsilon-\mu in)}{(\mu ax-\mu\text{in})} \quad (12)$$

where :υ *: is the new value ; υ: is the old value , μin: the minimum value The data has been normalized using a min–max scaler , μax: the maximum value.

4.5. Data splitting

The correct measurement of data mining techniques demand the use of test data that were not yet intended during the training stage previously. Through Cross Validation (CV), a section of the SM dataset is divided into

DOI: https://doi.org/10.54216/JCIM.140226
Received: January 30, 2024 Revised: April 10, 2024 Accepted: July 13, 2024

two segments on every fold. The data is split into two subsets: the training set is 80% of the data and here the validation set is 20% of the training data.

4.6. Predictor stage

4.6.1. Class labeling

This step is essential for the categorization phase. The class attribute is determined by the utilization of two distinct formulae. The volatility of a stock is determined by subtracting the previous closing price from the current closing price. If the result is positive, it is classified as "high volatility". If the result is negative, it is classified as "low volatility". Otherwise, it is classified as "stable".

$$ci = \begin{cases} x & if\ close_i > close_{i-1} \\ 0 & if\ close_i = close_{i-1} \\ -1 & if\ close_i < close_{i-1} \end{cases}$$

Where ci=Δ close price, i from 1 to total historical data.

4.6.2. Model architecture

The model comprises multiple layers. The initial layer is the Attention layer, which enables the model to concentrate on certain segments of the input data. The Attention is in this model, consisting of scaling (scaling=True hence dropping rate is 0.3). Next step is a Transmission layer, which is composed of a densely connected network composed of 256 neurons and using ReLU as an activation function. ReLU is a non-linear function that has x as the output when the input value x >0 and reaches 0 when x is < or = 0. The formula for ReLU activation function is shown below

$$\int(x) = \max(0, x) \quad (13)$$

Next, there is a dropout layer implemented with a dropout rate of 0.3 in order to mitigate the issue of overfitting. Then the data are sent to the fully connected neuron that has 128 units and the ReLU activation function. As result, another dropout layer with dropout rate 0.3 come next.

Finally, the data is transmitted through the output layer, which contains the various number of units equal to the total number of the unique classes in the problem. Softmax actrivation function is used to derive the results into probabilities. Another notable point is that the training and learning of the model's weight parameter in various deep learning models is being facilitated the utilization of the adaptive moment estimation (Adam) optimizer which has been proven to be this function from empirical perspective. Adam provides guru services in fine-tuning weight factors, particularly for those who would like to make forecasts that involve turbulence and instability. The model initialize the call function, where the shape of the input data by reshaped to be a two-dimensional and get passed to the attention layer. Afterwards, the data is reshaped to be flattened into the fully Connected Layers with dropout layers. To complete the procedure, we then pass the data through the output layer to obtain the required results.

The Self-Attention Structure consist:

Step1: Prepare the inputs and the Q, K, V matrices

Inputs: The basic inputs

q: The query matrix representing the attention query.

k: The key matrix representing the keys for computing attention.

v: The value matrix representing the value for attention computation.

Step 2: Assign values to the matrices' weights. The attention weights (attention weights) in the build () method are generated using the add weight () function and initialized with random normal. The trainable attribute is assigned a value of True in order to guarantee that the attention weights are modified during the training process.
Step 3: Implement scaling when the use scale parameter is set to True, the Q, K, and V matrices are multiplied by the attention weights in order to generate scaled values.
Step 4: Compute attentiveness scores the attention scores are computed by performing matrix multiplication between the query matrix (Q) and the inputs, utilizing the transpose of the matrix (transpose=True). The matmul () function is utilized for the purpose of executing matrix multiplication.

DOI: https://doi.org/10.54216/JCIM.140226
Received: January 30, 2024 Revised: April 10, 2024 Accepted: July 13, 2024

Step 5: Implement the softmax function. The softmax () function is used to convert the attention scores into a single attention weight. This weight represents the importance of distinct positions collected by different queries, and the total of the weight values is equal to 1.

Step 6: Compute the ultimate attention representation the value matrix (V) is multiplied by the attention scores (attention scores), and the resulting product is used to determine the weighted sum of the values, which gives us the final attention representation.

Step 7: Retrieve the ultimate attention representation and provide it as output.

4.7. Evaluating model

Accurate assessment of data mining techniques requires the use of test data that has not been previously seen during the training phase. In Cross Validation (CV), the SM dataset is divided into two sections in each fold. 80% of the data is allocated as the training set, whereas 20% of the training data is selected as the validation set. In order to assess the efficacy of the model. applied on three distinct sector businesses, including Tesla, NASDAQ, and S&P500, utilizing the Attention model, we utilize significant machine learning and deep learning algorithms to assess their performance using evaluation metrics such as: [13][33][34] [11]

- **Accuracy**= $\frac{TP+TN}{TP+TN+FP+FN}$ (%) (14)

Where TP and FP are quantify the true-positive rate and false-positive rate of the algorithm, respectively

Where FN denotes the false-positive rate of the algorithm.

- F1-measure is calculated as:

$$F1=\frac{2*Precision*Recall}{precision+Recall} \quad (15)$$

- Mean Squared Error (MSE) The quality of a predictor is assessed using a measure that is always non-negative, with numbers closer to zero indicating greater performance. The Mean Squared Error (MSE) is a statistical measure that quantifies the dispersion of predictions from a prediction model. It takes into account both the variance of the forecasts, which indicates how much the predictions differ from one data sample to another, and the bias, which measures the average deviation of the projected values from the observed values. The formula is presented in Equation [27] [35].

$$MSE = \frac{\sum(\boldsymbol{xi-xbi})^2}{\boldsymbol{n}} \quad (16)$$

Where (xbi) is estimated values, (xi) is ture values [6][36].

- Root mean square error (RMSE), is a statistical measure that calculates the average of the squared differences between anticipated values (ybi) and actual values (yi) [37] [6].

$$RMSE=\sqrt{\frac{\sum_i^n(yi-ybi)^2}{n}} \quad (17)$$

- MAPE: is calculated as the average of the relative errors (APE) for the n observation days. The Mean Absolute Percentage Error (MAPE) is computed using the following equation:

$$MAPE=\frac{\sum_i^n\left|\frac{(yi-ybi)}{yi}\right|}{n} * 100 \quad (18)$$

Where: MAPE is the Mean Absolute Percentage Error; n is the total number of observations or data points, yi represents the actual values, ybi represents the predicted values.

4.8. Results and discussion

This study involved conducting an experiment using three datasets of stock indexes: TESLA, NASDAQ, and S&P 500. The comparison of ATTENTION with other approaches in stock index prediction is presented in Table (2). In comparison to the standard neural network model, the ATTENTION model achieves superior performance with

DOI: https://doi.org/10.54216/JCIM.140226
Received: January 30, 2024 Revised: April 10, 2024 Accepted: July 13, 2024

a MAPE of 0.004, MSE of 0.004, and RMSE of 0.06. The suggested ATTENTION model consistently achieves an accuracy of above 90% on the stock indexes dataset. This demonstrates the effectiveness of our technique in analyzing prediction jobs on financial time series. Irrespective of the type of dataset. The experimental findings have demonstrated that the proposed model attained a remarkable accuracy of 99% for the Tesla dataset, as shown in Figure 6. Furthermore, the model exhibited great performance when applied to the S&P500 and NASDAQ datasets, as illustrated in Figures 8 and 9, respectively.

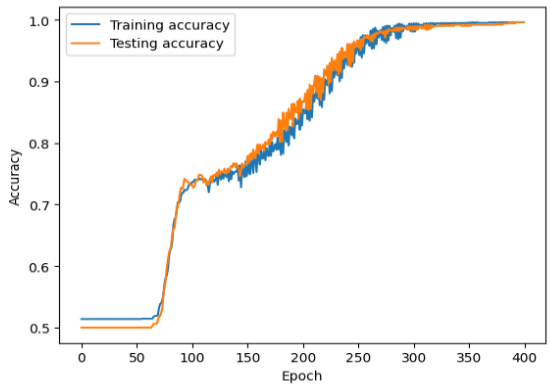


Figure 7. Shows the accuracy datasets using the TESLA datasets

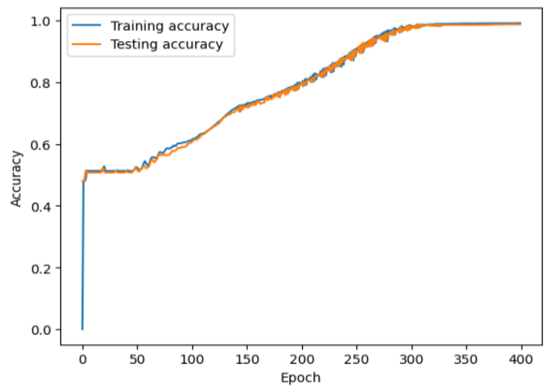


Figure 8. The accuracy datasets using the NASDAQ datasets

Figure 8 shows The result are excellent because they show increased accuracy and achieved 98% during training using 400 epochs.

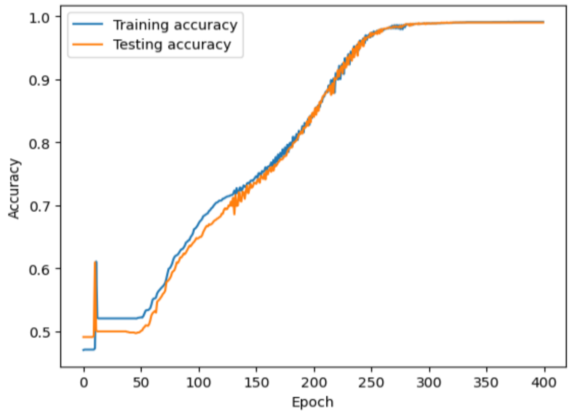


Figure 9. Shows the accuracy datasets using the S&P500 datasets

DOI: https://doi.org/10.54216/JCIM.140226
Received: January 30, 2024 Revised: April 10, 2024 Accepted: July 13, 2024

Figure 9 shows the result is excellent because they show increased accuracy and achieved 99%, during training using 400 epochs.

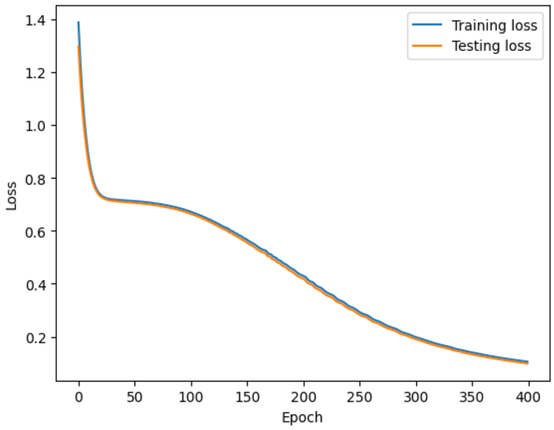


Figure 10. The loss for the train and test datasets using the TESLA dataset.

Figure 10 shows the results are excellent because they show reduced loss to 0.11 during training using 400 epoachs.

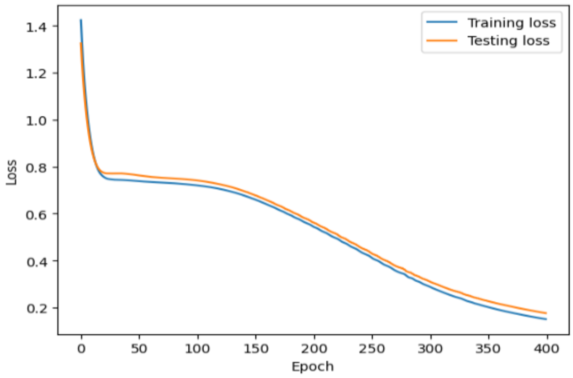


Figure 11. The loss for the train and test datasets using the NASDAQ dataset

Figure 11 shows the results are excellent because they show reduced loss to 0.15 during training using 400 epochs.

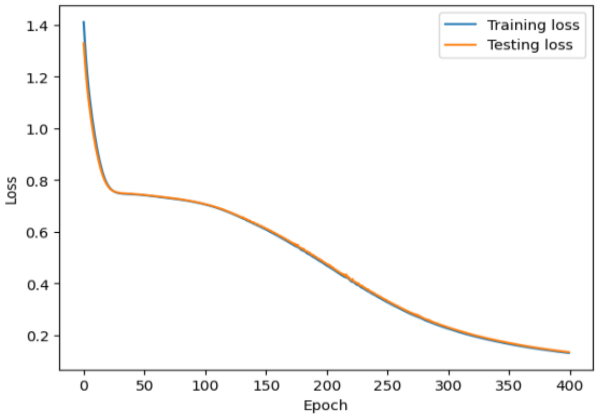


Figure 12. The loss for the train and test datasets using the S&P500 dataset

Figure 12 shows the results are excellent because they show reduced loss to 0.13 during training using 400 epochs.

DOI: https://doi.org/10.54216/JCIM.140226
Received: January 30, 2024 Revised: April 10, 2024 Accepted: July 13, 2024

Table 2: Comparative analysis of previous and proposed study.

Ref	Dataset	Method	Evaluation metrics
[1]	S&P500	RNN	MAPE=2.03
[9]	Tesla	Random Forest	Accuracy=91.93,F1=91
[11]	NASDAQ	DNA	MSE=1472.278
[16]	S&P500 NASDAQ	ANN	Accuracy=0.56, F1=0.68 Accuracy=0.65, F1=0.69
[17]	S&p500	GRU-CNN	Accuracy=56.3
[18]	NASDAQ	Bagging	Accuracy=0.84, F1=0.87
[19]	Tesla	Linear Regression	RMSE=44.6, MSE=1989.1 ,R2=0.85
[20]	S&p500 NASDAQ	LSTM	MAPE for S&P500=0.75 MAPE for NASDAQ=0.84
[21]	S&P500	CNN	MSE=50
[22]	NASDAQ S&P500	MBNN	RMSE for s&p500 =34.86 RMSE for NASDAQ=26.93 MAPE for s&p500 =64.2600 MAPE for NASDAQ=56.19
[23]	Tesla	ANN	Accuracy=0.94
Proposed result	Tsla NASDAQ S&P500	Attention	Accuracy=0.99 F1=0.99, R2=0.99 MSE=0.004 MAPE=0.004 RMSE=0.06 Accuracy=0.98 ,F1=0.98 R2=0.98, MSE=0.01 MAPE=0.01, RMSE=0.1 Accuracy=0.99 F1=0.98, R2=0.98, MSE=0.01 MAPE=0.01 RMSE=0.1

5. Conclusion

A new technique for predicting stock market trends has been proposed. Furthermore, a novel stock market prediction model has been presented, which utilizes the ATTENTION MECHANISM approach. We employ an attention technique to identify the most influential variable in stock price changes and assign them the highest weight. The experimental results demonstrate that the suggested forecasting model outperforms other models in terms of prediction accuracy. The Mean Absolute Percentage Error (MAPE), Mean Squared Error (MSE), and

DOI: https://doi.org/10.54216/JCIM.140226
Received: January 30, 2024 Revised: April 10, 2024 Accepted: July 13, 2024

Root Mean Squared Error (RMSE) provided by the prediction model reduce as the duration of the training period grows. Consequently, the suggested forecasting model presents a novel and promising approach for predicting stock market trends. In the future, utilizing the proposed prediction strategy, it will be possible to develop a novel selection method for identifying the most lucrative stocks with the least amount of risk. Moreover, the presented work can be used to create a sophisticated automated system that acts as an intelligent agent in stock markets. This system aims to maximize profits while minimizing risks by providing recommendations on when to buy and sell the most promising stocks .Academics and business professionals from large corporations are currently engaged in ongoing efforts to discover additional indicators and models that can reliably predict the state of the economic market. In the foreseeable future, it is possible to hypothesize that machine learning techniques, such as learning and mastering the market, will lead to a promising future.

Funding: “This research received no external funding”

Conflicts of Interest: “The authors declare no conflict of interest.”

References

[1] P. Patil, C. S. M. Wu, K. Potika, and M. Orang, “Stock market prediction using ensemble of graph theory, machine learning and deep learning models,” *ACM Int. Conf. Proceeding Ser.*, pp. 85–92, 2020, doi: 10.1145/3378936.3378972.

[2] P. He, Y. Sun, Y. Zhang, and T. Li, “COVID–19’s Impact on Stock Prices Across Different Sectors—An Event Study Based on the Chinese Stock Market,” *Emerg. Mark. Financ. Trade*, vol. 56, no. 10, pp. 2198–2212, 2020, doi: 10.1080/1540496X.2020.1785865.

[3] E. Blankespoor, E. deHaan, and I. Marinovic, “Disclosure processing costs, investors’ information choice, and equity market outcomes: A review,” *J. Account. Econ.*, vol. 70, no. 2–3, 2020, doi: 10.1016/j.jacceco.2020.101344.

[4] S. Singh, T. K. Madan, J. Kumar, and A. K. Singh, “Stock Market Forecasting using Machine Learning: Today and Tomorrow,” *2019 2nd Int. Conf. Intell. Comput. Instrum. Control Technol. ICICICT 2019*, pp. 738–745, 2019, doi: 10.1109/ICICICT46008.2019.8993160.

[5] K. Bhushanwar and K. Barua, “Analyzing the Impact of Global Influencing Features with Weighted Attention Model for Stock Market Forecasting,” *Int. J. Intell. Syst. Appl. Eng.*, vol. 11, no. 11s, pp. 352–367, 2023.

[6] L. Hang, D. Liu, and F. Xie, “A Hybrid Model Using PCA and BP Neural Network for Time Series Prediction in Chinese Stock Market with TOPSIS Analysis,” *Sci. Program.*, vol. 2023, 2023, doi: 10.1155/2023/9963940.

[7] D. Shah, H. Isah, and F. Zulkernine, “Stock market analysis: A review and taxonomy of prediction techniques,” *Int. J. Financ. Stud.*, vol. 7, no. 2, 2019, doi: 10.3390/ijfs7020026.

[8] H. Agarwal, G. Jariwala, and A. Shah, *Analysis and Prediction of Stock Market Trends Using Deep Learning*, vol. 121, no. May. Springer Singapore, 2020. doi: 10.1007/978-981-15-3369-3_39.

[9] Nagamalla, V. karkee, J. Kumar, R. "Integrating Predictive Big Data Analytics with Behavioral Machine Learning Models for Proactive Threat Intelligence in Industrial IoT Cybersecurity," Journal of International Journal of Wireless and Ad Hoc Communication, vol. 7, no. 2, pp. 08-24, 2023. DOI: https://doi.org/10.54216/IJWAC.070201

[10] A. A. R. Hussien and N. A. Z. Abdullah, “A Review for Arabic Sentiment Analysis Using Deep Learning,” *Iraqi J. Sci.*, vol. 64, no. 12, pp. 6572–6585, 2023, doi: 10.24996/ijs.2023.64.12.37.

[11] H. Chung and K. shik Shin, “Genetic algorithm-optimized multi-channel convolutional neural network for stock market prediction,” *Neural Comput. Appl.*, vol. 32, no. 12, pp. 7897–7914, 2020, doi: 10.1007/s00521-019-04236-3.

[12] M. C. Wu, S. Y. Lin, and C. H. Lin, “An effective application of decision tree to stock trading,” *Expert*

DOI: https://doi.org/10.54216/JCIM.140226
Received: January 30, 2024 Revised: April 10, 2024 Accepted: July 13, 2024

Syst. Appl., vol. 31, no. 2, pp. 270–274, 2006, doi: 10.1016/j.eswa.2005.09.026.

[13] Chaudhari, R. Gad, R. Gawali, P. Gite, M. A., D. "Hate Speech Detection on Social Media Using Machine Learning Algorithms," Journal of Journal of Cognitive Human-Computer Interaction, vol. 2, no. 2, pp. 56-59, 2022. DOI: https://doi.org/10.54216/JCHCI.020203

[14] C. Lohrmann and P. Luukka, "Classification of intraday S&P500 returns with a Random Forest," *Int. J. Forecast.*, vol. 35, no. 1, pp. 390–407, 2019, doi: 10.1016/j.ijforecast.2018.08.004.

[15] M. Nabipour, P. Nayyeri, H. Jabani, A. Mosavi, E. Salwana, and S. Shahab, "Deep learning for stock market prediction," *Entropy*, vol. 22, no. 8, 2020, doi: 10.3390/E22080840.

[16] P. V. Chandrika and K. S. Srinivasan, "Predicting stock market movements using artificial neural networks," *Univers. J. Account. Financ.*, vol. 9, no. 3, pp. 405–410, 2021, doi: 10.13189/ujaf.2021.090315.

[17] M. Zulqarnain, R. Ghazali, M. G. Ghouse, Y. M. M. Hassim, and I. Javid, "Predicting Financial Prices of Stock Market using Recurrent Convolutional Neural Networks," *Int. J. Intell. Syst. Appl.*, vol. 12, no. 6, pp. 21–32, 2020, doi: 10.5815/ijisa.2020.06.02.

[18] A. Shams and A. Shams, "Market Using Machine Learning Faria Prediction Market Prediction Using Machine Learning Stock Market Using Machine Learning Faria Prediction Market Prediction Using Machine Learning Abdulhamit Market Prediction Using Machine Learning Market Prediction Usi," *Procedia Comput. Sci.*, vol. 194, pp. 173–179, 2020, doi: 10.1016/j.procs.2021.10.071.

[19] H. Li, "Tesla stock prediction: a comparative study between four models," *Highlights Business, Econ. Manag.*, vol. 24, pp. 182–187, 2024, doi: 10.54097/y1yk0a33.

[20] G. Bathla, "Stock price prediction using LSTM and SVR," *PDGC 2020 - 2020 6th Int. Conf. Parallel, Distrib. Grid Comput.*, pp. 211–214, 2020, doi: 10.1109/PDGC50313.2020.9315800.

[21] H. S. Sim, H. I. Kim, and J. J. Ahn, "Is Deep Learning for Image Recognition Applicable to Stock Market Prediction ?," vol. 2019, 2019.

[22] U. Gurav and N. Sidnal, "Adaptive Stock Forecasting Model using Modified Backpropagation Neural Network (MBNN)," *Proc. Int. Conf. Comput. Tech. Electron. Mech. Syst. CTEMS 2018*, pp. 380–385, 2018, doi: 10.1109/CTEMS.2018.8769290.

[23] I. Suryani and D. C. P. Buani, "Stock price prediction using artificial neural network integrated moving average," *J. Phys. Conf. Ser.*, vol. 1641, no. 1, pp. 916–922, 2020, doi: 10.1088/1742-6596/1641/1/012028.

[24] Eyad, I. "Intelligent systems and AI techniques: Recent advances and Future directions," Journal of International Journal of Advances in Applied Computational Intelligence, vol. 1, no. 2, pp. 30-45, 2022. DOI: https://doi.org/10.54216/IJAACI.010202

[25] M. H. Guo *et al.*, "Attention mechanisms in computer vision: A survey," *Comput. Vis. Media*, vol. 8, no. 3, pp. 331–368, 2022, doi: 10.1007/s41095-022-0271-y.

[26] M. Biswas, A. Shome, M. A. Islam, A. J. Nova, and S. Ahmed, "Predicting stock market price: A logical strategy using deep learning," *ISCAIE 2021 - IEEE 11th Symp. Comput. Appl. Ind. Electron.*, pp. 218–223, 2021, doi: 10.1109/ISCAIE51753.2021.9431817.

[27] Z. Xie and Y. Wang, "Exploration of Stock Portfolio Investment Construction Using Deep Learning Neural Network," *Comput. Intell. Neurosci.*, vol. 2022, 2022, doi: 10.1155/2022/7957097.

[28] K. J. Kim, "Financial time series forecasting using support vector machines," *Neurocomputing*, vol. 55, no. 1–2, pp. 307–319, 2003, doi: 10.1016/S0925-2312(03)00372-2.

[29] J. Ayala, M. García-Torres, J. L. V. Noguera, F. Gómez-Vela, and F. Divina, "Technical analysis strategy optimization using a machine learning approach in stock market indices[Formula presented]," *Knowledge-Based Syst.*, vol. 225, 2021, doi: 10.1016/j.knosys.2021.107119.

[30] T. H. H. Aldhyani and A. Alzahrani, "Framework for Predicting and Modeling Stock Market Prices Based on Deep Learning Algorithms," *Electron.*, vol. 11, no. 19, 2022, doi:

DOI: https://doi.org/10.54216/JCIM.140226
Received: January 30, 2024 Revised: April 10, 2024 Accepted: July 13, 2024

10.3390/electronics11193149.

[31] E. El., "A New Particle Swarm Optimization Based Stock Market Prediction Technique," *Int. J. Adv. Comput. Sci. Appl.*, vol. 7, no. 4, pp. 322–327, 2016, doi: 10.14569/ijacsa.2016.070442.

[32] H. Chung and K. Shin, "Genetic algorithm-optimized multi-channel convolutional neural network for stock market prediction," *Neural Comput. Appl.*, vol. 32, pp. 7897–7914, 2020.

[33] S. Albahli, A. Awan, T. Nazir, A. Irtaza, A. Alkhalifah, and W. Albattah, "A deep learning method DCWR with HANet for stock market prediction using news articles," *Complex Intell. Syst.*, vol. 8, no. 3, pp. 2471–2487, 2022, doi: 10.1007/s40747-022-00658-0.

[34] D. K. Padhi, N. Padhy, A. K. Bhoi, J. Shafi, and S. H. Yesuf, "An Intelligent Fusion Model with Portfolio Selection and Machine Learning for Stock Market Prediction," *Comput. Intell. Neurosci.*, vol. 2022, 2022, doi: 10.1155/2022/7588303.

[35] M. Vijh, D. Chandola, V. A. Tikkiwal, and A. Kumar, "ScienceDirect Stock Closing Closing Price Price Prediction Prediction using using Machine Machine Learning Learning Techniques Techniques," *Procedia Comput. Sci.*, vol. 167, no. 2019, pp. 599–606, 2020, doi: 10.1016/j.procs.2020.03.326.

[36] N. N. M. Azam, M. A. Ismail, M. S. Mohamad, A. O. Ibrahim, and S. Jeba, "Classification of COVID-19 Symptoms Using Multilayer Perceptron," *Iraqi J. Comput. Sci. Math.*, vol. 4, no. 4, pp. 99–110, 2023, doi: 10.52866/ijcsm.2023.04.04.009.

[37] A. Kumar *et al.*, "Generative adversarial network (GAN) and enhanced root mean square error (ERMSE): deep learning for stock price movement prediction," *Multimed. Tools Appl.*, vol. 81, no. 3, pp. 3995–4013, 2022, doi: 10.1007/s11042-021-11670-w.

[38] Yahoo Finance. (2023, October 20). Historical Data. Retrieved from https://finance.yahoo.com

DOI: https://doi.org/10.54216/JCIM.140226
Received: January 30, 2024 Revised: April 10, 2024 Accepted: July 13, 2024