Prediction of Cryptocurrency Price using Time Series Data and Deep Learning Algorithms

Michael Nair[1], Mohamed I. Marie[2], Laila A. Abd-Elmegid[3]
Department of Management Information Systems, Higher Institute of Qualitative Studies, Heliopolis, Cairo, Egypt[1]
Associate professor, Department of Information Systems-Faculty of Computers and Artificial Intelligence,
Helwan University, Cairo, Egypt [2, 3]

***Abstract*—One of the most significant and extensively utilized cryptocurrencies is Bitcoin (BTC). It is used in many different financial and business activities. Forecasting cryptocurrency prices are crucial for investors and academics in this industry because of the frequent volatility in the price of this currency. However, because of the nonlinearity of the cryptocurrency market, it is challenging to evaluate the unique character of time-series data, which makes it impossible to provide accurate price forecasts. Predicting cryptocurrency prices has been the subject of several research studies utilizing machine learning (ML) and deep learning (DL) based methods. This research suggests five different DL approaches. To forecast the price of the bitcoin cryptocurrency, recurrent neural networks (RNN), long short-term memories (LSTM), gated recurrent units (GRU), bidirectional long short-term memories (Bi-LSTM), and 1D convolutional neural networks (CONV1D) were used. The experimental findings demonstrate that the LSTM outperformed RNN, GRU, Bi-LSTM, and CONV1D in terms of prediction accuracy using measures such as Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Squared Error (MSE), and R-squared score (R^2). With RMSE= 1978.68268, MAE=1537.14424, MSE= 3915185.15068, and R^2= 0.94383, it may be considered the best method.**

***Keywords*—Cryptocurrency; deep learning; prediction; LSTM**

I. INTRODUCTION

The fiat currency used in the present monetary system has various disadvantages, including government control over the money supply; transactions are often carried out via intermediaries like financial institutions, which results in expensive fees and prolonged transfer times, as well as the present ledgers used to record transactions being vulnerable to manipulation [1]. Hence Due to its decentralization, immutability, and security, cryptocurrencies have become a worldwide phenomenon that draws a sizable number of users. They are founded on confidence in technology infrastructure, enabling money transfer from any location with nearly negligible delay [2]. Throughout its limited life, the cryptocurrency market has expanded irrationally and astoundingly [3].

Bitcoin, a kind of electronic money, was originally launched in 2008 and was first used as an open-source in 2009 by a person named Satoshi Nakamoto [4]. As the first currency ever created, it has become the most significant currency [5]. Without a single administrator or central bank, it is decentralized digital money that may be transmitted between users on a peer-to-peer network without the involvement of mediators like banks [6].

Most cryptocurrencies use blockchain technology and feature attributes like decentralization, transparency, and immutability [7]. Blockchain allows for the permanent recording of network transactions [8], and each record is encrypted and carries the block's [9] cryptographic hash before it. A date, sender and recipient details, and the total amount of money transmitted are all included in each record. An extremely complex technology called a secure shell links transaction blocks [10]. This technology aims to store data that makes it difficult or impossible to alter, hack, or defraud the system [11].

From 2009 to 2017, the price of Bitcoin increased to over USD 20,000. As of December 2019, the daily average market volume was around USD 19.45 billion [12], and as of April 2021, the price of Bitcoin hit an all-time high of around $65000 [13]. Although investments have yielded rich returns, the constant price swings seen by most cryptocurrencies make them difficult and hazardous [14]. Consequently, it takes work to anticipate the price of cryptocurrencies.

Additionally, the sharp variations in bitcoin prices have emerged as a brand-new worldwide concern. Therefore, it is crucial to foresee changes in the price of Bitcoin [15]. Because of this, investors need a forecasting strategy to efficiently capture swings in the price of cryptocurrencies to reduce risk and boost profits [16] .

Cryptocurrency price prediction is a time series prediction problem in its early phases. In contrast, older methods were used to anticipate time series based on linear hypotheses and required information that could be categorized as trend or seasonal [3], such as sales forecasting. Due to the extreme volatility and lack of seasonality in the Bitcoin market, these strategies are unsuccessful. Based on its success in similar domains, deep learning is an attractive technological choice, given the difficulty of the challenge [17]. From this point on, DL methods are considered efficient for time series forecasting since they are noise-resistant, can accommodate data sequences natively, and can recognize non-linear temporal correlations on such sequences [18].

Estimating the price of the Bitcoin cryptocurrency is the aim of this research, and evaluating the forecasting accuracy of five different deep learning models, including LSTM, RNN, GRU, Conv1D, and Bi-LSTM. The research uses the (RMSE), (MAE), (MSE), and (R^2) as measurement techniques to assess the performance of DL models using the closing price of Bitcoin in USD.

The issue that motivated us to conduct this paper was the lack of a specific model with high accuracy that could be relied upon to predict the price of cryptocurrencies, which may have a significant impact on the increase in financial profits. It was vital to provide a solid approach to address this issue for investors that invest in these encrypted currencies. Deep learning methods were used as a consequence because they produced positive outcomes in various study domains.

This paper contributes to providing knowledge to everyone interested in this field in identifying deep learning techniques and their ability to deal with time series data to predict the prices of cryptocurrencies, where the results of the research proved that the use of deep learning method resulted in better results than the traditional machine learning techniques, and also to assist investors interested in trading cryptocurrencies in selecting the best deep learning model to predict prices, and to make the right decision to decrease their loss exposure and increase profitability during the trading process in this currency.

The paper is divided into seven sections: Section 2 is a literature review, Section 3 provides background knowledge, and Section 4 presents the model to guide our approach. Section 5 tests the suggested model; Section 6 presents the findings of the experiment; Section 7 conclusions and future work.

II. Literature Review

Bitcoin is a cryptocurrency and a kind of electronic money. It is a well-known cryptocurrency with a bright future [19], and it is a web-based trade technique that uses cryptographic tools to carry out financial transactions [20]. It is crucial to forecast the values of this currency because of the considerable price volatility of this encrypted money, which has the potential to impact investors negatively and international and commercial ties [21]. Numerous researches have been carried out to forecast time series and the value of bitcoin [10]. In contrast, deep learning models [13] and machine learning models [4] were employed to forecast the price of Bitcoin.

The prior research on predicting cryptocurrency prices will be examined in the following part, employing various ML and DL models for time series prediction, as shown in Table I.

TABLE I. Literature Review for Cryptocurrency Price Prediction Price using ML and DL

Author	**Year**	**Technique**	**Cryptocurrency**	**Dataset Source**	**Data Range**	**Prediction Methods**	**Performance Measures and results**	**Demerit**
HASAN et al [7]	2022	DL	Bitcoin Ethereum Monero	Investing.com	between Jan 22, 2015 to Feb 12, 2020	LSTM, RNN and Proposed method	The Proposed method has achieved the best performance when predicting Bitcoin price with MSE= 18.65, MAE= 2.15 and RMSE= 4.21	Not Explored time-series model such as GRU
NEMATA LLAH et al [10]	2022	DL	Bitcoin	Kaggle	between 1 Jan 2012 to 31 Mar 2021	RNN LSTM	MAPE and RMSE LSTM performs better than RNN	Not Explored time-series model such as GRU
Bitto et al [22]	2022	ML	Bitcoin, Ethereum, Litecoin and Tether token	Yahoo Finance	between 2015-1-1 to 2021-6-1	AR MA ARMA	MAE and RMSE. AR model giving better performs than others models with 97.21% For bitcoin, 96.04% for Ethereum, 95.8% for Litecoin and 99.91% accuracy for Tether-token	Not considered deep learning models for prediction
Ammer et al [12]	2022	DL	AMP, Ethereum, Electro-Optical System, and XRP	CoinMarketCap	between May 2015 through April 2022	LSTM	MSE, RMSE, NRMSE and R. LSTM achieved R = 96.73% for training And R= 96.09% for testing when predicting XRP price	Not Explored time-series model such as GRU
FAKHAR CHIAN et al [15]	2022	DL	Bitcoin	Yahoo Finance	between 05/02/2021 To 10/09/2021	proposed models based on CNN and LSTM	Model-9 achieved the best performance with MSE= 0.00151, RMSE= 0.0388, MAE= 0.02519, MedAE= 0.01747 and R2= 0.98219	Not Explored time-series model such as GRU
ZHANG et al [23]	2022	ML DL	Bitcoin	Data.Bitcoinity. Org, Blockchain.com , and CoinMarketCA P	between 05/02/2021 To 10/09/2021	LSSVM BP SDAE-B	SDAE-B model giving better performs than others models with MAPE= 0.016, RMSE= 131.643 and DA= 0.817	Not Explored time-series model such as LSTM and GRU

GURRIB et al [24]	2022	ML	Bitcoin	CoinMarketCAP	between 17 Jun 2016 to 21 Apr 2021	LDA SVM	LDA model giving better performs than SVM with accuracy of 0.585	Not considered deep learning models for prediction
CAVALLI et al [25]	2021	ML	Bitcoin	CoinMarketCAP	between 28 of Apr 28, 2013 to Feb 15, 2020	1D CNN LSTM	RMSE 1D CNN model giving better performs than LSTM	Not considered deep learning models for prediction
LIU et al [26]	2021	ML DL	Bitcoin	Coindesk.com BTC.com	between Jul 2013 to Dec 2019	BPNN SVR SDAE	SDAE model giving better performs than others models with MAPE= 0.1019, RMSE= 160.63 and DA= 0.5985	Not Explored time-series model such as LSTM and GRU
MARNE et al [27]	2020	ML DL	Bitcoin	Kaggle	between Jan 2014 to Jan 2019	SVM RNN LSTM	LSTM model giving better performs than others models with RMSE = 3.38	Not Explored time-series model such as GRU

III. METHODOLOGY

We will go through several related concepts in this section.

A. *Time-Series*

It is one of the most effective methods for forecasting situations with some degree of future uncertainty by analyzing past patterns and assuming that future trends will be similar. Time series forecasting is also based on data for efficient and effective planning to solve forecasting problems with a time component [3].

B. *Deep Learning Methods used for Bitcoin Price Prediction*

The approaches utilized in DL, a subfield of ML, are built on the structure and design of ANNs. Five DL algorithms were used in this study to forecast the price of Bitcoin. LSTM, GRU, BiLSTM, simple RNN, and the 1D CNN algorithm.

1) Recurrent Neural Network (RNN): Artificial neural networks were inspired by how the human brain processes information. The neural network comprises synthetic neurons, and its architecture determines its properties. Traditional neural networks do not have feedback loops, which is how RNNs vary from them. It is thus relevant anytime the input context affects how well a prediction is made. Each neuron's current state depends on its past state due to the recurrent nature of an RNN's layers, which leaves the neural network with a finite amount of memory. Sequential data may be input into a recurrent neural network, and both the networks In and Out may be sequences of variable lengths that pass through each cell consecutively [28]. Suppose there is an input neuron Xt, an invisible output status ht, and the prior invisible output status ht-1. In that case, the RNN has a single-layer recurrent module with a tanh squashing function. Fig. 1 [29] demonstrates that W represents the weighted matrix and yt for the result.

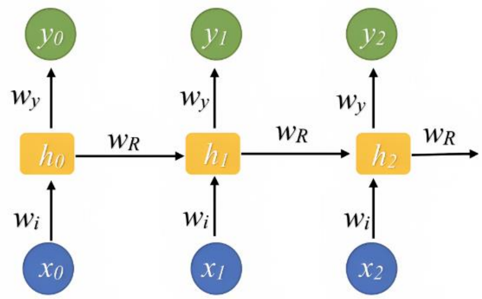

Fig. 1. The recurrent neural network simple architecture [29].

2) Long Short-Term Memory (LSTM): Recurrent neural networks with the ability to learn long term dependencies are called LSTMs. The recommended networks by Hochreiter and Schmid Huber [30] because the last state needed to be sufficiently recent and thus influenced the present state, the RNN model may inaccurately predict the current state [31]. From left to right, the LSTM is crafted to keep track of information throughout time and lessen the issue of vanishing gradient descent. Three interconnected layers in the LSTM, the input gate, forget gate, and output gate, control the data flow necessary to forecast the output of the network [32].

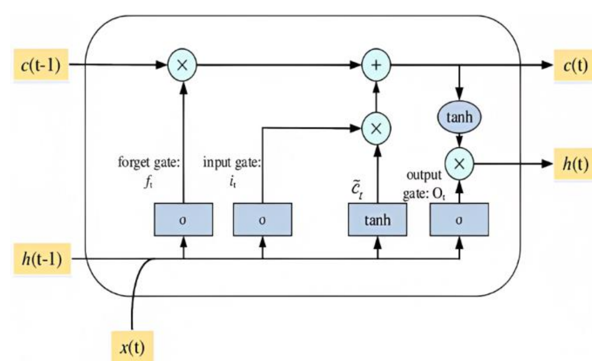

Fig. 2. Schematic diagram of LSTM [33].

Input gate: Information will initially pass through the input gate after importing the data. The switch decides whether or not to store the information based on the state of the cell.

Output gate: The amount of output information is determined by it.

Forget gate: It chooses whether to keep or forget the information obtained. [34, 35], as shown in Fig. 2.

3) Gate Recurrent Unit (GRU): Another RNN version is a GRU, which combines the three gated units into only two gated units: the gate for updating and resetting [36]. GRUs address the vanishing gradient issue of RNNs and the optimization of the structure of the LSTM model. The two gates may store relevant data in the memory cell while transmitting values to the network's later stages. GRU and LSTM are equal when evaluating performance across various test scenarios [37]. Fig. 3 depicts the organization of the GRU units.

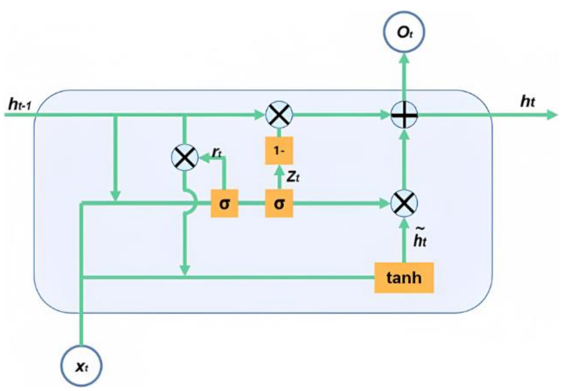

Fig. 3. GRU unit structure [38].

4) Bidirectional Long Short-Term Memory (BiLSTM): The BiLSTM model can extract contextual information from feature sequences by considering both forward and backward dependencies. Using a front LSTM, that processes the sequence in chronological order and a backward LSTM that processes the sequence in reverse order, BiLSTM allows looking ahead. The output is then produced by joining the LSTM's forward and reverse states [39, 40] as seen in Fig. 4.

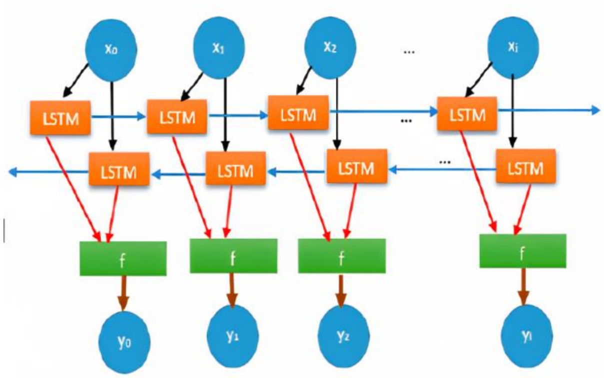

Fig. 4. BiLSTM architecture [11].

5) 1D Convolutional Neural Network (CONV1D) model: It is easy to find basic patterns in data using a convolutional neural network (CNN), which is then used to build more complex patterns in the top layers. A 1D CNN is helpful when extracting key features from tiny (fixed-length) segments of the whole dataset. The feature's location within the segment is irrelevant; this is correct for analyzing historical data and evaluating sensor data time series. Input, output, and hidden layers comprise a CNN; a feedforward neural network is created using the intermediary layers. Since their inputs and outputs are blind to the activation function and final convolution [31], as illustrated in Fig. 5, these are called hidden layers.

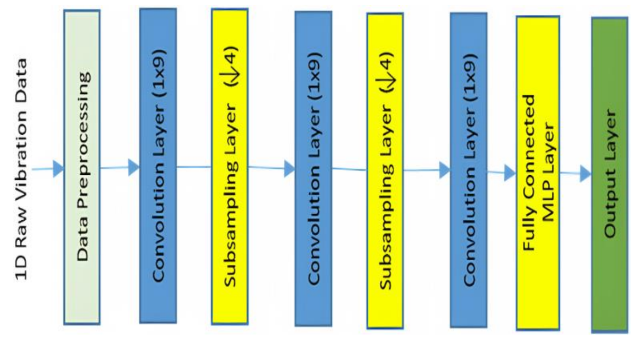

Fig. 5. 1D CNN architecture [41].

IV. The Proposed Model for Predicting Cryptocurrency Price Movement

This section's suggested model focuses on three key elements. Fig. 6 illustrates the three steps used to anticipate the movement of the cryptocurrency price: (1) Dataset; (2) Data pre-processing; and (3) Deep learning-based algorithms.

Table I displays the literature review, methods used, and limitations of each study, which show the inaccuracy, the use of primitive methods, or a small dataset are all examples of shortcomings. In our research, we used similar and different methods, such as RNNs, LSTMs, GRUs, CONV1D, Bi-LSTM, and different datasets with large sizes from the Kaggle website. Using all these methods helped improve the accuracy.

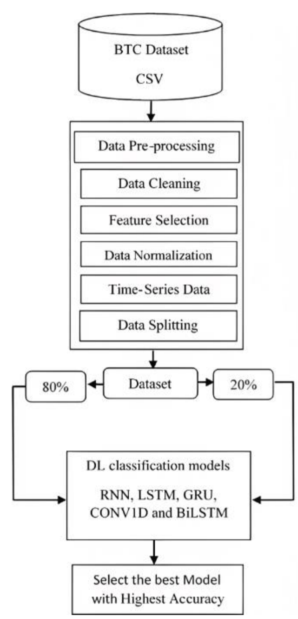

Fig. 6. The proposed model for predicting BTC price.

V. Proposed Model Testing

During this research, an experiment was conducted to test five DL models, (RNN), (LSTM), (GRU), (Bi-LSTM), and (CONV1D); the models are designed to predict BTC Price.

A. Dataset

The data used in this study were downloaded from the Kaggle website for Bitcoin Cryptocurrency in CSV format. The dataset contains a variety of columns, including, Open, high, low, close, and Adj close prices and the volume, from the period 2014-09-17 to 2022-02-01, as shown in Fig. 7, the sample data from the datasets of the Cryptocurrency used in the study, and the target variable in this research is only the (Close Prices) Bitcoin.

	Date	Open	High	Low	Close	Adj Close	Volume
0	2014-09-17	465.864014	468.174011	452.421997	457.334015	457.334015	21056800
1	2014-09-18	456.859985	456.859985	413.104004	424.440002	424.440002	34483200
2	2014-09-19	424.102997	427.834991	384.532013	394.795990	394.795990	37919700
3	2014-09-20	394.673004	423.295990	389.882996	408.903992	408.903992	36863600
4	2014-09-21	408.084991	412.425995	393.181000	398.821014	398.821014	26580100

Fig. 7. Historical prices for bitcoin before the preprocessing.

B. Data Pre-Processing

Data preparation is the first step of the experiment. Before data is supplied to the DL models, preprocessing is considered a crucial step that must be finished. Many stages were conducted during the processing, including data cleaning, feature selection, data normalization, time-series data, and data splitting.

- Data Cleaning: Replacing null or incorrect values with legitimate ones or eliminating the whole data point.
- Feature Selection: There are several variables in cryptocurrency data. Since only the Close Prices are the goal variable in our study, only relevant characteristics should be chosen, and extraneous features should be eliminated, as shown in Fig. 8.
- Data normalization or standardization: These processes are crucial for ensuring all data is on the same scale. In our models, we employ the MinMaxScaler function to normalize all data to a range of 0 to 1, which is crucial when working with Bitcoin data, which might have a broad range of values.
- Time-series Data: Cryptocurrency prices are time series data; as a result, it is crucial to transform the data into a time series format to recognize any patterns, trends, or seasonal impacts.
- Data Splitting: Divide the pre-processed data into training and testing sets. The prediction model will be developed using the training dataset, and its effectiveness will be assessed using the testing dataset.

	Date	Close
0	2021-02-06	39266.011719
1	2021-02-07	38903.441406
2	2021-02-08	46196.464844
3	2021-02-09	46481.105469
4	2021-02-10	44918.183594

Fig. 8. The target variable (Close Prices).

C. Deep Learning Models for Predicting BTC Price Movement

We propose in this section five DL algorithms. (1) RNN (2) LSTM (3) GRU (4) BiLSTM (5) CONV1D. The architecture of these models is shown in Tables II to VI.

1) Recurrent Neural Network (RNN) model: Simple RNN is the model's first layer; it has one simple RNN layer consisting of 50 filters that acquire data, process it, and then pass it on to the next layer. The results are in a 7 x 50 matrix using ReLU as an Activation Function. The second layer is another simple RNN, which generates a 1 x 20 matrix using ReLU as an Activation Function. Next, the last stage in the model consists of two fully connected layers, the first one with 50 nodes and the last one with one node, which is the model's output, and we used the Adam optimizer to calculate the learning rate, as shown in Table II.

TABLE II. Recurrent Neural Network (RNN) Model

Layer (type)	Output shape	Param #
simple_rnn (SimpleRNN)	(None, 7, 50)	2600
simple_rnn_1 (SimpleRNN)	(None, 20)	1420
dense_4 (Dense)	(None, 50)	1050
dense_5 (Dense)	(None, 1)	51
Total params : 5,121 Trainable params : 5,121 Non-trainable params : 0		

2) Long Short-Term Memory (LSTM) model: LSTM is the second DL model; LSTM is the model's first layer consisting of 50 filters that acquire data, process it, and then pass it on to the next layer. The results are in a 7 x 50 matrix using ReLU as an Activation Function. The second layer is another LSTM, which generates a 1 x 25 matrix using ReLU as an Activation Function. Next, the last stage in the model consists of two fully connected layers: the first one with 50 nodes and the last with one node, which is the model's output, and the Adam optimizer method, as shown in Table III.

TABLE III. Long Short-Term Memory (LSTM) Model

Layer (type)	Output shape	Param #
lstm_5 (LSTM)	(None, 7, 50)	10400
lstm_6 (LSTM)	(None, 25)	7600
dense_12 (Dense)	(None, 50)	1300
dense_13 (Dense)	(None, 1)	51
Total params : 19,351 Trainable params : 19,351 Non-trainable params : 0		

3) Gate Recurrent Unit (GRU) model: GRU is the third DL model, GRU is the model's first layer which generates a 1 x 50 matrix, and the last stage in the model is composed of two fully connected layers, the first one with 50 nodes and the last with one node which is the output of the model, and the Adam optimizer method as shown in Table IV.

TABLE IV. GATE RECURRENT UNIT (GRU) MODEL

Layer (type)	Output shape	Param #
gru (GRU)	(None, 50)	7950
dense_8 (Dense)	(None, 50)	2550
dense_9 (Dense)	(None, 1)	51
Total params : 10,551 Trainable params : 10,551 Non-trainable params : 0		

4) Bidirectional Long Short-Term Memory (Bi-LSTM) Model: Bi-LSTM is the fourth DL model. Bi-LSTM is the model's first layer consisting of 200 filters that acquire data, process it, and then pass it on to the next layer. The results are in a 207 x 200 matrix. A dropout layer is a regularization approach that prevents overfitting problems in deep learning by ensuring that no units are codependent with one another. Next, the last stage in the model is composed of two fully connected layers, the first using ReLU as an Activation Function with 20 nodes and the last with one node, which is the model's output, and the Adam optimizer method, as shown in Table V.

TABLE V. BIDIRECTIONAL LONG SHORT-TERM MEMORY (BI-LSTM) MODEL

Layer (type)	Output shape	Param #
bidirectional (Bidirectional)	(207, 200)	81600
dropout (Dropout)	(207, 200)	0
Dense_10 (Dense)	(207, 20)	4020
dense_11 (Dense)	(207, 1)	21
Total params : 85,641 Trainable params : 85,641 Non-trainable params : 0		

5) 1D Convolutional Neural Network (CONV1D) model: 1DCNN is the fifth DL model; 1DCNN is the model's first layer consisting of 64 filters that acquire data, process it, and then pass it on to the next layer. The results are in a 7 x 64 matrix which uses ReLU as an Activation Function. In order to simplify the output and avoid overfitting the data, the maximum pooling layer is used after a CNN layer; This indicates that the output matrix for this layer is 2 x 64 in size. The Max Pooling1D layer shrinks the input representation by taking the maximum value across all time dimensions. Next, the last stage in the model is composed of two fully connected layers, the first with 50 nodes, then using the Flatten layer; the Flatten layer transforms convolutional layer output into a single, one-dimensional vector that may be utilized as the input for a dense layer. The last dense layer has one node, the model's output, and the Adam optimizer method, as shown in Table VI.

TABLE VI. 1D CONVOLUTIONAL NEURAL NETWORK (CONV1D) MODEL

Layer (type)	Output shape	Param #
conv1d (Conv1D)	(None, 7, 64)	512
max_pooling1d(Maxpooling1D)	(None, 2, 64)	0
dense_6 (Dense)	(None, 2, 50)	3250
flatten (Flatten)	(None, 100)	0
dense_7 (Dense)	(None, 1)	101
Total params : 3,863 Trainable params : 3,863 Non-trainable params : 0		

VI. EXPERIMENTAL RESULTS AND DISCUSSION

The experiment's results will be discussed in this section.

A. Model Training

To find the best DL model, we trained utilizing DL models on the dataset in the first phase, splitting it into two groups of 80% training and 20% testing. Four assessment measures—RMSE, MSE, MAE, and R^2—were used to examine and contrast the DL models, as will discuss in Section 6(C).

B. Epochs

The number of training set iterations is called an "epoch." The model's capacity for generalization improves as epochs increase. However, if the number of epochs is excessively high, an overfitting issue is readily created, and the model's capacity for generalization is diminished [42]. Therefore, picking the appropriate number of epochs is crucial. In this research, we used 200 epochs.

Tables VII, to XI show the loss and val_loss for each epoch on the various DL models. As shown in Fig. 9 to 13, the model's loss for the training and validation phases decrease in each epoch, indicating that the model performs optimally. The model predicts the actual and prediction phases shown in Fig. 14 to 18.

TABLE VII. LOSS, VAL LOSS OF RNN MODEL

Epoch	Loss	Val_Loss
1/200	0.0858	0.0042
2/200	0.0039	0.0032
3/200	0.0041	0.0033
4/200	0.0031	0.0037
5/200	0.0034	0.0034

TABLE VIII. LOSS, VAL LOSS OF LSTM MODEL

Epoch	Loss	Val_Loss
1/200	0.0654	0.0153
2/200	0.0110	0.0068
3/200	0.0082	0.0102
4/200	0.0082	0.0065
5/200	0.0073	0.0059

TABLE X. LOSS, VAL LOSS OF GRU MODEL

Epoch	Loss	Val_Loss
1/200	0.0214	0.0039
2/200	0.0040	0.0032
3/200	0.0039	0.0029
4/200	0.0038	0.0031
5/200	0.0035	0.0026

TABLE XI. LOSS, VAL LOSS OF BI-LSTM MODEL

Epoch	Loss	Val_Loss
1/200	0.0242	0.0082
2/200	0.0104	0.0095
3/200	0.0098	0.0089
4/200	0.0092	0.0049
5/200	0.0087	0.0125

TABLE XII. LOSS, VAL LOSS OF CONV1D MODEL

Epoch	Loss	Val_Loss
1/200	0.0160	0.0045
2/200	0.0049	0.0051
3/200	0.0046	0.0031
4/200	0.0037	0.0046
5/200	0.0044	0.0029

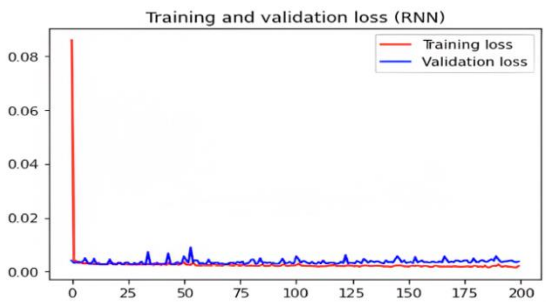

Fig. 9. RNN model loss for training and validation.

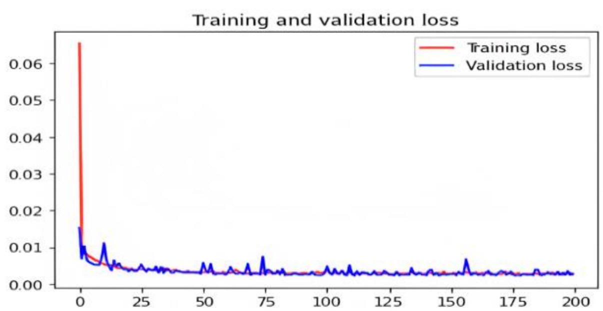

Fig. 10. LSTM model loss for training and validation.

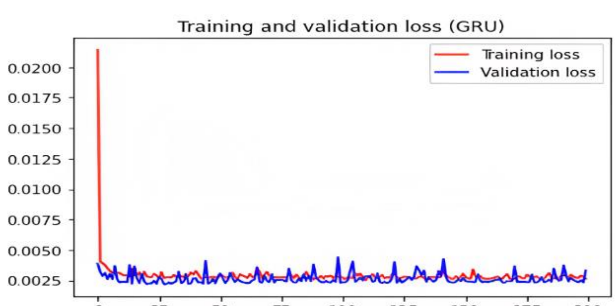

Fig. 11. GRU model loss for training and validation.

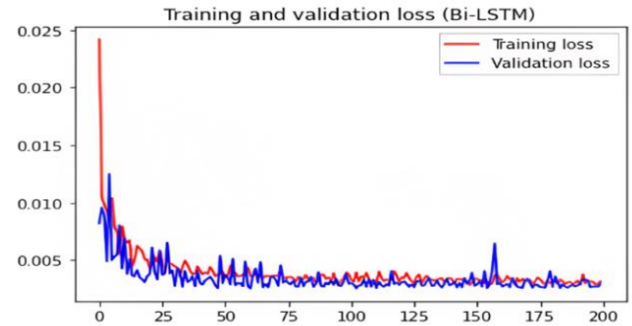

Fig. 12. Bi-LSTM model loss for training and validation.

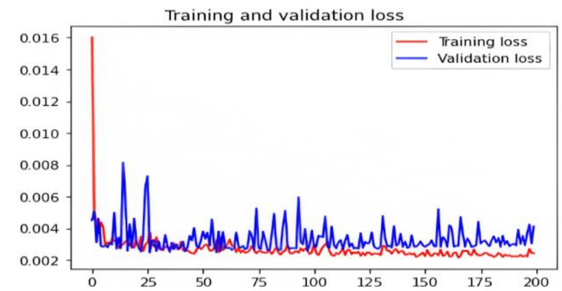

Fig. 13. CONV1D model loss for training and validation.

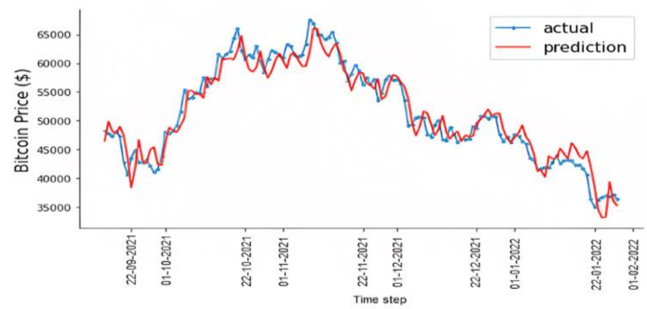

Fig. 14. BTC price prediction based on RNN model.

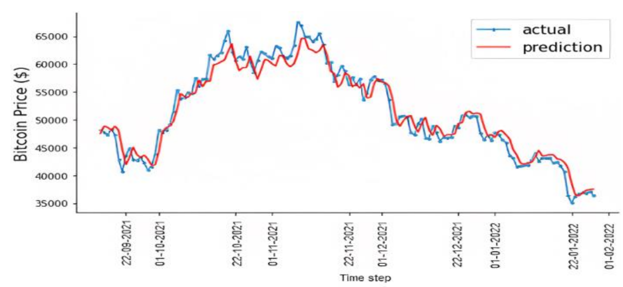

Fig. 15. BTC price prediction based on LSTM model.

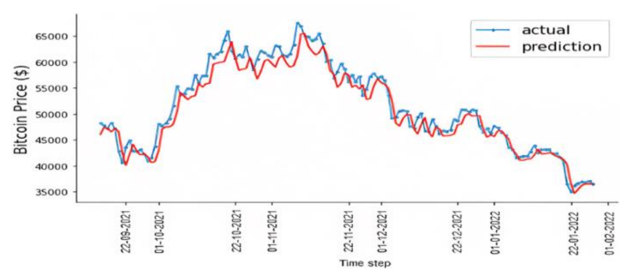

Fig. 16. BTC price prediction based on GRU model.

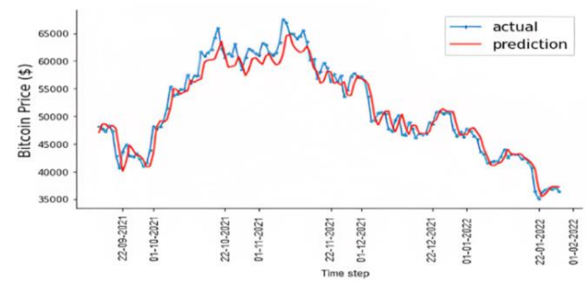

Fig. 17. BTC price prediction based on Bi-LSTM model.

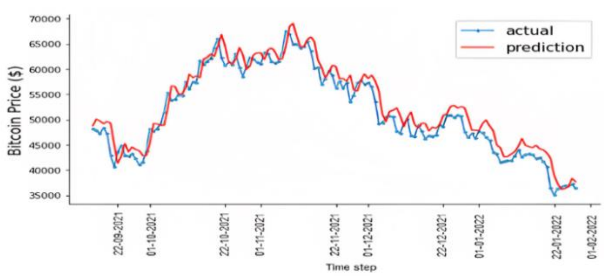

Fig. 18. BTC price prediction based on CONV1D model.

C. Evaluation Metrics

R-squared score (R^2), Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE) were used to assess the performance of the deep learning models.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(y_{t_i} - \hat{y}_{t_i}\right)^2}{n}} \quad (1)$$

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}\left(y_{t_i} - \hat{y}_{t_i}\right)^2 \quad (2)$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|y_{t_i} - \hat{y}_{t_i}\right| \quad (3)$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|y_{t_i} - \hat{y}_{t_i}\right| \quad (4)$$

D. Deep Learning Prediction Models Outcomes

This section covered the outcomes of the prediction models created utilizing DL models between 22-09-2021 and 01-02-2022. How well prediction models perform Tables XII, to XVI, show the testing and training results regarding RMSE, MAE, MSE, and R^2. Table XVII displays the outcomes of utilizing various models, with the model with the lowest error values chosen as the best model. The comparison between the actual and expected values for BTC price prediction models is shown in Fig. 19, as well.

TABLE XIII. TESTING AND TRAINING OUTCOMES FOR RNN MODEL

	RMSE	MAE	MSE	R^2
Training	1512.57653	1042.15566	2287887.78387	0.9739
Testing	2312.72885	1855.64295	5348714.75799	0.9232

TABLE XIV. TESTING AND TRAINING OUTCOMES FOR LSTM MODEL

	RMSE	MAE	MSE	R^2
Training	1908.46769	1476.50991	3642248.95956	0.95846
Testing	1978.68268	1537.14424	3915185.15068	0.94383

TABLE XV. TESTING AND TRAINING OUTCOMES FOR GRU MODEL

	RMSE	MAE	MSE	R^2
Training	2091.59478	1631.35273	4374768.76158	0.95011
Testing	2170.99032	1693.30095	4713198.99972	0.93238

TABLE XVI. TESTING AND TRAINING OUTCOMES FOR BI-LSTM MODEL

	RMSE	MAE	MSE	R^2
Training	1882.75025	1462.76491	3544748.50956	0.95957
Testing	2048.97955	1574.88476	4198317.23545	0.93977

TABLE XVII. TESTING AND TRAINING OUTCOMES FOR CONV1D MODEL

	RMSE	MAE	MSE	R^2
Training	2039.41352	1535.66479	4159207.51736	0.95257
Testing	2418.95978	1949.64524	5851366.42758	0.91606

Table XVII shows that the LSTM model, which has the lowest RMSE, MAE, and MSE values and the greatest R2 value, performs the best in forecasting BTC prices. Fig. 19, which demonstrate how closely the forecasts of the LSTM model match the actual prices, support this. The findings show that LSTM is a better predictor than RNN, GRU, Bi-LSTM, and CONV1D. The second and third-best models are the Bi-LSTM and GRU, with higher RMSE, MAE, and MSE values.

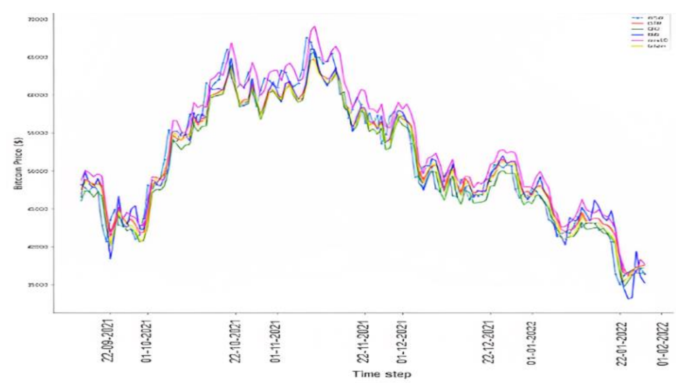

Fig. 19. Summary of BTC price prediction models between actual and predicting.

TABLE XVIII. SUMMARY OF DIFFERENT DL MODELS PREDICTION IN TERMS OF THE VARIOUS CRITERIA

Model	RMSE	MAE	MSE	R2
RNN	2312.72885	1855.64295	5348714.75799	0.92327
Conv1D	2418.95978	1949.64524	5851366.42758	0.91606
GRU	2170.99032	1693.30095	4713198.99972	0.93238
Bi-STM	2048.97955	1574.88476	4198317.23545	0.93977
LSTM	1978.68268	1537.14424	3915185.15068	0.94383

These models are reliable and appropriate based models are reliable and appropriate based on the assessment techniques and outcomes. It should be emphasized that these models contain many flaws that may affect how well they can forecast BTC values:

- As cryptocurrency values rely heavily on various factors, LSTMs, RNN, GRU, Bi-LSTM, and CONV1D may only be able to account for some of these dependencies, producing predictions that could be better.

- These models are vulnerable to overfitting, particularly when trained on small datasets, which can lead to subpar performance when used with new data.

VII. CONCLUSION AND FUTURE WORK

In this research, the market capitalization of the BTC cryptocurrency was utilized to forecast the price using five different deep learning techniques: LSTM, RNN, GRU, Bi-LSTM, and CONV1D. RMSE, MAE, MSE, and R2 values were used to assess the models' performance. The study's findings showed that the LSTM model, followed by the Bi-LSTM and GRU models, offered the best accurate forecasts for the price of the BTC coin. The study's findings show that deep learning algorithms are good at forecasting cryptocurrency values and that the LSTM model outperforms RNN, GRU, Bi-LSTM, and CONV1D.

To increase the precision of BTC predictions, the researcher plans to apply more deep learning algorithms or hybrid DL models in the future. The epoch size might also be increased to get a greater accuracy rate. Deep learning methods will also examine how emotion and tweets affect BTC pricing.

The research limitations can be represented in the following points:

- The prediction process focused on Bitcoin only. It did not apply the prediction to other cryptocurrencies, for example, Ethereum and Litecoin, which can correlate and impact the price of Bitcoin.

- Not considering another factor that can impact the rise and fall of a currency's price, such as comments on social media.

REFERENCES

[1] P. L. Seabe, C. R. B. Moutsinga, and E. Pindza, "Forecasting cryptocurrency prices using LSTM, GRU, and bi-directional LSTM: a deep learning approach,". Fractal and Fractional, vol. 7, no. 2, pp. 203, 2023.

[2] A. A. Oyedele, A. O. Ajayi, L. O. Oyedele, S. A. Bello, and K. O. Jimoh, "Performance evaluation of deep learning and boosted trees for cryptocurrency closing price prediction,". *Expert Systems with Applications, vol. 213*, pp. 119233, 2023.

[3] N. Latif, J. D. Selvam, M. Kapse, V. Sharma, and V. Mahajan, "Comparative Performance of LSTM and ARIMA for the Short-Term Prediction of Bitcoin Prices,". Australasian Accounting, Business and Finance Journal, vol. 17, no. 1, pp. 256-276, 2023.

[4] S. A. Basher, and P. Sadorsky, "Forecasting Bitcoin price direction with random forests: How important are interest rates, inflation, and market volatility?,". Machine Learning with Applications, vol. 9, pp. 100355, 2022.

[5] Z. Ye, Y. Wu, H. Chen, Y. Pan, and Q. Jiang, "A stacking ensemble deep learning model for bitcoin price prediction using Twitter comments on bitcoin,". Mathematics, vol. 10, no. 8, pp. 1307, 2022.

[6] N. Alsalmi, S. Ullah, and M. Rafique, "Accounting for digital currencies,". Research in International Business and Finance, vol. 64, pp. 101897, 2023.

[7] S. H. Hasan, S. H. Hasan, M. S. Ahmed, and S. H. Hasan, "A Novel Cryptocurrency Prediction Method Using Optimum CNN,". Computers, Materials & Continua, vol. 71, no. 1,pp. 1051-1063, 2022.

[8] E. Şaşmaz, and F. B. Tek, "Tweet sentiment analysis for cryptocurrencies,". In 2021 6th International Conference on Computer Science and Engineering (UBMK), pp. 613-618, IEEE, 2021.

[9] A. S. Salama, and A. M. Eassa, "IOT and cloud based blockchain model for COVID-19 infection spread control,". J Theor Appl Inf Technol, 100(1), 113-126, 2022.

[10] H. F. Nematallah, A. A. H. Sedky, and K. M. Mahar, "Bitcoin Price Trend Prediction Using Deep Neural Network,". Webology (ISSN: 1735-188X) Vol. 19, 2022.

[11] M. J. Hamayel, and A. Y. Owda, "A novel cryptocurrency price prediction model using GRU, LSTM and bi-LSTM machine learning algorithms,". AI, vol. 2, no. 4, pp. 477-496, 2021.

[12] M. A. Ammer, and T. H. Aldhyani, "Deep learning algorithm to predict cryptocurrency fluctuation prices: Increasing investment awareness,". Electronics, vol. 11, no. 15, pp. 2349, 2022.

[13] B. Agarwal, P. Harjule, L. Chouhan, U. Saraswat, H. Airan, and P. Agarwal, "Prediction of dogecoin price using deep learning and social media trends,". EAI Endorsed Transactions on Industrial Networks and Intelligent Systems, vol. 8, no. 29, e2-e2, 2021.

[14] E. Bouri, S. J. H. Shahzad, and D. Roubaud, "Co-explosivity in the cryptocurrency market,". Finance Research Letters, 29, pp. 178-183, 2019.

[15] S. Fakharchian, "Designing forecasting assistant of the Bitcoin price based on deep learning using the market sentiment analysis and multiple feature extraction,", 2022.

[16] F. Fang, C. Ventre, M. Basios, L. Kanthan, D. Martinez-Rego, F. Wu, and L. Li, "Cryptocurrency trading: a comprehensive survey,". Financial Innovation,vol. 8, no. 1, pp. 1-59, 2022.

[17] S. McNally, J. Roche, and S. Caton, "Predicting the price of bitcoin using machine learning,". In 2018 26th euromicro international conference on parallel, distributed and network-based processing (PDP), pp. 339-343, IEEE, 2018.

[18] K. Murray, A. Rossi, D. Carraro, and A. Visentin, "On Forecasting Cryptocurrency Prices: A Comparison of Machine Learning, Deep Learning, and Ensembles,". Forecasting, vol. 5, no. 1, pp. 196-209, 2023.

[19] X. Jiang, "Bitcoin price prediction based on deep learning methods,". Journal of Mathematical Finance, vol. 10, no. 1, pp. 132-139, 2019.

[20] E. Mahendra, H. Madan, S. Gupta, and S. V. Singh, "Bitcoin price prediction using deep learning and real time deployment,". In 2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN), pp. 264-268, IEEE, 2020.

[21] T. Awoke, M. Rout, L. Mohanty, and S. C. Satapathy, "Bitcoin price prediction and analysis using deep learning models,". In Communication Software and Networks: Proceedings of INDIA 2019 (pp. 631-640). Singapore: Springer Singapore, 2020.

[22] A. K. Bitto, I. Mahmud, M. H. I. Bijoy, F. T. Jannat, M. S. Arman, M. M. H. Shohug, and H. Jahan, "CryptoAR: scrutinizing the trend and market of cryptocurrency using machine learning approach on time series data,". Indonesian Journal of Electrical Engineering and Computer Science, vol. 28, no. 3, pp. 1684-1696, 2022.

[23] S. Zhang, M. Li, and C. Yan, "The Empirical Analysis of Bitcoin Price Prediction Based on Deep Learning Integration Method,". Computational Intelligence and Neuroscience, 2022, 2022.

[24] I. Gurrib, and F. Kamalov, "Predicting bitcoin price movements using sentiment analysis: a machine learning approach,". Studies in Economics and Finance, vol. 39,no. 3, pp. 347-364, 2022.

[25] S. Cavalli, and M. Amoretti, "CNN-based multivariate data analysis for bitcoin trend prediction,". Applied Soft Computing, 101, pp. 107065, 2021.

[26] M. Liu, G. Li, J. Li, X. Zhu, and Y. Yao, "Forecasting the price of Bitcoin using deep learning,". Finance research letters, vol. 40, pp. 101755, 2021.

[27] S. Marne, S. Churi, D. Correia, and J. Gomes, "Predicting Price of Cryptocurrency-A deep learning approach,". NTASU-9 (3), 2020.

[28] R. Karlemstrand, and E. Leckström, "Using Twitter attribute information to predict stock prices,". arXiv preprint arXiv:2105.01402, 2021.

[29] B. A. Chandio, A. S. Imran, M. Bakhtyar, S. M. Daudpota, and J. Baber, "Attention-based RU-BiLSTM sentiment analysis model for Roman Urdu,". Applied Sciences, vol. 12, no. 7, pp. 3641, 2022.

[30] K. Irie, R. Csordás, and J. Schmidhuber, "The dual form of neural networks revisited: Connecting test time predictions to training patterns via spotlights of attention,". In International Conference on Machine Learning (pp. 9639-9659). PMLR, 2022.

[31] A. P. Rodrigues, R. Fernandes, A. Shetty, K. Lakshmanna, and R. M. Shafi, "Real-time twitter spam detection and sentiment analysis using machine learning and deep learning techniques,". Computational Intelligence and Neuroscience, 2022, 2022.

[32] N. R. Bhowmik, M. Arifuzzaman, and M. R. H. Mondal, "Sentiment analysis on Bangla text using extended lexicon dictionary and deep learning algorithms,". Array, vol. 13, pp. 100123, 2022.

[33] B. Aditya Pai, L. Devareddy, S. Hegde, and B. S. Ramya, "A time series cryptocurrency price prediction using lstm,". In Emerging Research in Computing, Information, Communication and Applications: ERCICA 2020, Vol. 2, pp. 653-662. Springer Singapore, 2022.

[34] D. Liu, and A. Wei, "Regulated LSTM Artificial Neural Networks for Option Risks,". FinTech, vol. 1, no. 2, pp. 180-190, 2022.

[35] M. Bilgili, N. Arslan, A. ŞEKERTEKİN, and A. YAŞAR, "Application of long short-term memory (LSTM) neural network based on deeplearning for electricity energy consumption forecasting,". Turkish Journal of Electrical Engineering and Computer Sciences, vol. 30, no. 1, pp. 140-157, 2022.

[36] C. Bai, "AGA-GRU: An optimized GRU neural network model based on adaptive genetic algorithm,". In Journal of Physics: Conference Series (Vol. 1651, No. 1, p. 012146). IOP Publishing, 2020.

[37] K. Zarzycki, and M. Ławryńczuk, "Advanced predictive control for GRU and LSTM networks,". Information Sciences, vol. 616, pp. 229-254, 2022.

[38] G. Xu, W. Guo, and Y. Wang, "Subject-independent EEG emotion recognition with hybrid spatio-temporal GRU-Conv architecture,". Medical & Biological Engineering & Computing, vol. 61, no. 1, pp. 61-73, 2023.

[39] H. Elfaik, and E. H. Nfaoui, "Deep bidirectional LSTM network learning-based sentiment analysis for Arabic text,". Journal of Intelligent Systems, vol. 30, no. 1, pp. 395-412, 2020.

[40] K. Yousaf, and T. Nawaz, "A deep learning-based approach for inappropriate content detection and classification of youtube videos,". IEEE Access, vol. 10, pp. 16283-16298, 2022.

[41] L. Eren, T. Ince, and S. Kiranyaz, "A generic intelligent bearing fault diagnosis system using compact adaptive 1D CNN classifier,". Journal of Signal Processing Systems, vol. 91, pp. 179-189, 2019.

[42] G. Xu, Y. Meng, X. Qiu, Z. Yu, and X. Wu, "Sentiment analysis of comment texts based on BiLSTM,". Ieee Access, vol. 7, pp. 51522-51532, 2019.