2021 IEEE/ACIS 6th International Conference on Big Data, Cloud Computing, and Data Science (BCD) | 978-1-7281-7681-9/21/$31.00 ©2021 IEEE | DOI: 10.1109/BCD51206.2021.9581397

A Cryptocurrency Prediction Model Using LSTM and GRU Algorithms

Jongyeop Kim
Department of Information Technology
Georgia Southern University
Statesboro GA, U.S.A
Jongyeopkim@georgiasouthern.edu

Seongsoo Kim
Department of Information Technology
Georgia Southern University
Statesboro GA, U.S.A
SK13153@georgiasouthern.edu

Hayden Wimmer
Department of Information Technology
Georgia Southern University
Statesboro GA, U.S.A
Hwimmer@georgiasouthern.edu

Hong Liu
Department of Computer Science
Indiana University Kokomo
HLius@iu.edu

***Abstract*— This study aims to predict cryptocurrency prices using Long Short-Term Memory(LSTM) and Gated Recurrent Unit(GRU) for three different coins: BitCoin, Ethereum, and Litecoin. For the training data for prediction, two data sets with different statistical characteristics in terms of Kurtosis and Skewness are used. LSTM and GRU models are trained and tested on the same hyperparameter configuration while increasing the number of epochs from 1 to 30. The accuracy of each model is measured by Root Mean Square Error (RMSE) and MAE (Mean Absolute Error). As a result of comparing GRU and LSTM, in BTC and ETH, the GRU was more advantageous for the downward stabilization trend, and the LSTM was suitable for the upward stabilization trend. However, in case of low-priced LTC, LSTM and GRU showed the same performance in sample type A, and in the case of type B, GRU was more accurate.**

Keywords— Cryptocurrency, LSTM, GRU, Deep learning, Big Data.

I. INTRODUCTION

Starting from the first cryptocurrency proposed by Satoshi Nakamoto[1], over 5,200 cryptocurrencies [2], including Bitcoin and Ethereum[3], are being traded on the market nowadays. Because of this apparent phenomenal in the cryptocurrency market, many researchers are seeking to predict the price of cryptocurrency using deep learning algorithms such as Long Short-Term Memory(LSTM)[4, 5], Gated Recurrent Unit(GRU)[6, 7], and Convolutional Neural Network(CNN)[8]. These algorithms are especially suitable for time series forecasting that estimates cryptocurrency prices at a particular time interval. Due to the nature that algorithms handle the complexity of a sequence dependence among the input variable.

To create a proper machine learning model required for cryptocurrency data analysis, various factors should be considered, such as training data selection, choosing the ideal number of nodes, number of layers, hyperparameter settings, and model evaluation methodology. High accuracy of the deep learning model is achievable when these kinds of conditions are appropriately configured. Among those factors, the most crucial parts of making an efficient machine learning model are selecting the proper training data set and choosing the suitable hyperparameter configurations.

Choosing the right training data set is essential because it is used as input data of the model. For instance, if an insufficient data set is used as training data, no matter how well-designed the model is, errors are amplified during the training process and eventually have a model trained by distorted values. Therefore, in the end, this model will reduce the accuracy of future price predictions when test data sets are applied.

Most of the researchers choose training data one-day time interval. However, it is not a valid data set to predict the coin market because it could drastically change over a few minutes or a second. Therefore, our data set was collected every ten seconds from the coinmarket.com site using web scraping technology to get a more accurate result. Moreover, unlike other researchers, we also consider the characteristics of Kurtosis and Skewness to determine the machine learning model's profer configuration parameters depending on data types. This distribution reflects the information features that the fluctuation range of the cryptocurrency price is large or small, or the price is rising or falling. Thus, in our research, two types of data samples are considered: Skewness and Kurtosis as training data sets for three virtual currencies.

Other than choosing the proper data set to get an effective machine learning model, one more thing to consider: selecting a training data set with the appropriate parameter configuration.

There are two kinds of parameters involved: model parameter and model hyperparameter. The former is a parameter that can be assumed from the data, and the latter is a parameter that cannot be estimated from the data[9]. Proper setting of this hyperparameter is essential because it controls the learning process and leads to an underfitting or overfitting state [10] of a model. Overfitting occurs when the details and noise of the training dataset are overtrained and negatively affect the model's performance. Underfitting occurs when the model cannot be generalized due to insufficient training or is not trained with enough data[11].

978-1-7281-7679-6/21/$31.00 ©2021 IEEE
ICIS2021-Fall, October 13-15, 2021, Xi'an, China

The financial data price prediction using LSTM [12, 13] [5, 14-16] or GRU [7] [17] has been proposed in various directions by many researchers. However, we could not find a paper that studied how the GRU and LSTM models work according to the change in the number of Ephcos with the characteristics of the data sample. Our contribution in this study, we selected two data sets with statistically different features for each coin and trained on LSTM and GRU under the same hyperparameter configuration while increasing the number of Epochs for comparison purposes. The rest of the paper is organized as follows: Section 2 introduces preliminary works and reviews; Section 3 proposes methodology; Section 4 shows the experimental results; and lastly, in Section 5, concluding remarks are drawn along with future works.

II. Preliminary

In this section, the characteristics of the dataset applied to our deep learning model were explained in terms of skewness and Kurtosis, and reviews of LSTM and GRU structures, Overfitting, and Underfitting were presented.

A. *Data Sampling*

The Sample data used as training data is collected from the CoinMarketCap [2] for a certain period of time. Three coins, namely Bitcoin(BTC), Ethereum(ETH), and Lite Coin(LTC), were collected. BTC and ETH are the coins with the most transaction volume, and LTC has low currency prices. As shown in Figures 1 and 2, BTC type A sample is data in which the price is on a downward stabilization trend, and in the case of type B is data that upward stabilized movement. In addition, piece trend plots for Ethereume and Lite coins are provided in the Appendix.

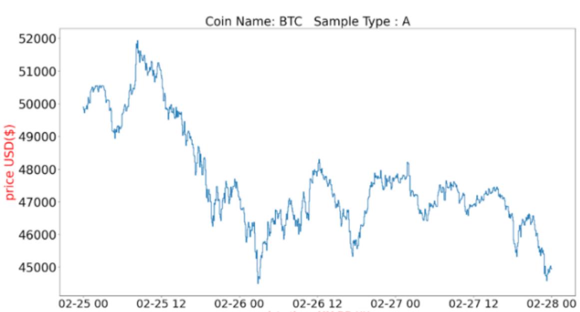

Figure 1. Bit Coin – Downward Stabilization trend

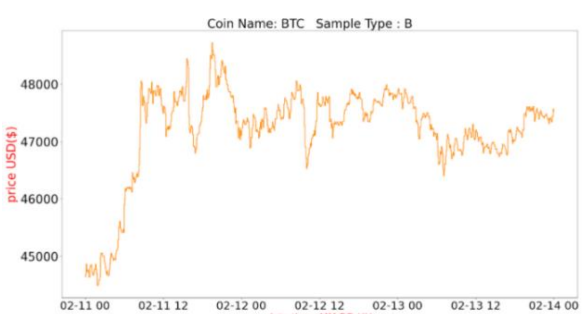

Figure 2. BitCoin – Upward Stabilization trend

B. *Skewness and Kurtosis of Sample*

Statistically, skewness measures the asymmetry of the probability distribution of a real-valued random variable about its mean [18, 19]. The skewness value can be positive, zero, negative, or undefined.

$$\tilde{\mu}_3 = E\left[\left(\frac{x-\mu}{\sigma}\right)^3\right] = \frac{\mu^3}{\sigma^3} = \frac{E[(x-\mu)^3]}{\left(E[(x-u^2)]\right)^{3/2}} = \frac{k_3}{k_2^{3/2}} \quad (1)$$

** where μ is the mean, σ is the standard deviation, E is the expectation operator, μ3 is the third central moment, and κt is the t-th cumulants.*

Kurtosis is a measure of the tailedness of the probability distribution of a real-valued random variable. Like skewness, Kurtosis describes the shape of a probability distribution and there are different ways of quantifying it for a theoretical distribution and corresponding ways of estimating it from a sample from a population [18]. Different measures of Kurtosis may have various interpretations.

$$kurt[x] = E\left[\left(\frac{x-\mu}{\sigma}\right)^4\right] = \frac{E[(x-u^4)]}{(E[(x-u)^2])^2} = \frac{u^4}{\sigma^4} \quad (2)$$

**where u^4 is the fourth central moment and σ is the standard deviation.*

We selected data sets that can represent those two statistical characteristics as input data of GRU and LSTM. As shown in Figures 3 and 4, in the case of Bitcoin, the sample type A has a Skewness: 0.788 a65nd Kurtosis: -0.153, respectively, and in type B, Skewness: 1.749 and Kurtosis: 2.941.

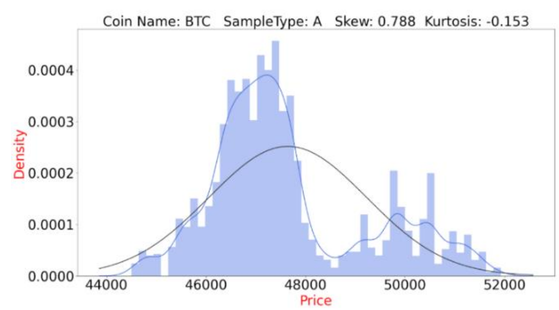

Figure 3. Bit Coin – Downward Stabilization trend

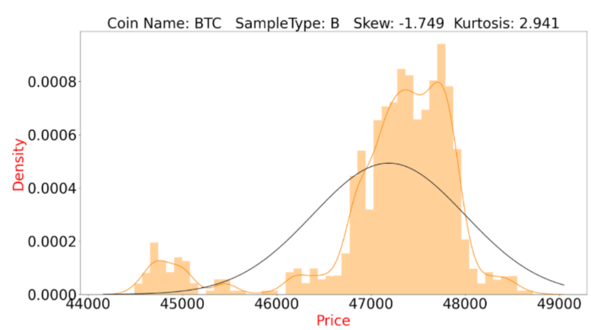

Figure 4. Bit Coin – Upward Stabilization trend

C. Long Short-Term Memory

LSTM is a novel recurrent network architecture conjunction with appropriate gradient-based learning algorithms introduced by Sepp Hochreiter and Jurgen Schmidhuber[20]. It is now widely used as a deep learning algorithm. It has four gates: one input gate, two-states gates, two output gates, and one forget gate. The forget gate is the gate that makes LSTM a particularly unique machine learning algorithm, among others. The following formula explains the behavior of LSTM by Olah [23].

$$f_t = \sigma\left(w_f \cdot [h_{t-1}, x_t] + b_f\right) \quad (1)$$

$$i_t = \sigma(w_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

$$\tilde{C}_t = tanh\ (w_C \cdot [h_{t-1}, x_t] + b_c) \quad (3)$$

$$C_t = f_t * C_{t-1} + it * \tilde{C}_t \quad (4)$$

$$o_t = \sigma(w_o \cdot [h_{t-1}, x_t] + b_o) \quad (5)$$

$$h\_t = o\_t * tanh\ (C\_t\ ) \quad (6)$$

Forget gate layer (1): the output of vector, f_t , determines either "remembered" or "forgotten"; Input gate layer(2): input layer decides values to update; Tanh layer(3): creates a vector of new candidate values and combines those two values (2) and (3). Update new cell state(4) from C_{t-1} to C. Multiply the old state by f_t , then add $i_t * \tilde{C}_t$. Put the cell state(5) through tanh and multiply it by the output of the sigmoid gate to produce result (6).

D. Gated Recurrent Network

The GRU is a newer generation of Recurrent Neural Network and it is similar to an LSTM. It only has two gates, a reset gate and update gate [21]. Compared to LSTM, GRU is faster in terms of speed because it has fewer Tensors, but it is not known which model is more efficient. In our study, there was a difference between accuracy and processing speed depending on the characteristics of the sample applied to each model.

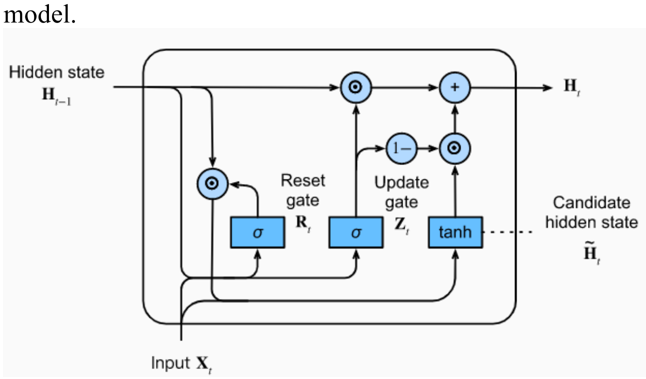

Figure 5. Hidden State in GRU model [22]

Suppose $X_t \in \mathbb{R}^{n\times d}$ (number of examples: n number of inputs : d) and previous hidden step $X_{t-1} \in \mathbb{R}^{n\times h}$ (number of hidden unit: h) then reset gate (1) $R_t \in \mathbb{R}^{n\times h}$ and update gate gate (2) $Z_t \in \mathbb{R}^{n\times h}$. Integrate the reset gate R_t it leads to the candidate hidden sate (3) [22].

$$R_t = \sigma(X_t W_{xr} + H_{t-1} W_{hr} + b_r) \ (1)$$

$$Z_t = \sigma(X_t W_{xz} + H_{t-1} W_{hz} + b_z) \ (2)$$

$$\tilde{H}_t = tanh(X_t W_{xh} + (R_t \Theta H_{t-1}) W_{hh} + b_h) \ (3)$$

**where $W_{xh} \in \mathbb{R}^{n\times h}$ and $W_{hh} \in \mathbb{R}^{h\times h}$ are weight parameter , $b_h \in \mathbb{R}^{1\times h}$ is the bias, symbol Θ is the Hadamard product.*

E. Overfitting and Underfitting

For the designed ML model to have an appropriate result value, the accuracy of the model must be verified. The model should be properly trained through the iterative training set in the dataset and hyperparameter. However, if you do not generalize to an insufficient data set or a new data set, it shows negative performance called undertraining. Conversely, overfitting can result if the model is overtrained by the noise included in the dataset [11]. The following exam shows Underfit(a), good fit(b), and Overfit(c).

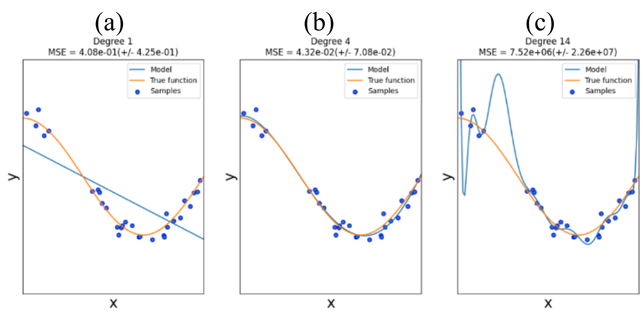

Figure 6. Underfitting vs. Overfitting

In this study, the dataset Type A and Type B data were used as training sets, and the RMSE was calculated by substituting the test set for increasing the number of epochs to the trained model.

III. Proposed Method

This section explains how we apply GRU and LSTM to three different currencies. As shown in Figure 6, (1) at the data extraction function extracts data from the MySQL database and loaded to the Pandas data frame by divided into type A and B for each Cryptocurrency(BTC, ETH and LTC). The data stored in the data frame is again divided into x_train and y_train, and the LSTM(3) and GRU(4) models are applied and trained according to the epoch parameter settings (3).

After the training is completed, RMSE and MAE values are obtained by applying the test data set to the trained model according to each epoch. Repeat this procedure until all epoch values are used.

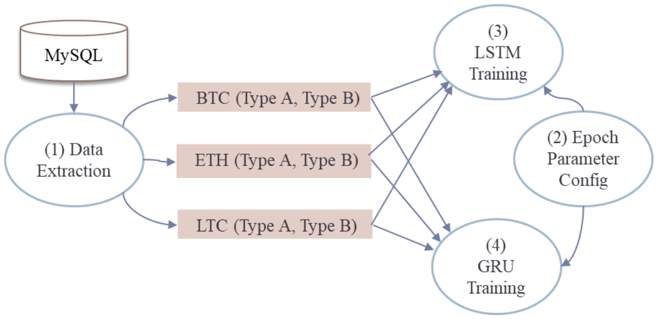

Figure 7. LSTM and GRU model System Flow Diagram

A. Data Collection Sampling

As shown in the table below, for each coin, data, Type A, Bitcoin:21694, Ethereume:21616, and Lite coin: 21569 were collected for three days from 2021-02-25 to February 28. In the case of Type A, the three coins are generally characterized by downward stability. In terms of skewness, they are all concentrated on the left of the figure. From the point of view of Kurtosis, it can be seen that the price fluctuation range of LTC is the largest.

- Sample Type A
  #DateTimeList = load_dateTime(CoinName, 'dateTime', '2021-02-25', '2021-02-28')

Coin / Dist.	Bit Coin	Ethereum	Lite Coin
Count	21694	21616	21569
Mean	47642.27	1512.18	177.83
Std	1585.20	62.39	8.86
Min	44497.44	1379.44	158.74
25%	46593.29	1468.56	172.56
50%	47248.61	1492.13	175.60
75%	48208.35	1540.86	179.49
Max	51927.28	1669.14	203.34
Skew	0.788	0.736	1.142
Kurtosis	-0.153	-0.408	0.874

Table 1: Characteristics of Sample Type A

- Sample Type B
  DateTimeList = load_dateTime(CoinName, 'dateTime', '2021-02-11', '2021-02-14'

Coins / Dist.	Bit Coin	Ethereum	Lite Coin
Count	21938	21938	21912
Mean	47188.38	1799.55	196.30
Std	808.29	33.93	13.640
Min	44488.96	1721.21	178.150
25%	47005.87	1774.68	185.620
50%	47361.74	1796.24	189.890
75%	47692.00	1825.49	204.530
Max	48718.08	1869.07	227.900
Skew	-1.749	0.012	0.777
Kurtosis	2.941	-0.640	-0.649

Table 2: Characteristics of Sample Type B

Sample Type B is collected for three days data from 2021-02-11 to 2021-02-14. The number of data collected is BitCoin: 21938, Ethereume:21938, and Litecoin:21912. The characteristics of these data are generally collected during periods of rising prices. In particular, in the case of bitcoin, it can be seen that the price fluctuation is very large with a kurtosis of 2.9s

B. Hyperparameter Configuration

The determination of hyperparameter values is essential as it has a decisive effect on the model's performance. The same parameters were used to compare GRU and LSTM under the same conditions. We applied two hidden layers and two dense layers determined by variable w,x. The batch is size 400. For the number of epochs, a total of 7 values from 1 to 30 were applied. Epoch = {1, 5, 10, 15, 20, 25, 30}. And the seed used fixed seed 30.

The pseudo-code below shows an example of the GRU model, and the LSTM uses the same hyperparameter configurations.

- GRU_Cryptocurrency

```
Input: w,x ← 60, y ← 25, z ← 1, b=400, e= [1, .., 30]
        GRU_x_train, GRU_y_Train
Output: An Array of Prediction P[ ]
GRU_Cryptocurrency( GRU_x_Train, GRU_y_Train, w, x , y, z, b,
e, seed )
{
   Set Seed ( seed )
   GRU_x_train , GRU_y_train
          = TdataCreat(T, r , len(T[]))
   GRU_model = sequential()
   GRU_model.add(GRU(w,
              input_shape=(GRU_x_train)))
   GRU_model.add(GRU(x))
   GRU_model.add(Dense(y))
   GRU_model.add(Dense(z))
   GRU_model.fit(GRU_x_train , GRU_y_train,
   batch_size=b, epochs=e)
   for i= r to len(T)
   {
      GRU_x_Test ←T[i – r, i]
   }
   P[ ]= GRU_model.predict(GRU_x_test)
   P[ ]= P[ ].inverse_transform

  GRU_Verity(P, D)
}
```

Pseudo Code 1: GRU_Cryptocurrency Prediction Model

C. Evaluating the Predictive model

The GRU and LSTM model trained whit sample data via hyperparameter configuration are ready for the prediction. For future price prediction, some portion of the data in the sample should be collected by the range parameter as a test data set and applied to each model. In this model, the range parameter r=10,000 is used to test data collecting. As shown in Tables(3,4,5), Type A and Type B data samples are trained as the Epoch value increases. Model accuracy of each model is evaluated in terms of Root Mean Square Error(RMSE) and Mean Absolute Error(MAE).

Figure 8, shows BTC prediction, red-colored line represent predicted by GRU, yellow line by LSTM, the blue line means real value. At the enlarged figure 8, when epoch =1, LSTM has more accurate value than GRU prediction. However, when the Epoch value 25, GRU reached an optimal point with RMSE:7.164 than LSTM:29.749.

Our model classified optimal Epoch values for each Cryptocurrency as follows:

- BTC Type A: LSTM (30), GRU (25)
  Type B: LSTM (20), GRU (15)
- ETH Type A: LSTM(15), GRU (30)
  Type B: LSTM(25), GRU (15)
- LTC Type A: LSTM(25), GRU (25)
  Type B: LSTM(30), GRU (5)

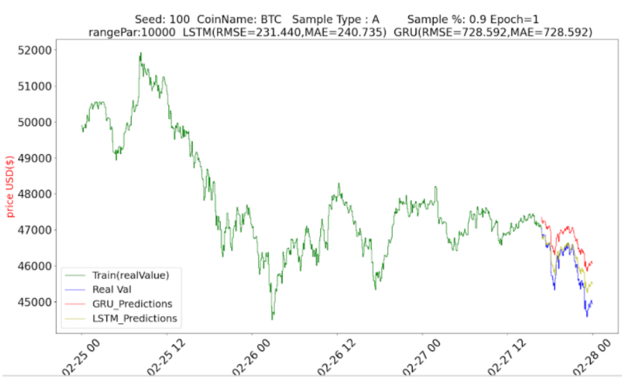

Figure 8. Prediction for BTC (Type A)

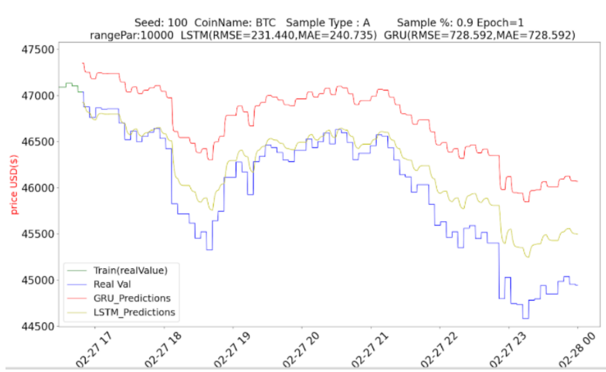

Figure 9. Prediction for BTC – enlarged (Type A, epo=1)

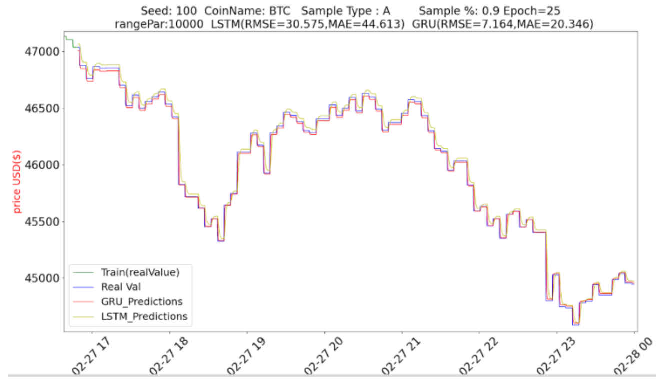

Figure 10. Prediction for BTC – enlarged (Type A, epo=25)

IV. EXPERIMENTAL RESULT AND ANALUSIS

We implemented LSTM and GRU. model on the Lambda Workstation connected to the Georgia Southern Lan with 1000M/bps. (GPU: 2x RTX 3090, CPU: i9-10980XE(18 core), Memory: 256GB, OS Ubuntu 12.04, Pycharm IDE)

A. Bitcoin Predictions

The price of a bitcoin is $47,188 on average, the highest among the three coins. For sample type A, the appropriate epoch value was 30 for LSTM and 25 for GRU. In the case of Type B, it was LSTM 20 and GRU 15.

Coins	Type &Epoch		LSTM		GRU	
	T*	E*	RMSE	MAE	RMSE	MAE
BTC	A	1	231.440	240.735	728.592	728.592
		5	43.159	58.395	39.816	45.479
		10	35.347	51.282	41.840	46.460
		15	54.631	65.408	13.003	22.677
		20	31.507	46.280	13.781	22.609
		25	29.749	43.982	7.164	20.346
		30	19.205	34.946	85.637	87.484
	B	1	25.054	49.229	209.640	209.651
		5	84.277	86.791	42.052	42.757
		10	33.467	36.289	27.273	29.435
		15	17.157	22.679	20.700	21.917
		20	15.223	20.944	45.045	46.255
		25	25.638	32.112	78.947	79.605
		30	33.960	36.255	28.272	29.032

(T*: Type of Sample, E*: Epoch)

Table 3: BTC- Model Evaluation Results

For BitCoin, if the price fluctuation range is higher, Sample type B, more trainings was required in both GRU and LSTM. When the initial value epoch is 1, LSTM has good efficiency, but as the epoch value increases, the GRU. First obtained good results, and the overtraining effect appeared as the epoch value increased thereafter, as shown in Table 3, Figures 4 and 5.

- Bit Coin: Sample Type A
  Skew : 0.788 Kurtosis: -0.153
  Optimal Epoch: LSTM (30), GRU(25)

- Bit Coin: Sample Type B
  Skew : -1.749   Kurtosis: 2.941
  Optimal Epoch: LSTM(20), GRU(15)

In BTC coin, when the price fluctuation range was small, the GRU. Performance was excellent, but when the price fluctuation range was immense, LSTM needed more epochs, but it was more efficient than GRU.

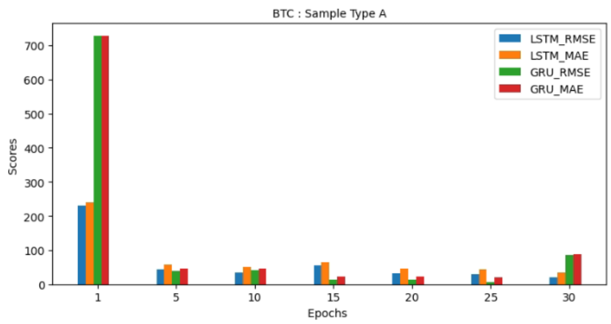

Figure 11. Bit Coin – Sample A Evaluation

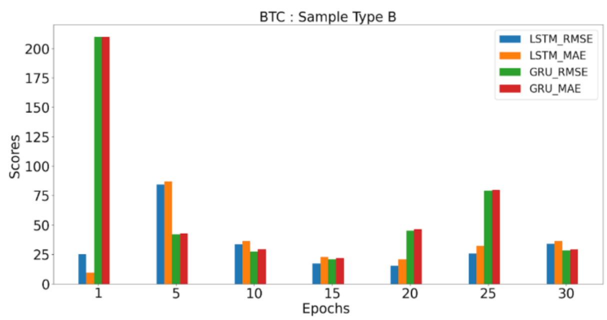

Figure 12. Bit Coin – Sample B evaluation

B. Ethereum Prediction

The price of Ethereum is $1799, with the median price of the three coins. In the case of sample type A, LSTM had good efficiency with epoch 15, but as training continued, the efficiency of GRU was higher with epoch 30.

- Ethereum: Sample Type A
  Skew : 0.736   Kurtosis: -0.408

- Ethereum: Sample Type B
  Skew : 0.012   Kurtosis: -0.640

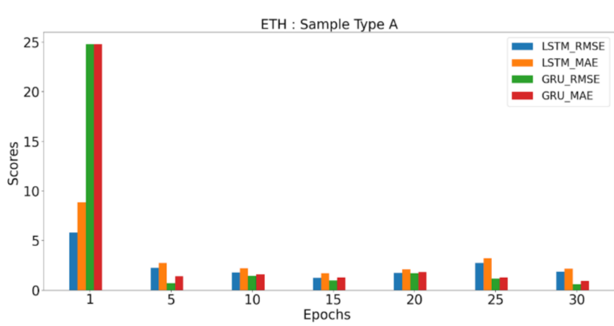

Figure 13. Ethereum – Sample A Evaluation

In Ethereum, the predicted values were more accurate than LSTM in type a and GRU in type B.

Coins	Type &Epoch		LSTM		GRU	
	T*	E*	RMSE	MAE	RMSE	MAE
ETH	A	1	5.776	8.824	24.756	24.756
		5	2.239	2.733	0.688	1.356
		10	1.761	2.169	1.409	1.553
		15	1.217	1.697	0.932	1.277
		20	1.73	2.09	1.70	1.81
		25	2.72	3.19	1.13	1.25
		30	1.85	2.15	0.58	0.90
	B	1	1.048	1.611	5.705	5.719
		5	1.232	1.533	1.187	1.292
		10	1.210	1.463	2.215	2.253
		15	1.365	1.612	0.553	0.651
		20	0.723	1.036	0.820	0.907
		25	0.314	0.715	0.670	0.758
		30	0.333	0.705	1.263	1.317

Table 4: ETH – Model Evaluation Results

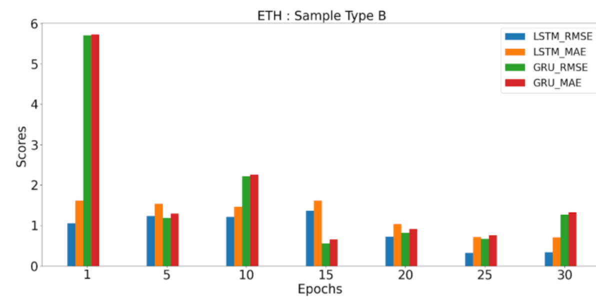

Figure 14. Ethereum – Sample B evaluation

C. Lite Coin Prediction

The price of a Lite Coin is $177 on average, the highest among the three coins

- Lite Coin: Sample Type A
  Skew : 1.142   Kurtosis: 0.874
- Lite Coin: Sample Type B
  Skew : 0.777   Kurtosis: -0.649

Coins	Type &Epoch		LSTM		GRU	
	T*	E*	RMSE	MAE	RMSE	MAE
LTC	A	1	1.39	1.46	4.04	4.04
		5	0.11	0.30	0.17	0.20
		10	0.06	0.23	0.23	0.28
		15	0.01	0.27	0.02	0.13
		20	0.04	0.25	0.08	0.13
		25	0.007	0.210	0.054	0.108
		30	0.415	0.453	0.091	0.138
	B	1	0.418	0.532	2.403	2.408
		5	0.799	0.836	0.102	0.152
		10	0.466	0.536	0.570	0.589
		15	0.821	0.863	0.283	0.309
		20	0.394	0.472	0.730	0.740
		25	0.338	0.408	0.532	0.548
		30	0.268	0.343	0.447	0.464

Table 5: LTC – Evaluation Results

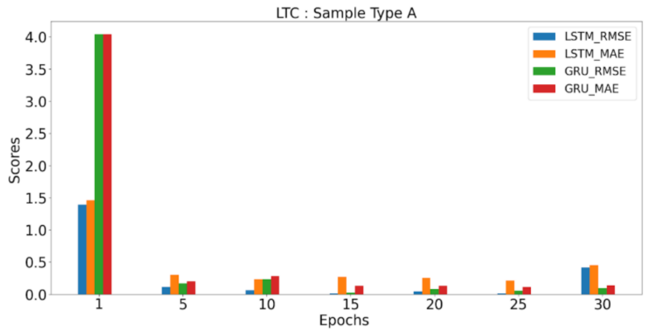

Figure 15. Lite Coin – Sample A Evaluation

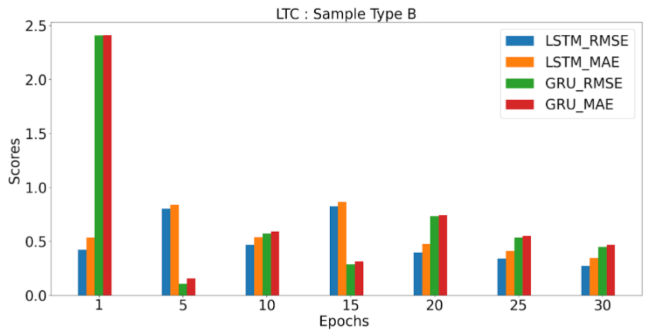

Figure 16. Lite Coin – Sample B Evaluation

V. Conclusion

This study is conducted to predict cryptocurrency prices for three coins: BTC, ETH, and LTC. Sample data with two different statistical characteristics were collected as a training data set to find appropriate algorithms for each data set. The trained model by GRU and LSTM on the same hyperparameter configuration varying epoch values are tested with the same test data which is chosen some portion of training data. The accuracy of each model was verified through RMSE, and an algorithm suitable for the characteristics of Type A and Type B was identified along with its Epoch value per coin. In BTC and ETH, the GRU was more advantageous for the downward stabilization trend, and the LSTM was suitable for the upward stabilization trend. However, in case of low-priced LTC, LSTM and GRU showed the same performance in sample type A, and in the case of type b, GRU was more accurate.

References

[1] Nakamoto, S.: 'Bitcoin: A peer-to-peer electronic cash system', in Editor (Ed.)^(Eds.): 'Book Bitcoin: A peer-to-peer electronic cash system' (Manubot, 2019, edn.), pp.

[2] CoinMarketCap: 'Today's Cryptocurrency Prices by Market Cap', 2021

[3] Wood, G.: 'Ethereum: A secure decentralised generalised transaction ledger', Ethereum project yellow paper, 2014, 151, (2014), pp. 1-32

[4] Kwon, D.-H., Kim, J.-B., Heo, J.-S., Kim, C.-M., and Han, Y.-H.: 'Time Series Classification of Cryptocurrency Price Trend Based on a Recurrent LSTM Neural Network', Journal of Information Processing Systems, 2019, 15, (3)

[5] Nelson, D.M., Pereira, A.C., and de Oliveira, R.A.: 'Stock market's price movement prediction with LSTM neural networks', in Editor (Ed.)^(Eds.): 'Book Stock market's price movement prediction with LSTM neural networks' (IEEE, 2017, edn.), pp. 1419-1426

[6] Rahman, M.O., Hossain, M.S., Junaid, T.-S., Forhad, M.S.A., and Hossen, M.K.: 'Predicting prices of stock market using gated recurrent units (GRUs) neural networks', Int. J. Comput. Sci. Netw. Secur, 2019, 19, (1), pp. 213-222

[7] Dutta, A., Kumar, S., and Basu, M.: 'A gated recurrent unit approach to bitcoin price prediction', Journal of Risk and Financial Management, 2020, 13, (2), pp. 23

[8] Gudelek, M.U., Boluk, S.A., and Ozbayoglu, A.M.: 'A deep learning based stock trading model with 2-D CNN trend detection', in Editor (Ed.)^(Eds.): 'Book A deep learning based stock trading model with 2-D CNN trend detection' (IEEE, 2017, edn.), pp. 1-8

[9] Brownlee, J.: 'What is the Difference Between a Parameter and a Hyperparameter', Machine Learning Mastery) Retrieved February, 2019, 24, pp. 2020

[10] Koehrsen, W.: 'Overfitting vs. Underfitting: A Complete Example', Towards Data Science, 2018

[11] Brownlee, J.: 'Overfitting and Underfitting With Machine Learning Algorithms', 2019

[12] Livieris, I.E., Kiriakidou, N., Stavroyiannis, S., and Pintelas, P.: 'An Advanced CNN-LSTM Model for Cryptocurrency Forecasting', Electronics, 2021, 10, (3), pp. 287

[13] Yamak, P.T., Yujian, L., and Gadosey, P.K.: 'A comparison between arima, lstm, and gru for time series forecasting', in Editor (Ed.)^(Eds.): 'Book A comparison between arima, lstm, and gru for time series forecasting' (2019, edn.), pp. 49-55

[14] Selvin, S., Vinayakumar, R., Gopalakrishnan, E., Menon, V.K., and Soman, K.: 'Stock price prediction using LSTM, RNN and CNN-sliding window model', in Editor (Ed.)^(Eds.): 'Book Stock price prediction using LSTM, RNN and CNN-sliding window model' (IEEE, 2017, edn.), pp. 1643-1647

[15] Ferdiansyah, F., Othman, S.H., Radzi, R.Z.R.M., Stiawan, D., Sazaki, Y., and Ependi, U.: 'A LSTM-Method for Bitcoin Price Prediction: A Case Study Yahoo Finance Stock Market', in Editor (Ed.)^(Eds.): 'Book A LSTM-Method for Bitcoin Price Prediction: A Case Study Yahoo Finance Stock Market' (IEEE, 2019, edn.), pp. 206-210

[16] Wu, C.-H., Lu, C.-C., Ma, Y.-F., and Lu, R.-S.: 'A new forecasting framework for bitcoin price with LSTM', in Editor (Ed.)^(Eds.): 'Book A new forecasting framework for bitcoin price with LSTM' (IEEE, 2018, edn.), pp. 168-175

[17] Shen, G., Tan, Q., Zhang, H., Zeng, P., and Xu, J.: 'Deep learning with gated recurrent unit networks for financial sequence predictions', Procedia computer science, 2018, 131, pp. 895-903

[18] Brown, S.: 'Measures of shape: Skewness and kurtosis', Retrieved on August, 2011, 20, pp. 2012

[19] Wikipedia: 'Skewness', 2021

[20] Hochreiter, S., and Schmidhuber, J.: 'Long short-term memory', Neural computation, 1997, 9, (8), pp. 1735-1780

[21] Phi, M.: 'Illustrated Guide to LSTM's and GRU's: A step by Step explanation ', 2018

[22] Aston Zhang, Z.C.L., Mu Li and Alexander J. Smola: 'Dive Into Deep Learning ', 2021, pp. 9.1 Gated Recurrent Units

[23] Olah, C.: 'Understanding lstm networks', 2015

APPENDIX

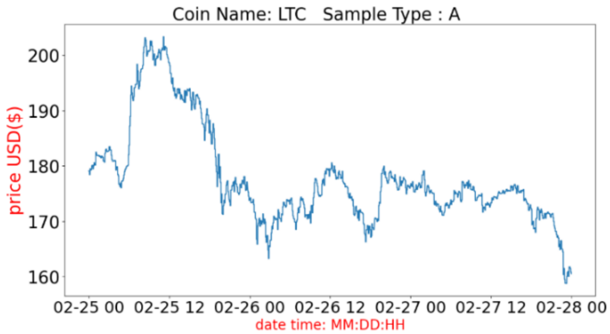

Figure 17. Lite Coin – Downward Stabilization trend

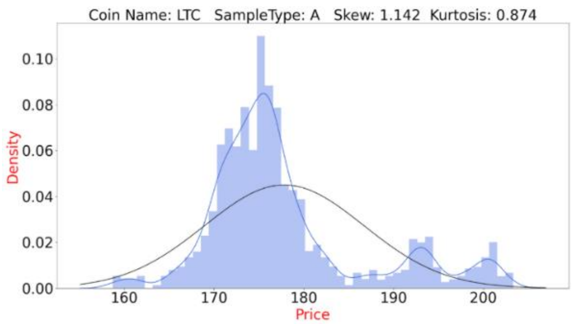

Figure 18. Lite Coin – Downward Skew & Kurtosis

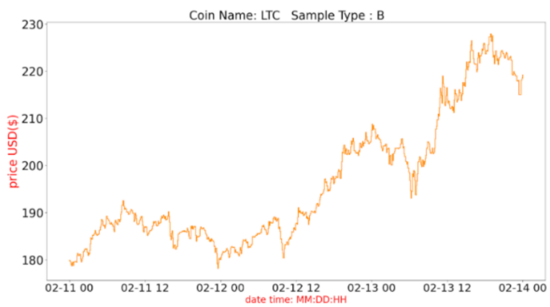

Figure 19. Lite Coin – Upward Stabilization trend

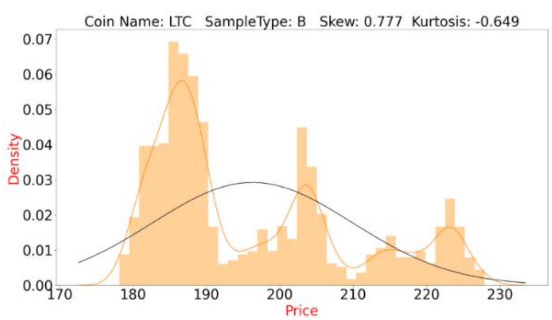

Figure 20. Lite Coin – Upward Skew & Kurtosis

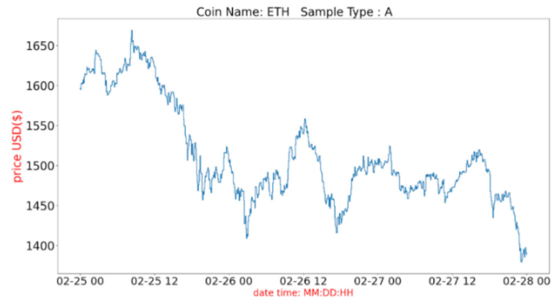

Figure 21. Ethereum – Downward Stabilization trend

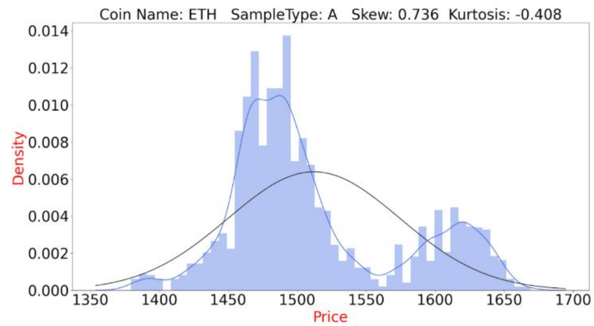

Figure 22. Ethereum– Downward Skew & Kurtosis

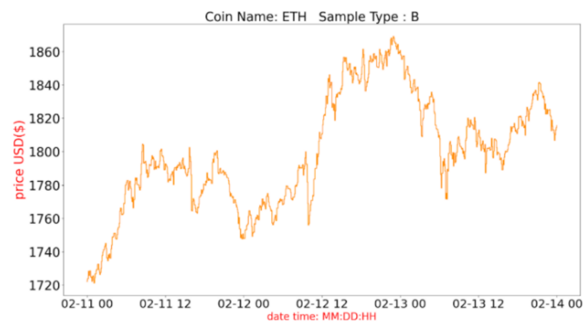

Figure 23. Ethereum – Upward Stabilization trend

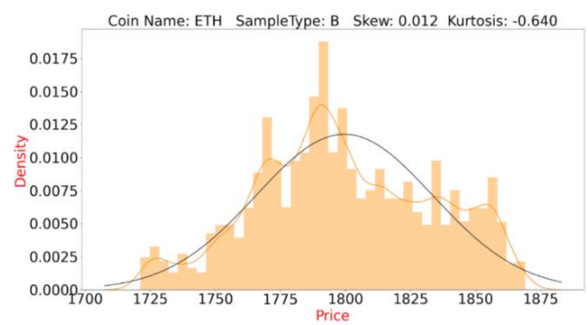

Figure 24. Ethereum – Upward Skew & Kurtosis