The Effect of Loss and Optimization Functions on Bitcoin Rate Prediction in LSTM

Berke Kaan Kirci
Department of Mathematics and Computer Sciences
Istanbul Kultur University
Turkey, Istanbul
berkekirci@gmail.com

Gozde Karatas Baydogmus
Department of Computer Engineering
Marmara University
Turkey, Istanbul
gkaratas@marmara.edu.tr

2022 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA) | 978-1-6654-6835-0/22/$31.00 ©2022 IEEE | DOI: 10.1109/HORA55278.2022.9799928

***Abstract*— In recent years, Bitcoin cryptocurrency has become a growing trend in the world. For this reason, researchers from many fields are examining various artificial intelligence models to predict Bitcoin rates. In particular, Deep Learning algorithms have been shown to outperform traditional models in predicting cryptocurrency rates. However, very few studies have examined the effect of parameters used in deep learning algorithms on the algorithm. Optimization and loss functions are very important, which affect the algorithm's ability to make a successful prediction. In this study, Long-Short Term Memory, a deep learning algorithm, is used to predict daily Bitcoin prices and the effect of optimization/loss functions on the accuracy rate is evaluated. Experimental results showed that the Long-Short Term Memory model made the best predictions as a result of working with the Adam optimization function and the Mean Square Error loss function.**

Keywords—bitcoin, LSTM, loss function, optimization function

I. INTRODUCTION

Bitcoin is a crypto currency that belongs to anyone, used for internet payments or investment purposes around the world. The fact that it does not belong to anyone means that the transactions are not tied to any country. Investment using Bitcoin can be made through various marketplaces. These marketplaces are called "bitcoin exchanges" and allow people to sell/buy Bitcoins using different currencies. Bitcoins are stored in a digital wallet that basically looks like a virtual bank account, a record of all transactions made with bitcoin, timestamp data in a place called Blockchain. As the crypto currency system became popular in the world, it started to attract attention in many areas. Such that El Salvador also assigned Bitcoin as its official currency unit.

Bitcoin buying/selling enthusiasts constantly follow the agenda and shape their investments about the future situation. It is not a fixed exchange rate like the main currencies of other countries, since it is in a situation that changes rapidly at every moment, it is tried to make future forecasts with various methods. In this sense, deep learning algorithms have come to the fore in recent years for examining data and making predictions on it. Through these algorithms and the parameters they contain, certain predictions are made by making analyzes on data taken from the real world.

In this study, the deep learning algorithm of the Bitcoin market, which has become an increasing trend in the world, is targeted with Long-Short Term Memory (LSTM) which is deep learning algorithm and the effect of recording and optimization functions in making the closest prediction to the real data is examined. As a result of the study, it is aimed that people who invest in Bitcoin can make predictions about their investments without being victimized.

In the next part of the study, related work are given in Section 2, information about the dataset, LSTM, loss functions and optimization functions are given in Section-3, experimental results are carried out in Section 4, and the results of the study are evaluated in Section 5. The aim of the study is to give an idea to researchers in estimating Bitcoin rates.

II. RELATED WORK

Various studies has done for this topic and still being done at this time because the topic is relevant. The topic changes and the artificial intelligence only improving overtime.If we briefly look at the done researches for this topic;

A study done by Rizwan and others in 2019 focuses on the prediction of the closest prices of bitcoin while comparing a RNN model GRU and LSTM model. Their model has not multi layered and no specific optimizers. LSTM has 52% total accuracy score and %8 RMSE [1]. Adyan Marendra Ramadhani, and this friends created RBM model that has 3 layers. This study also uses Adam and MSE optimizers and has 89% accuracy score [2]. Another study has done by Orhan G. Yalcin. The model he build, uses LSTM as, RMSprop and mean_squared_error as optimizers and with these parameters it has just one layer. He reached 87% accuracy score in the experimental results [3]. Next work has done by Edwin Sin and Lipo Wang. They used MLP model for Bitcoin rate prediction. Their model shown bad results compared to other studies [4]. Minakhi et al. Did research about Bitcoin rate prediction using deep learning algorithms. Research's main idea is to compare DNN models while predicting bitcoin prices by daily basis. For conclusion their research proves that LSTM and GRU has the best accuracy amongst the other models [5]. The next work has done by Zheshi Chen et al. Their model uses LSTM but ADAgrad as a optimizer. They also used MSE as a Loss optimizer for the study [6]. In 2019 researchers

did a study using DNN and their effects on the Bitcoin rate prediction. Their main idea of the research is to compare different such as DNN, LSTM etc. algorithms while predicting daily bitcoin prices. On their DNN research the best model that has the closest prices to bitcoin prices is the LSTM model [7].

III. PROPOSED WORK

A. System Specs

The computer that used for the study has I7-4720hq as processor, Nvidia GTX 950M as graphics card and 12GB Ram. All the codes written with Python programming languages and its libraries such as Pandas, Numpy and Sklearn.

B. Dataset

Dataset used in the research is collected from Yandex! finance API. Since the dataset comes from an API that updates itself every day, it stays up-to-date so that the results obtained in the study can be checked in real time.

Dataset used for our work has BTC/USD data. Also it includes bitcoin price data from 01.01.2016 till March. The dataset that we used has daily closing prices of BTC for US clocks and that means the dataset has over 2300 prices by March.

C. Deep Learning Model - LSTM

LSTM used for this study because Bitcoin dataset is sequential and the use of LSTM is ideal for such datasets. Our model contains 4 LSTM layers, 4 dropout layers, 32 epochs and as a batch size 64 batches. We used more layers than usual because layers disables overfitting considerable amount.

1) Optimization Function

For optimization functions, we picked five different optimizers for out accuracy tests. These loss optimizers were SGD, Adagrad, Adadelta, RMSprop and Adam.

- SGD: SGD is an optimization algorithms which is iterative and first-order.
- Adagrad: It is an optimization algorithm that keeps track of the sum of the squared gradients.
- Adadelta: Adadelta is an extension of Adagrad that seeks to reduce its aggressive learning rate.
- RMSprop: RMSprop fix the issue of AdaGrad that has by multiplying distort rate to the cumulative sum.
- Adam: Adam is the developed version of the RMSprop which has done by taking the first and second momentum of the gradient separately.

2) Loss Function

In this study we picked five random loss optimizers for accuracy tests. These loss optimizers were MSE, MAE, MAPE, Cosine Similarity and Huber. To get in to details;

- MSE: mediums the average of the squares of the errors.
- MAE: is medium of errors between paired observations.
- MAPE: is a median of the accuracy of a forecasting method.
- Cosine Similarity: is a median of similarity between two sequences of numbers.
- Huber: is a combination of the mean squared error function and the absolute value function.

D. Improved Model

Our main idea for this model is to compare BTC prediction optimizers and calculate the every optimizer accuracy for optimum accuracies. We used the dataset that given in the Dataset section and with LSTM model using different optimizers we predicted the BTC price for 30 days. The accuracy rate was calculated as given by equation (1).

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \quad (1)$$

In this equation TP is true positive, TN is true negative, FP is false positive and FN is false negative.

IV. EXPERIMENTAL RESULTS

Combinations of optimization and loss functions have been tried in various ways in the study and a general conclusion has been made. Accordingly, the performances of the optimization functions were initially examined and all other parameters were left as standard.

TABLE I. PERFORMANCE COMPARISON BETWEEN OPTIMIZERS ACCURACY RATES

Optimizers	Accuracy (%)
Adam	93.54
SGD	92.49
RMSprop	87.13
ADAgrad	59.68
ADAdelta	43.25

Table I shows the accuracy rates of the analyzed optimization functions. These rates are shown in Figure-1 for a better understanding of the difference between them.

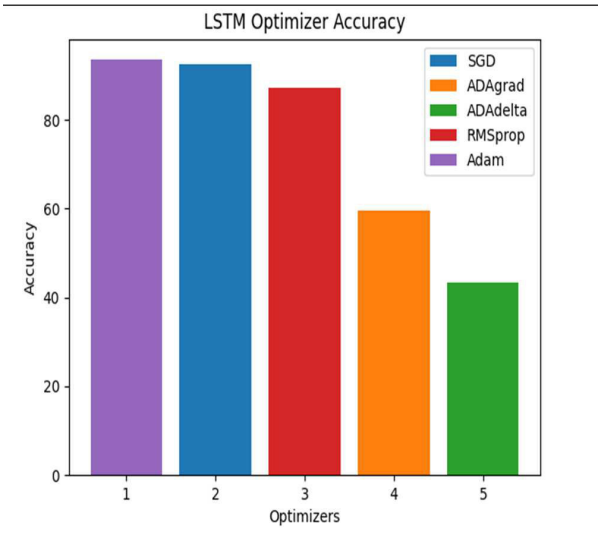

Figure 1. *Optimizer Functions Accuracy*

It is seen that the most successful result is achieved with the Adam optimization function when the data is examined.

After the optimization function was determined as Adam, the accuracy rates of working with loss functions were evaluated. Table II shows the accuracy rates of the loss functions working with the Adam optimization function. In order to interpret the results, the accuracy rates are also added in the form of graphics as in Figure 2.

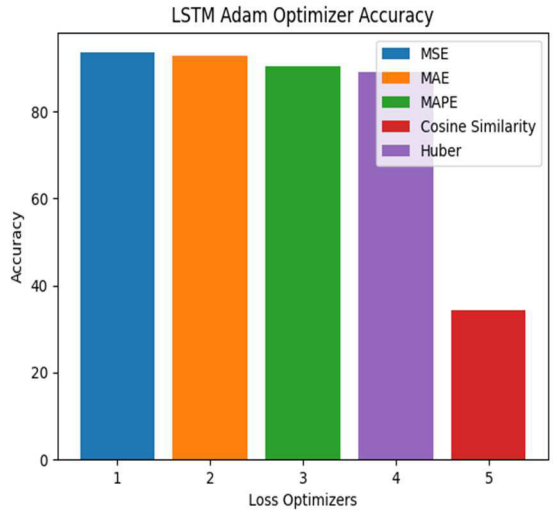

Figure 2. *Adam Optimizer Accuracy Rates*

It is seen that the MSE and MAE algorithms are the most successful functions in working with the Adam function when the results are examined.

In such studies, it is very important to examine the running time of the algorithm along with the accuracy rates. An algorithm that takes days to run may seem inefficient, even if its accuracy is high. Therefore, the running time of operations with optimization functions is shown in Figure 3.

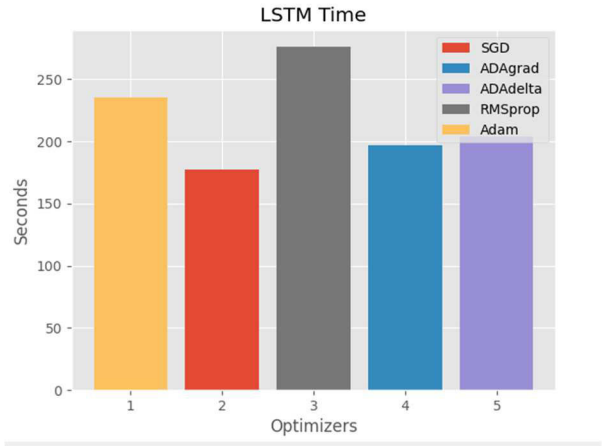

Figure 4. *Optimizer functions time complexity*

In addition, the working time of the loss functions examined using the Adam function is shown in Figure 4.

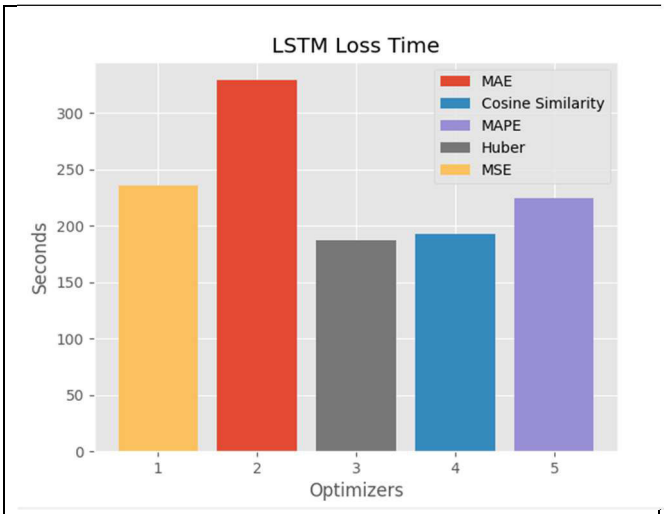

Figure 3. *Loss Functions Time Complexity with Adam optimizer*

When these figures are examined, it is seen that the Adam optimization function and the loss function MSE, which has the highest accuracy, have an acceptable running time.

V. Conclusion

In this study, we examined the effect of optimization and loss functions in a deep learning-based Bitcoin price prediction model using Bitcoin cryptocurrency information. For this, experiments were carried out with various optimization and loss functions using the LSTM deep learning algorithm. The accuracy rate and time complexity of each function were examined separately. As a result of the study, it was seen that the Adam optimization function was more successful than the other functions. Thereupon, the evaluation was carried out by examining its work with various loss functions. And as a result, it has been observed that the use of Adam and MSE functions in the LSTM algorithm highly affects the prediction rate. Through this study, researchers can done easier operations on optimization and loss function selection in other deep learning algorithms.

In future studies, it is aimed to prepare a detailed hyper-parameter study on popular deep learning algorithms and to prepare an article that will evaluate the effect on Bitcoin rate increase.

References

[1] Rizwan, M., Narejo, S., & Javed, M. (2019, December). Bitcoin price prediction using deep learning algorithm. In *2019 13th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics (MACS)* (pp. 1-7). IEEE.

[2] Ramadhani, A. M., Kim, N. R., Lee, T. H., & Ryu, S. E. (2018). Bitcoin Price Forecasting Using Neural Decomposition and Deep Learning. *Journal of the Korea Industrial Information Systems Research*, *23*(4), 81-92.

[3] Cryptocurrencies, T. O., & Yalçin, O. G. (2019). Sciences & Humanities.

[4] Sin, E., & Wang, L. (2017, July). Bitcoin price prediction using ensembles of neural networks. In *2017 13th International conference on natural computation, fuzzy systems and knowledge discovery (ICNC-FSKD)* (pp. 666-671). IEEE.

[5] Awoke, T., Rout, M., Mohanty, L., & Satapathy, S. C. (2021). Bitcoin price prediction and analysis using deep learning models. In *Communication Software and Networks* (pp. 631-640). Springer, Singapore.

[6] Chen, Z., Li, C., & Sun, W. (2020). Bitcoin price prediction using machine learning: An approach to sample dimension engineering. *Journal of Computational and Applied Mathematics*, *365*, 112395.

[7] Ji, S., Kim, J., & Im, H. (2019). A comparative study of bitcoin price prediction using deep learning. *Mathematics*, *7*(10), 898.