Avrupa Bilim ve Teknoloji Dergisi
Özel Sayı 37, S. 72-75, Haziran 2022

Araştırma Makalesi

www.ejosat.com ISSN:2148-2683

European Journal of Science and Technology
Special Issue 37, pp. 72-75, June 2022

Research Article

Forecasting Cryptocurrency Prices Using Long Short-Term Memory

Fırat Erdoğdu[1], Ufuk Cebeci[2*]

[1] Istanbul Technical University, Management Faculty, Industrial Engineering Department, İstanbul, Türkiye (ORCID: 0000-0003-4367-6206)
[2] Istanbul Technical University, Management Faculty, Industrial Engineering Department, İstanbul, Türkiye (ORCID: 0000-0002-7589-9239)

(5th International Symposium on Innovative Approaches in Smart Technologies– 28-29 May 2022)

(**DOI:** 10.31590/ejosat.1134210)

ATIF/REFERENCE: Erdoğdu, F. & Cebeci, U. (2022). Forecasting Cryptocurrency Prices Using Long Short-Term Memory. *European Journal of Science and Technology*, (37), 72-75.

Abstract

Since the 1950s a discipline called 'Artificial Intelligence' has been gaining significant popularity. The curiosity about creating computers that can think and produce information like human beings has allowed scientists and computer engineers to contribute to this field. Many components such as robots, softwares and algorithms have been produced due to this purpose. Like various disciplines, Artificial Intelligence has been branched into several sub-disciplines. One of these branches is named 'Machine Learning'. Machine Learning has different types of sub-branches such as Supervised Learning, Unsupervised Learning and Deep Learning. Deep Learning is the main Machine Learning technique used in this study. The ability to cope with complex situations allows Deep Learning models to be used in different application areas widespread. Predicting cryptocurrency prices can be counted as one of these applications. Because of investors' desire to observe the cryptocurrency prices trend and reduce the investment risk using an effective method is becoming crucial. For this purpose, we created a Long Short-Term Memory which is a type of Deep Learning with the appropriate parameters via Python programming language. The dataset which is used to feed this model was obtained from the internet. After running the algorithm with this dataset, the validity of the model is calculated by a statistical tool called Mean Square Error. To visualize the effectiveness of the model's output, a Python programming language library known as Matplotlib was chosen. Also, after the reviewing results of the model required interpretations and information about future studies will be explained by us in the Conclusion chapter.

Keywords: *Cryptocurrency, Deep Learning, Long Short-Term Memory, Forecasting*

Uzun Kısa-Süreli Bellek ile Kripto Para Fiyatlarının Tahmini

Öz

Yapay Zeka 1950'lerden buy ana önemli popülerlik kazanan disiplinlerden birisidir. İnsanlar gibi düşünen ve bilgi üreten bilgisayarların tasarlanmak istemesi bilim insanları ve bilgisayar mühendislerinin bu alana katkı yapmasına olanak tanımıştır. Robotlar, yazılımlar ve algoritmalar gibi araçlar bu amaç doğrultusunda üretilmiştir. Birçok diğer disiplin gibi Yapay Zeka'da farklı alt disiplinlere ayrılmıştır. Bu dallardan biri 'Makine Öğrenmesi' olarak adlandırılmıştır. Makine Öğrenmesi'nin Gözetimli Öğrenme, Gözetimsiz Öğrenme ve Derin Öğrenme gibi alt disiplinleri vardır. Derin Öğrenme bu çalışmada kullanılan Makine Öğrenmesi tekniğidir. Kompleks durumlarla başa çıkma özelliği Derin Öğrenme modellerinin farklı alanlarda yaygın olarak kullanılmasına olanak tanır. Kripto para fiyatlarını tahmin etmek bu uygulamalara örnek olarak verilebilir. Yatırımcıların kripto para fiyatlarındaki trendi gözlemleme ve yatırım riskini azaltma isteği sebebiyle etkin bir yöntem kullanmak oldukça önemlidir. Bu amaçla Derin Öğrenme'nin bir çeşidi olan uygun parametrelere sahip Python programlama dili ile yazılmış bir Uzun Kısa-Süreli Bellek modeli bizler tarafından oluşturulmuştur. Modeli beslemek için kullandığımız veri seti internetten alınmıştır. Bu veri seti ile algoritma çalıştırıldığı zaman modelin geçerliliğini hesaplamak için ise Ortalama Kare Yöntemi denilen bir istatistiksel araca başvurulmuştur. Modelin çıktılarını görselleştimesi içinse Python dili için yazılmış Matplotlib kütüphanesi tercih edilmiştir. Bunun yanı sıra, modeli değerlendirdikten sonra gerekli yorumlar ve geleceğe yönelik çalışmalar hakkında bilgi tarafımızca sonuçlar kısmında sağlanacaktır.

Anahtar Kelimeler: *Kripto Para, Derin Öğrenme, Uzun Kısa-Süreli Bellek, Tahmin*

1. Introduction

After World War II, many countries decided to accelerate their developments to become victorious against their opponents. Some important and era changer scientific researches were conducted during these years. Artificial Intelligence (AI) was one of the prominent fields that were created as a result of the efforts. Especially, studies conducted by famous computer scientist and mathematician called Alan Turing can be counted as pioneers of artificial intelligence technology. The article called 'Computing Machinery and Intelligence' by Alan Turing is published in 1950 and has been considered a revolutionary work by many people. [1] After the publication of this article, several works maintained by various scientists and experts have started to increase dramatically. However, in the 70s some opposing views arose. These views included negative consequences that were thought to have a high probability of occurrence. As a result of the negative criticisms studies related to Artificial Intelligence were halted for a while and developments related to this field could not appear significantly. But, when the 1980s arrived the negative approach toward Artificial Intelligence has started to change. Many studies have begun to arise and new methods have been created during these years. This trend continued in the 1990s. Some inventions such as Deep Blue showed the world that machines can behave like human beings. For example, the mentioned computer 'Deep Blue' could be able to defeat Garry Kasparov, one of the greatest chess players, in a chess game. This ability can be assessed as incredible, but it is not exactly like that. Because most Artificial Intelligence products and applications can handle only one task. To improve these technologies and applications, AI is divided into a few sub-disciplines. Machine Learning (ML) is one of the branches of AI. Machine Learning is a method that aiming computers to learn and produce output on their own. Like AI, Machine Learning has also a few sub-branches. One of the components of these branches is called 'Deep Learning'. Deep Learning is a Machine Learning Method that imitates the human brain. As the human brain uses its nerve cells for transporting impulses, in Deep Learning models artificial neural networks are executing the same operation. Some components of Deep Learning models such as Layers, Nodes and Activation Functions are assisting the model to cope with the data handling process. Thus, Deep Learning models become able to solve complex problems and produce a valid output. This ability is the main reason for the widespread utilization of Deep Learning models.

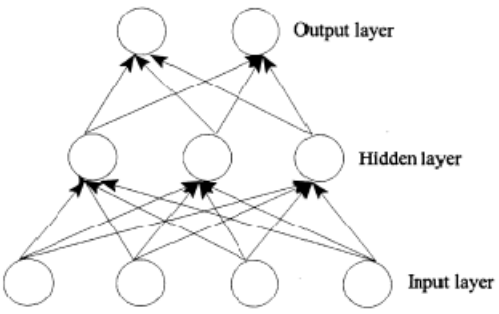

Figure 1. General Structure of a DL Model [2]

On the other hand, it is seen that cryptocurrencies have been gaining popularity in the last decade. Cryptocurrencies are virtual money types that were mainly created for escaping government control policies on economical activities. Some types such as Bitcoin, Litecoin and Ether can be given as essential examples of cryptocurrencies. They caught the attention of many people when they were first introduced to the market area but, their popularity skyrocketed especially in 2017. Since this year, many investors put their money into cryptocurrencies to obtain more profit. At this point, establishing some methods to reduce the investment risk becomes more crucial for related people. Deep Learning is one of the best alternatives for this purpose. The Ability of processing datasets for complex issues, allows Deep Learning Model to produce valid outcomes. So in this study, we will try to create a Deep Learning model that can predict cryptocurrency values according to the date of each element of the dataset after being trained by a historical dataset.

2. Material and Method

2.1. Material

For establishing the price prediction Deep Learning model, a proper dataset is required for the training of the model. As mentioned in the abstract, the appropriate dataset which is containing relevant information about cryptocurrency values is obtained from a website called 'https://min-api.cryptocompare.com/'. On this website, users can find data tables of cryptocurrencies which is including some necessary information such as date, highest-lowest price and closed value. These types of information can be obtained through a link provided by this website. This link was used while gathering the dataset for the price forecasting model. Moreover, users can obtain the dataset for different types of currencies such as the US Dollar, Euro and Canadian Dollar. For our model, we have chosen Euro as the cryptocurrency price to be converted into and Bitcoin as the correspondent cryptocurrency type. To see the dataset more clearly, the table below should be consulted:

Table 1. First Five Elements of Our Dataset

time	high	low	open	volumefrom	volumeto	close
2020-12-30	36960.45	34936.99	35231.98	132.96	4791836.13	36920.20
2020-12-31	37191.61	35448.55	36920.20	144.15	5262525.40	36626.94
2021-01-01	37288.09	35899.74	36626.94	99.59	3644271.18	36701.49
2021-01-02	40728.34	35192.06	36701.49	211.87	8156208.77	39585.56
2021-01-03	42233.31	34462.09	39585.56	245.68	9734913.75	40353.36

We have a total of 501 components in the dataset. The list of the dataset is very long, so that is why only the first five data are printed above. Each data is consisting of seven titles and titles except time and close are our model's input while 'close' is the target. After obtaining the results 'Time' expression will be used for creating graphics to understand the model validity better.

2.2. Method and Application

For establishing the model, we decided to follow four steps. These steps are including obtaining real-time data on cryptocurrency prices, splitting the dataset into training and test parts, conducting a prediction operation using the Deep Learning technique and assessing the validity of the model. The data for the model was gathered from the relevant website, so the first step was accomplished. To execute the remaining step we decided to use the Python programming language which we launched from Google COLAB. After importing the dataset to the COLAB interface we split the whole dataset into 'training' and 'test' groups. %80 of the dataset was decided to be the training set and the remaining chapter was used as the test set. Normally, Python does not contain any element to execute this operation. So, external libraries such as Scikit-learn and Keras were used for our model. Here, the graphic of our dataset can be seen below:

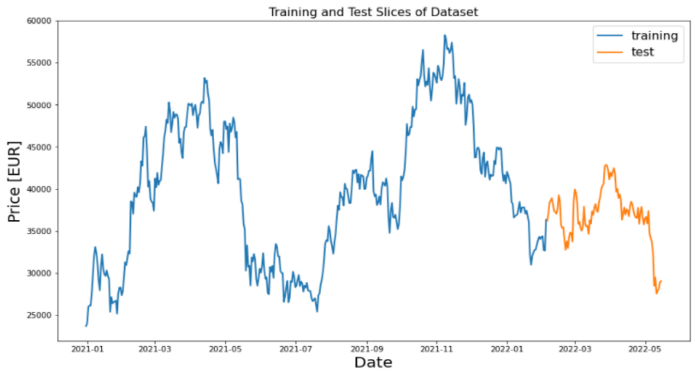

Figure 2. Splitting Dataset into Training and Test Sets

The third step is the most important chapter of our study. Because we have built the prediction algorithm in this chapter. The first movement conducted here is data normalization. Data normalization is a process that aiming reduction of unnecessary components such as repeated data. [3] In addition to data normalization, a transaction called 'data extracting' was conducted immediately after the previous step. Because we need only input variables for training the model we should have to eliminate the unnecessary ones. After arranging the data the conditions required to build a Deep Learning model become appropriate. At this point, selecting the correct Deep Learning technique is a very significant factor. For our model, we decided to use Long Short Term Memory (LSTM) technique for establishing a price prediction algorithm. The main reason for selecting the LSTM method is its ability to include feedback connections. Thus, LSTM layers can use the information stored in the previous layers instead of the current ones. [4] Instead of the used Deep Learning technique some model components such as the number of neurons, type of the activation function and model error calculation technique are playing an important role in the study. Well, if the number of neurons is lower the model cannot produce the correct output. On the other hand, if the number of neurons is higher the model will become too complex and hard to be executed. So, before deciding the neuron number trying different parameters is recommended. Furthermore, the other element called 'Activation Function' can impact the model mechanism and output. Functions such as sigmoid, tanh and linear are the most common ones to be used in Deep Learning models. It is hard to predict which one is better, so trying each one would be preferred by developers. In this study, all of them were tried before comparing outputs. Also, there is another parameter called 'Epoch' is included in Deep Learning models. Epoch is an equivalent term to 'tour'. So, the number of epochs is affecting the performance and output of the Deep Learning structures. For a clear output, all of those parameters are stored in a Python function placed in our code block. Also, to evaluate the error for the model a statistical method called 'Mean Squared Error' was chosen. Mean Squared Error (MSE) is measuring the mean of the difference between calculated output and the actual value. Thus, possible negative outcomes will be eliminated with the help of this tool. [5] Besides the error calculation method, optimizers are another element of a Deep Learning algorithm. Optimizers are a type of function that schedules attributes of neural networks like weights and learning rate. The main purpose of optimizers is to reduce the overall loss and improve the efficiency of artificial neural networks. [6] Because optimality can only be achieved after implementing a way. If we go back to the model, the parameters were set up according to many trials conducted by us. The number of neurons is chosen as 100 because the error has increased otherwise. Also, the amount of epochs is fixed to an amount of 30. As the number of neurons, the number of epochs was chosen after many trials. The next step after setting the parameters is running the code of our model. Running the code makes us able to obtain the result. The result and the error rate which is illustrating our model's validity will be provided in the Results chapter.

3. Results and Discussion

3.1. Results

After 30 epochs, the results of our model could be produced as mentioned in the previous chapter. We decided to use the Mean Absolute Error method for a correct assessment. The error was calculated as %0.029 approximately. This value can be classified as a low value. Furthermore, a graphic created with the assistance of an external Python library called 'Matplotlib' can present a visualization for a better understanding of our model's accuracy level. After the forecasting operation, we gathered real-time data when the time has come. Thus, a graphic that includes the comparison of the actual and predicted data could be created. Below, this graphic can be seen:

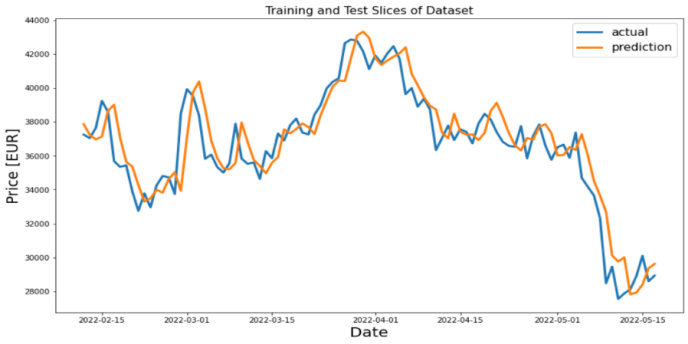

Figure 3. Comparison of Real Data and Prediction Results (EUR)

Also, to ensure the model validity we decided to use the same model for a different type of currency: the US Dollar. This time the error rate was evaluated as %0.028. It is similar to the previous operation, but the error rate is a little bit smaller. So, we

can determine that the consequences are similar, and the validity of our model is registered. Another graphic presented below is illustrating the comparison between real data and predicted ones like the first one, but this time it is using US Dollar as the base:

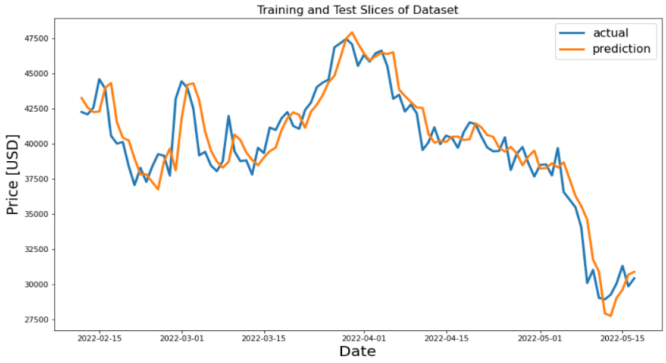

Figure 4. Comparison of Real Data and Prediction Results (USD)

According to these two graphics, our model is generally seen as successful. The difference between the two lines is appropriate. This means that we are successful to escape from overfitting and underfitting problems.

4. Conclusions and Recommendations

As mentioned in the introduction, Deep Learning is an effective approach to solving complex problems. With the assistance of its components such as layers, nodes, and activation functions Deep Learning models' ability to handle hard issues will result in their popularity to become higher. In this study, we provided a Deep Learning technique called Long Short-Term Memory for producing forecasted cryptocurrency prices. We gathered historical data from a particular website and fed this dataset into our model. After conducting some necessary data transactions, we formed an algorithm via Python programming language. According to error rate that calculated countless experiments, we finally found out the true parameters. Moreover, using a statistical validation calculation tool (MSE) allowed us to determine the effectiveness of our model. Also, for better understanding we used visualization methods in our model and created relevant graphics for better grip. The produced results turned out as we wanted. The error rate for two trials was calculated at low levels and we did not observe some problems such as underfitting and overfitting. Thus, the inference emphasizing that the model is valid can be made due to this outcome.

On the other hand, some other studies were conducted before our work. In one of them, the researchers used different types of Deep Learning techniques for cryptocurrency price prediction. They used the GRU method besides the LSTM. According to their outputs, GRU performs at higher speeds, but when the data amount increases LSTM is producing less error. [7] So, the reliability of the LSTM models is proven again via this study. But this does not mean that other methods should not be used. Techniques such as Recurrent Neural Networks, Convolutional Neural Networks and Restricted Boltzmann Machines should be used in the future for clear results. Also, another Artificial Intelligence methods can be chosen for this type of studies for observing the different outcomes.

In short, it is proven that the Long Short-Term Memory method is a valid technique for predicting cryptocurrency prices in this study. If similar types of studies become more widespread, people who want to reduce investment risk can benefit from these developments.

References

[1] Haenlein, M., & Kaplan, A. (2019). A brief history of artificial intelligence: On the past, present, and future of artificial intelligence. California management review, 61(4), 5-14.

[2] Basheer, I. A., & Hajmeer, M. (2000). Artificial neural networks: fundamentals, computing, design, and application. *Journal of microbiological methods*, *43*(1), 3-31.

[3] Doherty, K., Adams, R. G., & Davey, N. (2004). Non-Euclidean norms and data normalisation. In *Proceedings of the European Symposium* on Artificial Neural Networks Bruges (Belgium), 28-30 April 2004, d-side publi., ISBN 2-930307-04-8, pp. 181-186

[4] Yu, Y., Si, X., Hu, C., & Zhang, J. (2019). A review of recurrent neural networks: LSTM cells and network architectures. *Neural computation*, *31*(7), 1235-1270.

[5] Frost, J., (n.d). *Mean Squared Error (MSE).* Retrieved from https://statisticsbyjim.com/regression/mean-squared-error-mse/

[6] Choi, D., Shallue, C. J., Nado, Z., Lee, J., Maddison, C. J., & Dahl, G. E. (2019). On empirical comparisons of optimizers for deep learning. *arXiv preprint arXiv:1910.05446*.

[7] 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.