

Vedecký príspevok/ Scientific Article
Recenzované/ Review: 7. 12. 2021
<https://doi.org/10.24040/eas.2021.22.2.67-81>

Dogecoin price prediction – can be a determinism supposed?

Predikcia vývoja ceny Dogecoinu – je možné predpokladať determinizmus?

Juraj Medzihorský

Abstract: Pump-and-dump schemes are a phenomenon of many cryptocurrencies, including Dogecoin. We have identified repeating shapes on the graph of its price. It is obvious that the price development during last period replicates the development from the previous period. The aim of the article is to create a predictive model based on these replications. Using linear regression, we have developed the model with 87% accuracy on 3-month test set. Therefore, we claim that Dogecoin price development shows a significant degree of determinism. Our model also predicted the significant recent short-term rise (on 28 October 2021), despite the fact it was caused by Elon Musk's tweet. We have shown that timing of this rise corresponds to the timing of another (also related to Elon Musk's tweet) during a previous period. Thanks to the determinism in price development, we can predict the beginning of the next large Dogecoin pump most probably on 27 November 2021. This pump should be supported by Elon Musk's tweet, as well.

Key words: Dogecoin. Price prediction. Pump-and-dump scheme. Time series analysis.

JEL Classification: G17. G14. C22.

Introduction

Cryptocurrencies have become a subject of increasing interest to investors, professionals, and the public in recent years. In addition to the most famous Bitcoin, there are more than 7,000 altcoins on the market. That is more than a five-fold increase in the last four years (Statista, 2021). Total crypto market capitalization as of November 2021 is reaching 3 trillion USD (Yahoo, 2021).

Several coins show signs of pump-and-dump scheme. During the accumulation phase, a group of individuals gradually accumulate a certain amount of selected financial asset, then organises an artificial increase in the price of the asset (e.g. through the spread of disinformation on social media) and sells it at an increased price; consequently, there will be a significant

decrease. Such scheme can be considered a form of fraud (Kamps and Kleinberg, 2018). The alternative approach to this scheme is based on trade-oriented manipulation, where the only instrument of the manipulator is trading (not disinformation) i.e., the pump must be triggered by its own purchase at the time of the start of the pump, that is followed by investors attempting to capture the induced price increase. Li et al. (2021) even describes this phenomenon as ubiquitous in the cryptocurrency market, pointing out that, in addition to members of organised groups, only those investors - who buy the asset quickly enough - benefit. Morgia et al. (2021) has observed about 900 pumps since 2017. It is clear from the essence of the scheme that it should mainly concerns assets with low market capitalisation that are volatile enough and reflect market sentiment.

Price prediction of any financial asset based only on its past development is problematic or may not be accurate, especially when a longer period is predicted, since the price development of financial assets is traditionally stochastic, not deterministic (returns on financial assets are characterized by white noise, their prices by random walk). The existence of pump-and-dump schemes can be an important element in this context, especially in the case of predicting the pump itself. Ozdamar et al. (2021) show that there can be a significant correlation between past and future returns precisely for the cryptocurrency market.

The aim of the article is to create the predictive model of Dogecoin price by identifying a possible determinism in its development. The main aim can be further broken down into the partial aims, as follows:

- C1: Identify deterministic trends in price development;
- C2: Create the auxiliary model for time contexts;
- C3: Create the predictive model, and verify it on a test set;
- C4: Predict a possible future pump.

In accordance with the main aim, we set the hypothesis, as follows. H: The current Dogecoin price development is more determined by its previous development than it would be stochastic one.

The rest of paper is structured, as follows. Theoretical background in Section 1, Material and Methods in Section 2, Results and discussion in Section 3, and Conclusion in the final section..

1. Theoretical background

Pump-and-dump schemes

Pump-and-dump schemes can relate to any financial asset e.g., shares. A recent example is GameStop, which price has significantly risen thanks to a group of individuals joined in WallStreetBets - forum on the social medium Reddit. In the case of cryptocurrencies, which are essentially deregulated, and thanks to large number of unknown altcoins, the possibilities for manipulation are even more evident (Morgia et al., 2021). We can say that the development of social media is one of the main 'catalysts' for the number and effectiveness of the pumps. On the other hand, pumps which are based solely on trade-oriented manipulation, may not be effective if the influence of a manipulator on the price is not higher at the time of purchase than at the time of sale (Friedman, 1953; Li et al., 2021).

Multiple economists (e.g. Hamrick et al., 2019; Morgia et al., 2021) pointed to the social media Telegram and Discord, where such groups are organized. Tandon et al. (2021) pointed to a correlation between Bitcoin price increases and Elon Musk's tweets, but they surprisingly do not mention a similar relationship for Dogecoin. They claim that one person cannot control the market as they show that such impulses affect the price only in the short term.

Depending on how the pump is triggered and how communication takes place in groups on social media, the pumps can be divided as follows (Hamrick, 2019):

- Obvious pumps – Groups have a name containing 'pump' or 'dump'. First, some signals are sent, then timing and a market, finally the name of a coin. The last step is done only at the announced pump time. Often, only a picture of a coin is sent, therefore the name cannot be detected using automated tools. Some members of these groups have premium access so they will get the information sooner.
- Target pumps – Groups do not have a name apparently related to pump-and-dump schemes. All information about a pump is sent at the same time. If the membership is paid, messages are sent a little earlier than in freely available groups.
- Copied pumps – it is only a forwarding the original information (with time lag).

Dogecoin

Dogecoin is one of the meme coins (i.e., cryptocurrencies that are based on a certain imitation, social idea or were created as a parody). Its creators Billy Markus and Jackson Palmer wanted to create a comical alternative to Bitcoin. From a technical point of view, Dogecoin was based on the Luckycoin platform, which itself is based on the Litecoin platform. Although the

amount of Dogecoin was initially limited (it had a potential of being a deflationary asset), it later became unlimited (Chohan, 2021). Its current market capitalisation USD 34.5 billion ranks it the 9th among cryptocurrencies (CoinMarketCap, 2021). Recently, it has been the subject of tweets by Elon Musk (also known as ‘Dogefather’), who has vowed not to sell any Dogecoin which he owns. If big investors sell off a larger amount of Dogecoin, he will support Dogecoin. Moreover, his company SpaceX is planning a mission to the Moon called DOGE-1, accepting payments in Dogecoin for lunar cargo (Sheetz, 2021).

2. Material and methods

The object of the analysis is the time series of Dogecoin price (DOGE/USD) in the period from 28/1/2021 to 10/11/2021 (note: DD/MM/YYYY format is used in the article) which covers two pumps and two accumulation phases. That enables us to analyse their relations. Longer development (e.g. during years) is not analysed that can be an important limitation of the study. The source database is Coindesk (2021). The beginning of the period is the last day before the start of January pump. We have divided the total period into two separate components: 28/1/2021 – 10/4/2021 (Period 1) and 11/4/2021 – 10/11/2021 (Period 2) - according to the date when the April pump started. We identify a pump primarily from the development of the price (i.e., from its effects, not from messages of organized groups on social media). In our analysis, we define pump as a multiple price increase over a short period of time (in the order of days). Since we will predict the price development during Period 2 through development during Period 1, we will divide both of them into training and test sets (Period 1 in a ratio of 60:40, Period 2 to correspond to that according to the results of the time model). The main method used in our article is linear regression model. Regarding mathematical notation, the model is as follows:

$$P2_i = \alpha + \beta P1_i + \varepsilon_i \quad (1)$$

where $P2_i$ is the i^{th} price during Period 2,

$P1_i$ is the i^{th} price during Period 1 corresponding to $P2_i$ according to time model

α , β , ε are intercept, regression coefficient, error terms, respectively.

To analyse the time context, we apply a separate model:

$$T2_j = \alpha + \beta T1_j + \varepsilon_j \quad (2)$$

where $T2_j$ is the time (expressed as a day in row) in Period 2,
 $T1_j$ is the time (expressed as a day in row) in Period 1,
 rest of symbols like in Equation 1.

T1 and T2 represent a time in days that passed since the beginning of the January pump, April pump, respectively. In the training set, we are going to search for points in both periods that correspond to each other (based on the assumption that development in Period 2 simply replicates the development in Period 1).

Logarithmic price values (using natural logarithm) are used throughout the analysis, except for the calculation of the mean quadratic error, prediction error, and the prediction in Appendix 3. Since the pump prediction is a partial aim of the article, and past pumps enter the model significantly, we use 24-hour highs instead of closing prices. This approach is also in line with the results of Ozdamar et al. (2021), who pointed out a correlation of the maximum daily earnings from previous month with the development in the next week in their study.

For model evaluation, we use the mean quadratic error δ and prediction error γ . We consider them to be the most appropriate measures for this purpose since the aim of the model is prediction. Similarly, Agarwal et al. (2021) selected δ , Tandon et al. (2021) γ . Square of the mean quadratic error for n-observations can be noted as follows:

$$\delta^2 = \frac{1}{n} \sum_{i=1}^n (\text{PredictedPrice}_i - \text{ActualPrice}_i)^2 \quad (3)$$

and a prediction error (rate) as follows:

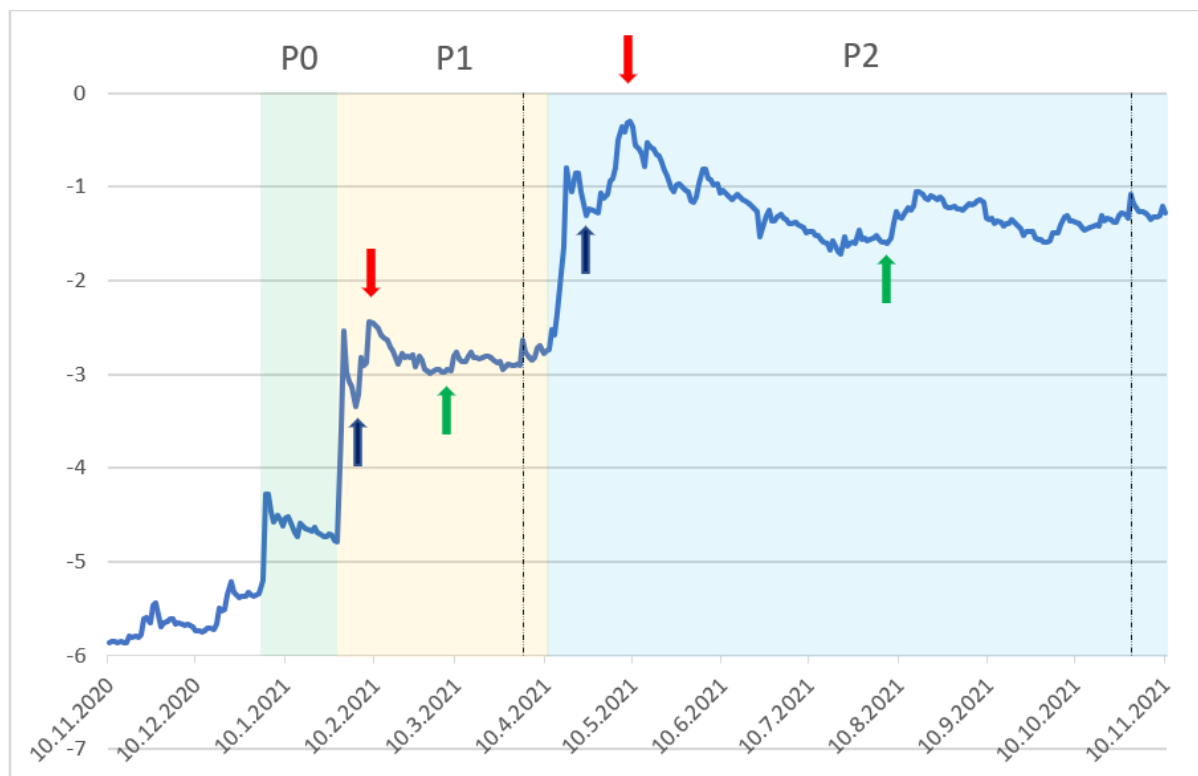
$$\gamma = \frac{1}{n} \sum_{i=1}^n \frac{|\text{PredictedPrice}_i - \text{ActualPrice}_i|}{\text{ActualPrice}_i} \quad (4)$$

The accuracy of the prediction is then defined as $1-\gamma$.

3. Results and discussion

A certain degree of determinism in the development of the Dogecoin price is evident from graphs. On Graph 1, we can see repetitive shapes both in the pump phase and the accumulation phase, especially when comparing Period 1 and 2; however, some relation with Period 0 is obvious, as well. Development in Period 2 more or less replicates development in Period 1. The

gradual extension of periods and shapes is also characteristic for this development. Elon Musk has published several messages on Twitter about Dogecoin. On the graph, we can see two selected tweets – at the end of the accumulation phase in Period 1, and the recent tweet, which, as we will demonstrate later, corresponds to timing of the first one.



Graph 1 Dogecoin price development over the past 12 months

Source: Own processing from Coindesk (2021), Bambrough (2021), and Ciaccia (2021).

Notes: P0, P1, P2 represents Period 0, 1, 2, respectively; the dashed line represents selected Elon Musk’s messages on Twitter; arrows indicate the events listed in Table 1.

Table 1 Time relations

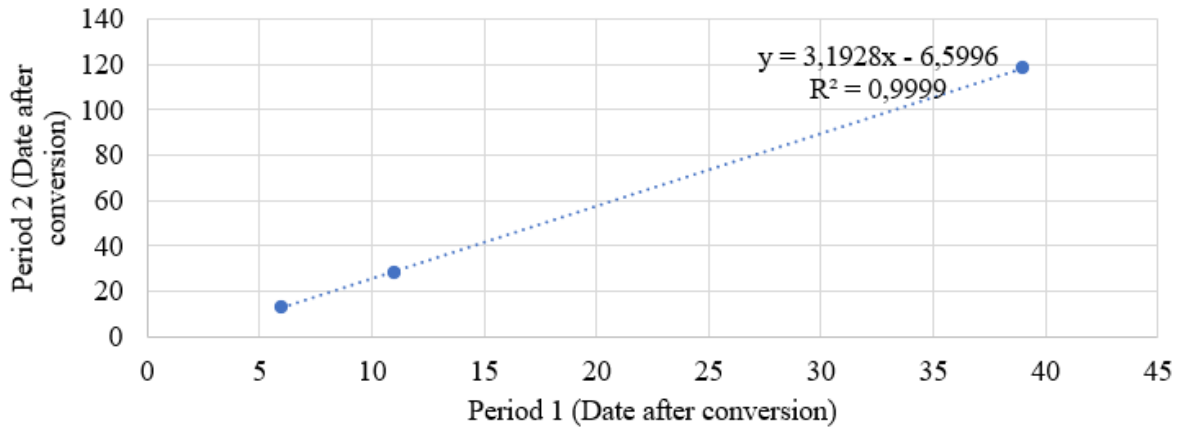
| Date of the event in Period 1 | Conversion 1 | Date of corresponding event in Period 2 | Conversion 2 | Ratio Conversion 1 / Conversion 2 |
|-------------------------------|--------------|---|--------------|-----------------------------------|
| 3/2/2021 | 6 | 24/4/2021 | 13 | 2.17 |
| 8/2/2021 | 11 | 9/5/2021 | 28 | 2.55 |
| 8/3/2021 | 39 | 7/8/2021 | 118 | 3.02 |

Source: Own processing¹.

For further research, we apply Equation 2 (see Section 2). To create the auxiliary, time model, we have determined the key time points (see Graph 1 and Table 1) which correspond to

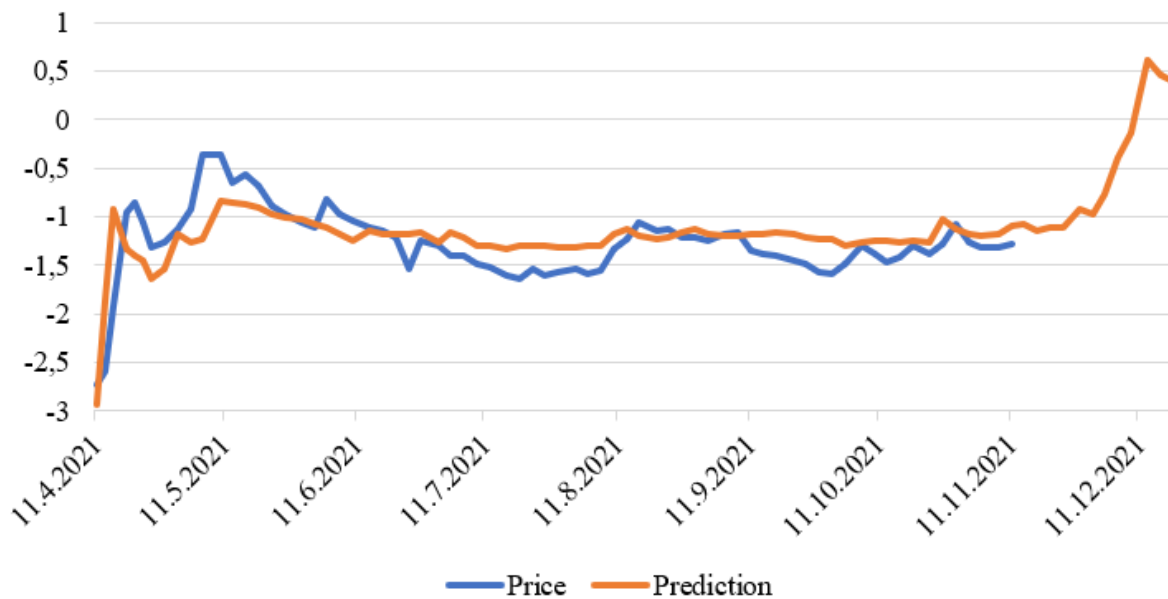
¹ Note: Conversion represents the number of days between the beginning of a period and the date of an event (e.g. the first day of the Period 1 – 28/1/2021 - would be numbered as zero after conversion; zero value would represent 11/4/2021 – beginning of Period 2, as well).

each other (these points are selected from the training set). According to values presented in Table 1, it is clear that we cannot apply a simple multiplication, but to model the time (rate of extension of shapes when we compare Period 1 and Period 2 is not stable but growing).



Graph 2 Time model
Source: Own processing

The time model is presented in Graph 2. Significant limitation of the auxiliary model is the fact that it is based on 3 observations. On the other hand, the regression line passes through these points perfectly, i.e., R^2 is nearly 1. Using the time model, we can match dates from both periods except the first 5 days in Period 1. For them, we have used fixed value 2.17 for computation.

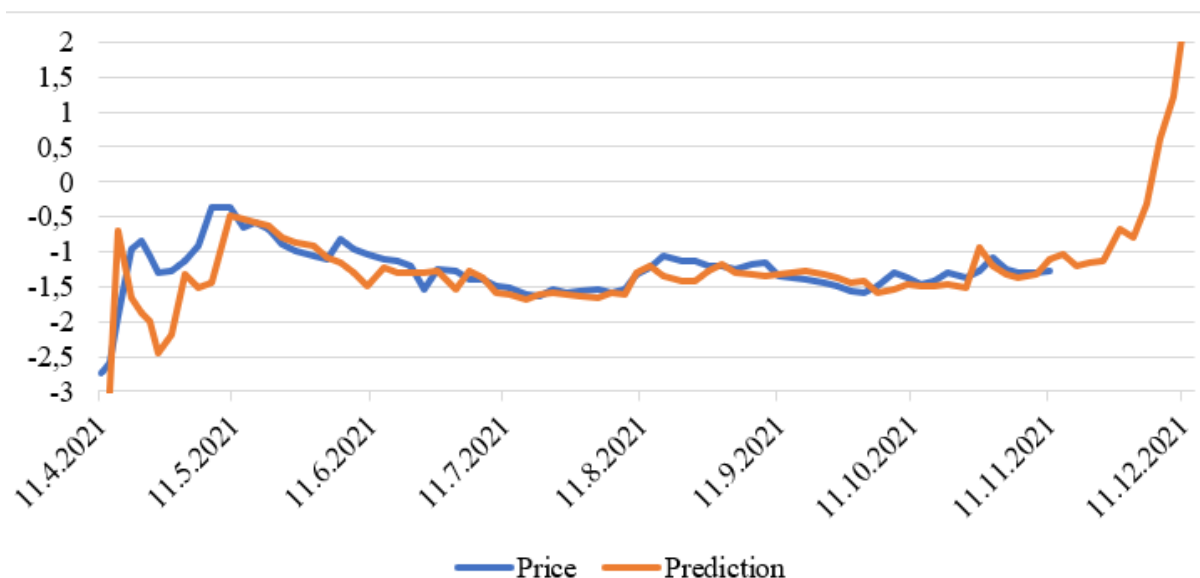


Graph 3 Original predictive model
Source: Own processing from Coindesk (2021).

If we consider the prices in Period 2 that correspond to the prices in Period 1 (according to the time model) as the dependent variable, and the prices in Period 1 as the independent variable (in accordance with Equation 1), we can make a predictive model (see Graph 3, and Appendix 1). Regarding accuracy, the model has achieved $\delta = 0.34$ and $\gamma = 0.28$ in the training set; $\delta = 0.17$ and $\gamma = 0.16$ on the test set (these values are calculated from prices, not logarithmic prices). There is a higher values of both error metrics in the training set than in the test set (than unusual) and this can be justified by a more volatile development (and, in the case of δ , by higher prices, as well) during the pump phase than in the accumulation phase.

As we can see in the graph, the model predicts quite well with only some inaccuracies. However, it predicts much lower volatility than actually occurs. We have partially confirmed the significant determinism in the development of the Dogecoin price. We would like to emphasize that there is almost an exact prediction of the increase, which occurred on 28/10/2021 (the prediction was on 26/10/2021), and which was caused by Elon Musk's tweet (see Graph 1). Elon Musk tweeted about Dogecoin on 1/4/2021, as well. Exactly 2/4/2021 corresponds to 26/10/2021 according to our time model. The question arises whether it is possible to predict when Elon Musk is going to tweet about Dogecoin.

The importance of the predictive model lies not only in measuring its error rate on the test set, but especially in the prediction of the future i.e., data that we do not have yet. The model predicts that the next large pump will start on 27/11/2021. Furthermore, if Elon Musk's tweets timing is also predictable, we can predict the next tweet at the beginning of December 2021 (that will support the next pump after its start).



Graph 4 Final prediction model

Source: Own processing from Coindesk (2021).

Since our original model exhibits certain inaccuracies, we apply maximal weekly values for Period 1 and the corresponding maximal values from Period 2 to eliminate inaccuracies within the week; in order to refine the value of the regression coefficient and the intercept (the auxiliary predictive model with weekly data is in Appendix 2). Ozdamar et al. (2021) recommend in their study methodologically a similar approach – an averaging of maximal daily returns through several (2, 3, 4, and 5) days. Through application of the refined values of α and β to the original (daily) data, the final prediction model is created (see Graph 4). Mean quadratic error values for the final model decreased in comparison to the initial model ($\delta = 0.13$ and $\gamma = 0.27$ on the training set; $\delta = 0.05$ and $\gamma = 0.13$ on the test set). The model is based only on the past development, using basic method of linear regression, and at the same time it is a prediction for a longer period of time (an almost 3-month test set). Therefore, the accuracy of the model with an average error rate of 13% is more than satisfactory. Therefore, we are able to confirm that the price of Dogecoin shows significant signs of deterministic, not stochastic development in a longer period of time. Stochasticity remains in short-term development. However, we should stress that prediction of the future price after the next pump is more realistic in an original than final model.

Finally, we have compared the prediction error rate of our model with other studies, where more advanced techniques were used. Agarwal et al. (2021) achieved $\delta = 0.02$ using machine learning. Their results, using only historical prices, are not very different from those when they included sentiment on Twitter. However, δ values cannot be compared between models unless

they are based on the same data. In this case, at least a partial comparison is possible as they studied the same cryptocurrency. Tandon et al. (2021), using the extended Dickey-Fuller test and ARIMA models, reached $\gamma = 0.0395$ for the Bitcoin price model, but for a significantly shorter period of test set than ours.

Deterministic or not, the price development of any cryptocurrency can be affected by the development of entire crypto market. There exist significant correlations between prices of cryptocurrencies. We do not apply such relations in the analysis, that can be one of the most important limitations of our model.

Conclusion

Long-term prediction of price of financial assets based only on historical data shows usually relatively low accuracy due to the stochasticity of the price development. Lots of cryptocurrencies are influenced by pump-and-dump schemes, which can make price prediction even more complicated, because only members of organized groups are informed in advance about a pump. However, if we look at the price development of Dogecoin, we can find repetitive shapes in graphs, which could facilitate a prediction, including a pump prediction. We have identified that the price development in Period 2 replicates the development in Period 1. The recognized repeating shapes tend to lengthen over time. Therefore, we first created the auxiliary, time model. Our final predictive model achieved 13% error rate on a relatively long 3-month test set. We are able to confirm that the overall development in Period 2 can be well predicted only through development in Period 1. As a result, the price development of Dogecoin is characterized by a significant degree of determinism. Furthermore, our time model suggests some relation in timing between two separate Elon Musk's tweets.

References

- [1] Agarwal, B., Harjule, P., Chouhan, L., Saraswat, U., Airan, H., & Agarwal, P. (2021). Prediction of dogecoin price using deep learning and social media trends. *EAI Endorsed Transactions on Industrial Networks and Intelligent Systems*, 8(29), 1-12. [dx.doi.org.10.4108/eai.29-9-2021.171188](https://doi.org/10.4108/eai.29-9-2021.171188)
- [2] Bambrough, B. (2021, October 28). Memecoin Price Wars: Elon Musk Helps Dogecoin Suddenly Surge Back Above Shiba Inu As Bitcoin And Ethereum Plummet. *Forbes*. Retrieved from: <https://www.forbes.com/sites/billybambrough/2021/10/28/memecoin-price-wars-elon-musk-helps-dogecoin-suddenly-surge-back-above-shiba-inu-as-bitcoin-and-ethereum-plummet/?sh=45240b3b7e14>

- [3] Chohan, U. (2021, February 12). *A History of Dogecoin*. SSRN. Retrieved from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3091219
- [4] Ciaccia, C. (2021, April 2). *Elon Musk says he's going to put Dogecoin on 'the literal moon'*. *Space - Future US Inc*. Retrieved from: <https://www.space.com/elon-musk-spacex-dogecoin-on-the-moon>
- [5] Coindesk. (2021). *Dogecoin*. Retrieved from: <https://www.coindesk.com/price/dogecoin/>
- [6] CoinMarketCap. (2021). *Doge Price Live Data*. Retrieved from: <https://coinmarketcap.com/currencies/dogecoin/>
- [7] Friedman, M. (1953). *Essays in Positive Economics*. Chicago, IL: University of Chicago Press.
- [8] Hamrick, J.T., Rouhi, F., Mukherjee, A., Feder, A., Gandal, N., Moore, T., & Vasek, M. (2019, January 3). *An Examination of the Cryptocurrency Pump and Dump Ecosystem*. SSRN. Retrieved from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3303365
- [9] Kamps, J., & Kleinberg, B. (2018). To the moon: defining and detecting cryptocurrency pump-and-dumps. *Crime Science*, 7(18), 1-18. doi.org/10.1186/s40163-018-0093-5
- [10] Li, T., Shin, D., & Wang, B. (2021, February 10). *Cryptocurrency Pump-and-Dump Schemes*. SSRN. Retrieved from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3267041
- [11] Morgia, L., Mei, A., Sassi, F., & Stefa, J. (2021, May 3). *The Doge of Wall Street: Analysis and Detection of Pump and Dump Cryptocurrency Manipulations*. Ithaca, NY: Cornell University. Retrieved from: <https://arxiv.org/abs/2105.00733v1>
- [12] Ozdamar, M., Akdeniz, L., & Sensoy, A. (2021). Lottery-like preferences and the MAX effect in the cryptocurrency market. *Financial Innovation*, 7(74), 1-27. doi.org/10.1186/s40854-021-00291-9
- [13] Sheetz, M. (2021, May 9). *SpaceX accepts Dogecoin as payment to launch 'DOGE-1 mission to the Moon' next year*. *CNBC*. Retrieved from: <https://www.cNBC.com/2021/05/09/spacex-accepts-dogecoin-payment-for-doge-1-mission-to-the-moon.html>
- [14] Statista. (2021). *Number of cryptocurrencies worldwide from 2013 to November 2021*. Retrieved from: <https://www.statista.com/statistics/863917/number-crypto-coins-tokens/>
- [15] Tandon, C., Revankar, S., Palivela, H., & Parihar, S. (2021). How can we predict the impact of the social media messages on the value of cryptocurrency? Insights from big data analytics. *International Journal of Information Management Data Insights*, 1, Article, No. 100035. doi.org/10.1016/j.jjime.2021.100035

[16] Yahoo. (2021). *Cryptocurrencies hit market cap of \$3 trillion for the first time as Bitcoin and Ether reach record highs*. Retrieved from: <https://finance.yahoo.com/news/cryptocurrencies-hit-market-cap-3-063245377.html>

Authors' addresses: Ing. Juraj Medzihorský, Faculty of Economics Matej Bel University in Banská Bystrica, Tajovského 10, 975 90 Banská Bystrica, Slovakia, email: juraj.medzihorsky@umb.sk

Appendices

Appendix 1 Original model

SUMMARY OUTPUT

| <i>Regression Statistics</i> | |
|------------------------------|----------|
| Multiple R | 0,699027 |
| R Square | 0,488639 |
| Adjusted R Square | 0,475855 |
| Standard Error | 0,345083 |
| Observations | 42 |

| ANOVA | | | | | |
|------------|-----------|-----------|-----------|----------|-----------------------|
| | <i>df</i> | <i>SS</i> | <i>MS</i> | <i>F</i> | <i>Significance F</i> |
| Regression | 1 | 4,551639 | 4,551639 | 38,22268 | 2,62E-07 |
| Residual | 40 | 4,763286 | 0,119082 | | |
| Total | 41 | 9,314925 | | | |

| | <i>Coefficients</i> | <i>Standard Error</i> | <i>t Stat</i> | <i>P-value</i> | <i>Lower 95%</i> | <i>Upper 95%</i> |
|------------------|---------------------|-----------------------|---------------|----------------|------------------|------------------|
| Intercept | 1,322889 | 0,420257 | 3,147807 | 0,003105 | 0,473517 | 2,172261 |
| Beta-coefficient | 0,887685 | 0,143581 | 6,18245 | 2,62E-07 | 0,597496 | 1,177874 |

Source: Own processing.

Appendix 2 Model based on weekly maximums

SUMMARY OUTPUT

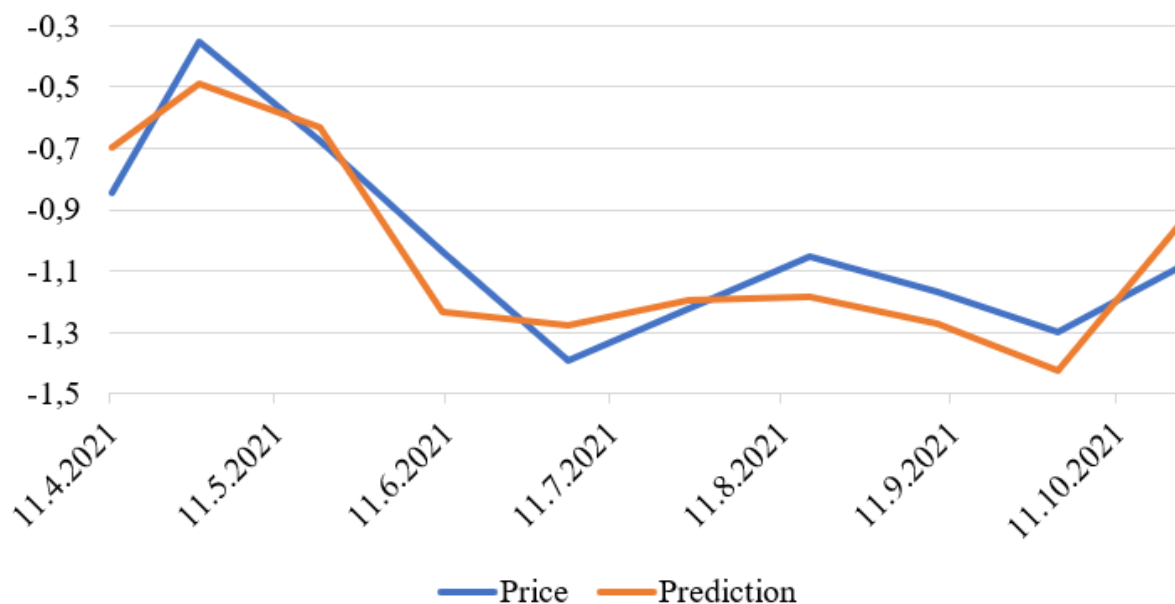
| <i>Regression Statistics</i> | |
|------------------------------|----------|
| Multiple R | 0,931092 |
| R Square | 0,866932 |
| Adjusted R Square | 0,833665 |
| Standard Error | 0,153904 |
| Observations | 6 |

| <i>ANOVA</i> | | | | | |
|--------------|-----------|-----------|-----------|----------|-----------------------|
| | <i>df</i> | <i>SS</i> | <i>MS</i> | <i>F</i> | <i>Significance F</i> |
| Regression | 1 | 0,617262 | 0,617262 | 26,05977 | 0,006959 |
| Residual | 4 | 0,094746 | 0,023686 | | |
| Total | 5 | 0,712008 | | | |

| | <i>Coefficients</i> | <i>Standard Error</i> | <i>t Stat</i> | <i>P-value</i> | <i>Lower 95%</i> | <i>Upper 95%</i> |
|------------------|---------------------|-----------------------|---------------|----------------|------------------|------------------|
| Intercept | 4,815908 | 1,125583 | 4,278588 | 0,012862 | 1,690787 | 7,941028 |
| Beta-coefficient | 2,174614 | 0,425988 | 5,104877 | 0,006959 | 0,991883 | 3,357345 |

Source: Own processing.

Note: These values of intercept and Beta-coefficient are used for final model, as well.



Source: Own processing.

Appendix 3 Price prediction after the beginning of the next expected pump



Source: Own processing.

Notes: Dates are predicted through our time model. Values are predicted through our original predictive model, which predicts more realistic future prices than our final model. Prices are in USD (without the logarithmic scale specifically in this graph).