# A Multi-window Bitcoin Price Prediction Framework on Blockchain Transaction Graph

Xiao Li[1(✉)] and Linda Du[2]

[1] The University of Texas at Dallas, Richardson, TX 75080, USA
xiao.li@utdallas.edu

[2] The University of Texas at Austin, Austin, TX 78712, USA
yingfan.du@mccombs.utexas.edu

**Abstract.** Bitcoin, as one of the most popular cryptocurrency, has been attracting increasing attention from investors. Consequently, bitcoin price prediction is a rising academic topic. Existing bitcoin prediction works are mostly based on trivial feature engineering, that is, manually designed features or factors from multiple areas. Feature engineering not only requires tremendous human effort, but the effectiveness of the intuitively designed features can not be guaranteed. In this paper, we aim to mine the abundant patterns encoded in Bitcoin transactions, and propose $k$-order transaction graphs to reveal patterns under different scopes. We propose features based on a transaction graph to automatically encode the patterns. The Multi-Window Prediction Framework is proposed to train the model and make price predictions, which can take advantage of patterns from different historical periods. We further demonstrate that our proposed prediction method outperforms the state-of-art methods in the literature.

**Keywords:** Bitcoin · Blockchain · Transaction · Machine learning

## 1 Introduction

Bitcoin blockchain [24][1], the first application of blockchain, has been attracting increasing attention from various areas. *Bitcoin* is the cryptocurrency traded in the Bitcoin blockchain, which is a reward to the miners for successfully appending a block. Bitcoin can be traded with regular currency in financial markets like many other financial products, e.g. stocks, gold and crude oil [28]. Different from other products, bitcoin has highly volatile prices [1,5]. This provides investors with a great opportunity to earn a fortune from the striking difference in prices. Thus, bitcoin is becoming a popular financial asset, and attracts huge amounts of investment [30].

[1] In this paper, the terms "Bitcoin blockchain" or "Bitcoin" refer to the whole Bitcoin blockchain system and "bitcoin" refers to the cryptocurrency.

This work is partially supported by NSF 1907472.


W. Wu and H. Du (Eds.): AAIM 2021, LNCS 13153, pp. 317–328, 2021.
https://doi.org/10.1007/978-3-030-93176-6_27

Bitcoin price forecasting models are eagerly desired to provide the suggestions on whether the bitcoin price will rise or fall [10,29] to help investors decide whether and when they should buy or sell bitcoins. However, bitcoin price forecasting models usually require well-designed features to reveal the reason of bitcoin price change, which is a challenging task. The basic features of blockchain are the indexes reflecting the transaction information of Bitcoin blockchain, such as average degree of addresses, number of new addresses and total coin amount transferred in transactions [2]. Maesa et al. try to analyze the latent features of Bitcoin blockchain from the perspective of users transferring graph [20]. Mallqui et al. [21] include international economic indicators that were used to reflect the features of the global financial market, such as S&P500 future, NASDAQ future, and DAX index, which are features from a financial perspective. CerdaR et al. [8] and Yao et al. [29] introduce public opinion features into bitcoin price prediction through mining the sentiment from social media like Twitter and news articles.

Existing work has created features covering many aspects, including blockchain network, financial market information, and even public opinions. However it is still unclear what features or factors are useful, and how these features impact the price of bitcoin. Manually discovering or creating the features not only relies on heuristics but also consumes huge labour resource. In this paper, we try to develop a bitcoin prediction model that can directly learn features from the Bitcoin blockchain transactions without directly incorporating tedious information outside the blockchain, e.g. financial market information, and public sentiment. Instead, if the external factors beyond the Bitcoin blockchain, such as public sentiment or news, contribute to the bitcoin price change, they will eventually be reflected by the changes in the transactions and structure of the Bitcoin blockchain. In other words, if the external factors influence the action of users, the different actions taken by users will be reflected by the changes in the transactions in the Bitcoin blockchain. In this paper, we argue that the structure of Bitcoin blockchain encodes abundant transaction pattern information that can interpret the factors behind the bitcoin price change.

To capture these transaction patterns, we propose a blockchain transaction graph.

The blockchain transaction graph encodes the patterns of transactions which reflects market trend and status. As mentioned in [4], if the input addresses of a transaction is more than the output addresses, then the transaction is gathering bitcoins, indicating some users are buying bitcoins. On the other hand, if the input addresses of a transaction is less than the output addresses, then the transaction is splitting the bitcoins, indicating some users are selling bitcoins. Therefore by discovering these transaction patterns with a Bitcoin transaction graph and proposed prediction framework, we can leverage valuable information that can hardly be managed by manual feature engineering.

The main contributions of the paper can be summarized as follows:

- We propose a ***k-order Transaction Subgraph*** based on a transaction graph, to represent the transaction feature of blockchain.
- We proposed a transaction graph based feature to encode the implicit patterns behind the transactions, which is further fed to a novel machine learning

based ***Multi-Window Prediction Framework*** that can effectively learn the features of different historical periods.
– We evaluate the proposed method empirically using historical bitcoin prices and the results demonstrate superiority over recent state-of-the art methods.

The remainder of this paper is organized as follows: In Sect. 2, we review related recent literature. Section 3 proposes a transaction graph and describes how the subgraph feature is extracted. Next, in Sect. 4, we propose the Multi-Window Prediction Framework. In Sect. 5 we evaluate the proposed feature and the prediction framework. Finally, in Sect. 6, we conclude.

## 2 Related Work

The key issue of bitcoin price prediction is to discover and analyze determinants of bitcoin price. Various determinants have been studied including Google Trends [16,22], Wikipedia [16], Bitcoin tweets [6,22], social media or public opinions [7,8,29], and so on. Some papers consider both traditional features in the market as well as economical features of a digital currency [3,11]. Pieters and Vivanco [26] study the 11 bitcoin markets and present that standard financial regulations can have a non-negligible impact on the market for Bitcoin. Both Georgoula et al. [13] and Kristoufek [17] study the difference between long-term and short-term impact of the determinants on bitcoin price. Kristoufek [17] stresses that both time and frequency are crucial factors for bitcoin price dynamics since the price of bitcoin evolves overtime.

The structural information of the Bitcoin blockchain has also been used to mine determinants of the price of bitcoin. Akcora et al. [4] propose a Bitcoin graph model, upon which chainlets is proposed to represent graph structures in the Bitcoin.

In their further work [2], they propose occurrence matrix and amount matrix to encode the topological features of chainlets. In this paper, we also adopt the concept of occurrence matrix to encode the topological features. However, we design a different graph representation model to reveal the topological features of the Bitcoin blockchain.

There are also several theoretical [18,19,27] and empirical studies [5,15,23] that have looked at Bitcoin transactions focusing on the volume-return causality in the Bitcoin market. These studies focus on trading volumes or number of unique Bitcoin transactions and employ regression techniques. In this paper, we take our analyses further and extract patterns from the Bitcoin transactions using graph models.

Various machine learning methods can be adopted to learn the patterns from the features and forecast the price of bitcoin [10,31]. Felizardo et al. [9,12] compare several popular machine learning methods adopted in the bitcoin price prediction task. Methods include using a Hidden Markov Models to tackle the volatility of cryptocurrencies and predicting future movements with Long Short Term Memory networks (LSTM) [14] and using hybrid methods between AutoRegressive Integrated Moving Average (ARIMA) and machine learning [25].

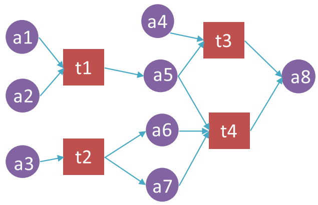



**Fig. 1.** A simple transaction graph

## 3 Transaction Graph and Subgraph Occurrence Pattern

In order to mine the blockchain transaction features, we define ***transaction graph*** to extract the blockchain transaction information.

**Definition 1.** (Transaction Graph): *A transaction graph is a directed graph $G = (A, T, E)$, where $A$ is the set of addresses in the blockchain, $T$ is the set of transactions in the blockchain, and $E$ is the set of direct links from $a_i$ to $t_k$, indicating $a_i$ is one of the inputs of $t_k$, or from $t_k$ to $a_j$, indicating $a_j$ is one of the outputs of $t_k$, where $a_i, a_j \in A$ and $t_k \in T$.*

Figure 1 presents an example of a transaction graph, which contains 8 addresses and 4 transactions.

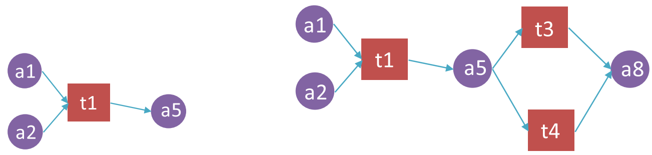
### 3.1 k-Order Transaction Subgraph

To specify characteristics of each transaction in the transaction graph, we define the $k$-order transaction subgraph of each transaction. The $k$-order transaction subgraph of a transaction $t_i$ is a graph $G^k_{t_i}$ that contains only $t_i$ and the transactions that spend the output of $t_i$ in the next $k-1$ steps, along with the corresponding addresses that connect to these transactions. The formal definition is given as Definition 2.

**Definition 2.** ($K$-order transaction subgraph): *The $K$-order transaction subgraph of a transaction $t_i$ is a graph $G^k_{t_i} = (A^k, T^k, E^k)$, where $T^k = \{t_j |\ \exists\ a_n \in A^{k-1},\ (a_n, t_j) \in E$ and $\exists (t_l, a_n) \in E^{k-1}$ for $t_l \in T^{k-1}\}$, $A^k = \{a_n | a_n \in A^{k-1}$ or $(t_j, a_n) \in E$ where $t_j \in T^k\}$. Specially, if $k = 1$, $G^1_{t_i} = (A^1, T^1, E^1)$, where $A^1 = \{a_n | (a_n, t_i) \in E$ or $(t_i, a_n) \in E\}$, $T^1 = \{t_i\}$ and $E^1 = \{(a_n, t_i)$ or $(t_i, a_n)$ $| a_n \in A^1\}$.*

If $k = 1$, then the $k$-order transaction subgraph of $t_i$ contains only $t_i$ along with its input addresses and output addresses. When $k$ increases, the $k$ order transaction subgraph will trace further along the bitcoin flow output by transaction $t_i$. Figure 2(a) and 2(b) shows the 1-order and 2-order transaction subgraph of transaction $t_1$ in Fig. 1, respectively.

The $k$-order transaction subgraphs have different patterns. Here we consider different patterns as different numbers of inputs and outputs addresses of the $k$-order transaction subgraphs.




(a) 1-order transaction subgrpah of $t_1$, $G^1_{t_1}$

(b) 2-order transaction subgrpah of $t_1$, $G^2_{t_1}$

**Fig. 2.** The 1 order nd 2-order transaction subgraph of $t_1$ in Fig. 1

The input addresses of a $k$-order transaction subgraph $G^k_{t_i}$ are the addresses that input to the first transaction in $G^k_{t_i}$. The output addresses of $G^k_{t_i}$ are the addresses that accepts the outputs of the last transactions in $G^k_{t_i}$. The input and output addresses are formally defined in Definition 3.

**Definition 3.** (Input and Output addresses of $K$-order transaction subgraph): *The input and output addresses of $K$-order transaction subgraph $G^k_{t_i}$ is $\mathcal{I}_{G^k_{t_i}}$ and $\mathcal{O}_{G^k_{t_i}}$, respectively.* $\mathcal{I}_{G^k_{t_i}} = \{a_n | \exists (a_n, t_j) \in E^k, t_j \in T^k$ *and* $\forall t_k \in T^k, (t_k, a_n) \notin E^k\}$. $\mathcal{O}_{G^k_{t_i}} = \{a_n | \exists (t_k, a_n) \in E^k, t_k \in T^k$ *and* $\forall t_j \in T^k, (a_n, t_j) \notin E^k\}$.

In Fig. 2(a), the addresses $a_1$ and $a_2$ are the input addresses of $G^1_{t_1}$, and the address $a_5$ is the output address of $G^1_{t_1}$. For higher order transaction subgraphs, the input and output addresses may be more complicated. For example, in Fig. 2(b), the input addresses of $G^2_{t_1}$ are $\{a_1, a_2\} = \mathcal{I}_{G^2_{t_1}}$, and the output addresses are $\{a_8\} = \mathcal{O}_{G^2_{t_1}}$.

Based on the concept of $\mathcal{I}_{G^k_{t_i}}$ and $\mathcal{O}_{G^k_{t_i}}$, we now further define the ***pattern*** of a transaction subgraph. The pattern of a $k$-order transaction graph of transaction $t_i$ is denoted as $G^k_{(m,n)} = \{G^k_{t_i} || \mathcal{I}_{G^k_{t_i}}| = m, |\mathcal{O}_{G^k_{t_i}}| = n\}$, where $m$ and $n$ are the number of input addresses and output addresses of $G^k_{t_i}$ respectively.

For a given transaction graph generated from a blockchain transaction record during a specific period $T$, we can obtain a $k$ order transaction subgraph $G^k_{t_i}$ of each transaction $t_i \in T$. The obtained transaction subgraphs may belong to different patterns. For the example in Fig. 2, $G^2_{t_1}$ belongs to the pattern $G^2_{(2,1)}$, while $G^2_{t_2}$ belongs to the pattern $G^2_{(1,1)}$.

We believe these patterns contain valuable information revealing the characteristics of each corresponding blockchain transaction in a period. In addition, the patterns obtained under different order $k$ can reveal different levels of latent information. The benefit of denoting the pattern based on the number of input addresses and out addresses is that the patterns can be easily encoded into matrices, and therefore can be adopted as the features of the current transaction graph.

By summarizing the patterns of all $k$-order transaction graph $G^k_{t_i}$ of every transaction $t_i$ in a transaction graph $G$, two key characteristics can be obtained 1) what kinds of patterns occur in the transaction graph, and 2) how many times these patterns occur. We extend the concept of occurrence matrix in literature [2] to a $k$ order pattern occurrence matrix, denoted as $OC^k$, where the entry of $OC^k$ is $OC^k_{(m,n)} = |G^k_{(m,n)}|$. The entry of pattern occurrence matrix $OC^k_{(m,n)}$ denotes the number of $k$-order transaction graphs that belong to the pattern $G^k_{(m,n)}$.

Finally we concatenate $OC^k$ for $k = 1, 2, 3, .., s$ as the feature $v$ of the transaction graph $G$ we obtain from the blockchain transaction record. Now the *Bitcoin Price Prediction* problem can be specified in detail: use the feature vector $v$ that is calculated from the transaction graph based on Bitcoin historical data in time period $[t-i, t]$, to predict bitcoin price at some future time $t+h$, $P_{t+h}$. Formally, we define the *price prediction task* as Definition 4.

**Definition 4.** (Bitcoin Price Prediction): *Given time $t' = t + \Delta t$, where $\Delta t \geq 0$, and Bitcoin historical data in time period $[t-s, t]$, where $s \in N^+$. Let $P_t$ denote the price of bitcoin at time $t$. the bitcoin price prediction problem is to predict the price at time $t'$, e.g. $P_{t'}$.*

## 4 Multi-window Prediction Framework

Transactions in the blockchain are time sequential, meaning the blockchain may shows different patterns at different periods of time. How much the future price is influenced by historical patterns and how far back we should look to discover these patterns are empirical questions. To answer these questions more systematically, we propose the Multi-Window Prediction Framework. This framework uses the features from different lengths of historical data to construct different submodels and incorporates the results from every submodel to form a final result. By taking advantage of all the submodels, this framework can boost the accuracy of our predictions.

Figure 3 illustrates the $Multi-Window\ Prediction\ Framework$. $M_1$ to $M_s$ are $s$ submodels that are trained separately on different windows of time with length $s$. For example, $M_1$ is the model trained by the features extracted from the past 1 day, and $M_2$ is the model trained by the features extracted from the past 2 days. When making price forecasts for a specific day $t' = t + \Delta t$ ($\Delta t \geq 0$),

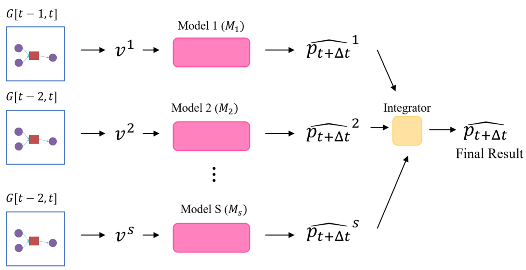



**Fig. 3.** Overview of the Multi-window Prediction Framework

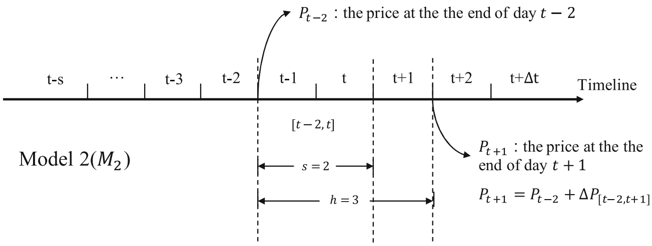



**Fig. 4.** Illustration of settings for submodel 2 ($M_2$) to predict $P_{t+1}$

each submodel will first output its individual prediction. The integrator will then combine the results into one final result.

The accuracy of the final result depends on the performance of each submodel. Next, we describe how each model is trained and makes future price predictions. In this paper, we predict the daily end price of bitcoin. The end price of day $t'$ is denoted as $P_{t'}$. After extracting features from a historical period, say $[t-s,t]$, it is natural to directly predict $P_{t'}$. However, it is more reasonable to predict the price difference between $P_{t'}$ and $P_{t-s}$, denoted as $\Delta P_{[t-s,t']}$, and then derive the predicted $P_t$ as $\hat{P_t} = P_{t-s} + \Delta \hat{P_{[t-s,t']}}$. The reason is twofold: 1) we know the historical price $P_{t-s}$, and it should be considered to improve the prediction; 2) whatever features extracted from $[t-s,t]$ represents the characteristics only during $[t-s,t]$ in the bitcoin market, and these are the characteristics that bring changes to the price. Thus, it is more reasonable to use the features to interpret the price change rather than the exact price. Therefore, in this paper, we construct data sample pairs as $(\mathbf{x}, y)$, where $\mathbf{x}$ is the feature vector extracted from a historical period $[t-s,t]$, and $y = \Delta P_{[t-s,t']} = P_{t'} - P_{t-s}$. Each submodel will be retrained if it aims to predict a different future time. We denote the distance from the future time to be predicted as $h = t' - (t-s)$. Figure 4 illustrates an example of the parameters setting for submodels making predictions.

The integrator will combine the results from each submodel with different weights, which can be a simple linear function as follows:

$$\hat{P_{t'}} = r_1 * \hat{P_{t'}}^1 + r_2 * \hat{P_{t'}}^2 + ... + r_s * \hat{P_{t'}}^s \tag{1}$$

where $r_1 + r_2 + ... + r_s = 1$.

In this paper, we elaborately design the weights. Let $W_i = [r_1, r_2, r_3, ..., r_i]$. Specially, if the historical window size is 1, which indicates that we only employ one model to make the prediction, $W_1 = [r_1] = [1.0]$. As the historical window size increases, $i > 1$, $W_i$ is defined as Eq. 2:

$$\begin{aligned} W_{i+1}[k] &= W_i[k] \ (k = 1, ..., i-1) \\ W_{i+1}[i] &= W_i[i] * \alpha \\ W_{i+1}[i+1] &= W_i[i] * (1-\alpha) \end{aligned} \tag{2}$$

where $\alpha$ controls the speed of decay of weights corresponding to results from submodels with data further back in history. Equation 2 maintains the property that $\sum_{r_j \in W_i} r_j = 1$ for $i > 0$.

## 5 Experimental Results

In this section, we present the evaluation of our proposed transaction graph based blockchain feature and Multi-Window Prediction Framework.

### 5.1 Data Preparation

To conduct the bitcoin price prediction task, we collect Bitcoin blockchain historical data and bitcoin price historical data. The Bitcoin blockchain data is downloaded from Google Bigquery public dataset crypto_Bitcoin[2] whose data is exported using Bitcoin etl tool[3]. The bitcoin price data is collected from Coindesk[4].

We select two historical periods for the experiments.

- **Interval 1:** From August 19th, 2013 to July 19th, 2016. The timestamps are divided daily. This period contains 1065 days, the first 80% days are used to train the model and the latter 20% is reserved for testing.
- **Interval 2:** From April 1st, 2013 to April 1st, 2017. The timestamps are divided daily. This period contains 1461 days, the first 70% days are used to train the model and the latter 30% is reserved for testing.

*Interval 1* and *Interval 2* are identical to the datasets used in the literature [21], which will be used as a benchmark in the next sections. In this paper, we predict bitcoin daily closing price during the above periods.

For the evaluation metric, we adopt Mean Absolute Percentage Error (MAPE) to show the error between predicted prices and real prices. The MAPE is defined as $MAPE = \frac{1}{N}\sum_{i=1}^{N}\frac{|\hat{p_i} - p_i|}{p_i}$, where $\hat{p_i}$ is the predicted bitcoin price, while $p_i$ is the real realized price.

### 5.2 Performance of Difference Submodels

Table 1 shows each submodel, $M_1$ to $M_4$, where each submodel adopts the same training strategy and machine learning prediction model. They only differ by the length of the historical window of time used when extracting the features. $s$ is the length of the window of time, where $s = 1$ means the model extracts features from the past 1 day. $h$ is the future time that the model aims to predict, where $h = 1$ means the model predicts the price the next day. Due space

[2] Dataset ID is bigquery-public-data: crypto_Bitcoin at https://cloud.google.com/bigquery.

[3] https://github.com/blockchain-etl/Bitcoin-etl.

[4] https://www.coindesk.com/.

**Table 1.** MAPE of Submodels (SVM Prediction) for Predicting Future Price

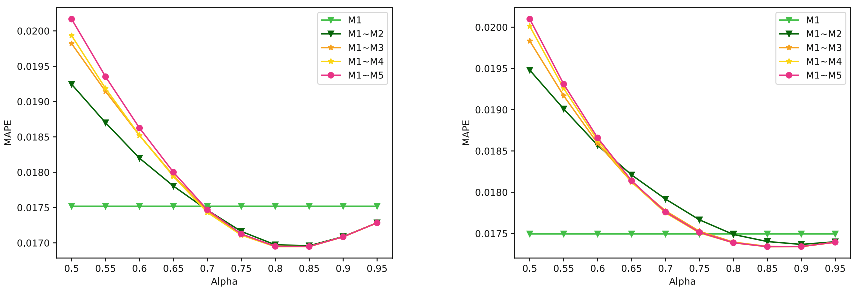
| Submodels | Interval 1 | | | | | Interval 2 | | | | | Year 2017 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | h = 1 | h = 2 | h = 3 | h = 4 | h = 5 | h = 1 | h = 2 | h = 3 | h = 4 | h = 5 | h = 1 | h = 2 | h = 3 | h = 4 | h = 5 |
| M1 (s = 1) | 1.75% | 2.59% | 3.15% | 3.77% | 4.31% | 1.74% | 2.57% | 3.21% | 3.78% | 4.29% | 4.73% | 7.09% | 8.36% | 10.30% | 12.20% |
| M2 (s = 2) | – | 2.61% | 3.16% | 3.76% | 4.29% | – | 2.58% | 3.20% | 3.78% | 4.29% | – | 7.01% | 8.36% | 10.05% | 11.90% |
| M3 (s = 3) | – | – | 3.17% | 3.76% | 4.29% | – | – | 3.20% | 3.76% | 4.27% | – | – | 8.24% | 9.91% | 11.70% |
| M4 (s = 4) | – | – | – | 3.75% | 4.29% | – | – | – | 3.78% | 4.28% | – | – | – | 9.85% | 11.60% |

constraints, we only show the results where each submodel adopts the Support Vector Machine (SVM) algorithm, which is the best in our record. We find that including more historical information in our models does not necessarily result in better performance in terms of MAPE. For example, $M_2$ at $h = 2$ obtains a worse prediction than $M_1$ at Interval 1 and Interval 2, despite the fact that $M_2$ considers one further day back than $M_1$. One can identify additional similar cases in Table 1. Therefore, we expect to achieve a higher MAPE by taking into consideration all the different submodels.

### 5.3 Performance of Combined Model

Figure 5 shows the effects of combining the submodels to produce the final prediction. $M_1$ means only submodel $M_1$ is adopted, $M_1 \sim M_2$ means the results from both submodels $M_1$ and $M_2$ were both used, $M_1 \sim M3$ means the results from submodels $M_1$, $M_2$ and $M_3$ were used, and so on. When $\alpha > 0.7$ in Interval 1, and $\alpha > 0.75$ in Interval 2, we can see the combined models outperform the single model (only $M_1$). When $\alpha = 0.85$ the Multi-Window Prediction Framework can produce the most accurate prediction with the lowest MAPE value.

Table 2 shows the specific MAPE values when $\alpha = 0.85$ and $h = 1$. We can observe that $M_1 \sim M_4$ produces the best results. Therefore we can conclude that 4-day historical information seems to be sufficient for predicting bitcoin price with our proposed Multi-Window Prediction Framework. The results reflect the high volatility of bitcoin price where the current price does not relate much to




(a) Interval 1: MAPE of submodels (b) Interval 2: MAPE of submodels

**Fig. 5.** MAPE of Multi-Window Prediction Framework when combining different submodels and alpha in Interval 1 and Interval 2 ($h = 1$), all using SVM

**Table 2.** MAPE of Multi-Window Prediction Framework when combining different submodels (SVM, $\alpha = 0.85$, $h = 1$)

| Submodels incorporated | Interval 1 | Interval 2 |
|---|---|---|
| M1 | 1.75% | 1.74% |
| M1+M2 | 1.70% | 1.73% |
| M1+M2+M3 | 1.70% | **1.72**% |
| M1+M2+M3+M4 | **1.69**% | 1.72% |
| M1+M2+M3+M4+M5 | 1.70% | 1.72% |
| $Mallquietal. - SVM$ [21] | 1.91% | 1.81% |

historical prices too far back and, instead, is more highly influenced by very recent characteristics in the Bitcoin blockchain.

### 5.4 Comparison with Benchmark

Mallqui et al. [21] study a similar bitcoin price prediction task. Mallqui et al. utilize several machine learning methods to forecast bitcoin price based on the proposed features including historical price, volume of trades and financial indicators. Since the SVM model performs the best in [21], we adopt the SVM prediction model for comparison, denoted as $Mallquietal. - SVM$. The result of $Mallquietal. - SVM$ on both Interval 1 and Interval 2 are shown in Table 2. Our proposed combined model $M1 \sim M4$ outperforms $Mallquietal. - SVM$.

## 6 Conclusion

In this paper, we propose a transaction graph based machine learning method to forecast the price of bitcoin. The $k$-order transaction graphs of the transactions are proposed to reveal the transaction patterns in the Bitcoin blockchain. The occurrence matrix is then defined to encode the information patterns and we further represent them as features of the Bitcoin blockchain. We also propose the Multi-Window prediction framework to learn the transaction patterns from multiple blockchain historical periods. Results of comparative experiments show that the method we propose outperforms recent state-of-art methods, further demonstrating the effectiveness of our method.

## References

1. Aalborg, H.A., Molnár, P., de Vries, J.E.: What can explain the price, volatility and trading volume of bitcoin? Financ. Res. Lett. **29**, 255–265 (2019)
2. Abay, N.C., et al.: ChainNet: learning on blockchain graphs with topological features. In: Wang, J., Shim, K., Wu, X. (eds.) 2019 IEEE International Conference on Data Mining, ICDM 2019, Beijing, China, 8–11 November 2019, pp. 946–951. IEEE (2019)

3. Aggarwal, A., Gupta, I., Garg, N., Goel, A.: Deep learning approach to determine the impact of socio economic factors on bitcoin price prediction. In: 2019 Twelfth International Conference on Contemporary Computing, IC3 2019, Noida, India, 8–10 August 2019, pp. 1–5. IEEE (2019)
4. Akcora, C.G., Dey, A.K., Gel, Y.R., Kantarcioglu, M.: Forecasting bitcoin price with graph chainlets. In: Phung, D., Tseng, V.S., Webb, G.I., Ho, B., Ganji, M., Rashidi, L. (eds.) PAKDD 2018. LNCS (LNAI), vol. 10939, pp. 765–776. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-93040-4_60
5. Balcilar, M., Bouri, E., Gupta, R., Roubaud, D.: Can volume predict bitcoin returns and volatility? A quantiles-based approach. Econ. Model. **64**, 74–81 (2017)
6. Balfagih, A.M., Keselj, V.: Evaluating sentiment classifiers for bitcoin tweets in price prediction task. In: 2019 IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA, 9–12 December 2019, pp. 5499–5506. IEEE (2019)
7. Burnie, A., Yilmaz, E.: An analysis of the change in discussions on social media with bitcoin price. In: Piwowarski, B., Chevalier, M., Gaussier, É., Maarek, Y., Nie, J., Scholer, F. (eds.) Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2019, Paris, France, 21–25 July 2019, pp. 889–892. ACM (2019)
8. Cerda, G.C., Reutter, J.L.: Bitcoin price prediction through opinion mining. In: Amer-Yahia, S., et al. (eds.) Companion of The 2019 World Wide Web Conference, WWW 2019, San Francisco, CA, USA, 13–17 May 2019, pp. 755–762. ACM (2019)
9. Chen, C., Chang, J., Lin, F., Hung, J., Lin, C., Wang, Y.: Comparison of forcasting ability between backpropagation network and ARIMA in the prediction of bitcoin price. In: 2019 International Symposium on Intelligent Signal Processing and Communication Systems, ISPACS 2019, Taipei, Taiwan, 3–6 December 2019, pp. 1–2. IEEE (2019)
10. Chen, Z., Li, C., Sun, W.: Bitcoin price prediction using machine learning: an approach to sample dimension engineering. J. Comput. Appl. Math. **365** (2020)
11. Ciaian, P., Rajcaniova, M., d'Artis Kancs: The economics of bitcoin price formation. Appl. Econ. **48**(19), 1799–1815 (2016). https://doi.org/10.1080/00036846.2015.1109038
12. Felizardo, L., Oliveira, R., Del-Moral-Hernandez, E., Cozman, F.: Comparative study of bitcoin price prediction using WaveNets, recurrent neural networks and other machine learning methods. In: 6th International Conference on Behavioral, Economic and Socio-Cultural Computing, BESC 2019, Beijing, China, 28–30 October 2019, pp. 1–6. IEEE (2019)
13. Georgoula, I., Pournarakis, D., Bilanakos, C., Sotiropoulos, D.N., Giaglis, G.M.: Using time-series and sentiment analysis to detect the determinants of bitcoin prices. In: Proceedings of the 9th Mediterranean Conference on Information Systems, MCIS 2015, Samos, Greece, 2–5 October 2015, p. 20. AISeL (2015)
14. Hashish, I.A., Forni, F., Andreotti, G., Facchinetti, T., Darjani, S.: A hybrid model for bitcoin prices prediction using hidden Markov models and optimized LSTM networks. In: 24th IEEE International Conference on Emerging Technologies and Factory Automation, ETFA 2019, Zaragoza, Spain, 10–13 September 2019, pp. 721–728. IEEE (2019)
15. Koutmos, D.: Bitcoin returns and transaction activity. Econ. Lett. **167**, 81–85 (2018)
16. Kristoufek, L.: Bitcoin meets google trends and Wikipedia: quantifying the relationship between phenomena of the internet era. Sci. Rep. **3**(1), 1–7 (2013)
17. Kristoufek, L.: What are the main drivers of the bitcoin price? Evidence from wavelet coherence analysis. PloS One **10**(4), e0123923 (2015)

18. Kyle, A.S.: Continuous auctions and insider trading. Econometrica: J. Econometric Soc. 1315–1335 (1985)
19. Llorente, G., Michaely, R., Saar, G., Wang, J.: Dynamic volume-return relation of individual stocks. Rev. Financ. Stud. **15**(4), 1005–1047 (2002)
20. Maesa, D.D.F., Marino, A., Ricci, L.: Uncovering the bitcoin blockchain: an analysis of the full users graph. In: 2016 IEEE International Conference on Data Science and Advanced Analytics, DSAA 2016, Montreal, QC, Canada, 17–19 October 2016, pp. 537–546. IEEE (2016)
21. Mallqui, D.C.A., Fernandes, R.A.S.: Predicting the direction, maximum, minimum and closing prices of daily bitcoin exchange rate using machine learning techniques. Appl. Soft Comput. **75**, 596–606 (2019)
22. Mittal, A., Dhiman, V., Singh, A., Prakash, C.: Short-term bitcoin price fluctuation prediction using social media and web search data. In: 2019 Twelfth International Conference on Contemporary Computing, IC3 2019, Noida, India, 8–10 August 2019, pp. 1–6. IEEE (2019)
23. Naeem, M., Bouri, E., Boako, G., Roubaud, D.: Tail dependence in the return-volume of leading cryptocurrencies. Financ. Res. Lett. **36**, 101326 (2020)
24. Nakamoto, S.: Bitcoin: a peer-to-peer electronic cash system (2009)
25. Nguyen, D.-T., Le, H.-V.: Predicting the price of bitcoin using hybrid ARIMA and machine learning. In: Dang, T.K., Küng, J., Takizawa, M., Bui, S.H. (eds.) FDSE 2019. LNCS, vol. 11814, pp. 696–704. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-35653-8_49
26. Pieters, G., Vivanco, S.: Financial regulations and price inconsistencies across bitcoin markets. Inf. Econ. Policy **39**, 1–14 (2017)
27. Schneider, J.: A rational expectations equilibrium with informative trading volume. J. Financ. **64**(6), 2783–2805 (2009)
28. Vassiliadis, S., Papadopoulos, P., Rangoussi, M., Konieczny, T., Gralewski, J.: Bitcoin value analysis based on cross-correlations. J. Internet Bank. Commer. **22**(S7), 1 (2017)
29. Yao, W., Xu, K., Li, Q.: Exploring the influence of news articles on bitcoin price with machine learning. In: 2019 IEEE Symposium on Computers and Communications, ISCC 2019, Barcelona, Spain, 29 June–3 July 2019, pp. 1–6. IEEE (2019)
30. Yermack, D.L.: Is bitcoin a real currency? An economic appraisal. Econ. Innov. eJournal (2013)
31. Yogeshwaran, S., Kaur, M.J., Maheshwari, P.: Project based learning: predicting bitcoin prices using deep learning. In: Ashmawy, A.K., Schreiter, S. (eds.) IEEE Global Engineering Education Conference, EDUCON 2019, Dubai, United Arab Emirates, 8–11 April 2019, pp. 1449–1454. IEEE (2019)