Computational Economics (2021) 57:267–280
https://doi.org/10.1007/s10614-020-10057-7

The Determinants of Bitcoin's Price: Utilization of GARCH and Machine Learning Approaches

Ting-Hsuan Chen[1] · Mu-Yen Chen[2] · Guan-Ting Du[1]

Accepted: 4 October 2020 / Published online: 9 November 2020

Abstract

This study explores the determinants of Bitcoin's price from 2010 to 2018. This study applies Generalized Autoregressive Conditional Heteroskedastic model to investigate the Bitcoin datasets. The experimental results find the Bitcoin price has positive relationship to the exchange rates (USD/Euro, USD/GBP, USD/CHF and Euro/GBP), the DAX and the Nikkei 225, while a negative relationship with the Fed funds rate, the FTSE 100, and the USD index. Especially, Bitcoin price is significantly affected by the Fed funds rate, followed by the Euro/GBP rate, the USD/GBP rate and the West Texas Intermediate price. This study also executes the decision tree and support vector machine techniques to predict the trend of Bitcoin price. The machine learning approach could be a more suitable methodology than traditional statistics for predicting the Bitcoin price.

Keywords Generalized Autoregressive Conditional Heteroskedastic Model (GARCH) · Decision tree · Support vector machine · Bitcoin price

1 Introduction

Virtual currency, also known as cryptocurrencies or digital currencies, are currencies that can be process transactions online without a third party. Virtual currency that has been devised for anonymous payments made entirely independently of governments and banks. Bitcoin has the highest market value and transaction volume among all virtual currency currently in circulation. In the public media community, there is ongoing a lively debate on the extreme price volatility of Bitcoin in the economy, and actually function as a payment tool such as US dollar or Euro etc.

✉ Mu-Yen Chen
mychen119@gs.ncku.edu.tw

1 National Taichung University of Science and Technology, Taichung, Taiwan

2 Department of Engineering Science, National Cheng Kung University, Tainan, Taiwan

Bitcoin is first proposed by Nakamoto (2009). It is a two-way virtual currency can be exchanged for legal tender at a specific exchange rate, and is a Peer to Peer (P2P) trading platform using Stigmergy software. This approach gave rise to virtual currencies and blockchain technology. Bitcoin is largest virtual currency in terms of total market value and recognition. Bitcoin is characterized by wide exchange rate fluctuations against other conventional currencies, raising questions about its utility as a store of value or investment.

The number of bitcoins is capped at 21 million to avoid inflation. Bitcoin requires increasingly laborious computations to produce, and the cost of producing a bitcoin through the mining process is rising. This increase in input prices have an influence on Bitcoin prices. New bitcoins are introduced into the market when miners process blocks of transactions and the rate at which new coins are introduced is designed to slow over time. Their growth rate has slowed from 6.9% in 2016 to 4.0% in 2018.

This issue has raised considerable interest in identifying factors that affect the fluctuation of the Bitcoin price. This study investigates what drives the Bitcoin's price to understand the reason of its fluctuation. This study both applies the time-series statistical model and machine learning approach from July 2010 to December 2018 period.

This paper is organized into the following sections: Sect. 2 introduces Bitcoin and reviews the relevant literature. Section 3 presents the research methodologies. Section 4 describes the experimental design. Section 5 presents the empirical results. Finally, Sect. 6 concludes the findings and the future work.

2 Literature Review

2.1 Bitcoin and Government Attitudes Towards Bitcoin

The initial Bitcoin structure was designed to allow for bilateral transactions without the need for a third-party organization. Peer-to-peer (P2P) networks allow traders to directly complete bilateral transactions, thus reducing middlemen fees and ensuring user anonymity, but also removing third-party supervision and guarantees provided (Conti et al. 2018). While Bitcoin still has some problems that need to be addressed, an emerging consensus suggests that Bitcoin has potential to serve as a legitimate financial product in the future.

Tsukerman (2015) suggests two reasons for the increased global attention on Bitcoin. First, Bitcoin can be used as a currency or store of value. Second, its value fluctuates strongly. Based on these two characteristics, many investors have adopted Bitcoin as a commodity for speculation. While Grinberg (2012) raises important concerns about Bitcoin, many scholars see Bitcoin as having positive impact on micropayment infrastructure and virtual networks. However, user anonymity makes Bitcoin well suited to illegal activities such as money laundering, tax evasion, drug trafficking, etc. While there is no clear definition of Bitcoin, each country has different views as to its value. Cofnas (2014) proposed that, in the current market, Bitcoin can be used either as a currency or a financial product depending on how it is used (i.e., either for a medium exchange for goods and services or as a medium for speculation). Damodaran (2017) identifies Bitcoin as a currency, but one which generates no cash flow. Urquhart (2016) calls Bitcoin an emerging investment target that is still in its infancy but, if Bitcoin becomes more widely used in the future, and more arguments for Bitcoin can be made, Bitcoin can be established as an investment target thus increasing its efficiency of use.

The literature also notes that Bitcoin could be a buffer. Wu and Pandey (2014) point out that Bitcoin efficiency improves by combining with other investment targets. Brière, Oosterlinck, and Szafarz (2015) note that Bitcoin offers advantages through diversification. Brière et al. (2015) used Bitcoin in a portfolio with other assets and found a significant relationship with gold and global inflation bonds. Harper (2013) also suggested that the limited supply of Bitcoin provides an obvious hedge against inflation, so investors can use Bitcoin and inflation-related commodities to protect their assets. But, because Bitcoin is still an emerging product, scholars have reservations of using Bitcoin as a hedge against other traditional assets.

Government around the world attitudes are distinct towards Bitcoin. Four different attitudes towards Bitcoin: legal, illegal, restricted and unrepresented. Germany was the first country to recognize Bitcoin as legal and to include it in its national regulatory system, while Thailand is believed to be the first country to ban any and all Bitcoin transactions. Other countries, such as Italy, the US and Japan recognize the legality of Bitcoin, but do so in different ways. Although the US does not recognize Bitcoin as having the characteristics of a currency, it does regard it as a legitimate financial asset, and was the first country to introduce Bitcoin futures. Japan, Italy and the UK recognize Bitcoin as a currency usable for trading. In addition, more and more physical shops have begun to accept Bitcoin as payment. In Russia, initial coin offerings (ICO) aside from Bitcoin are prohibited, but the underlying blockchain technology is allowed and seen as a potential foundation for a more stable monetary mechanism in the future. To all appearances, Bitcoin is strictly prohibited, but the government is unable to limit the private development of blockchain technologies. Finally, some countries, such as South Korea, recognize the use of Bitcoin as a currency, but heavily regulate it.

One of the key characteristics of Bitcoin is its decentralization, preventing government regulation, and many governments have yet to promulgate specific regulations for Bitcoin (Anders 2014; Grinberg 2012; Plassaras 2013). In recent years, this environment has resulted in its widespread use in tax evasion, money laundering and illegal transactions. Governments are actively formulating needed regulations to prevent such activity. In China, for example, Ju et al. (2016) note that Bitcoin-based capital outflows dried up upon the government's announcement of its intention to regulate Bitcoin, demonstrating that such intervention could also be effective against other virtual currencies.

Government action can not only reduce the negative aspects of Bitcoin, but is also a key factor affecting the fluctuation of Bitcoin value in response to external influences. Technologies which are easier to use are more likely to be widely accepted (Daştan and Gürler, 2016), thus future government regulations of Bitcoin and other virtual currencies will benefit national economic and commercial development (Franklin 2016).

2.2 Related Literature

Previous studies explore the relationships between Bitcoin price and golden price (Baur et al. 2018; Dyhrberg 2016), and other economic indicators (Brière et al. 2015). Dyhrberg (2016) used the GARCH model to study wither Bitcoin has the same properties as gold, and cited macroeconomic variables previously found to affect gold, namely the use of dollar-to-euro and dollar-to-sterling exchange rates as independent variables. His results indicated that Bitcoin prices react more strongly to the dollar-to-sterling exchange rate, indicating that different countries and regions exhibit different specific impacts on Bitcoin prices. Takaishi and Adachi (2018) also used the USD/CHF, USD/JPY and Euro/GBP exchange rates as independent variables to examine the time series related properties of Bitcoin. Comparison with other exchange rates indicated that Bitcoin does feature time series correlations. Baur et al. (2018) qualitatively explored the relationship between Bitcoin prices and USD/Euro and USD/GBP exchange rates. Although their results show that Bitcoin prices are not directly related to exchange rates, a negative correlation was found with other dollar-denominated financial assets, suggesting that fluctuations in the Bitcoin rate of return may be related to any dollar-based exchange rate financial assets. Zhu et al. (2017) used the Vector Error Correction (VEC) model to find a significant negative relationship between Bitcoin prices and the Fed funds rate.[1] This is because, when Bitcoin is used as a financial asset, rising Fed interest rates will allow funds to flow back into the US market, thus reducing investment in Bitcoin. Bouri, Molnár, Azzi, Roubaud, and Hagfors (2017) used the FTSE 100 Index, the US Dollar Index, the DAX and the Nikkei 225 to form a portfolio with Bitcoin, using daily and weekly observations. They found that the US Dollar Index had no significant correlation with the Bitcoin price either daily or weekly. They explained the positive correlation found with the FTSE 100 and the DAX as the result of, despite not being able to directly hedge Bitcoin, investors can use Bitcoin to increase portfolio diversity. Finally, a significant negative relationship was found with the Nikkei 225, but more obviously in daily observations, indicating that investors in Japan tend to use Bitcoin as a portfolio hedge. Baur et al. (2018) also found a negative correlation between the US Dollar Index and Bitcoin. The Bitcoin price is negatively correlated with financial assets denominated in USD, and is affected by fluctuations in the US dollar. Dyhrberg (2016) used the GARCH model to study Bitcoin functions like gold, comparing gold prices against gold futures and Bitcoin prices. He found that, similar to gold, the Bitcoin return rate depends on trader demand, and this correlation is not affected by currency appreciation. Therefore, in terms of rate or return, over the long run Bitcoin functions similarly to gold. Brière et al. (2015) also found a significant positive relationship with gold when studying the relationship between Bitcoin prices and other economic indicators, indicating that, similar to gold, Bitcoin is arbitrage-protected. Brière et al. (2015) also compared oil and Bitcoin prices, finding no significant correlation. But Bouri et al. (2017) found a weak negative

[1] Banks could borrow money for short periods (typically overnight) to make up transitory cash shortfalls. The interest rate that is paid on these borrowed reserves is called the federal funds rate.

Table 1 Descriptive statistics

	Mean	Median	S.D.	Max	Min	N
Constant	1540.2	329.15	2495.3	18,848	0.0495	1955
USD/Euro	0.8144	0.8032	0.0749	0.9627	0.6745	1955
USD/GPB	0.6758	0.6472	0.0654	0.8249	0.5826	1955
USD/CHF	0.9497	0.9576	0.0461	1.0604	0.7232	1955
USD/JPY	101.33	104.25	14.768	125.62	75.82	1955
Euro/GPB	0.8313	0.8433	0.0520	0.9267	0.6937	1955
Fed funds rate	0.4596	0.15	0.5898	2.4	0.04	1955
FTSE100 Index	6498.3	6557.3	667.43	7877.5	4944.4	1955
US Dollar Index	87.628	85.5	8.3429	103.3	73.12	1955
DAX	9476.7	9631.4	2247.8	13,559	5072.3	1955
Nikkei 225	15,443	15,733	4741.4	24,270	8160	1955
Gold price	1353.8	1293.5	184.58	1895	1049.4	1955
Gold future	1354.2	1293.6	184.73	1888.7	1050.8	1955
West Texas Intermediate	73.989	76.02	23.046	112.86	26.55	1955

relationship between oil and Bitcoin in the same portfolio, indicating that they can be used together, but not to particularly good effect.

3 Research Methodology

3.1 Generalized Autoregressive Conditional Heteroskedastic Model (GARCH)

The GARCH model is a time series model most commonly used to predict future price fluctuations. This study interprets GARCH results under different orders to explore determinants of price trends affecting Bitcoin.

In this study, the GARCH model was constructed using the following Eq. (1) to observe the effect of each variable on the Bitcoin price. Following prior literature, this study also uses the Bitcoin price as the main dependent variable to explore factors that influence Bitcoin price fluctuation.

$$\begin{aligned} Bitcoin\ price_t = {} & \alpha + \beta_1 \frac{USD}{EUR}_{t-1} + \beta_2 \frac{USD}{GBP}_{t-1} + \beta_3 \frac{USD}{CHF}_{t-1} + \beta_4 \frac{USD}{JPY}_{t-1} \\ & + \beta_5 \frac{EUR}{GBP}_{t-1} + \beta_6 FFR_{t-1} + \beta_7 FTSE_{t-1} + \beta_8 USdollar_{t-1} + \beta_9 DAX30_{t-1} \\ & + \beta_{10} Nikkei225_{t-1} + \beta_{11} Goldcash_{t-1} + \beta_{12} Goldfutures_{t-1} + \beta_{13} Oil_{t-1} + \varepsilon_{t-1} \end{aligned} \tag{1}$$

In Eq. (1), t is the current period, and t-1 is the lag period for the USD/Euro, USD/GPB, USD/CHF, USD/JPY, and Euro/GPB exchange rates, along with the Fed funds rate, the FTSE100 Index, the US Dollar Index, the DAX, the Nikkei 225, current gold price, gold futures and West Texas Intermediate (WTI) oil price are used to predict the current price of Bitcoin so as to determine the relative impact of each on the Bitcoin's price. Table 1 presents the descriptive statistics for all variables.

3.2 Support Vector Machine (SVM)

SVM is a supervised machine learning method that can be used for prediction. SVM divides a data set into two more precise categories. Previously, classification could only be done by drawing a line in the middle, but this could result in non-uniform data distributions. To achieve more efficient data distributions, SVM finds an optimal decision boundary to maximize the margins, thus resulting in an even data distribution.

Equation (2) presents the mathematical model by which SVM extends both margins, where $w^t x+b=0$ represents the most basic classification margin in a data set, and y_i represents the classification result, either 1 or − 1, and I is the number of data.

$$\min\frac{1}{2}w^2 s.t., y_i\left(w^t x_i + b\right) \geq 1, i = 1, \ldots n \tag{2}$$

To maximize the margins (i.e., to find the optimal margin), the study uses the KKT condition to derive Eq. (3) according to the optimization principle:

$$\begin{aligned} & max_\alpha \sum_{i=1}^{n} \alpha_i - \frac{1}{2}\sum_{i=1}^{n} \alpha_i\alpha_j y_i y_j x_i^t x_j \\ & s.t., \alpha_i \geq 0, i = 1, \ldots n \\ & \sum_{i=1}^{n} \alpha_i y_i = 0 \end{aligned} \tag{3}$$

In the research, polynomial kernel function was selected and implemented in the SVM model. The first consideration of choosing polynomial kernel function is the because it has fewer parameters to select and has significant performance to other kernel functions for the hybrid SVM model with stochastic optimization technique (Harish et al. 2015). Second, Fujibuchi and Kato (2007) found the different kernel function - linear, polynomial, Gaussian radial basis function kernel function (RBF) showed the similar accuracies for the heterogeneous microarray data when no noise was added. However, the standard linear and RBF kernels decreased the accuracies with increasing noise levels in the training datasets. Third, Sebtosheikh and Salehi (2015) claimed that the prediction by SVMs using small training datasets, they recommended to use the polynomial kernel function. The empirically results showed the polynomial kernel function outperformed than the RBF kernel function. Fourth, Chiang et al. (2020) used the second-order polynomial SVM model and found it can obtain a good approximation of the limit state without high computational complexity.

3.3 Decision Tree

Decision trees are a means of data mining which can be applied to classification forecasting, and uses a tree branch pattern (i.e., starting from a root at the top and diverging into branches) to identify the most relevant factors to be used as decision indicators.

Decision trees are divided into two categories. The first is a classification tree for qualitative data, e.g., gender, education, etc. The second is a regression tree that can analyze quantitative data. This study uses the C4.5 model which is more advanced decision tree algorithm. The C4.5 algorithm generates a decision tree from a set Y of cases. If Y satisfies a stopping criterion, the tree for Y is a leaf associated with the most frequent class in Y. The tree for Y has test T as its root with one subtree for each outcome Ti that is constructed by applying the same procedure recursively to the cases in Di. The classification results can be used for predictive analysis.

Decision trees permits us to visualize trees at different pruning stages. The branch diagram proceeds downward from the top, and items closer to the root have a greater influence on the subject of investigation, allowing us to easily identify key factors.

4 Empirical Design

4.1 Data Source

The data in this study are collected from bitcoinity.org, investing.com, Federal Reserve Economic Data (FRED) and the World Gold Council website. Bitcoinity.org is used to receive Bitcoin prices. Other variables including stock market indices, oil prices, exchange rates, interest rates and gold prices are obtained from Investing.com, FRED and the World Gold Council website. The sample time period ranges from July 19, 2010 to December 31, 2018. The earliest date bitcoin was traded is July 19, 2010. All data are obtained as daily values. Unavailable and non-trading days are excluded from the analysis. The final sample are 1955 observations.

4.2 Parameters Setting

SVM works by finding a line that best separates the data into the two groups. A key parameter is the type of Kernel to use. In this research, the experiments is adopted Polynomial Kernel that will separate the classes using a curved or wiggly line, the higher the polynomial, the more wiggly exponent value. Accordingly, this research is set the parameter of complexity, which controls how flexible the process for drawing the line to separate the classes can be, equal to 1 to allow violations of the margin. In the SVM model, the searching range of C and γ were set between 1 and 100 (Hsu et al. 2010).

In the decision tree classification, the tree produced through the growing and pruning phase of C4.5. Furthermore, the experiments setup the C4.5 parameters, the confidence factor, CF (default value of 25%), the minimum numbers of split-off cases, MS (default value of 2), and the batch size is 100.

In order to perform the experiments, the database is separated in two groups; a training set and a test set. The training data and testing data set undergo 10-fold cross validation, where ten runs are performed in each fold and the result of the fold is the average result of those ten runs.

4.3 Performance Metrics

In this study, model evaluation is conducted using performance indicators Precision, Recall and Accuracy, derived as follows:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{4}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{5}$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{6}$$

where TP = True positive; FP = False positive; TN = True negative; FN = False negative.

TP (True Positive) refers to data correctly placed in the appropriate position. FP (False Positive) refers to data that is eventually determined to be positive but was originally identified as negative. Recall is TP/(TP + FN); Precision is TP/(TP + FP).

5 Empirical Analysis

5.1 GARCH Model

Time series are characterized by non-stationary data, thus this study needs to use a unit root test to process the original non-stationary data into a fixed state for empirical analysis. Table 2 shows the results of the ADF unit root test, presenting the verification values of each variable after the original sequence and the first-order difference. The numerical value shows that each variable is significant at the 5% level in the original sequence check, indicating that this study cannot reject the null hypothesis of the sequence with a single root and non-stationary data. However, given the first-order difference, each variable rejects the null hypothesis that the sequence has a single root, indicating that the processed data is fixed, allowing us to continue empirical analysis.

This study uses ARCH-LM test for testing the volatility of the series. It is found that GARCH model has successfully captured the volatility. The null hypothesis of ARCH-LM test is rejected since the *p* value had lower than 1% of significance level. F-statistic test also rejected the null hypothesis at the same condition. The ARCH-LM test on the residuals of this model showed that ARCH effect exists, implying the conditional heteroscedasticity was present in Bitcoin price series.

In this paper, the GARCH model is used for empirical analysis, with the results summarized in Table 3. The order is set at (1,1), (1,2), (1,3). When the GARCH

Table 2 ADF Unit root test

Variables	Values	
	Origin	First difference
Bitcoin price	– 2.640916	– 10.27300 ***
USD/Euro	– 2.298625	– 45.10295 ***
USD/GPB	– 2.106370	– 13.47230 ***
USD/CHF	– 4.600922	– 11.42138 ***
USD/JPY	– 0.688731	– 44.52244 ***
Euro/GPB	– 1.845967	– 44.29262 ***
Fed funds rate	– 1.789769	– 19.61394 ***
FTSE100 Index	– 1.789769	– 19.61394 ***
US Dollar Index	– 2.416499	– 45.38820 ***
DAX	– 2.202357	– 18.53821 ***
Nikkei 225	– 2.201568	– 46.61062 ***
Gold price	– 2.723796	– 45.08760 ***
Gold future	– 2.846214	– 45.96046 ***
West Texas Intermediate	– 1.965975	– 31.38048 ***

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

model order is set to (1,1), we find 1% to 5% significance for a positive correlation of Bitcoin prices to the USD/Euro, USD/GPB, USD/CHF, and Euro/GBP exchange rates, along with the DAX, the Nikkei 225 and the current price of gold, while a significant inverse correlation is found between the Bitcoin price against the Fed funds rate, the FTSE 100 and the US dollar index. Relationships are similarly interpreted for order (1,2), while order (1,3) shows an additional significant positive relationship between Bitcoin and the West Texas crude price.

This study uses Akaike information criteria (AIC) to evaluate the suitability of each of the three order models and select the optimal model. The smaller the AIC value, the better the fit. Using AIC values, the experimental results found that the GARCH(1,2) model provided the best fit.

5.2 Support Vector Machine

Table 4 shows the classification and prediction results using SVM. We use lagged independent variables mentioned in Eq. (1) as inputs in this model. The model achieves accurate prediction outcomes as high as 80.6138% for Bitcoin prices. True Positive (TP) Rate, False Positive (FP) rate, and Precision are shown in this confusion matrix. TP rate is the ratio of all true results. FP rate is the ratio of all false results. Precision is the ratio of all true results to all predicted values. It can be clearly seen that the TP rate is 80% when the value is true, and FP rate is nearly 20% when the value is negative, the precision is also 81% and the recall rate is also 80%. This study classifies according to below (L) and above (H) the median Bitcoin price. In the confusion matrix, letters a (L) and b (H) are assigned to be the class values

Table 3 GARCH model

Variables	(1,1)	(1,2)	(1,3)
Constant	– 5457.879 (– 2.006268)	– 5513.692 ** (– 2.393612)	– 5510.703 ** (– 2.348255)
USD/Euro	21950.28 *** (5.989018)	– 21742.07 *** (– 6.566146)	– 21716.34 *** (– 5.851174)
USD/GPB	24488.07 *** (12.42582)	– 24241.88 *** (– 13.88814)	– 24213.72 *** (– 14.48764)
USD/CHF	23396.19 *** (26.50957)	– 23357.82 *** (26.85870)	– 23359.90 *** (26.51297)
USD/JPY	7.003273 (0.966764)	– 7.808649 (– 1.423688)	– 6.358561 (– 1.172701)
Euro/GPB	9970.488 *** (9.015034)	– 10183.02 *** (– 10.10241)	– 10221.25 *** (– 9.855308)
Fed funds rate	– 537.0104 *** (– 3.640583)	– 546.0177 *** (– 4.053550)	– 596.5021 *** (– 4.490357)
FTSE100 Index	– 3.277636 *** (– 14.50106)	– 3.261566 *** (– 16.73503)	– 3.280698 *** (– 22.66119)
US Dollar Index	– 797.2181 *** (– 15.61174)	– 792.8051 *** (– 17.01582)	– 798.2959 *** (– 15.23798)
DAX	0.922097 *** (10.29341)	– 0.913540 *** (– 11.18072)	– 0.927608 *** (– 22.75940)
Nikkei 225	0.828132 *** (21.29426)	– 0.818434 *** (– 22.44492)	– 0.836599 *** (– 16.00484)
Gold price	4.903433 * (1.787356)	– 4.819889 * (– 1.843007)	– 4.958377 * (– 1.194922)
Gold future	– 1.340952 (– 0.47436)	– 1.407367 (– 0.527844)	– 1.270723 (– 0.432898)
West Texas Intermediate	5.464968 (1.402763)	– 5.572220 (– 3.647464)	– 5.850481 * (– 1.754774)
AIC	15.83987	15.81310	15.83149
N	1995	1995	1995

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

and provides expected class values in rows and predicted class values for each column. This study finds that 123 of the 977 (L) data points were misclassified, while only 256 of the 978 (H) data points were misclassified.

5.3 Decision Tree

The C4.5 algorithm improves ID3 with regard to the splitting rule and the calculation method (Quinlan, 1993). It uses gain-ratio index instead as a measurement method to segment attributes and thus can reduce the influence of ID3 drawback that segmentation nodes prefer too many sub-trees.

This study used the classification regression tree C4.5 model for classification prediction. Table 5 shows that this model achieves accurate prediction outcomes as high as

Table 4 Confusion matrix and statistics using support vector machine

	Precise		80.6138%	
	TP rate	FP rate	Precision	Recall
	0.876	0.262	0.769	0.874
	0.738	0.126	0.854	0.738
weight	0.806	0.194	0.812	0.806
Confusion matrix				
a	b			
854	123	a = L		
256	722	b = H		

Note: TP rate (True Positive Rate) is the ratio of all true results. FP rate (False Positive Rate) is the ratio of all false results. Precision is the ratio of all true results to all predicted values; Recall is the ratio of true results to the value of all predicted results

Table 5 Confusion matrix and statistics using C4.5 classification

	Precise		98.0051%	
	TR rate	FP rate	Precision	Recall
	0.977	0.017	0.983	0.977
	0.983	0.023	0.978	0.983
weight	0.980	0.020	0.980	0.980
Confusion matrix				
a	b			
955	22	a = L		
17	961	b = H		

98% for Bitcoin prices, and the values calculated in the confusion matrix are composed of the classification and prediction results. When this study classify according to below (L) and above (H) the median Bitcoin price, this study can see it should be classified as L. Only 22 of the 977 (L) data points were misclassified, achieving an accuracy rate as high as 90%, while only 17 of the 978 (H) data points were misclassified, also achieving an accuracy rate of 90%. Therefore, this study can see that the decision tree model is highly effective for classification.

In addition, Fig. 1 illustrates the Bitcoin price built into the decision tree model. From the figure, this study can see that the Fed funds rate is the root factor, making it the primary factor in predicting Bitcoin prices, followed by the Euro/GBP rate, the USD/GBP rate, and the West Texas oil price, making these the most influential factors for the Bitcoin price. The experimental results obtain a precision of above 90%, and implying that the decision tree model is highly effective for classification.

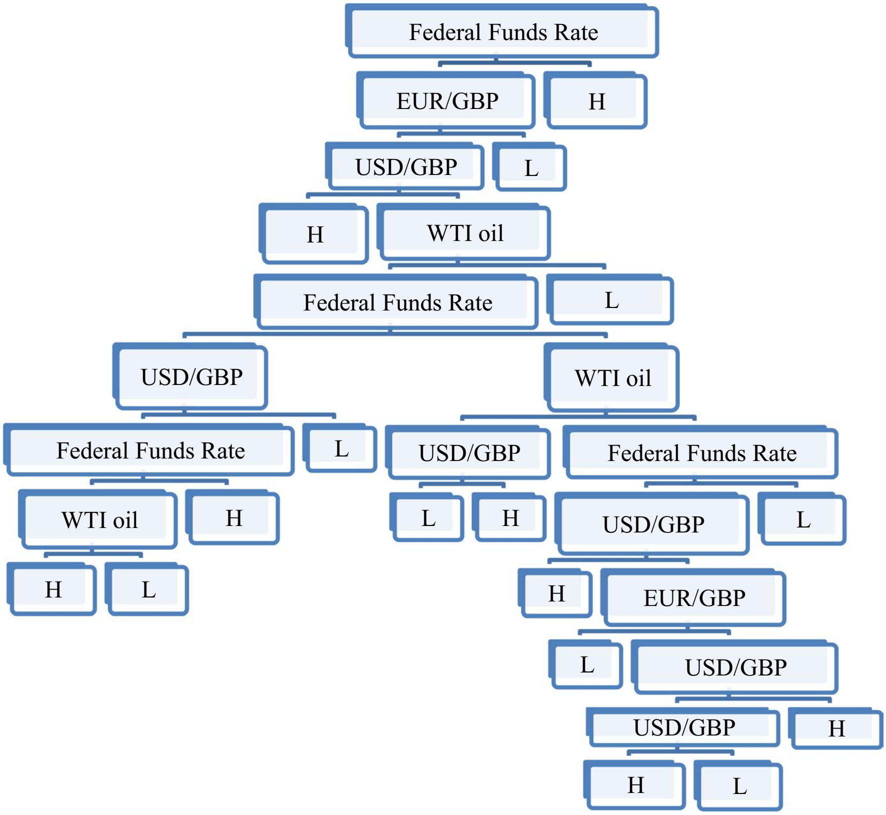

Fig. 1 Decision Tree formulated using C4.5 for Bitcoin price

6 Conclusion

With the development of Fintech, Bitcoin has emerged as a focal point of attention as a financial investment target. The rise of virtual currencies offers investors greater opportunities to diversify their portfolios. Currently, some national governments have accepted Bitcoin as a currency, and the current literature supports a currency role for Bitcoin, which can also be used for investment and hedging. Therefore, the paper seeks to combine previous research to identify key impact factors influencing the Bitcoin price, thus providing investors with deeper investment criteria reference when considering Bitcoin as an investment.

This research focuses on impact factors including various exchange rates, interest rates, stock indices, and the prices of gold and oil. Three models are applied in this study, including generalized autoregressive conditional heteroskedasticity (GARCH) time series model, support vector machine methods (SVM), and decision tree algorithm.

In the traditional GARCH financial model, this study can see that the USD/Euro, USD/GBP, USD/CHF, and the Euro/GBP exchange rates all have a positive impact on the Bitcoin price. According to the previous literature, exchange rates will have varying impacts. The Nikkei 225 index has a positive impact, as is the gold price, while the Fed funds rate and the FTSE 100 are both negative. These results confirm previous findings that these two cannot be combined in a portfolio with Bitcoin. Support vector machine is also applied. SVM has the highest forecasting precision on the Bitcoin's price. One reason that SVM could minimize an upper bound of the generalization error. Therefore, SVM is less affected by the over-fitting problem, and could obtain a high generalization performance.

This study further uses current high-precision machine learning techniques to find that the Fed funds rate has the greatest impact on the Bitcoin price in the decision tree model, due to the politicization of US monetary policy in recent years causing investors to lose confidence in the government and turn towards unregulated Bitcoin. The next most influential factors are the Euro/GBP and the USD/GBP exchange rates, due to investors seeking a hedge from the uncertainty caused by Brexit.

Based on the presented empirical analysis, this study finds that, aside from the actual Bitcoin price, potential Bitcoin investors can use the abovementioned USD/Euro, USD/GBP, USD/CHF, and Euro/GBP exchange rates, along with the Fed funds rate, the FTSE 100, the US dollar index, the DAX, the Nikkei 225 and the price of gold as reference indicators.

References

Anders, N. (2014). *The troublesome history of the Bitcoin exchange mtgox*. https://anders.io/the-troublesome-history-of-the-bitcoin-exchange-mtgox/

Baur, D. G., Dimpfl, T., & Kuck, K. (2018). Bitcoin, gold and the US dollar—A replication and extension. *Finance Research Letters, 25,* 103–110.

Bouri, E., Molnár, P., Azzi, G., Roubaud, D., & Hagfors, L. I. (2017). On the hedge and safe haven properties of Bitcoin: Is it really more than a diversifier? *Finance Research Letters, 20,* 192–198.

Brière, M., Oosterlinck, K., & Szafarz, A. (2015). Virtual currency tangible return: Portfolio divesification with Bitcoin. *Journal of Asset Management, 16*(6), 365–373.

Chiang, H. S., Sangaiah, A. K., Chen, M. Y., & Liu, J. Y. (2020). A novel artificial bee colony optimization algorithm with SVM for bio-inspired software-defined networking. *International Journal of Parallel Programming, 48,* 310–328.

Conti, M., Kumar, E. S., Lal, C., & Ruj, S. (2018). A survey on security and privacy issues of bitcoin. *IEEE Communications Surveys & Tutorials, 20*(4), 3416–3452.

Damodaran, A. (2017). *The Bitcoin bom asset, currency, commodity or collectible?* http://aswathdamodaran.blogspot.com/2017/10/the-bitcoin-boom-asset-currency.html.

Daştan, İ., & Gürler, C. (2016). Factors affecting the adoption of mobile payment systems: An empirical analysis. *Emerging Markets Journal, 6*(1), 17–24.

Dyhrberg, A. H. (2016). Bitcoin, gold and the dollar—A GARCH volatility analysis. *Finance Research Letters, 16,* 85–92.

Franklin, M. (2016). A profile of BitCoin currency: An exploratory study. *International Journal of Business & Economics Perspectives, 11*(1), 80–92.

Fujibuchi, W., & Kato, T. (2007). Classification of heterogeneous microarray data by maximum entropy kernel. *BMC Bioinformatics, 8*(267), 1–10.

Grinberg, R. (2012). Bitcoin : An innovative alternative digital currency. *Hastings Science and Technology Law Journal, 4,* 159–208.

Harish, N., Mandal, S., Rao, S., & Patil, S. G. (2015). Particle swarm optimization based support vector machine for damage level prediction of non-reshaped berm breakwater. *Applied Soft Computing, 27,* 313–321.

Harper, J. (2013). What is the value of Bitcoin. *CATO At Liberty,* https://www.cato.org/blog/what-value-bitcoin.

Hsu, C.W., Chang, C.C., & Lin, C.J. (2010). *A practical guide to support vector classification, National Taiwan University*. https://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf

Ju, L., Lu, T., & Tu, Z. (2016). Capital flight and Bitcoin regulation. *International Review of Finance, 16*(3), 445–455.

Nakamoto, S. (2009). *Bitcoin: a peer-to-peer electronic cash system*. https://bitcoin.org/bitcoin.pdf.

Plassaras, N. A. (2013). Regulating digital currencies: Bringing Bitcoin within the reach of the IMF. *Chicago Journal of International Law, 14,* 377–407.

Quinlan, J. R. (1993). *Programs for machine learning*. San Fransisco: Morgan Kaufmann.

Sebtosheikh, M. A., & Salehi, A. (2015). Lithology prediction by support vector classifiers using inverted seismic Attributes data and petrophysical Logs as a new approach and Investigation of training data set size effect on its performance in a heterogeneous carbonate reservoir. *Journal of Petroleum Science and Engineering, 134,* 143–149.

Takaishi, T., & Adachi, T. (2018). Taylor effect in Bitcoin time series. *Economics Letters, 172,* 5–7.

Tsukerman, M. (2015). The block is hot a survey of the state of Bitcoin regulation and suggestions for the future. *Berkeley Technology Law Journal, 30*(385), 1127–1170.

Urquhart, A. (2016). The inefficiency of Bitcoin. *Economics Letters, 148,* 80–82.

Wu, C. Y., & Pandey, V. K. (2014). The value of Bitcoin in enhancing the efficiency of an investor's portfolio. *Journal of Financial Planning, 27*(9), 44–52.

Zhu, Y., Dickinson, D., & Li, J. (2017). Analysis on the influence factors of Bitcoin's price based on VEC model. *Financial Innovation, 3*(1), 1–13.