Forecasting the Dynamics of Cryptocurrency Rates Based on Logistic Regression

Abdel-Badeeh M. Salem[1], Rostyslav Yurynets[2], Zoryna Yurynets[3], Grzegorz Konieczny[4], Paulina Kolisnichenko[4]

[1] Ain Shams University, 1 Elsarayat St., Cairo, 11517, Egypt
[2] Lviv Polytechnic National University, 12 Stepan Bandery str., Lviv, 79013, Ukraine
[3] Ivan Franko Lviv National University, Svobody Avenue 18, Lviv, 79000, Ukraine
[4] WSHIU Akademia Nauk Stosowanych, 28 Czerwca 1956 r. str. 217/219, Poznan, 61-485, Polska

Abstract

This study examines the simulation of cryptocurrency price changes to predict their future trajectory using discriminant analysis tools. The research focuses on identifying key quantitative financial risk variables that significantly influence the growth or decline of cryptocurrencies. Using logit regression, the study assesses the probability of categorizing cryptocurrency prices into growth or decline groups. The study shows the effect on the cryptocurrency price of the NASDAQ Composite stock index, S&P 500 stock index, Nikkei 225 stock index, the value of the SSE PLC company's stock, and the value of the company's Intel Corporation's stock. The research uses a synthesizing and deductive approach to systematize the factors influencing cryptocurrency rates and employs logistic regression analysis to uncover the elements that make up the logit model influencing cryptocurrency dynamics. Based on the developed logit regression model, this research provides investors with a valuable tool for selecting optimal investment alternatives in cryptocurrency projects.

Keywords

forecasting, machine learning methods, logistic regression, cryptocurrency, multicollinearity, method of principal components

1. Introduction

In a rapidly changing global economic environment, the study of the formation and forecasting of cryptocurrency rates is extremely interesting and relevant. The investment attractiveness of cryptocurrencies comes to the fore, while many fundamental factors can affect the volatility of cryptocurrency rates.

The general interest in digital currencies and technology trends is driving the rapid growth in their number and distribution. Cryptocurrencies are a cheap, convenient and technologically advanced way to conduct settlement transactions around the world, as well as a promising form of investment. One of the most pressing issues for market participants when conducting financial transactions with cryptocurrencies is effective forecasting of price dynamics.

Stock market forecasting has always been considered a challenging task that has attracted the attention of both academics and investors. The complexity of the task can be attributed to the many factors and uncertainties that interact in markets, including economic and political conditions, as well as human behaviour. The ability to consistently predict market price movements is difficult, but not impossible. According to scientific research, market price movements are not random but behave in a highly nonlinear and dynamic manner [1]. Previous studies have also shown that it is not necessary to be able to predict the exact value of the future

COLINS-2024: 8th International Conference on Computational Linguistics and Intelligent Systems, April 12–13, 2024, Lviv, Ukraine
abmsalem@yahoo.com (Abdel-Badeeh M. Salem); rostyslav.v.yurynets@lpnu.ua (R. Yurynets); zoryna_yur@ukr.net (Z. Yurynets); rektor@wshiu.pl (G. Konieczny); paulina.kolisnichenko@wshiu.pl (P. Kolisnichenko)
0000-0003-0268-6539 (Abdel-Badeeh M. Salem); 0000-0003-3231-8059 (R. Yurynets); 0000-0001-9027-2349 (Z. Yurynets); 0000-0002-6337-9738 (G. Konieczny); 0000-0001-6730-1236 (P. Kolisnichenko)

price to profit from financial forecasts. Predicting the direction of market movement versus its value can lead to higher profits [2].

The availability of many types of data and numerous resources makes cryptocurrencies a good research subject from which to gain insights into market behaviour by applying machine learning techniques.

At the current moment, the task of searching for and developing special tools that allow us to foresee and predict the adjustments of cryptocurrency exchange rates seems extremely urgent. In the scientific publication space, research questions are mostly considered, aimed either at expert assessment of the current and upcoming prospects of development of the considered market or at the use of special methods of exchange technical analysis, revealing the features and trends of exchange rate fluctuations of cryptocurrencies. At the same time, the use of special methods of economic and mathematical modelling, involving the use of progressive tools and mechanisms, seems appropriate in the framework of research activities.

The logistic regression model informs the investors about the alleged success of the dynamics of cryptocurrency rates based on the selected factors.

The approach to forecasting the dynamics of cryptocurrency rates is based on the use of the major factors: trading volume (per day), the price of cryptocurrency, cryptocurrency capitalization, the NASDAQ Composite stock index, the S&P 500 stock index, the Nikkei 225 stock index, the value of the SSE PLC share (SSE. United Kingdom. GBR), the value of Intel Corporation (INTC, USA) shares, the cost of Brent oil futures (LCOH8, USD/barrel), the cost of copper futures (MCUcl, USD/ton), positive and negative news for investing related to the features of cryptocurrency as an asset (expert assessment on a 5-point scale).

The research is founded on the following methods: synthesis and deductive approach to systematization of factors that affect dynamics of cryptocurrency rates; logistic regression analysis to disclose the elements of creating the logit model of the influence of factors on the dynamics of cryptocurrency rate.

2. Related Works

Cryptocurrencies were created and recognised as a new electronic method of currency exchange. Cryptocurrency trading is regarded as one of the most popular and promising types of profitable investments. This financial market is characterised by significant volatility and strong price fluctuations over time. Currently, cryptocurrency forecasting is usually considered one of the most difficult forecasting problems due to the large number of factors and significant volatility of cryptocurrency prices, which leads to complex dependencies [2-3].

Due to the growing popularity of cryptocurrencies, new empirical data appear very quickly; therefore, the literature that analyses the properties of volatility in the cryptocurrency market [4, 5], as well as between cryptocurrencies and other financial assets [10-11], is growing.

There are studies in the literature that use machine learning techniques to forecast cryptocurrency prices and directions for improving forecasting accuracy. In particular, Derbentsev et al. [2] simulated the short-term dynamics of the three most capitalized cryptocurrencies using several complex forecasting models. They evaluated the predictive performance of an artificial neural network, a random forest, and a binary autoregressive tree model. Their experimental results showed that the first model and the third model have an average directional motion prediction accuracy of 63%, which is significantly higher than the "naive" model.

Chowdhury et al. [3] applied machine learning forecasting models on the index and component of cryptocurrencies for forecasting. Their main goal was to predict nine major cryptocurrencies. They used various machine learning models including gradient-boosted trees, ANN, nearest neighbor, and robust ensemble learning models. Ensemble models and gradient-boost trees demonstrated the best predictive performance, which was competitive with, and sometimes better than, similar state-of-the-art models proposed by other authors.

Pintelas et al. [10] conducted a study evaluating complex deep-learning models for cryptocurrency price forecasting. Their research revealed significant limitations of deep learning models for obtaining reliable predictions. Based on the experimental analysis, the authors emphasized the need to adopt better algorithmic approaches for the development of efficient and reliable cryptocurrency models.

Patel et al. [9] proposed a hybrid cryptocurrency forecasting approach that focuses on such cryptocurrencies as Monero and Litecoin. The model is based on a recurrent neural network architecture that applies LSTM and GRU layers. Experiments have shown that the proposed model outperforms traditional LSTM networks, showing some promising results.

The authors [10] used machine learning unique techniques to solve the important problem, both a multiple regression method based on highly correlated features and a deep learning mechanism that applications a conjugate gradient mechanism combined with a linear search to guess the BTC price. Paper [11] analyzes the price movements of such cryptocurrencies as Ethereum, Bitcoin and Ripple. The authors use artificial intelligence systems involving a fully linked artificial neural network (ANN) and a long-term recurrent memory neural network (LSTM) and find that ANN relies more on long-term history, while LSTM relies more on short-term dynamics, which means that LSTM is more effective in extracting significant information from historical memory than ANN. A study [12] on predicting the daily price of Bitcoin cryptocurrency based on high measurement data shows that logistic regression and linear discriminant analysis achieve an accuracy of 66%. On the other hand, the sophisticated machine learning algorithm outperforms the benchmark for daily price forecasting, with statistical methods and machine learning algorithms having the highest accuracy of 66% and 65.3%, respectively. The study [13] examines the use of support vector machines (SVM), neural networks (NN) and random forests (RF). The obtained calculation results indicate that machine learning and sentiment analysis are successful in forecasting changes in the cryptocurrency markets and that NN outperforms other models.

In [14], both linear and non-linear components of the time series of the stock exchange data set were used for forecasting using a hybrid model. In nonlinear time series forecasting, CNN and Seq2Seq LSTM have been successfully combined for the dynamic modelling of short-term and long-term dependent patterns. The study [15] analyzed the social factors that are increasingly used for online transactions worldwide using a multi-linear regression model and analyzed two cryptocurrencies with a large capital market, BTC and LTC. The authors [15] found that R2 values were 44% for LTC and 59% for BTC. In [16], two different LSTs were used for LTC and BTC. In [16], two different LSTM models were used (the standard LSTM model and the LSTM model with AR (2)). This study presents a forecasting system using an LSTM model to forecast daily bitcoin prices. Research [16] showed that the AR (2) model is better than LSTM.

To predict future cryptocurrency prices, Akyildirim et al. [17] predicted the 12 most liquid cryptocurrencies using machine learning classification algorithms such as support vector machines, logistic regression, artificial neural networks, and random forests. Plakandaras and others. [18] applied different methodologies – such as ordinary least squares (OLS), support vector regression (SVR), and least absolute compression and selection operator (LASSO) methods – in the field of machine learning to predict cryptocurrency prices.

Kaminskyi et al. [19] assessed the investment risk of 327 cryptocurrencies with a capitalization of over $1 million, using the criteria of capitalization and historical returns. Using five approaches, including variability indicators and the Hurst exponent, the study aimed to categorize cryptocurrencies into risk clusters using the Kohonen self-organizing map technique. The analysis resulted in the identification of three distinct risk clusters, adding valuable insights to the cryptocurrency investment analysis literature.

The study [20] proposes an integrated risk assessment for alternative investments based on eight approaches, including volatility, losses, asymmetry, sensitivity, interdependence, risk/return coupling, long-term memory and liquidity risk. Using cluster analysis with risk attitude, five distinct risk-assessed clusters were identified. Naive diversified portfolios were then constructed for each cluster, highlighting the benefits for investment portfolio management.

This adds important insights to the literature on integrated risk assessment in alternative investments.

The paper [21] explores short-term cryptocurrency forecasting using machine learning (ML) and analyses the methodological principles and advantages of ML algorithms. It estimates the 90-day dynamics of Bitcoin, Ethereum, and Ripple using the Binary Autoregressive Tree model (BART), Neural Networks (Multilayer Perceptron, MLP), and an ensemble of Classification and Regression Trees models (Random Forest, RF). Computational experiments confirm the viability of these ML models for forecasting short-term financial time series.

3. Methods

The mathematical formulation of the estimation of the price increase (decrease) of cryptocurrencies involves the study of the variation of a defined value under the influence of the change in the values of a defined value under the influence of certain factors, which is a classical econometric problem. Machine learning methods, which have been extensively developed and substantiated, are often used in economic process research.

To assess the rate of increase (decrease) of cryptocurrency, it is necessary to establish a relationship between a certain list of factors and the fact of growth or decline of cryptocurrency. The growth or decline of a cryptocurrency can be indicated by only two values of a binary variable, usually 1 and 0. Therefore, we need to build a model to predict the value of the binary variable.

Traditional multiple regression may not yield the desired results because the values of the dependent variable may not fall within the [0, 1] range, making interpretation difficult. However, the task of constructing a regression dependence for such an assessment can be presented not as a prediction of the values of a binary variable, but as a modelling of some continuous variable that acquires a value from the interval [0, 1]. Such problems can be described using linear probability models, or logit and probit models. Due to the way these models are constructed, the predictive values that the researched variable acquires can not only correspond to the value of 1 and 0 but also be interpreted as an increase or decrease in the price of cryptocurrency.

This paper delves into the modeling of cryptocurrency price movements to predict future states, addressing the evaluation of a qualitative variable through various quantitative factors. Discriminant analysis tools can be used to select the most informative financial risk variables. Logit regression makes it possible to determine the growth group of cryptocurrencies, as well as to estimate the probability of assigning the price of cryptocurrency to one or another growth group.

The logit model allows the estimation of the probability that the analyzed (dependent) variable will acquire the value of one at the given specific factor values, which serves as an estimate of the share of "units" at the those factor value.

The logit model looks like this:

$$p(x) = P(Y = 1 | X = x) = (1 + \exp(x^T w))^{-1},$$

where w are the unknown parameters to be estimated.

The logistic function has such a property that its values range from zero to one for any values of the argument.

It is also necessary to highlight the following positive points: logit analysis takes into account models of non-linear dependence, and logit analysis can unambiguously interpret the resulting indicator of growth or decline in the price of cryptocurrency. Acquiring values limited to the interval from 0 to 1 determines the nominal value of the realization of growth or decline in the price of cryptocurrency.

The following data were used to build a model for estimating the rate of growth or decline of the cryptocurrency price.

Table 1
A fragment of the matrix of factor values for evaluating the growth or decline of the cryptocurrency price

y	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}
1	4100	72	102	7810	2860	21600	1160	54.3	69.6	2.97	0
1	5300	93	118	8110	2935	22410	1150	51	72.1	2.98	0
1	8100	143	332	7835	2880	21620	1120	46.5	59.8	2.72	1
0	13000	210	1850	7950	2920	21150	1115	48	66.2	2.79	1
0	11900	214	475	7965	2925	20480	1095	45.9	58.3	2.67	0
0	9800	195	313	8140	2985	21950	1235	50.8	64.5	2.68	0
1	7500	134.5	290	8155	3010	22790	1320	52.2	64.55	2.73	0
0	8300	147	640	8120	2800	18860	1480	51.8	33.75	2.57	-2
1	5400	129	745	8570	2830	19600	1225	60.2	28	2.43	0
1	9900	179	650	10450	3215	22610	1322	50.6	43.15	2.95	2
0	11400	213	960	11850	3550	23410	1255	52.1	44.4	3.07	0
1	12800	235	1450	11580	3445	23600	1325	53.5	42.45	3.2	1
1	18700	335	965	12390	3655	26800	1415	50	50	3.53	1
1	23550	433	1640	12800	3695	26620	1525	47	51.35	3.56	0
1	28500	522	2700	12890	3735	27500	1530	49	51.45	3.55	1
0	40400	738	3050	13130	3815	28250	1605	51.5	56.1	3.63	0
1	39500	732	3350	13 800	3900	29100	1465	58.7	59.8	3.92	3
0	50 300	935	4100	13350	4160	28050	1520	55.25	67.8	4.68	-3
1	31 800	562	2620	14550	4345	27850	1490	55.8	70.5	4.44	2
0	51 500	971	2280	15350	4530	29950	1655	53.55	71.4	4.41	-1
1	42 000	782	2550	14500	4340	29620	1580	53.6	78	4.23	0
1	50 500	987	2630	14380	4330	28000	1575	53.8	84	4.42	2
0	65 200	1240	2470	15860	4685	29750	1650	50.3	81.5	4.4	0
0	55 800	1070	5000	15300	4570	27800	1610	49.5	70.2	4.28	-1
0	46 300	876	2400	15400	4750	29200	1650	54.3	80	4.43	-3
0	42 500	790	3780	14400	4520	27500	1580	53	87	4.58	-3
0	46 300	885	1940	14450	4550	27750	1780	48.3	108.7	4.79	0
0	42 000	785	2470	13500	4430	27300	1815	47.9	106.5	4.72	-1

Prepared data, a fragment of which is presented in the table. 1, we will use to estimate the unknown parameters of the machine learning model:

$$P(y_i = 1|x_i) = F(w_0 + w_1x_{i1} + w_2x_{i2} + \cdots + w_{11}x_{i11}) + \varepsilon_i, \qquad i = 1,2, \dots n$$

where $P(y_i = 1 \mid x_i)$ – the probability that the i-th value of the binary variable is 1 given $\boldsymbol{x}_i$; $F(z) = \dfrac{1}{1+e^{-z}}$ – logistic function; ε_i – random component;

x_1 is the price of cryptocurrency
x_2 is cryptocurrency capitalization
x_3 is trading volume (per day)
x_4 is the NASDAQ Composite stock index
x_5 is the S&P 500 stock index
x_6 is the Nikkei 225 stock index
x_7 is the value of the SSE PLC share (SSE. United Kingdom. GBR)
x_8 is the value of Intel Corporation (INTC, USA) shares
x_9 is the cost of Brent oil futures (LCOH8, USD/barrel)
x_{10} is the cost of copper futures (MCUcl, USD/ton)
x_{11} is positive and negative news for investing related to the features of cryptocurrency as an asset (expert assessment on a 5-point scale)

There are some ways to find logistic regression coefficients. In practice, the maximum likelihood approach is quite often used in economic research. It is used in statistics to obtain estimates of the main parameters of the general population based on the sample data. The basis of the approach is the likelihood function, which shows the probability density (probability) of the co-occurrence of the results of the sample $Y_1, Y_2,\ldots,Y_n$:

$$L(Y_1,Y_2,\ldots,Y_k;\Theta)=p(Y_1;\Theta)\cdot\ldots\cdot p(Y_n;\Theta)$$

According to the maximum likelihood method, the value $\Theta=\Theta(Y_1,\ldots,Y_n)$ that maximizes the function L is taken as the estimate of the unidentified parameter.

Finding the estimate is simplified if we maximize not the function L itself, but the natural logarithm ln(L), since the maximum of both functions is reached at the same value of Θ:

$$L^*(Y;\Theta)=\ln(L(Y;\Theta))\rightarrow\max$$

In the case of a binary independent variable, which we have in logistic regression, the exposition can be continued in the following way. We denote by P_i the probability of the occurrence of a unit: $P_i = \mathrm{Prob}(Y_i=1)$. This probability will depend on $\boldsymbol{X_i}$, where $\boldsymbol{X_i}$ is a row of the matrix of regressors, and $\boldsymbol{W}$ is a vector of regression coefficients:

$$P_i = F(X_i), \qquad F(z) = \frac{1}{1+e^{-z}}$$

The log-likelihood function is equal to:

$$L(\boldsymbol{Y},\mathbf{W}) = \prod_{y_i=1}^{n} F(\boldsymbol{X_i W})^{Y_i}\,[\mathbf{1}-F(\boldsymbol{X_i W})]^{1-Y_i}$$

Usually, instead of the function L, its logarithm is used, which does not change the essence of the problem, but allows you to get rid of the product:

$$L^* = \ln L = \sum_{\boldsymbol{i=1}}^{n} \boldsymbol{Y_i}\ln F(\boldsymbol{X_i W}) + (\mathbf{1}-\boldsymbol{Y_i})\ln(\mathbf{1}-F(\boldsymbol{X_i W}))$$

Here, for brevity, the following notations are adopted:

$\boldsymbol{W} = (W_0, W_1,\ldots,W_m)^T$,

$\boldsymbol{X}_i = (1, X_{i1},\ldots,X_{im})$,

$\boldsymbol{X}_i\boldsymbol{W} = W_0 + W_1X_{i1} + W_2X_{i1} +\ldots+ W_mX_{im}$

To maximize the function L, the Newton-Raphson method can be applied. It consists in performing the following iterations, starting from some initial value of the $\boldsymbol{W}$ parameters:

$$\boldsymbol{W}_{t+1} = \boldsymbol{W}_t - \frac{\partial \ln L(\boldsymbol{W}_t)}{\partial \boldsymbol{W}}\left[\frac{\partial^2 \ln L(\boldsymbol{W}_t)}{\partial \boldsymbol{W}\,\partial \mathbf{W}'}\right]^{-1}$$

where

$$\frac{\partial \ln L(\boldsymbol{W})}{\partial \boldsymbol{W}} = \big(f_0(\boldsymbol{W}),\ f_1(\boldsymbol{W}),\ldots,\ f_m(\boldsymbol{W})\big)$$

$$f_0(\boldsymbol{W}) = \sum_{i=1}^{n} F(\boldsymbol{X_i W}) - \sum_{\{i:Y_i=1\}}^{n} 1$$

$$f_j(\boldsymbol{W}) = \sum_{\boldsymbol{i=1}}^{n} F(\boldsymbol{X_i W})X_{ij} - \sum_{\{i:Y_i=1\}}^{n} X_{ij}\,,\ j = 1,2,\ldots,m$$

$$\frac{\partial^2 \ln L(\boldsymbol{W}_t)}{\partial \boldsymbol{W}\,\partial \mathbf{W}'} = \begin{pmatrix} \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} F(\boldsymbol{X_i W})(\mathbf{1}-F(\boldsymbol{X_i W}))\,, & \ldots & \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} F(\boldsymbol{X_i W})\big(\mathbf{1}-F(\boldsymbol{X_i W})\big)X_{im}, \\ \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} F(\boldsymbol{X_i W})\big(\mathbf{1}-F(\boldsymbol{X_i W})\big)X_{i1}, & \ldots & \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} F(\boldsymbol{X_i W})\big(\mathbf{1}-F(\boldsymbol{X_i W})\big)X_{im}X_{i1}, \\ \ldots\ldots\ldots & & \ldots \\ \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} F(\boldsymbol{X_i W})\big(\mathbf{1}-F(\boldsymbol{X_i W})\big)X_{im}, & \ldots & \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} F(\boldsymbol{X_i W})\big(\mathbf{1}-F(\boldsymbol{X_i W})\big)X_{im}X_{im} \end{pmatrix}$$

Initial values can be defined as a vector of linear regression parameters:

$$\boldsymbol{W}^{(\text{поч})} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{Y}$$

Any gradient methods can be used to calculate logistic regression coefficients: conjugate gradient method, variable metric methods, etc.

Data analysis showed the presence of multicollinearity. Therefore, it is advisable to apply the approach of principal components. The main idea of the approach is to replace highly correlated variables with a set of new variables between which there is no correlation. At the same time, the new variables are linear combinations of the original variables

$$\boldsymbol{z} = \boldsymbol{x} \cdot \boldsymbol{a}$$

First, you need to normalize all explanatory variables:

$$x_{ij}^* = \frac{x_{ij} - \overline{x_j}}{\sigma_{x_j}}, \quad i = 1,2, \dots, n; \;\; j = 1,2, \dots, m$$

Next, calculate the correlation matrix

$$r = \frac{1}{n}(X^{*T}X^*)$$

It is necessary to find the characteristic numbers of the matrix r from the equation

$$|r - \lambda E| = 0$$

at what

$$\sum_{k=1}^{m} \lambda_k = m$$

where E is a unit matrix of size $m \times m$.

The eigenvalues λ_k $(k = 1,2, \dots, m)$ are ordered by the absolute level of the contribution of each main component to the total variance.

It is necessary to calculate the a_k eigenvectors when deciding to use the system of equations

$$(r - \lambda E)a = 0$$

Finding the main components-vectors takes place according to the formula:

$$z_k = X^* \cdot a_k$$

As a result of calculations, characteristic numbers and eigenvectors were found:

$$\lambda = \begin{Vmatrix} 4E\text{-}04 & 0.002 & 0.025 & 0.068 & 0.132 & 0.15 & 0.311 & 0.53 & 0.951 & 1.198 & 7.632 \end{Vmatrix}$$

$$a = \begin{Vmatrix}
0.721 & 0.083 & -0.121 & -0.011 & 0.536 & 0.099 & -0.177 & -0.080 & 0.014 & 0.002 & -0.352 \\
-0.678 & -0.105 & -0.070 & 0.043 & 0.593 & 0.123 & -0.168 & -0.068 & 0.008 & 0.004 & -0.351 \\
-0.019 & -0.001 & 0.160 & -0.066 & -0.458 & 0.218 & -0.681 & -0.350 & -0.081 & 0.161 & -0.307 \\
0.106 & -0.690 & 0.243 & 0.352 & -0.091 & -0.291 & 0.249 & -0.160 & 0.096 & 0.147 & -0.345 \\
-0.068 & 0.680 & 0.344 & 0.447 & -0.064 & -0.247 & 0.148 & 0.025 & 0.024 & 0.003 & -0.356 \\
-0.051 & 0.096 & 0.034 & -0.768 & -0.052 & -0.458 & 0.152 & -0.043 & 0.180 & 0.113 & -0.338 \\
0.000 & 0.032 & 0.063 & -0.164 & -0.167 & 0.710 & 0.561 & -0.101 & -0.046 & -0.059 & -0.326 \\
0.011 & -0.008 & 0.055 & -0.027 & 0.036 & 0.086 & 0.007 & 0.499 & -0.348 & 0.785 & -0.043 \\
0.031 & -0.171 & 0.284 & -0.076 & -0.093 & 0.088 & -0.228 & 0.734 & 0.121 & -0.453 & -0.238 \\
-0.039 & 0.023 & -0.830 & 0.199 & -0.313 & -0.104 & 0.016 & 0.186 & 0.022 & -0.036 & -0.352 \\
-0.005 & 0.057 & -0.019 & 0.085 & -0.007 & 0.207 & -0.056 & 0.080 & 0.901 & 0.338 & 0.102
\end{Vmatrix}$$

Having determined all the main components and discarding those that correspond to small values of the characteristic roots, we find the relationship of the dependent variable Y with the main components $z_1 - z_5$. To do this, we can construct a logistic regression model. Thus, the main parameters of the obtained logit model are as follows:

As can be seen from the parameters, the five-factor logit model provides high reliability, which is confirmed by the calculated chi-square value (18.43) and the almost zero probability of not rejecting the null hypothesis.

The analytical expression of the built model will look like this:

$$P(y_i = 1|z_i) = (1 + e^{1.96 - 5.88z_1 - 3.4z_2 - 11.88z_3 - 5.66z_4 - 1.68z_5})^{-1}$$

Table 2
Parameters of the logit model

y - Parameter estimates (Spreadsheet1) Distribution : BINOMIAL, Link function: LOGIT The modelled probability that y = 0						
	Level of	Column	Estimate	Standard	Wald	p
Intercept		1	1.9630	1.190108	1.602472	0.172828
z1		2	-5.8844	3.879610	2.300523	0.129330
z2		3	-3.3996	2.339770	2.111124	0.146232
z3		4	-11.8788	9.161676	1.681095	0.194779
z4		5	-5.6562	3.865900	2.140670	0.143439
z5		6	-1.6775	1.153813	2.113859	0.145971
Scale			1.0000	0.000000		

The adequacy of the built model can be determined by the McFadden likelihood ratio index using the formula

$$LRI = 1 - \frac{\ln(L(\boldsymbol{b}))}{\ln(L(b_0))} = 0.84 ,$$

where

$$L(\boldsymbol{y},\boldsymbol{b}) = \prod_{i=1} F(\boldsymbol{x}_i\boldsymbol{b})^{y_i}(1-F(\boldsymbol{x}_i\boldsymbol{b}))^{1-y_i} ,$$

$\ln L(\boldsymbol{b})$ is the maximum value of the log-likelihood function, which is reached at the point whose coordinates are equal to the estimates of the model parameters $\boldsymbol{b} = (b_0, b_1, \ldots, b_m)$, a $\ln L(b_0)$ is the value of the log-likelihood function calculated under the assumption that $b_l = b_2 = \ldots = b_m = 0$. The calculated value of the likelihood ratio index McFadden testifies to the adequacy of the constructed model.

The obtained expression can be used to estimate the rate of growth or decline of the cryptocurrency price at different values of the factors.

Let's estimate with the help of the built model the rate of growth or decline of the cryptocurrency price, information about which is given in Table 3.

Table 3
The value of factors for evaluating the indicator of growth or decline in the price of cryptocurrency

№ п/п	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	y
1	46 000	800	4000	12000	4000	20000	1500	55	100	3	2	0.53
2	60 000	900	4000	10000	5000	29000	1800	60	110	3	1	0.85
3	67 000	1000	2000	18000	6000	31000	2000	60	110	4	-1	0.95

We will introduce the procedure of calculation of the growth or decline rate of the cryptocurrency price for the first case.

The value of the main components is as follows

z_1	z_2	z_3	z_4	z_5
-1.93	1.14	0.71	0.49	-0.91

$$P(y_1 = 1|z_i) = \left(1 + e^{1.96 - 5.88\cdot(-1.93) - 3.4\cdot 1.14 - 11.88\cdot 0.71 - 5.66\cdot 0.49 - 1.68\cdot(-0.91)}\right)^{-1} = 0.53$$

The calculations show the possibility of price growth only in the second and third cases.

The rate of growth or decline of the cryptocurrency price at different values of the factors was evaluated.

Figure 1 shows the effect on the cryptocurrency price of the factor x_4 (NASDAQ Composite stock index) for certain values of the factor x_5 (S&P 500 stock index) and fixed values of other factors. That is, the growth of the NASDAQ Composite and S&P 500 stock indexes at fixed values of other indicators leads to an increase in the price of the considered cryptocurrency.

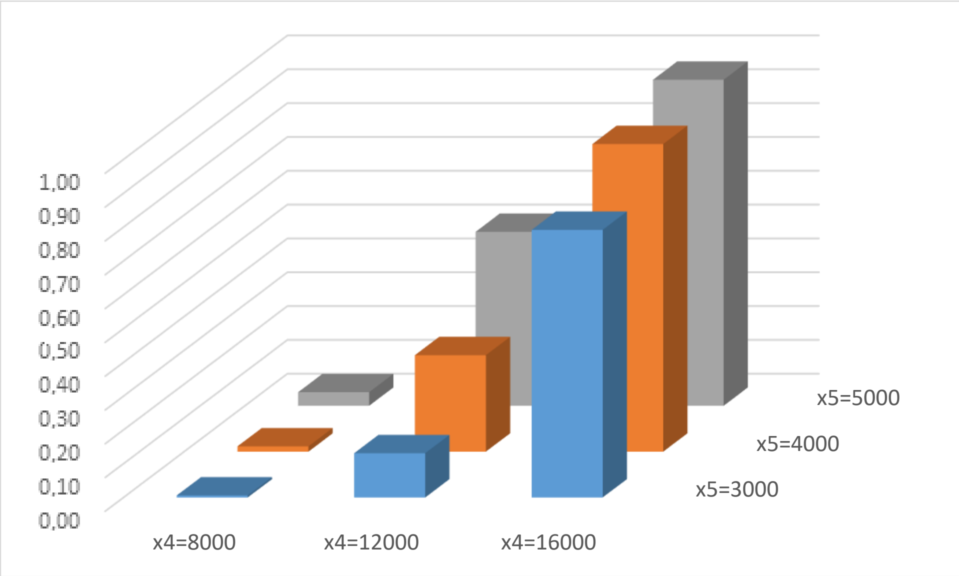

Figure 1. Dependence of the cryptocurrency price on the NASDAQ Composite and S&P 500 stock indices

Figure 2 shows the effect on the cryptocurrency price of the factor x_4 (NASDAQ Composite stock index) for certain values of the factor x_6 (Nikkei 225 stock index) and fixed values of other factors. That is, the growth of the NASDAQ Composite and Nikkei 225 stock indices at fixed values of other indicators leads to an increase in the price of cryptocurrency. The economic factors constantly fluctuate, impacting both traditional and crypto markets.

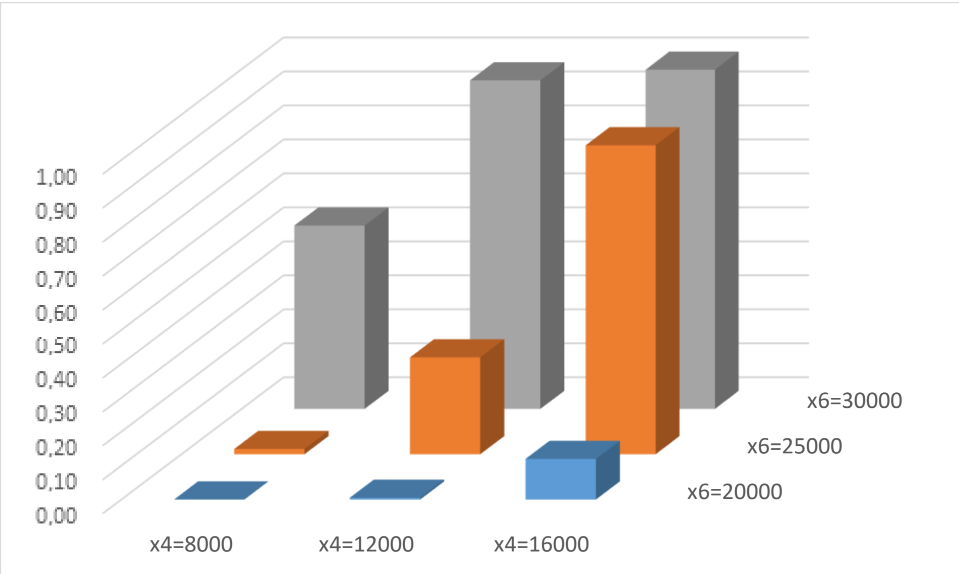

Figure 2. Dependence of the cryptocurrency price on the NASDAQ Composite and Nikkei 225 stock indices

Knowing the connections between the cryptocurrency market and the stock market will be very helpful in managing investors' portfolios and how much of their investment money will be allocated to cryptocurrency for their safe and profitable investment plan [31,32].

Figure 3 shows the effect on the cryptocurrency price of the factor x_7 (the value of the SSE PLC company's stock) for certain values of the factor x_8 (the value of the company's Intel Corporation's stock) and fixed values of the other factors. That is, the increase in the value of the shares of SSE PLC and Intel Corporation, with fixed values of other indicators, leads to an increase in the price of cryptocurrency.

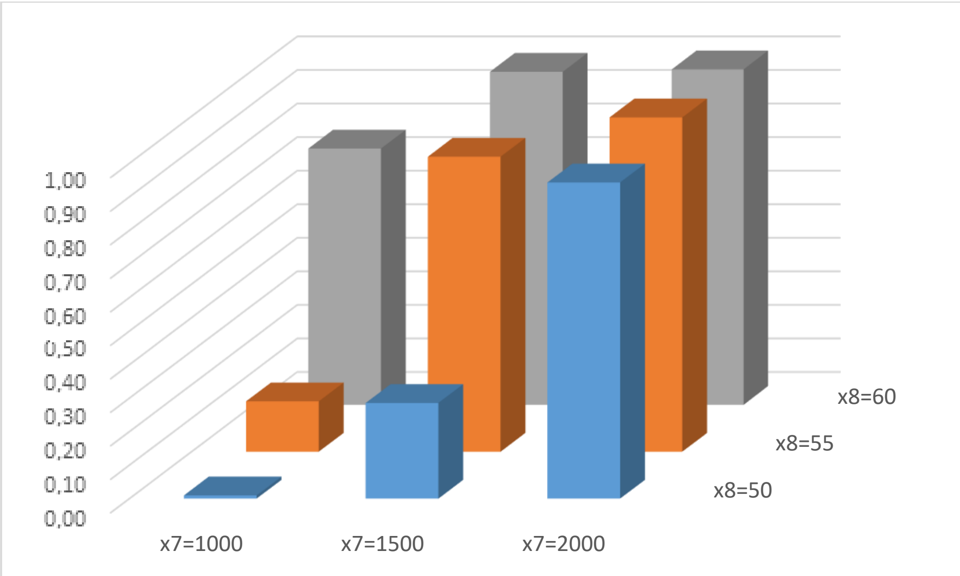

Figure 3. Dependence of the price of cryptocurrency on the value of the shares of SSE PLC and Intel Corporation

The study analyzed data from a specific period and found a positive and significant correlation between the cryptocurrency price and some traditional indicators, including shares (SSE PLC and Intel Corporation).

Positive performance in stock markets tends to boost investor confidence and overall market sentiment. Investors, when optimistic about the economy and financial markets, may be more willing to take on risk, including investing in higher-risk assets like cryptocurrencies.

As stock markets rise, individuals and institutional investors may experience a "wealth effect," feeling wealthier due to the increased value of their stock portfolios [33]. This increased wealth contributes to a higher propensity to invest in various asset classes, including cryptocurrencies. Positive movements in stock indices may lead institutional investors to allocate funds to cryptocurrencies as part of a diversified strategy.

Increased institutional interest and adoption in cryptocurrency markets can result in investment movements. It's important to note that these relationships are complex and can change over time.

4. Experiment, results and discussion

The logistic regression model confirms the possibility of combining quantitative and qualitative factors. An analytical expression of the constructed model is presented, and its adequacy is checked. Since there was a close correlation between the independent factors in the studied model, the method of principal components was used to remove the phenomenon of multicollinearity. The resulting expression was used to estimate the rate of growth or decline of the cryptocurrency price at different values of the factors. Our research confirms the influence of the considered eleven factors on the change in the price of cryptocurrency. The obtained data indicate that the selected factors can significantly affect the price of cryptocurrency.

It's crucial to conduct thorough research and consider factors when analyzing the relationship between stock indices and cryptocurrency prices. Additionally, market conditions can change rapidly, so staying informed about the latest developments is essential.

As cryptocurrencies gain more recognition and acceptance, they become part of the broader financial landscape. Positive sentiments in traditional markets can contribute to increased acknowledgment and acceptance of cryptocurrencies, driving demand. Additionally, market conditions can change, so investors should conduct thorough research and analysis based on the current environment.

5. Conclusion

The use of machine learning methods facilitates the exploration of relationships between various economic indicators and cryptocurrency price changes, providing solutions for analyzing and forecasting the growth or decline index of cryptocurrency price. A key aspect of modern cryptocurrency analysis is predicting the future course of cryptocurrencies, which, in addition to identifying trends in the development of the cryptocurrency market, allows effective strategic decisions to be made at all levels of management and increases the profitability indicators of an individual trader's work.

The research examines the impact of the NASDAQ Composite stock index, S&P 500 stock index, Nikkei 225 stock index, SSE PLC company's stock, Intel Corporation's stock on cryptocurrency prices. In particular, positive sentiment in traditional stock markets often influences the cryptocurrency market by increasing demand for higher-risk assets such as cryptocurrencies. If investors have confidence in the overall health of the global economy, they may be more inclined to invest in both traditional assets and cryptocurrencies.

The study confirms that the proposed methodology which incorporates a complex combination of factors and expert judgement, reliably identifies cryptocurrency market trends. This approach provides a basis for making effective financial and organisational decisions with a high probability of success.

References

[1] A. Chen, M. Leung, H. Daouk, Application of Neural Networks to an Emerging Financial Market: Forecasting and Trading the Taiwan Stock Index, Computers & Operations Research, 30, 2014, pp. 901–923, available at: https://doi.org/10.1016/S0305-0548(02)00037-0.

[2] V. Derbentsev, A. Matviychuk, V.N. Soloviev, Forecasting of Cryptocurrency Prices Using Machine Learning, In Advanced Studies of Financial Technologies and Cryptocurrency Markets, Springer: Berlin/Heidelberg, Germany, 2020, pp. 211–231.

[3] R. Chowdhury, M. Rahman, M. Rahman, M. Mahdy, Predicting and Forecasting the Price of Constituents and Index of Cryptocurrency Using Machine Learning, ArXiv, 2019, available at: https://arxiv.org/ftp/arxiv/papers/1905/1905.08444.pdf.

[4] J. Mba, S. Mwambi, A Markov-switching COGARCH approach to cryptocurrency portfolio selection and optimization, Financial Markets and Portfolio Management, 34(2), 2020, pp. 199–214, available at: https://doi.org/10.1007/s11408-020-00346-4.

[5] S. Aras, Stacking hybrid GARCH models for forecasting Bitcoin volatility. Expert Systems with Applications, 2021, 174(114747), available at: https://doi.org/10.1016/j.eswa.2021.114747.

[6] L. Garcia-Jorcano, S. Benito, Studying the properties of the Bitcoin as a diversifying and hedging asset through a copula analysis: Constant and time-varying, Research in International Business and Finance, 54(101300), 2020, available at: https://doi.org/10.1016/j.ribaf. 2020.101300.

[7] M. Umar, S. K. A. Rizvi, B. Naqvi, Dance with the devil? The nexus of fourth industrial revolution, technological financial products and volatility spillovers in global financial system, Technological Forecasting and Social Change, 163(120450), 2021, available at: https://doi.org/10.1016/j.techfore.2020.120450.

[8] G. Uzonwanne, Volatility and return spillovers between stock markets and cryptocurrencies, Quarterly Review of Economics and Finance, 82, 2021, pp. 30–36, available at: https://doi.org/10.1016/j.qref.2021.06.018.
[9] M.M. Patel, S. Tanwar, R. Gupta, N. Kumar, A Deep Learning-based Cryptocurrency Price Prediction Scheme for Financial Institutions, Journal of Information Security and Applications, 55(102583), 2020, available at: https://doi.org/10.1016/j.jisa.2020.102583.
[10] M. Saad, A. Mohaisen, Towards characterizing blockchain-based cryptocurrencies for highly-accurate predictions, In Proceedings of the IEEE INFOCOM—IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), 2018, available at: https://doi.org/10.1109/infcomw.2018.8406859.
[11] W. Yiying, Z. Yeze, Cryptocurrency Price Analysis with Artificial Intelligence, In Proceedings of the 5th International Conference on Information Management (ICIM), 2019, pp. 97–101, available at: https://doi.org/10.1109/INFOMAN.2019.8714700.
[12] Z. Chen, C. Li, W. Sun, Bitcoin price prediction using machine learning: An approach to sample dimension engineering, Journal of Computational and Applied Mathematics, 365(112395), 2019, available at: https://doi.org/10.1016/j.cam.2019.112395.
[13] F. Valencia, A. Gómez-Espinosa, B. Valdés-Aguirre, Price Movement Prediction of Cryptocurrencies Using Sentiment Analysis and Machine Learning, Entropy, 21(589), 2019, available at: https://doi.org/10.3390/e21060589.
[14] Y. Zhao, Z. Chen, Forecasting stock price movement: New evidence from a novel hybrid deep learning model, Journal of Asian Business and Economic Studies, 2021, available at: https://doi.org/10.1108/JABES-05-2021-0061.
[15] A. Jain, S. Tripathi, H.D. Dwivedi, P. Saxena, Forecasting Price of Cryptocurrencies Using Tweets Sentiment Analysis, In Proceedings of the 11th International Conference on Contemporary Computing (IC3), 2018, pp. 1–7, available at: https://doi.org/10.1109/IC3.2018.8530659.
[16] C.-H. Wu, C.-C. Lu, Y.-F. Ma, R.-S. Lu, A New Forecasting Framework for Bitcoin Price with LSTM, In Proceedings of the IEEE International Conference on Data Mining Workshops (ICDMW), 2018, pp. 168–175, available at: https://doi.org/10.1109/ICDMW. 2018.00032.
[17] E. Akyildirim, A. Goncu, A. Sensoy. Prediction of cryptocurrency returns using machine learning, Annals of Operations Research, 297, 2021, pp. 3–36, available at: https://doi.org/10.1007/s10479-020-03575-y.
[18] V. Plakandaras, E. Bouri, R. Gupta, Forecasting bitcoin returns: Is there a role for the US–China trade war, Journal of Risk, 23(3), 2021, pp. 1-17 .
[19] A. Kaminskyi, I. Miroshnychenko, K. Pysanets. Risk and return for cryptocurrencies as alternative investment: Kohonen maps clustering, Neuro-Fuzzy Modeling Techniques in Economics, 2019, 8, pp. 175-193, available at: https://doi.org/10.33111/nfmte.2019.175.
[20] A. Kaminskyi, D. Butylo, M. Nehrey. Integrated approach for risk assessment of alternative investments. International Journal of Risk Assessment and Management, 2022, 24(2-4), pp. 156-177, available at: https://doi.org/10.1504/IJRAM.2021.126413.
[21] V. Derbentsev, H. Velykoivanenko, N. Datsenko. Machine learning approach for forecasting cryptocurrencies time series. Neuro-Fuzzy Modeling Techniques in Economics, 2019, 8, 65-93, available at: http://doi.org/10.33111/nfmte.2019.065.
[22] M. Bublyk, V. Vysotska, K. Varenina, V. Schuchmann, M. Shevchenko, Big Data Analysis of Cryptocurrency Exchange Rate Changes Dynamics, 18th IEEE International Scientific and Technical Conference on Computer Sciences and Information Technologies, 195025, 2023, available at: https://doi.org/10.1109/CSIT61576.2023.10324204.
[23] O. Prokipchuk, L. Chyrun, M. Bublyk, V. Yakimtsov, R. Kovalchuk, Intelligent system for checking the authenticity of goods based on blockchain technology, CEUR Workshop Proceedings, 2917, 2021, pp. 618–665 .
[24] R. Yurynets, Z. Yurynets, D. Dosyn, Y. Kis, Risk Assessment Technology of Crediting with the Use of Logistic Regression Model, CEUR Workshop Proceedings, 2362, 2019, pp 153-162, URL: http://ceur-ws.org/Vol-2362/paper14.pdf.