Article

Expected Income of New Currency in Blockchain Based on Data-Mining Technology

Xingzhou Li [1,*] and Xin Zeng [2]

[1] School of Mathematics, University of Manchester, Manchester M13 9PL, UK
[2] School of Finance, NanKai University, Tianjin 300071, China; 1513399@mail.nankai.edu.cn
* Correspondence: xingzhou.li@outlook.com or xingzhou.li@student.manchester.ac.uk;
   Tel.: +44-1366-100-4820

check for updates

Abstract: In order to realize the understanding of expected returns after issuance of blockchain new currency initial coin offerings (ICO) and maximize investment returns, in this study, the Semantic Orientation Pointwise Mutual Information (SO-PMI) algorithm is used to create a customer emotional dictionary of blockchain new currency, and collect users' online comments based on blockchain currency before ICO. The Support Vector Machine (SVM) algorithm is used to construct an evaluation model, analyze and judge users' comments, make accurate prediction of the expected return of ICO issuing new currency, improve investment operations, and maximize the return of investment. The results show that the combination of the SO-PMI and SVM algorithms can accurately evaluate the price after the issuance of new currency, and then realize the judgment of expected return and obtain the expected return of investment. It can be seen that the combination of algorithms based on data-mining technology is applied to the study of the expected return of new currency issuance in blockchain, which achieves the goal of revenue anticipation and greatly reduces the investment risk of new currency issuance in blockchain.

Keywords: data mining; algorithms; currency; blockchain

1. Introduction

With the rapid development of science and technology, there are now many ordinary people involved in investment. In particular, some relatively small individual investors release their opinions and attitudes through the network platform, and actively communicate with other people [1]. These individual ordinary investors express their views and attitudes through various social platforms on the internet, such as WeChat, micro-blog, Facebook, Poster Bar, and Forum. Some investors even release their emotions directly on the internet and expose their psychological state. Investors of blockchain-encrypted currency often use the exchange on the network to infer the possible trend of new currency in blockchain, and then change their investment operation and make corresponding actions [2]. Speech at the network level will have a significant impact on many people. When the number of people who discuss and communicate about the intensified currency reaches a relatively large scale, there will be the so-called "herd effect". Many people will lose their ability to think independently and begin to follow the crowd, which will affect the mood and attitude of other investors, increase the turmoil of encrypted currency in blockchains, and lead to some traces [3]. Unlike traditional financial investment methods, behavioral finance shows that most investors' personal emotions, attitudes, and states have an obvious impact on their own investment operations, which is a core factor [4]. As a very influential platform that most people will be involved with and participate in, the network brings together a lot of investors. Most investors will express their opinions more or less on the internet, and will analyze and study the opinions of others, so that they can obtain valuable information. The exchange and

discussion of this information have a great impact on everyone's investment strategy at a certain level, and there will be indirect convergence of investment behavior. Sharing mechanisms of internet platforms will facilitate the spread of investors' decisions at an extremely efficient and rapid rate. The dissemination of information will have a joint impact on the behavior of other investors, and then on the price of the blockchain currency [5].

At present, the blockchain currency is still a newly emerging area, and there are many levels of non-standard immature. Research shows that there are many perceptual investors, and personal emotions have a greater impact on their investment behavior. At present, individual investors actively participate in most of the investments in blockchain currencies in China [6]. Compared with investment institutions, small and medium-sized individual investors have a great disadvantage in experience and rationality and their investment operations are often more casual. They are slow to respond to market information, easy to miss opportunities, and lack independent analysis and judgment ability [7].

As we all know, mankind has a strong ability to learn and use knowledge [8]. Therefore, how to combine data-mining technology with the expected return of new currency issuance in blockchain, improve investment operation and realize the expected return has become the main research direction of new currency issuance in blockchain in the future.

2. Literature Review

For the existence of blockchain currency, active researchers believe that it will essentially change people's payment behavior, and will have an impact on the world's economic situation. However, negative researchers believe that the emergence of encrypted currency will bring a lot of uncertainty, and will have a negative impact on the region and the global economy. The result must be a complete failure [9]. The fundamental reason is that most people do not know how digital money exists and how it survives and operates [10]. Fortunately, there are many studies on Bitcoin. Bitcoin can be used as a reference for analysis and research to discover the nature and operation of encrypted currency [11]. Zhong (2017) believed that Bitcoin was a point-to-point electronic cash trading system. Therefore, the purpose of the emergence of Bitcoin is to deal with the problems in the previous electronic cash-trading system. The strategy used is the point-to-point network, that is, the method mode of P2P [12]. The essence of the Bitcoin trading process is not the process that most people think a Bitcoin enters another person's account by transferring money. It is essentially a cycle of continuous building blocks [13].

Vitalik Buterin (2018) published a white paper on the future development of Yitaifang. At the same time, the Yitaifang project is also running at the same time. The purpose of this project is to use initial coin offerings (ICO) to obtain development funds. Vitalik believed that the value and significance of Bitcoin are to verify the blockchain model of workload proof (POS), which can help most people form a consensus on the issue of a digital currency trading sequence [14]. With the help of Bitcoin's own advantages, Yitaifang itself has made greater improvements. At the same time, Yitaifang will be designed as a blockchain platform with development function, which achieves the goal that users can freely program and develop their own projects in the Yitaifang system. Yitaifang mainly borrows the existence of an intelligent contract in blockchain to meet the development of all functions in Yitaifang network [15]. Because of the present situation of Yitaifang, it has made a great contribution to the development and actual landing of encrypted currency and blockchain, and at the same time led to the rapid development of ICO. For workload proof, this model was first proposed by Adam Back (2017) in his study on Hashcash. Initially, this mechanism was mainly to achieve the function of filtering in the mail. The main purpose of arrival in Bitcoin is to deal with the trouble of encrypted currency "double flowers", that is, double payment. It uses the hash's cryptographic method to obtain the result through calculation and then gets the right or ownership. It deals with the biggest problem in the process of using electronic cash, that is, the single currency is used twice [16]. The system set by Bitcoin's own network system development module determines that the time difference between new blocks will be maintained at 15 min. Out of the initial 210,000 blocks, each block will receive a return of 50 BTC upon

determination. The subsequent 210,000 blocks mean that each block will receive a return of 25 BTC after it is determined. By analogy, each time the return is 50% of the previous one. Yu Jiang (2017) proposed that the essence of Bitcoin is an open-source computer program, which uses mathematical problems to calculate the results, thus generating Bitcoin. In other words, Bitcoin is the return on data processing, which is also the most outstanding contribution of Zhong Bencong. However, not all encrypted digital money is obtained through mining [17].

At present, there are relatively few domestic studies on tightening money. Ling Qing (2016) analyzed and discussed the public's concern and sentiment about Bitcoin and the price fluctuation of Bitcoin. It is concluded that there is a positive correlation between the investor's sentiment and the price of Bitcoin, and it is regarded as the characteristic of speculation. For Bitcoin, a significant change in trend will deprive it of its role as a payment currency [18]. However, the improvement and progress of digital encryption currencies have changed a lot compared with the previous situation. At the same time, its entity application capability has not only the function of payment but also the function of financial investment and equity realization. Chen Hao (2016) analyzed and studied the possible causes of price fluctuation of Bitcoin, and concluded that the price of Bitcoin has a significant inverse proportional relationship with its circulation speed [19]. However, the results of this study are completely opposed to the conclusion of Matonis (2018) on the price fluctuation of Bitcoin. Meanwhile, researchers have also found that the price fluctuation of Bitcoin is more affected by speculation than by supply and demand. It must be admitted that in the current market, the speculative operation of investors is the main reason for the existence of the price of the encrypted currency and the actual application of blockchain encrypted currency is still relatively unsatisfactory [20]. However, the use of encrypted currency scenarios is increasing steadily compared with previous ones.

In this study, based on Semantic Orientation Pointwise Mutual Information (SO-PMI) and Support Vector Machine (SVM) algorithms, the expected trend of ICO issuing new currency is predicted. The SO-PMI algorithm is used to create a blockchain new currency customer sentiment dictionary and collect users' network comments before ICO based on blockchain currency. The SVM algorithm is used to build an evaluation model to analyze and judge the user's comments, to accurately predict the expected return of ICO's new currency issuance, and to improve the investment operation.

3. Semantic Orientation Pointwise Mutual Information (SO-PMI) Algorithm to Create a New Currency User Emotion Dictionary of Blockchain

3.1. SO-PMI Algorithm

In the field of finance and economics, there is no directly available emotion dictionary, so the first step is to construct an emotion dictionary based on the finance and economics corpus. In general, news related to profits will drive up stock prices, while news related to corporate losses will weaken investors' enthusiasm for buying. There are common words in the field of finance and economics, which contain important emotional colors. These words can accurately express the feelings of investors, and their number is extremely limited. Therefore, in order to more accurately classify financial communication speech, it is necessary to build an emotional dictionary in this field. In the SO-PMI algorithm, a group of commendatory words and a group of derogatory words are selected as benchmark words. If the mutual information between the points of $word_1$ and P(words) is subtracted from the mutual information between the points of $word_1$ and N(words), the emotional tendency of $word_1$ according to the difference can be judged. Therefore, the SO-PMI algorithm is suitable for this study.

The SO-PMI algorithm applies the PMI method to the calculation of Semantic Orientation (SO) to achieve the purpose of finding emotional words. The basic idea of sentiment tendency analysis based on SO-PMI is as follows. A group of positive affective reference words and a group of negative affective reference words are selected. Positive-Word and Negative-Word are used to express the two groups of words. The emotional inclination of the selected emotional words must be very obvious and can represent their domains [21].

The basic steps of the algorithm are as follows. Positive-Word and Negative-Word are selected as the benchmark dictionaries of affective dictionaries. Then, the SO-PMI algorithm is used to expand the dictionary, and new words discovered by the algorithm are added to Positive-Word or Negative-Word set one after another until a more complete set of words is obtained. The detailed procedure of the algorithm is as follows. Firstly, PMI is used to calculate the semantic similarity of each word in a new word and a reference word. The basic idea is to calculate the probability of two words appearing simultaneously in a text. The greater the probability of simultaneous occurrence, the higher the similarity between the two words and the closer the degree of association is [22]. The formulas for calculating the PMI values of the two words are as follows:

$$\text{PMI}(\text{word}_1, \text{word}_2) = \log_2 \frac{\text{p}(\text{word}_1, \text{word}_2)}{\text{p}(\text{word}_1)\text{p}(\text{word}_2)} \tag{1}$$

Among them, $\text{PMI}(\text{word}_1, \text{word}_2)$ denotes the probability that two words word_1, word_2 appear in a text at the same time, while $\text{PMI}(\text{word}_1, \text{word}_2)$ denotes the probability that two words appear alone. The higher the similarity of the two words in the set is, the greater the value of $\text{PMI}(\text{word}_1, \text{word}_2)$ is, and vice versa. According to the knowledge of mathematical statistics, the results calculated by the same 1og function can be divided into three situations:

$$\text{PMI}(\text{word}_1, \text{word}_2) \begin{cases} > 0; \text{ The two words are related. The greater the correlation value,} \\ \quad \text{the greater the corrleation degree.} \\ = 0; \text{ These two words appear, neither related nor multually exclusive.} \\ < 0; \text{ Two words are multually exclusive.} \end{cases} \tag{2}$$

$\text{PMI}(\text{word}_1, \text{word}_2)$ and $\text{PMI}(\text{word}_1)$ and $\text{PMI}(\text{word}_2)$ are calculated by the document frequency method. The total number of documents defined is N, while $\text{PMI}(\text{word}_1)$ and $\text{PMI}(\text{word}_2)$ represents the number of documents contained.

$$\text{p}(\text{word}_1) = \frac{\text{f}(\text{word}_1)}{\text{N}} \tag{3}$$

$$\text{p}(\text{word}_2) = \frac{\text{f}(\text{word}_2)}{\text{N}} \tag{4}$$

$$\text{p}(\text{word}_1, \text{word}_2) = \frac{\text{f}(\text{word}_1, \text{word}_2)}{\text{N}} \tag{5}$$

Another type of PMI formula can be derived from the above formulas:

$$\text{PMI}(\text{word}_1, \text{word}_2) = \log_2 \frac{\text{N} * \text{f}(\text{word}_1, \text{word}_2)}{\text{f}(\text{word}_1)\text{p}(\text{word}_2)} \tag{6}$$

Through the above calculation, the similarity of the two words can be obtained, which can measure the degree of a close relationship between words. It is found that the above algorithm can be used to expand the dictionary of professional fields. In this study, it is necessary to construct an emotional dictionary in the field of finance and economics. First, it is necessary to select some of the most representative and commonly used positive emotional words and negative emotional words as the benchmark words. The characteristics of these emotional words should be very obvious, and the PMI value of the new words can be subtracted from that of the reference words to obtain the difference. According to this difference, the emotional tendency of new words can be determined [23]. As the

critical value of emotional attitudes segmentation, emotional judgment will be made according to the magnitude of the value obtained:

$$\text{PMI}(\text{word}_1)\begin{cases} > 0; \text{ Positive affective tendency}: \text{ the higher the absolute value is, the stronger the affection is.} \\ = 0; \text{ Neutral affective tendency.} \\ < 0; \text{ Negative emotional tendency}: \text{ the higher the absolute value is, the stronger the negative emotion is.} \end{cases} \tag{7}$$

Through the SO-PMI method, a dictionary can be constructed based on a small amount of financial vocabulary to form a dictionary with comprehensive coverage. The storage structure of the emotional dictionary is shown in Figure 1.
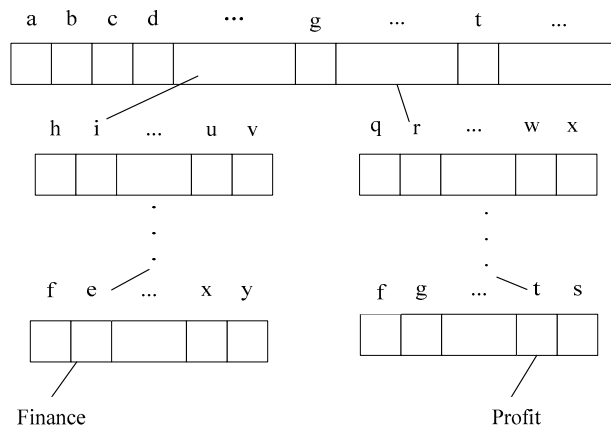
Figure 1. Dictionary method based on TRIE index tree.

3.2. Construction of User Emotional Dictionary Based on SO-PMI

The process of generating benchmark words is as follows.

Firstly, the corpus with currency-related labels issued by blockchains for a certain period of time is obtained from the social network. After text pre-processing, word frequency is counted and sorted. These words are screened from high to low in terms of word frequency, and the words with clear emotional orientation are selected as the benchmark words for dictionary generation.

Positive-Word: green, growth, rise, high, profit, good, support, potential.

Negative-Word: red, falling, low, loss, bad, regulatory, hacker, selling, attack.

The selected benchmark words are very common words in the news related to the new currency market in the blockchain. Investors can quickly and clearly judge the emotional tendency of releasing news through these words.

The steps of creating an emotional dictionary in the financial field are as follows.

The improved SO-PMI algorithm is used to construct an emotional dictionary in the financial field. The detailed method is described as follows:

Step1: Text preprocessing of data, including Chinese word segmentation, de-noising words, de-stop words.

Step2: The candidate words are selected and matched with the words already existing in the emotional dictionary in the financial field to determine whether they already exist in the dictionary. If it already exists in the current dictionary, it is necessary to stop the step. If not, it is necessary to move on to the next step.

Step3: The selected candidate words are calculated with the positive and negative emotional reference words in the established dictionary of financial and economic fields respectively, and the mutual information value of the emotional tendency points, namely SO-PMI value, is obtained. According to the SO-PMI equation, the emotional tendency is judged.

Step4: According to the emotional tendency judged in step 3, the candidate words are included in the corresponding domain emotional dictionary. If it is a neutral word, it should be discarded.

Step5: End.

3.3. Storage Structure of Emotional Dictionary

The storage structure of emotion dictionary directly affects the performance of the system. The Chinese word segmentation dictionary mechanism based on the TRIE index tree is a word segmentation mechanism that represents the key tree in the form of multiple linked lists of trees. The structure is usually divided into two levels, namely the initial hash table and the TRIE index tree node [24]. Chinese word segmentation dictionary mechanism based on whole word dichotomy is one of the most widely used Chinese word segmentation dictionary mechanisms. The structure is usually divided into three levels, namely, the first-word hash table, the word index table and the dictionary body. The first two levels are indexes, as shown in Figure 2. The first-word hash function is based on the national standard location code of Chinese characters. A hash operation can be used to locate the position of Chinese characters in the table. Because of the variable length of words, and to be able to access words randomly, it is necessary to build index table of words to meet this requirement. In a dictionary with words as the basic unit and arranged in ascending order according to the first words, the whole words in the dictionary can be searched quickly by dichotomy by combining it with the text and index table.
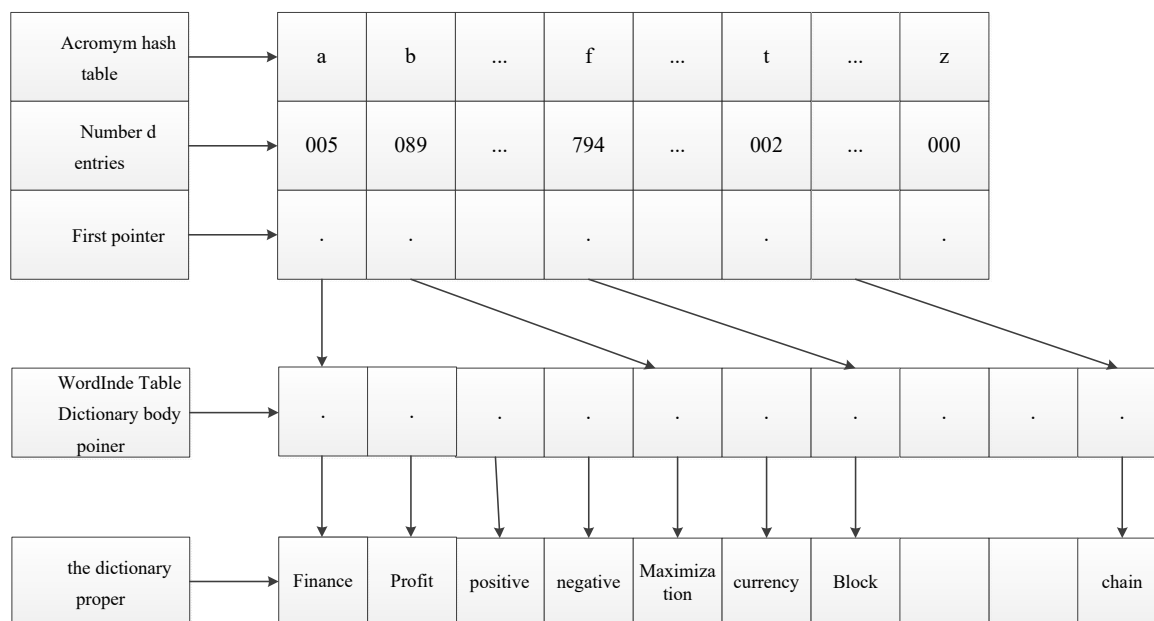

Figure 2. Word segmentation storage mechanism based on binary-seek-by word.

4. Data-Mining Support Vector Machine (SVM) Algorithms for Expected Income Analysis of New Currency in Blockchain

4.1. Data Mining

Data mining can analyze data more comprehensively and scientifically summarize the law of development. The key steps are preparing data, searching rules and revealing rules. In the process of social development, data mining is the key content, which is widely used in the fields of electricity, medicine, and agriculture. At present, the continuous strengthening of data mining needs to apply machine learning scientifically based on a large data background, constantly improve the operational ability, and promote the sustainable development of data mining operations. The data-mining operation steps are shown in Figure 3.
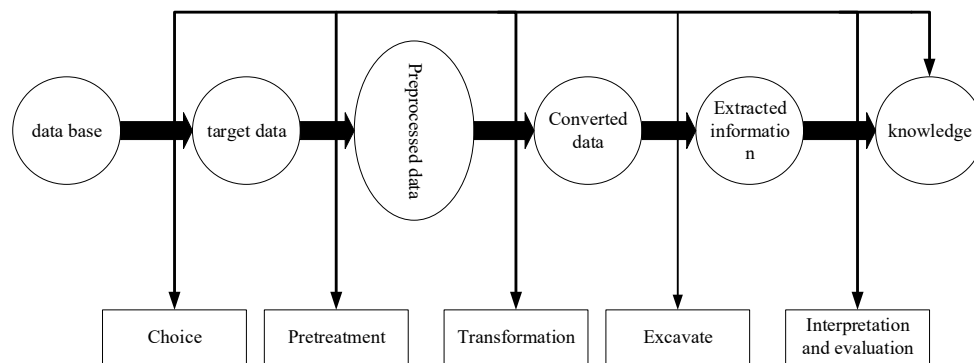
Figure 3. Data-mining process steps.

The functions of data mining mainly include: (1) classification: according to a large number of existing data, the classified training set is determined, and data mining classification technology is used to analyze the training set and establish a classification model; (2) valuation: the valuation function deals with the output of continuous values, while the classification function deals with the output of discrete variables; (3) prediction: through classification or estimation value model, unknown variables in the future are predicted. However, it takes a period of time to verify the accuracy of prediction; (4) association rules: through data analysis and processing, it is more likely to determine which things are combined together; (5) clustering: the difference between clustering and classification is that clustering does not rely on pre-set categories; (6) visualization: data-mining results are represented

4.2. Initial Coin Offering (ICO) Issues New Currency

An ICO issued new coins to realize the transaction realization of blockchain goods, thus recovering costs and gaining profits. Therefore, an ICO is a piece of good news for users who hold blockchain currencies. However, issuing new currency through ICO only shows that money can be freely traded. The return after ICO depends on whether the price of money rises or falls, and all depends on market sentiment and recognition. The rise and fall are totally influenced by market supply and demand, and there are huge arbitrage opportunities and space. Therefore, it is necessary to anticipate earnings before ICO issues new currency. Figure 4 shows the process of issuing new currency in the blockchain.
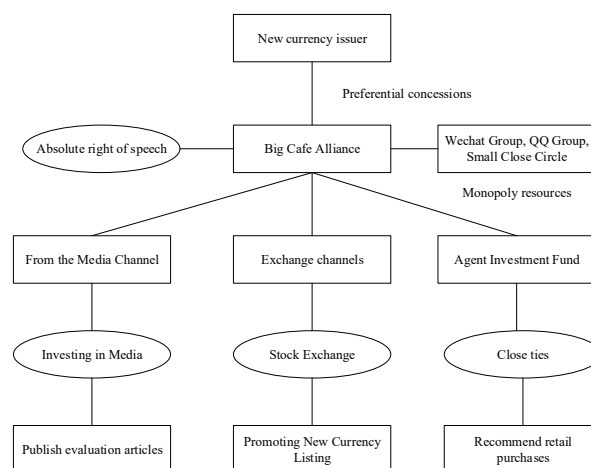

Figure 4. Flow chart of issuance of new initial coin offering (ICO) in blockchain.

According to statistics, there are 650 tokens on the market. There are two main ways to transfer these tokens. One is exchange trading, and the other is P2P address transfer. Exchanges have different classifications, such as centralized and decentralized exchanges, and French currency exchange [25].

The most distinctive feature of the token market is the cross-market nature of token trading. The value of tokens in different exchanges is relatively linked. The blockchain architecture model is shown in Figure 5.
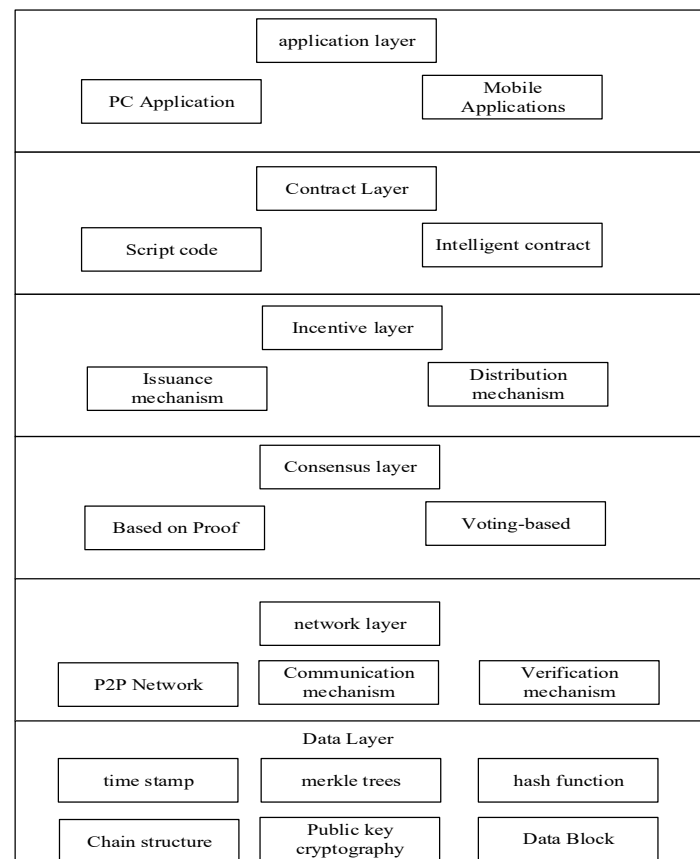

Figure 5. Blockchain framework.

4.3. Establishment of the Revenue Prediction Model for SVM Algorithms

The SVM algorithm can analyze the data more comprehensively and summarize the law of development scientifically. The key steps are preparing data, searching rules, and revealing rules [26]. In the process of social development, data mining is the key content, which is widely used in the fields of electricity, medicine and agriculture. Through the previous step, a training text of the new currency distribution corpus in the completed blockchain can be obtained. After that, the training is carried out using the Lib Short Text software package. In this study, the language model is chosen as the binary language model, i.e., extracting several words at a time, the former word and current word and the latter word. Meanwhile, the complexity of training and the requirement of the accuracy of results are considered. When calculating the feature weight of words, it is characterized by whether words appear or not. At the same time, the commonly used stop words are removed. In the aspect of classifier selection, a multi-classification SVM algorithm is selected, and the linear kernel function is used for kernel function.

For training, feature selection is based on emotional words and negative words modifying emotional words, as well as their location and number in the text. Through training, the classification model of affective orientation is obtained. Figure 6 shows the emotional classification SVM training flow.
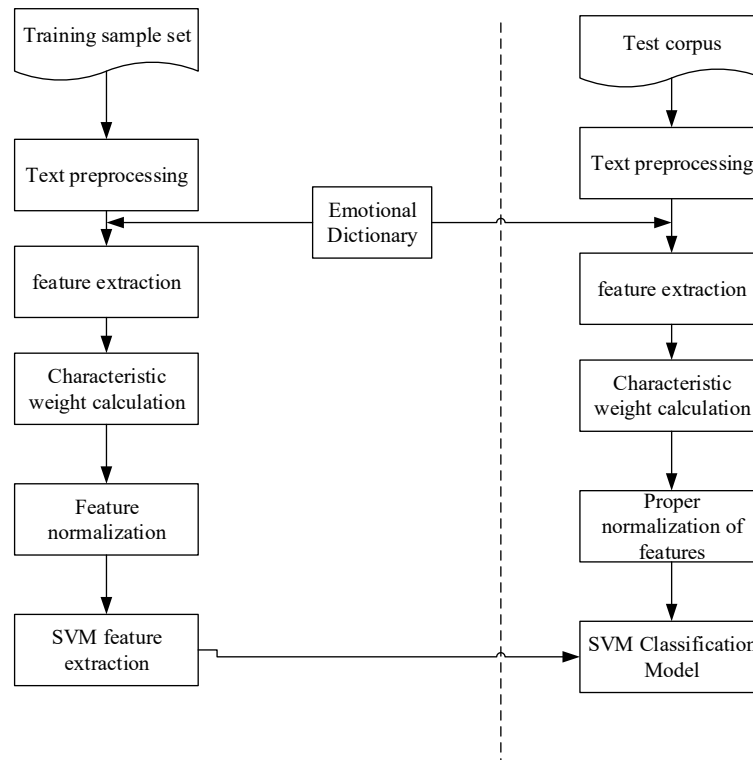
Figure 6. Flow chart of Support Vector Machine (SVM) algorithm for emotional classification.

A user's personal comments on blockchain currency posted on the Internet are all connected by punctuation marks. Therefore, a sentence can be expressed in the form of several sentences by punctuating short sentences, namely $S_1, S_2 \ldots S_n$. The emotional classification of independent financial news requires the following steps. Firstly, the emotional word w_i in each sentence is extracted from the emotional dictionary. If the emotional word w_v is modified by an adverb, the emotional value of the word will be affected by the adverb, and the emotional value of the emotional word needs to be weighted. The formula for calculating the weighted emotional value is as follows:

$$O_{w_i} = S_{W_I} \times M_{W_i} \tag{8}$$

In the former formula, S_{W_i} denotes the weight of the emotional word w_i in the sentence, and M_{W_i} denotes the weight of the degree adverb used to modify the emotional word.

When a negative word appears in a sentence, it may change the emotional tendency of the sentence. According to the method of judging whether the negative words are effective or not, the weight of emotion is calculated. The formula is as follows:

$$O_{w_i} = S_{W_I} \times M_{W_b} \tag{9}$$

In the former formula, S_{W_i} denotes the weight of the emotional word w_i in the sentence and M_{W_b} denies the weight of the negative word.

When a sentence contains several emotional words, the formula for calculating the emotional tendency of the sentence is as follows:

$$O_{S_i} = \sum_{i=1}^{k} O_{W_i} \tag{10}$$

Therefore, the formula of d$_i$ affection with n sentences can be obtained as follows:

$$O_{d_i} = \sum_{i=1}^{n} O_{S_i} \tag{11}$$

When the language d$_i$ contains m emoticons w$_e$, and the emoticons may be one or more, the weights of all emoticons need to be calculated.

$$O_{d_i}^t = O_{d_i} + \sum_{i=1}^{m} W_{e_i} \tag{12}$$

Radial basis kernel function:

$$k(x_i + x_j) = \exp(-\frac{\|x_i - x_j\|^2}{\sigma^2}) \tag{13}$$

At the network level, the value of the final emotional propensity value of the newly issued currency of ICO in the blockchain will appear in the following three situations:

$$O_{d_i}^t = \begin{cases} > 0 \text{ expresses positive affective tendency.} \\ = 0 \text{ expresses neutral emotions.} \\ < 0 \text{ expresses negative affective tendency.} \end{cases} \tag{14}$$

After the above calculation steps, the user's feelings about the blockchain currency can be classified and then the investor sentiment analysis can be completed, so as to judge the probability of the rise and fall of the new currency issued in the blockchain and make a judgment on the expected revenue.

As can be seen from Figure 7, the first step of the traditional product development and promotion mode is to do a good job of product quality and highlight its own value. After having a good sales effect in the market, a brand and platform can be created through social trust in ourselves. Finally, their strength and influence are used to convey their own rules to the industry. However, the encrypted currency has its own unique features. Because the existence of encrypted currency can initially penetrate the incentive and support mechanism, it can become a closed ecosystem. At the same time, because of the security and openness of blockchain-encrypted currency, there is trust at the beginning, which will directly improve project and product thinking from the beginning to ecological thinking.
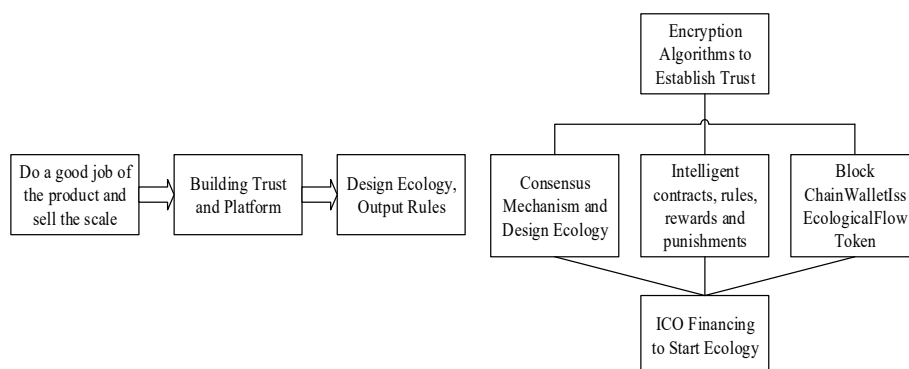

Figure 7. Comparisons between blockchain currencies and traditional business models.

5. Analysis of New Currency Income in Blockchain

With the development of computer and Internet technology, blockchain-encrypted currency has been widely understood and accepted and has become a hot research topic in the field of technology and finance. The flexible and equal architecture of blockchains and the trustworthiness of data storage

make them key technologies in future value internet. Bitcoin, the symbol of the blockchain encryption currency, has risen hundreds of thousands of times since it was invented. More and more countries have recognized the status of blockchain encrypted currency. In real life, it already has the ability to pay as a currency. Therefore, more and more people choose blockchain-encrypted currency as an investment direction. Blockchain currency has its own characteristics compared with other currencies, stocks, gold, funds and so on. It is not affected by inflation, and it is very safe and will not be lost. It will never be interfered with by third-party institutions and absolute exclusive rights. Wherever there is a network, people can manage their assets and trade. The cost is very low, and people can trade across platforms. The most important thing is that the encrypted currency is a world-class currency, which is recognized by most countries. It is precisely because there are so many advantages of a new currency in the blockchain that more and more people participate in the investment of encrypted currency. Therefore, ICO is very important for users who hold new coins, because ICO can realize converting their tokens into real money. The price of new ICO coins after issuance depends on the market's approval mood for new coins, so the income after issuance of new coins is uncertain. Therefore, it is necessary to study the expected income of blockchain currencies. Investors of blockchain currencies can be divided into several levels, including short-term, long-term and medium-term, as shown in Table 1.

Table 1. Blockchain-encrypted currency investor classification.

Investment Preference	Demand Group	Ways of Return
Short-term	Short-term speculators	Trading Spread
Metaphase	Value Investors	ICO token Trading Spread + ICO special token Selling Revenue
Long-term	Strategic investors, brokerage companies, industry veterans, ticketing companies, etc.	ICO token Trading Spread+, Selling Income + Dividend Income+, Non-expendable equity income

The model based on the data-mining algorithm must satisfy the need for accurate prediction of the issuance income of new currency in the blockchain. In this study, based on the expected return of ICO issuing new currency in blockchain, an emotional dictionary is constructed by combining the SVM algorithm and SO-PMI algorithm through collecting network speech in specific fields, thus realizing the judgment of users' attitudes and emotions towards blockchain currency. Users' attitudes and emotions will largely reflect whether they will invest in the issuance of new currency in the blockchain. Based on this rule, the possible trend of new coins in blockchain can be judged by arithmetic operation, and the ICO income of new coins held by oneself can be judged. Through the evaluation of data-mining technology, investors have an advance understanding of expected returns. The analysis is shown in Figure 4. In this study, the feelings, attitudes and causes of a large number of ordinary encrypted users are collected and analyzed to predict the trend of cryptocurrency. Other relevant research is carried out mainly based on the basic information of the cryptocurrency and the influence of policy system factors, so as to study the trend of the cryptocurrency.

In this study, 90 samples of ICO-listed cryptocurrencies are selected, and Bitcoin is used to benchmark the algorithm. Then, the ICO trend of the Litecoin and Etheric coins is studied.

As can be seen from Figure 8, based on the SVM algorithm and SO-PMI algorithm studied in this study, the overall trend of the predicted data is basically consistent with the real data in the prediction analysis of the trend of blockchain-encrypted currency Bitcoin cash. Although there are some differences in price, because of the forecast, it is within a reasonable range, which can meet the purpose of this study to achieve the expected return. Therefore, the combination of two algorithms based on data-mining technology can reduce investment risk, improve reasonable operation and achieve expected returns.
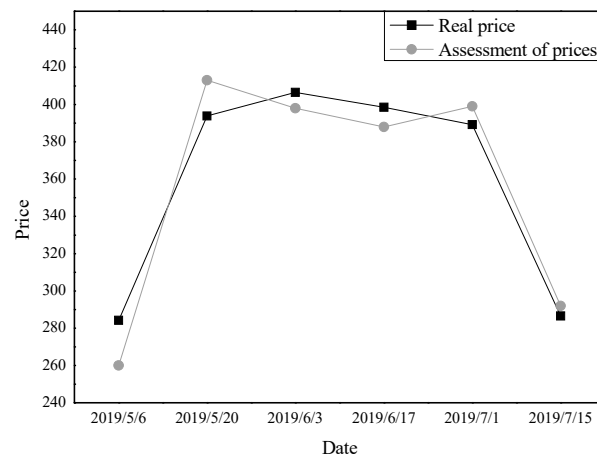
Figure 8. Bitcoin trend analysis based on SVM and Semantic Orientation Pointwise Mutual Information (SO-PMI) algorithms.

From Figure 9, it can be seen that SVM and SO-PMI algorithms show high accuracy and self-regulation ability in evaluating the closing prices of the ICO blockchain-encrypted currencies Yitaifang and Litecoin in the first four days of listing. The errors of real price and forecast value are within the acceptable range, and the trend is basically the same. Therefore, based on the data-mining algorithm, the expected revenue of blockchain new coins can be achieved.
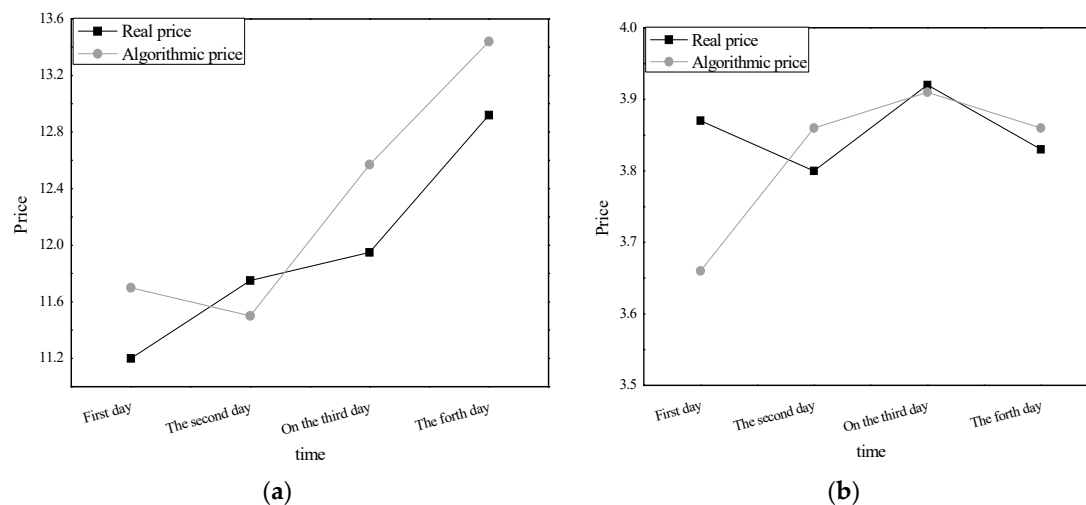

Figure 9. Analysis and research on ICO price trend of encrypted currency Yitaifang and Litecoin. (**a**) Etheric coin (**b**) Litecoin.

6. Conclusions

In this study, data-mining technology is applied to solve the problem of the expected return of currency in the blockchain. The SVM and SO-PMI algorithms are combined to realize the judgment of the price of the newly issued currency in ICO blockchain, and then the corresponding prediction is made according to the trend to realize the analysis of the expected return of a new currency in the blockchain. In the process of data mining, the combination of the two algorithms can accurately analyze the collected resources and make accurate judgments. The greatest difficulty of the expected earnings of blockchain-encrypted currency ICO lies in the fact that the transaction of encrypted currency is not limited by national boundaries, and there are investors from all over the world trading at every moment. Therefore, to realize the accurate expectation of the earnings of ICO, it needs large enough information data. The strategy here is to extract the communication information of the global

mainstream-encrypted currency community and relevant blockchain professionals' social platform release information. Then, the algorithm of this paper is used to analyze and study, master the market sentiment, and thus study the encrypted currency income.

Blockchain technology has good application prospects in digital currency and token financing. The latter provides the enterprise and project team with great imagination in business. Although large-scale commercial applications have not yet been born, it is believed that with the improvement of market attention and investment in technology research and development the era of blockchain large-scale commercial applications will come. For this reason, both the market and the regulatory authorities should maintain their continuous attention and exploration on blockchain technology and the crypto-token financing mode. Meanwhile, it is hoped that regulators will actively educate and guide the market, and really use blockchain technology in value innovation.

In this study, based on the emotional analysis of users, the return of money is predicted. In the SO-PMI algorithm of data mining, the TRIE index tree and its segmentation dictionary mechanism serve for the analysis of users' emotions. Therefore, in this study, first of all, how to build user sentiment analysis based on the SO-PMI algorithm is explained, and then the benefits of a blockchain new currency is predicted. It shows that the data-mining techniques combined with the SVM algorithm and SO-PMI algorithm can solve the problem of the expected return of a new currency in ICO blockchain to the greatest extent. In this study, the main contribution is to verify the algorithm model in the expected trend of the cryptocurrencies Litecoin and Etheric, so as to illustrate the value of the algorithm and the contribution of this work. The algorithm of this study plays a very important role in the daily life of stock speculation, futures and other financial activities, which can help individuals or institutions to study the future trend. Aiming at the problem of the expected return of a blockchain currency, the SVM and SO-PMI algorithms can be combined to solve it. Although good results have been achieved, due to the price trend of blockchain currencies, and many other factors like limited personal research ability, there are still many aspects to be explored in the study of the expected return of blockchain currencies.

References

1. Yli-Huumo, J.; Ko, D.; Choi, S.; Park, S.; Smolander, K. Where is current research on blockchain technology?—A systematic review. *PLoS ONE* **2016**, *11*, e0163477. [CrossRef]
2. Zheng, Z.; Xie, S.; Dai, H.N.; Chen, X.; Wang, H. Blockchain challenges and opportunities: A survey. *Int. J. Web Grid Serv.* **2018**, *14*, 352–375. [CrossRef]
3. Eyal, I.; Sirer, E.G. Majority is not enough: Bitcoin mining is vulnerable. *Commun. ACM* **2018**, *61*, 95–102. [CrossRef]
4. Ozercan, H.I.; Ileri, A.M.; Ayday, E.; Alkan, C. Realizing the potential of blockchain technologies in genomics. *Genome Res.* **2018**, *28*, 1255–1263. [CrossRef] [PubMed]
5. Wang, R.; Lin, Z.; Luo, H. Blockchain, bank credit and SME financing. *Qual. Quant.* **2019**, *53*, 1127–1140. [CrossRef]
6. Kano, Y.; Nakajima, T. A novel approach to solve a mining work centralization problem in blockchain technologies. *Int. J. Pervasive Comput. Commun.* **2018**, *14*, 15–32. [CrossRef]
7. Crosby, M.; Pattanayak, P.; Verma, S.; Kalyanaraman, V. Blockchain technology: Beyond bitcoin. *Appl. Innov.* **2016**, *2*, 71.
8. Chuen, D.L.K.; Guo, L.; Wang, Y. Cryptocurrency: A new investment opportunity? *J. Altern. Investig.* **2017**, *20*, 16–40. [CrossRef]
9. Yeoh, P. Regulatory issues in blockchain technology. *J. Financ. Regul. Compliance* **2017**, *25*, 196–208. [CrossRef]

10. Hayes, A.S. Cryptocurrency value formation: An empirical study leading to a cost of production model for valuing bitcoin. *Telemat. Inform.* **2017**, *34*, 1308–1321. [CrossRef]

11. Giungato, P.; Rana, R.; Tarabella, A.; Tricase, C. Current trends in sustainability of bitcoins and related blockchain technology. *Sustainability* **2017**, *9*, 2214. [CrossRef]

12. Risius, M.; Spohrer, K. A blockchain research framework. *Bus. Inf. Syst. Eng.* **2017**, *59*, 385–409. [CrossRef]

13. Mengelkamp, E.; Notheisen, B.; Beer, C.; Dauer, D.; Weinhardt, C. A blockchain-based smart grid: Towards sustainable local energy markets. *Comput. Sci. Res. Dev.* **2018**, *33*, 207–214. [CrossRef]

14. Kuo, T.T.; Kim, H.E.; Ohno-Machado, L. Blockchain distributed ledger technologies for biomedical and health care applications. *J. Am. Med. Inform. Assoc.* **2017**, *24*, 1211–1220. [CrossRef] [PubMed]

15. Mamoshina, P.; Ojomoko, L.; Yanovich, Y.; Ostrovski, A.; Botezatu, A.; Prikhodko, P.; Izumchenko, E.; Aliper, A.; Romantsov, K.; Zhebrak, A.; et al. Converging blockchain and next-generation artificial intelligence technologies to decentralize and accelerate biomedical research and healthcare. *Oncotarget* **2018**, *9*, 5665. [CrossRef] [PubMed]

16. Yoo, S. Blockchain based financial case analysis and its implications. *Asia Pac. J. Innov. Entrep.* **2017**, *11*, 312–321. [CrossRef]

17. Puthal, D.; Malik, N.; Mohanty, S.P.; Kougianos, E.; Das, G. Everything you wanted to know about the blockchain: Its promise, components, processes, and problems. *IEEE Consum. Electron. Mag.* **2018**, *7*, 6–14. [CrossRef]

18. Truby, J. Decarbonizing Bitcoin: Law and policy choices for reducing the energy consumption of Blockchain technologies and digital currencies. *Energy Res. Soc. Sci.* **2018**, *44*, 399–410. [CrossRef]

19. Yang, R.; Yu, F.R.; Si, P.; Yang, Z.; Zhang, Y. Integrated blockchain and edge computing systems: A survey, some research issues and challenges. *IEEE Commun. Surv. Tutor.* **2019**, *21*, 1508–1532. [CrossRef]

20. Eyal, I. Blockchain technology: Transforming libertarian cryptocurrency dreams to finance and banking realities. *Computer* **2017**, *50*, 38–49. [CrossRef]

21. Aste, T.; Tasca, P.; Di Matteo, T. Blockchain technologies: The foreseeable impact on society and industry. *Computer* **2017**, *50*, 18–28. [CrossRef]

22. Yuan, Y.; Wang, F.Y. Blockchain and cryptocurrencies: Model, techniques, and applications. *IEEE Trans. Syst. ManCybern. Syst.* **2018**, *48*, 1421–1428. [CrossRef]

23. Gatteschi, V.; Lamberti, F.; DeMartini, C.; Pranteda, C.; Santamaria, V. Blockchain and smart contracts for insurance: Is the technology mature enough? *Future Internet* **2018**, *10*, 20. [CrossRef]

24. Park, J.; Park, J. Blockchain security in cloud computing: Use cases, challenges, and solutions. *Symmetry* **2017**, *9*, 164. [CrossRef]

25. Seshia, S.A.; Hu, S.; Li, W.; Zhu, Q. Design Automation of Cyber-Physical Systems: Challenges, Advances, and Opportunities. *IEEE Trans. Comput. Aided Des. Integr. Circuits Syst.* **2017**, *36*, 1421–1434. [CrossRef]

26. Huang, Z.; Soljak, M.; Boehm, B.O.; Car, J. Clinical relevance of smartphone apps for diabetes management: A global overview. *Diabetes Metab. Res. Rev.* **2018**, *34*, e2990. [CrossRef]