How does social media affect cryptocurrency value? A case study of bitcoin in Thailand.

Sasiphan Nitayaprapha

School of Science and Technology, University of the Thai Chamber of Commerce

Abstract- Emergence of crypto currencies has developed technological revolution in global financial markets. What factors determine crypto valuation have an important effect on business and investment. The paper outlines the ongoing research on the interactions between social media and bitcoin value in Thai crypto market. The research model is constructed to examine the dynamic interactions between social media and bitcoin value. The data mining technique using in this research is vector error correction models. The expected results will show how bullish forum posts affect future bitcoin values, who has more effects on bitcoin – silent majority or vocal minority. Research findings will reveal whether social media sentiment is important in determining bitcoin's value. This research provides new insights into Thai digital currency market as well as the economic impact of social media.

Index Terms- bitcoin, cryptocurrencies, social media, fintech, vector error correction model, text mining.

I. INTRODUCTION

Bitcoin, invented in 2008, is the first cryptocurrency captured the attention of business world. In 2017, the world market capitalization of bitcoins surpassed US$73 billion. The New York Stock Exchange has created a bitcoin index. Additionally, the well-known payment gateway PayPal and retailers like Dell, Google, Microsoft, Newegg, and Overstock accept bitcoin. Hundreds of bitcoin ATMs operate on four continents. In fact, bitcoin has disrupted the existing payment and monetary systems [1].

Thus, the important question is "What determines bitcoin's monetary value?". The factors influencing bitcoin's price, on major bitcoin exchanges, has important practical and theoretical implications. Crypto investors should be able to estimate future price swings and speculate the expected return. Also, policy makers need to know the forces behind bitcoin to establish proper regulations and to deal with financial stability risks [2]. For IS researchers, identifying factors affecting bitcoin's value can advance IS theory by revealing the roles of different parties in new financial technology innovation.

This research aims to find out whether and to what extent social media impact bitcoin's value. Economics literature suggests that the worth of currency depends on variables as macroeconomic conditions, a nation's monetary policies, interest and inflation rates [3]. Yet, bitcoin is a virtual currency with no central bank or government backing, its value is not reflected by the traditional explanatory variables. The researcher, therefore, draws from the theory on the relationship between equity value and social media [4] and hypothesize that social media exerts an effect on the bitcoin value. It has been obvious that information on social media is unobtainable from traditional media. The discussions on social media are more abundant and timely compared to traditional media.

Therefore, studying the interaction between content on social media and bitcoin's value could offer investors and regulators, an indicator of future bitcoin's value. In addition, bitcoin provides an opportunity to understand the economic value of social media. Bitcoin investors emotion is increasingly visible online. Bitcoin's online discussions are abundant, and in diverse forms. Thus, bitcoin is suit for testing new IS theories. Previous research considers social media both as a whole and regarding the mixed signals from various users. The same as the rather, this study investigates the effects of groups of users with different levels of activity [5], 1) the vocal minority or the active users contributing most content, and 2) the silent majority or the inactive users contributing less content.

The research questions are as follows:
1. Is there a predictive relationship between bitcoin value and social media?
2. Does social media information created by vocal minority and silent majority has the same effect?

To answer these two questions, the researcher collects dataset from three main bitcoin Facebook pages, and conduct sentiment analyses of content. VECM (vector error correction model) is used to empirically test the relationship between bitcoin value and social media. This research has practical and theoretical contributions. The understanding of the role of social media sentiment on bitcoin's monetary value offers a new perspective on bitcoin's diffusion. In addition, the understanding of different influences of various social media contributed to the prior findings of Gao et al. [6]. admissions in reputed varsity. Now, here we enlist the proven steps to publish the research paper in a journal.

II. Theoretical Background

2.1 Sentimental Analysis

Sentimental Analysis is the method used to recognized a content's tone or polarity or sentiment. The polarity of content can be in positive, neutral or negative. The sentiment analysis has been used as decision making tools in many areas, such as finance, marketing, or politics.

For example, [7] uses S-Sense sentiment analysis application is used to investigate social media to indicate the viewpoint of Thai people for the general topics, whether it is positive or negative. In addition, in financial domain, [8] uses sentiment analysis to make the stock prediction. The study of [9] suggests that the Natural Language Processing (NLP) is an important technique to interpret the sentiment of content. NLP involves with the text classification and text segmentation. There are two approaches to score the sentiment of context, coarse sentiment and fine-grained sentiment analysis [10]. The coarse sentiment or binary sentiment analysis is the simple classification that distinguishes sentiment into positive, neutral, or negative polar. For instance, [11] categorizes their research dataset into three classes: positive, objective or neutral, and negative. The fine-grained sentiment analysis is more complex and precise classification since it breaks sentiment down into many subgroups. For example, [12] uses fine-grained sentiment approach to categorize research dataset into five groups: very negative, negative, neutral, positive, and very positive. Additionally, there are two data mining techniques for the sentiment analysis, lexicon based and machine learning based techniques [13].

2.2 VECM Analysis

This research uses VECM to study the dynamic relationship between bitcoin market variables and social media metrics. VECM is used to capture the interdependencies across time-series. These models extend the VAR system. VECM assumes that there are long-term common trends among the time series [28].

The research selects VECM rather than multiple regression [14], a more traditional method, for several reasons. Firstly, VECM allows for modeling the recursive relationship between interdependent variables without creating model restrictions by separating variables as endogenous and exogenous. Additionally, the data analyst does not need a knowledge about the mechanisms influencing variables, which is required by structural models. Secondly, VECM allows for autocorrelation and cross-correlation. Therefore, allows for better understanding the dynamic relationships among the variables. Thirdly, VECM as a timeseries model, which can be estimated using Granger causality. This allows to establish the causality between variables. In particular, for this research, VECM allows to test whether the values of social media variables in the past are useful for predicting the values of bitcoin. In this research, the models include bitcoin variables as daily bitcoin price, transaction, and trading volume. The models also include social media activities: number of forum posts categorizing as positive (POS) and negative (NEG). In formal, a

VECM with p variables, k lags, and co-integration order r has the following form:

$$\Delta Y_t = \sum_{j=1}^{k-1} \Gamma_j \Delta Y_{tt-k} + \alpha \beta' Y_{tt-1} + \mu + \epsilon_t$$

Where Δ is the first difference operator. Y_t is a p x 1 vector with order of integration 1. μ is a p x 1 constant vector denoting the linear trend. Γ_j is a p x p matrix indicating short-term relationships among variables. ϵ is the residual vector. Additionally, k is the lag length. α is a p x r matrix representing the speed with which the variables adjust to the long-term equilibria. β is a p x r matrix presenting the long-term relationships between the co-integrating vectors. The difference between the VECM and the VAR model, with first-differenced variables, is the added $\beta' Y_{t-1}$, the error correction term.

III. Research Model and System Architecture

3.1 Research Model

System development is being carried out to obtain dataset from three main Thai bitcoin Facebook pages (Bitcoin Addict Thailand, Bitcoin Thailand Community TH, and Siam Bitcoin), to categorize the news' sentiment, and to report the news along with their sentiment. User interface for displaying and reporting will be properly designed. The main functions of the system are as follows (see figure 1: Research Model):
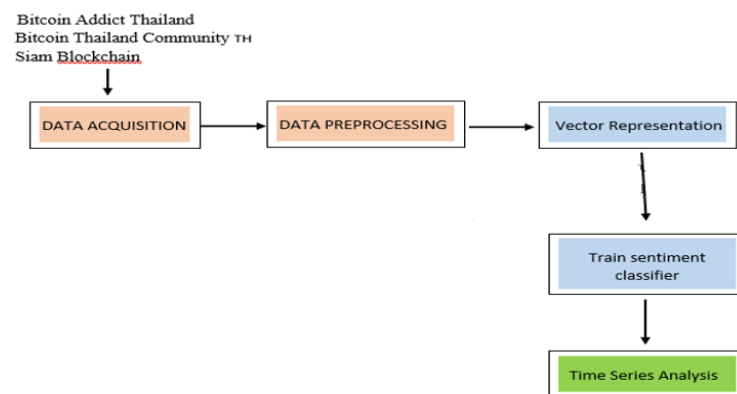

figure 1: Research Model

A. Data Acquisition: This function enables users to extract the daily posts about bitcoin, the source, and the date and time of the posts.
B. Data Preprocessing: This function is used to export posts from the system to the database in the text format, as well as to carry out data Cleansing.
C. Vector Representation: This function allows for feature transformation.
D. Train sentiment classifier: This function indicates the tone of the posts. Each post will be categorized into three classes (positive, neutral, and negative).

E. Time series Analysis: This function uses VECM to investigate the relationship between bitcoin value and the post's sentiment.

3.2 System Architecture

The system architecture is composed of five parts: Crypto Pages which are the data sources for the system, Back-end, Database, Client, and Front-end:

(1) Crypto Pages: This is the input sources of the system. The system uses web crawler to retrieve research dataset from three main bitcoin Facebook pages: Bitcoin Addict Thailand, Bitcoin Thailand Community TH, and Siam Bitcoin through the ETL Process at the back-end side.

(2) Back-end: The back-end system carries out ETL (Extract, Transform, Load) process. ETL is used to retrieve the news from the Facebook pages, as well as to carry out data Cleansing, Feature Transformation, and Model Inference function.

(3) Database: The retrieved data will be recorded in the database.

(4) Client: The client or computer can use the application via web browser.

(5) Front-end: This front-end part includes HTML, CSS Bootstrap, D3.js, and JavaScript.

The system development comprises of two parts: Web-based System Development and Data Modelling, For the data modelling part, after research dataset has been prepared, the data will be modelled using VECM technique. The system will be able to use in both a workstation and a mobile platform. The data will be gather in the duration of four months.

IV. CONCLUSION

This study aims to analyze the relationship between social media and bitcoin valuation. Though, there are several studies on the financial sentiment analysis in crypto currencies, the majority of them have rarely integrated with the Thai financial news. Thus, this research is set out to explore the economic impact of social media on bitcoin valuation. Future results will show whether social media sentiment is an important factor indicating bitcoin value. Also the study will point out which social user group has more impact on bitcoin's valuation, the silent majority or the minority vocal. Research findings will have implications for crypto currency investors as well as policymakers. The predictive relationship will offer information about bitcoin's demand and fluctuations. The investors can discern bitcoin value. Also, the regulators could be able to address the potential risks association with bitcoin's valuation. Although this study focuses on bitcoin, it has broader implications in crypto currencies and business practices in online social media.

This research has several limitations. The data is planned to collected from Facebook pages for the duration of four months. Further studies might include data from other social sites such as twitter. Subsequent text mining analyses of user social networks may create new insights of crypto currency valuation.

REFERENCES

[1] R. Böhme, N. Christin, B. Edelman, and T. Moore, "Bitcoin: economics, technology, and governance," Journal of Economic Perspectives, 2015, 29, 2, pp. 213–238. doi:10.1257/jep.29.2.213.

[2] B. Jopson, "Regulators say bitcoin poses financial stability risks," 2016, Financial Times, June 21.

[3] R. A. Meese and K. Rogoff, "Empirical exchange rate models of the seventies: Do they fit out of sample?," Journal of International Economics, 1983, 14, 1–2, pp.3–24. doi:10.1016/0022-1996(83)90017-X.

[4] X. Luo and J. Zhang, "How do consumer buzz and traffic in social media marketing predict the value of the firm?," Journal of Management Information Systems, 2013, 30, 2, pp.213–238. doi:10.2753/MIS0742-1222300208.

[5] A. Chen, Y. Lu, P. Y. Chau, and S. Gupta, "Classifying, measuring, and predicting users' overall active behavior on social networking sites," Journal of Management Information Systems, 2014, 31, 3, pp.213–253. doi:10.1080/07421222.2014.995557.

[6]. G. Gao, B. Greenwood, J. McCullough, and R. Agarwal, "Vocal minority and silent majority: How do online ratings reflect population perceptions of quality?," MIS Quarterly, 2015, 39, 3, pp.565–589. doi:10.25300/MISQ/2015/39.3.03.

[7] C. Haruechaiyasak, A. Kongthon, P. Palingoon, and K. Trakultaweekoon, "S-sense: a sentiment analysis framework for social media sensing," In Sixth International Joint Conference on Natural Language Processing, 2013, pp.3.

[8] X. Li, H. Xie, L. Chen, J. Wang and X. Deng," News impact on stock price return via sentiment analysis," Knowledge-Based Systems, 2014, 69, pp.14–23.

[9] K. Sriphaew, T. Hiroya, and O. Manabu, "Sentiment Analysis for Thai Natural Language Processing," In Proceedings of the 2nd Thailand-Japan International Academic Conference, TJIA 2009, pp. 123–124.

[10] R. Kohtes, "From valence to emotions: How coarse versus fine-grained online sentiment can predict real-world outcomes," Anchor Academic Publishing, 2014.

[11] S. Baccianella, A. Esuli, and F. Sebastiani, "Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining," LREC, 2010, Vol. 10, pp. 2200–2204.

[12] R. Socher, A. Perelygin, J. Wu, J. Chuang, C. D. Manning, et al., "Recursive deep models for semantic compositionality over a sentiment treebank," In Proceedings of the Conference on Empirical Methods in Natural Language Processing, 2013, pp. 1631–1642.

[13] Q. T. Ain, M. Ali, A. Riaz, A. Noureen, M. Kamran, B. Hayat, et al., "Sentiment analysis using deep learning techniques: A review," International Journal of Advanced Computer Science and Applications (IJACSA), 2017, 8(6), pp.424–433.

[14] W. Antweiler and M. Z. Frank, "Is all that talk just noise? The information content of internet stock message boards," Journal of Finance, 2004, 59, 3, pp.1259–1294. doi:10.1111/ j.1540-6261.2004.00662.x.

AUTHOR

Asst.Prof.Dr.Sasiphan Nitayaprapha, Ph.D. in Informatics (MBS., The University of Manchester), School of Science and Technology, University of the Thai Chamber of Commerce, sasiphan_nit@utcc.ac.th, , annona.reticulata@gmail.com (+662)6974826