



A Leading Indicator Approach with Data Mining Techniques in Analysing Bitcoin Market Value

Özerk Yavuz^{1*},

^{1*} Haliç University, Faculty of Management, Department of Management Information Systems, İstanbul, Turkey, (ORCID: 0000-0002-1371-688X), ozerkyavuz@halic.edu.tr; ozerk@alumni.bilkent.edu.tr; dr.ozerk@gmail.com

(International Conference on Design, Research and Development (RDCONF) 2021 – 15-18 December 2021)

(DOI: 10.31590/ejosat.1039530)

ATIF/REFERENCE: Yavuz, Ö. (2021). A Leading Indicator Approach with Data Mining Techniques in Analysing Bitcoin Market Value. *European Journal of Science and Technology*, (32), 20-26.

Abstract

In the last decade as a result of the changes in business landscape new payment systems have evolved. Some of the Consumers, business stakeholders, investors and individuals turned to different types of payment systems and virtual currencies for various reasons. Peer to peer architected Bitcoin which uses a blockchain mechanism is one of these approaches that found place in our lives. In this study, a leading indicator focused data mining methodology has been followed in analyzing Bitcoin market value and bitcoin valuation. Several classification and clustering algorithms applied to the data following a literature review, pre-processing of the data and conceptual framework formation. Finally performances of these supervised and unsupervised machine learning techniques with rules discovered have been compared, assessed and presented for this type of problem and research domains.

Keywords: Digital Currency, Payment Systems, Bitcoin, Blockchain, Clustering, Classification, Supervised Learning, Unsupervised Learning, Data Mining, Machine Learning, Quantitative Analysis

Bitcoin Piyasa Değerinin Analizinde Veri Madenciliği Teknikleriyle Bir Öncü Gösterge Yaklaşımı

Öz

Günümüzde son yıllarda iş ortamındaki değişikliklerin bir sonucu olarak yeni ödeme sistemleri gelişti. Bunun sonucunda tüketiciler, iş paydaşları, yatırımcılar ve bazı bireyler, çeşitli nedenlerle farklı ödeme sistemlerine ve sanal para birimlerine yöneldi. Bir blok zinciri mekanizması kullanan eşler arası mimarili Bitcoin, hayatımızda yer bulan bu yaklaşımlardan biridir. Bu çalışmada, Bitcoin piyasa değeri ve bitcoin değerlemesinin analizinde öncü gösterge odaklı bir veri madenciliği metodolojisi izlenmiştir. Literatür taraması, verilerin ön işlenmesi ve kavramsal çerçeve oluşturulmasının ardından verilere çeşitli sınıflandırma ve kümeleme algoritmaları uygulanmıştır. Son olarak, uygulanan bu denetimli ve denetimsiz makine öğrenmesi tekniklerinin keşfedilen kurallar ile performansları, bu tür problem ve araştırma alanları için karşılaştırılmış, değerlendirilmiş ve paylaşılmıştır.

Anahtar Kelimeler: Dijital Para Birimi, Ödeme Sistemleri, Bitcoin, Blockchain, Kümeleme, Sınıflandırma, Denetimli Öğrenme, Denetimsiz Öğrenme, Veri Madenciliği, Makine Öğrenimi, Kantitatif Analiz

* Corresponding Author: Asst. Prof. Dr. Özerk Yavuz, Haliç University, İstanbul, Turkey, ORCID: 0000-0002-1371-688X, ozerkyavuz@halic.edu.tr, ozerk@alumni.bilkent.edu.tr, dr.ozerk@gmail.com

1. Introduction

Today as a result of the changes in business landscape new payment systems have evolved. Some of the Consumers, business stakeholders, investors and individuals turned to different types of payment systems and virtual currencies for various reasons. Peer to peer architected Bitcoin which uses a blockchain mechanism is one of these approaches that found place in our lives. Bitcoin is a type of digital currency, using a peer to peer, blockchain structure. This system integrates, decentralized and digital forms and lack of a central authority or administrator. By applying online payment and currency system structures that bitcoin offers individuals can exchange between bitcoin and conventional currencies and store, sent money or make investments. If the etymology of the bitcoin word is examined, it is found that the name bitcoin is composed of bit and coin words as it was first published in a whitepaper in 2008 which dates back to the same years. Bit refers to a unit of data in computer whereas coin is a rounded small form of metal mostly found in silver or copper like color [1,2,3,4,5,6,8]. In the last decades with the advancements in information technology, finance field and business landscape. New payment, money transfer and investment options, technologies and trends found place in our lives. There is no doubt that for a sustainable and growing business landscape that contains huge amounts of transactions importance of reliable, secure, trustable payment systems is necessary. Among them Bitcoin and blockchain technologies have become one of these choices that found place in many's lives consequently in the near future [1,2,3,4,5,6,8].

In comparison to conventional money transfer approaches, bitcoin transfers do not need any additional intermediaries [1,2]. Bitcoin transactions are assessed and evaluated via respective nodes of the network applying security measures in an encrypted way and saved in blockchain in the form of a ledger which can be publicly shared [1,8]. As Rainer puts forth Bitcoin is a protocol which utilizes online communications, enables virtual currency usage composed of electronic payments. Bitcoin usage has started in 2009 and gained attention during the recent years. If the history of bitcoin is analysed it is seen that this technology dates back to 2008 which was developed and marketed as open source by Satoshi Nakamoto [1,3]. Shortly after it gained interest and attention from several stakeholders and individuals for various purposes as money storage, payment, money transfer or even for investment. It can be said that these types of transitions and transformations became available following the developments in information technology and business landscape. New trends in e-commerce, consumer behavior patterns, investor decisions have been a catalyzer in this process.

Human beings can turn to different forms of payment, money transfer and investment options in their lifetime. If their decions and behaviors associated with payment, money transfer and investment are analyzed it can be inferred that they are more risk aversive, considers options with higher trust with security and are in a tendency to reduce risks associated with these decisions while working on maximizing their rewards, gains and benefits associated to these options and decisions. In this search, alternatives that offers and provides solutions for the satisfaction of these individualistic, cognitive and intrinsic needs would benefit the most. In building a long term relationship and enhancing the lifetime value with the customers, individuals or investors it is important to respond to these expectations and

needs. Maslow is one of the pioneers in categorizing some of these needs which are humanistic in nature in literature [36]

Raineer asserts that Bitcoin technology structured by developers mainly influenced by business landscape. It has the mechanisms that constitute logs of transactions distributed across network participants and enables honest participation rewarding, early adapters' bootstrap acceptance and protected to concentrations of power [3]. Its structure includes currency transactions, money creation and constitutes a history of transactions that are being made [1,2,3]. It eliminates some of the barriers, burrocraics and procedures that the conventional currencies and currency transfers have. It enhances flexibility, privacy and decreases the amount of regulatory and governmental control [1,3]. As Jennifer asserts Bitcoin is considered a convertible virtual currency with an equivalent value in real currency or it can be considered as a real currency substitute. This type of virtual currency and real currencies can be converted or exchanged from one another [2]. While centralized virtual currencies have a centralized repository and control, decentralized virtual currencies have no central repository or admins [1,2]

As it is seen in many money storage, money transfer and investment choices and technologies, bitcoin and blockchain technologies have several unique characteristics that they posses. Some of these are, Bitcoin is lack of a central authority or regularity body. It is based on a peer to peer network structure compared to a centralized architecture. The ledger which records the transactions is distributed publicly and not stored in a central server unit. The ledger is administred and maintained by miners which have equal rights regardless of a single authoritarian system point or user. Different individuals can become miner. Ledger extensions and addition of new blocks to the ledger in the blockchain is decided through a contention basis. Until the addition of a new block to the blockchain the respective miner is not known, bitcoin issuance is decentralized and issued as a reward following the creation of a new administred block, any individual can create a new bitcoin wallet address which is used in money sending and receiving transactions. Individuals can freely share and commit a transaction with the peer-to-peer network free from additional approvals. In literature bitcoin and blockchain technologies are frequently cited with many unique benefits and advantages that they provide with some of the concerns that these technologies posses [1, 2, 3, 4, 8]. Therefore it may be expected that governmental leaders and regulatory bodies can work on new rules and regulations to enhance the public's trust by securing money storage, money transfer and investment options that are highly been carried out via these new technologies and respond to concerns associated with these technologies in the near future [1,2,3,4]. In Bitcoin and Blockchain technologies, a special encrypted structure is being used in money sending and storing processes among peers of the peer-to-peer network. Transactions are being stored in blockchains where assessed and evaluated. Miners are rewarded based on the transactions that are being carried out or based on the mathematical problems that are being solved by respective participants' computers. Also nodes which represent the individuals in peer to peer network can exchange their moneys among real currencies and bitcoin [1, 2, 3, 4].

In the near future globalization and e-commerce driven virtual currencies and payment systems are likely to be popularized. Additionally bitcoin like virtual currencies became to be popularized among investors as well. Benefits, advantages, flexibilities, consumer and investor confidence that come with

these technologies and technological structure can drive this interest and lead to more momentum. It is important to be customer, investor focused and centric in providing and developing these services, which would provide customer, investor and individual confidence, meet customer value expectations and lead to a long term solution preference, a long term customer or investor relation and lifetime value [25, 30, 40, 41, 42].

Consumers, Investors, individuals and other stakeholders should make benefit- risk analysis in evaluating these approaches and strengths, weaknesses, opportunities and threats associated with bitcoin, blockchain network structures should always be considered as in conventional methods [3, 4, 5, 7, 8]. As in Kaminsky's leading indicator approach data mining driven historical data analysis for exploratory and confirmatory understanding in such problem domains may give several benefits and enhance understanding of the bitcoin markets which would lead to better and more effective consumer, investor, business leader, governmental leader, individual decision making. As cited in literature these approaches are usefull in decision support mechanisms and may bring many insights [25, 30, 37, 38, 39, 40, 41, 42].

2. Material and Method

2.1. Researh Method

Research process is a structured set of steps in order to conduct a research activity in an organized way. This is also common in data mining driven quantitative studies. Initially a literature review has been done which is later followed with developing a quantitative research design and methodology in the form of data mining process [9,10,11,12,13, 25, 30]. As put forth by Özerk, data mining process can be defined as the process of extracting meaningful patterns, knowledge and insights from large amounts of data and data sets. Data mining based methodologies can be followed for several reasons as exploratory, confirmatory, knowledge discovery, prediction or forecasting. It may provide an in-depth understanding of the topic of interest for stakeholders. [25, 30]. It is mostly used with the aim of knowledge and insight discovery and prediction from quantitative and measurable data. Data mining process includes understanding and analysis of the situation and business problem which is later followed with the examination and pre-processing of data. It is important to form a conceptual framework or model based on analysis and literature. Following the model development, testing of the model with classification and clustering algorithms takes place. Finally findings are assessed and discussed [14,15, 16,17,18,19].

In JRip class labels are examined in an ordered way and an initial set of rules for the class labels are generated using incremental reduced error. JRip executes by considering all the cases of a particular instance as a class and discover and generate rules for the independent variables for the respective class label. After that it continues with the next class labels, following the same process up to all rules for all class labels are discovered and generated [25 ,30, 34]. JRip is a rule learner alike in principle to the rule learner Ripper [25, 30]. RIPPER rule learning algorithm is an extended version of learning algorithm IREP (Incremental Reduced Error Pruning). It constructs a rule set in which all positive examples are covered, and its algorithm performs efficiently on large, noisy datasets. Before building a rule, the current set of training examples are partitioned into two subsets,

e-ISSN: 2148-2683

a growing set (usually 2/3) and a pruning set (usually 1/3). The rule is constructed from examples in the growing set. The rule set begins with an empty rule set and rules are added incrementally to the rule set until no negative examples are covered. After growing a rule from the growing set, condition is deleted from the rule in order to improve the performance of the rule set on the pruning examples [25, 30].

The PART algorithm combines two common data mining strategies; the divide-and-conquer strategy for decision tree learning with the separate-and-conquer strategy for rule learning. The tree building algorithm splits a set of examples recursively into a partial tree. The first step chooses a test and divides the examples into subsets. PART makes this choice in exactly the same way as C4.5. Then the subsets are expanded in order of their average entropy starting with the smallest. The reason for this is that subsequent subsets will most likely not end up being expanded and the subset with low average entropy is more likely to result in a small sub tree and therefore produce a more general rule [25, 30]. OneR, generates a one-level decision tree that is expressed in the form of a set of rules that all test one particular attribute. OneR is a method that often comes up with quite good rules for characterizing the structure in data [25, 30]. Pseudo code for 1R is as follow.

For each attribute,

For each value of that attribute, make a rule as follows:

Count how often each class appears

Find the most frequent class

Make the rule assign that class to this attribute-value.

Calculate the error rate of the rules.

Choose the rules with the smallest error rates [25, 30].

A Multilayer Perceptron is a version of the original perceptron model proposed by Rosenblatt in the 1950s and considered as a type of neural networks (Rosenblatt, 1958). A perceptron (artificial neuron) is a function of several input perceptrons which is formed as a combination of input weights to the hidden layer perceptrons. As stated by Ramchoun in literature multilayer perceptron has one or more hidden layers between its input and output layers, the neurons are organized in layers, the connections are always directed from input layers to output layers and the neurons in the same layer are not interconnected [25, 30]. In this approach hidden layer is a function of the nodes in the previous layer, and the output nodes are a function of the nodes in the hidden layer. In Bayesian Network there are no deterministic rules which allow to identify a subscriber as a risk indicator. Graphical models such as Bayesian networks supply a general framework for dealing with uncertainly in a probabilistic setting and thus are well suited to tackle the problem of prediction. Every graph of a Bayesian network codes a class of probability distributions. The nodes of that graph comply with the variables of the problem domain. Arrows between nodes denote allowed (causal) relations between the variables. These dependencies are quantified by conditional distributions for every node given its parents [22]. A Bayesian network B over a set of variables U is a network structure Bs, which is directed acyclic graph (DAG) over U and set of probability tables $B_p = \{p(u|pa(u)) | u \in U\}$ where $pa(u)$ is the set of parents of u in Bs. A Bayesian network represents probability distributions [25, 30].

In Gauss processes which base on the Notion of the Gaussian distribution are named following the contributions of Carl Friedrich Gauss to scientific body of knowledge [31,32,33]. As it is indicated in literature, a Gaussian process is a stochastic process, such that every finite collection of those random variables has a multivariate normal distribution [31,32,33], Gaussian Processes is considered as a type of supervised learning which is a generalization of the Gaussian probability distribution. As it is generally cited in literature a probability distribution describes random variables which are scalars or vectors on the other hand a stochastic process is mainly concerned with the properties of functions. In this context Gaussian distribution is characterized by mean and covariance matrix on the other hand a Gaussian process is concerned with a mean and a covariance function. The mean is a function of α and the covariance is a function in the form of Covariance(α, β) which shows the anticipated covariance between the values of the respective function Θ in α and β points. The $\Theta(\alpha)$ function for a problem domain which is focused in data modeling is considered to be a sample from the related Gaussian distribution [25, 30, 31,32,33].

2.1.1. Data Gathering and Processing

In the research a research model composed of Opening Value, Highest Value, Lowest Value, Closing Value, Bitcoin Volume, USD Volume, Date, Symbol and Quarter. Public secondary data published online has been used as data set composed of 365 instances and 9 attributes for the understanding of bitcoin market value.

Table 1. List of Attributes

Opening Value	Numeric
Highest Value	Numeric
Lowest Value	Numeric
Closing Value	Numeric
Bitcoin Volume	Numeric
USD Volume	Numeric
Date	Nominal
Quarter	Nominal
Symbol	Nominal

2. Findings

Data mining may give many advantages in building transitions and relations between input and output layers by clustering, classification, association rules and other approaches as indicated by Özerk in literature. These approaches are mainly focused in transformation of independent variables to dependent variables in input and output spaces with mapping functions. These mappings can be done in a more hierarchically customizable layered structures as in multilayer perceptron or in more simplistic designs or fashions. As indicated in literature machine learning is a form of data mining approach in which models are developed and trained from data. In supervised learning, learning is formed from a training set of instances, which base on input and output pairs which are called attributes and class labels. In unsupervised learning system is fed by input however there is no output class labels used in training [19,20,21,22,23,24,25,30]. Among several machine learning approaches classification is a supervised machine learning method which base on a previously labeled set of instances used

for training. Later based on the prior trainings the model learns to classify the new instances. On the other hand, clustering is an unsupervised method which aims to group different objects into groups of similar objects. A cluster is a collection of objects that are similar to each other within the same cluster whereas opposite or dissimilar to the objects in other clusters [19, 24, 25, 26, 27, 30]. Among these machine learning approaches Jrip, part, OneR method, Multilayer Perceptron (Neural Networks) and Bayesian Networks have been chosen as the data mining techniques.

Actual and predicted values of Bitcoin market closing values can be found in Figure 1. Predicted Bitcoin Values have been calculated using Gaussian Processes approach which is a stochastic process in which every finite collection of those random variables has a multivariate normal distribution and that base on the notion of the Gaussian distribution [25, 30, 31,32,33]. The mean is a function of α and the covariance is a function in the form of Covariance(α, β) which shows the anticipated covariance between the values of the respective function Θ in α and β points. The $\Theta(\alpha)$ function for a problem domain which is focused in data modeling is considered to be a sample from the related Gaussian distribution [25, 30, 31,32,33].

Figure 1. Comparison of Actual and Predicted Bitcoin Closing Values in 2019

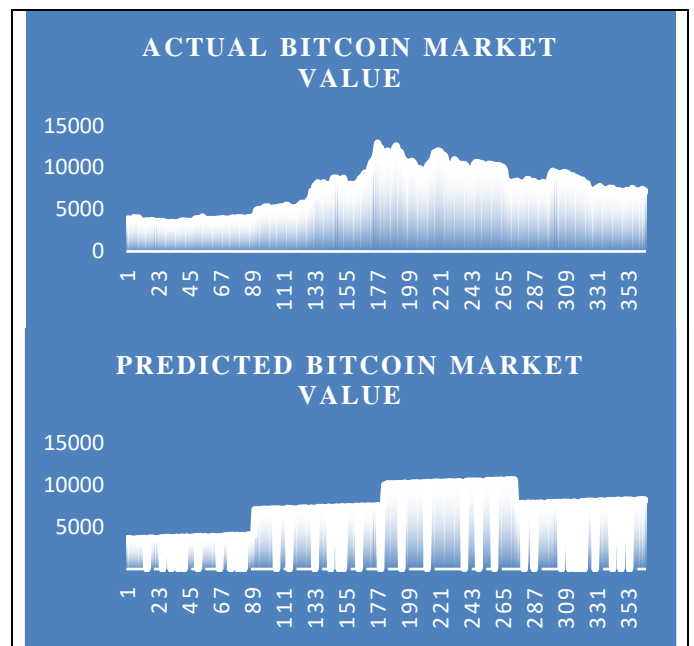


Table 2. Comparison of the data mining methods

Method	RMSE	Precision	Correctly Classified %	Incorrectly Classified %
OneR Method	0,08	0,985	98,38	1,62
Multilayer Percept.	0,12	0,95	95,16	4,84
Bayesian Networks	0,19	0,93	92,74	7,26
J48	0,12	0,97	96,77	3,23
JRip	0,12	0,97	96,77	3,23
Part	0,12	0,97	96,77	3,23

In this research, OneR Method, Multilayer Perceptron, Bayesian Networks, J48, JRip, Part data mining approaches have been examined. 66 percent of the data has been used for the training of the model whereas remaining part of the data set has been used for the testing. In assessing the performances of the

algorithms, some of the performance indicators in data mining have been considered. From these indicators the root mean square error, refers to the quantification of the average dispersion of a set of observations from a known value [28,30] whereas precision indicates the proportion of predicted positive values that are correctly real positives in other words proportion of relevant instances among all retrieved instances [25,29,30]. Among different data mining approaches OneR had the values (RMSE=0.08; Precision=0.98; Correct Classification Rate=98.38%; Incorrect Classification Rate=1.62). Multilayer Perceptron had the values (RMSE=0.12; Precision=0.95; Correct Classification Rate=95.16%; Incorrect Classification Rate=4.84). Bayesian Networks had the values (RMSE=0.19; Precision=0.93; Correct Classification Rate=92.74%; Incorrect Classification Rate=7.26; J48 had the values (RMSE=0.12; Precision=0.97; Correct Classification Rate=96.77%; Incorrect Classification Rate=3.23). JRip had the values (RMSE=0.12; Precision=0.97; Correct Classification Rate=96.77%; Incorrect Classification Rate=3.23). Part had the values (RMSE=0.12; Precision=0.97; Correct Classification Rate=96.77%; Incorrect Classification Rate=3.23). Among all the algorithms, OneR had the most correct classification rate with 98.38 percent in model building and a precision of 0.98. It also had the lowest RMSE with a value of 0.08 [25,30]. No significant difference has been found in the testing performances of the models. Comparison of data mining methods used can be seen in Table 2. Rules discovered with data mining approaches are listed in Table3.

Table 3. Rules Discovered by Data Mining Algorithms

Lower opening values indicate Q1 whereas Higher opening values indicate Q3 and Q2 quarters
Lower day-highs indicate Q1 quarter, higher day-highs indicate Q2 quarter
Day high value of 10498 and Q3 are in the same cluster on the other hand day-high value of 6307 and Q4 are in other cluster
If day-high is lower than 4190 then Q1 quarter
Lower closing value indicates Q1 quarter however higher closing value indicates Q2 quarter
If closing value is less than 4113 then it is Q1 quarter
If day high is greater than 9600 then Q3 quarter
Early dates refers to early quarters whereas late days refer to late quarters
10.09.2019, 9011.06 opening value, 9264.76 day-high value, 8729.11 day-low value, 9012.70 closing value, 12203405130511520 USD Volume and Quarter Q4 are in one cluster whereas 04.03.2019, 4150.91 opening value, 4237.53 day-high value, 4075.21 day-low value, 4167 closing value, 13894190276058918 USD Volume and Quarter Q1 are in another cluster
Closing value of 10184, 13672213744657206 USD Volume and Q3 are in the one cluster on the other hand 6170 closing value, 12308265209762638 USD volume, Q4 quarter are in another cluster

3. Conclusions and Recommendations

Today as a result of the changes in business landscape new payment systems have evolved. Some of the Consumers, business stakeholders, investors and individuals turned to different types of payment systems and virtual currencies for various reasons. Peer

to peer architected Bitcoin which uses a blockchain mechanism is one of these approaches that found place in our lives. Bitcoin is a type of digital currency, using a peer to peer, blockchain structure. In the last decades with the advancements in information technology, finance field and business landscape. New payment, money transfer and investment options, technologies and trends found place in our lives. There is no doubt that for a sustainable and growing business landscape importance of reliable, secure, trustable payment systems is necessary. Among them Bitcoin and blockchain technologies have become one of these choices that found place in many's lives consequently in the near future [1,2,3,4,5,6,8].

In comparison to conventional money transfer approaches, bitcoin transfers do not need any additional intermediaries [1,2]. Bitcoin transactions are assessed and evaluated via respective nodes of the network applying security measures in an encrypted way and saved in blockchain in the form of a ledger which can be publicly shared [1,8]. As Rainer puts forth Bitcoin is a protocol which utilizes online communications, enables virtual currency usage composed of electronic payments. Bitcoin usage has started in 2009 and gained attention during the recent years. If the history of bitcoin is analysed it is seen that this technology dates back to 2008 which was developed and marketed as open source by Satoshi Nakamoto [1,3].

In the analysis part of the research machine learning algorithms focused on classification and clustering efforts in the form of supervised and unsupervised machine learning and have been applied. The model has been trained and built using 66 percent of the data whereas remaining part of the data have been used for testing of the model for each algorithm. Finally associations between different constructs in conceptual framework and model have been assessed. Among all the algorithms, OneR had the most correct classification rate in model building performance with 98.38 percent and a precision 0.98. It also had the lowest RMSE with a value of 0.08. For model testing performances, no significant difference has been detected for the respective data mining algorithms.

Based on the analysis results and performance indicators of machine learning such as classification rate, RMSE, precision, OneR method can be considered as an effective and efficient method in analyzing bitcoin market value, its antecedents and bitcoin market closing rate in a time series fashion. Additionally actual and predicted bitcoin closing market values have been compared. Predicted market closing values of Bitcoin have been calculated by applying Gaussian processes as suggested in literature.

As a result of the data mining driven analysis results some rules and insights have been discovered. These findings shortly as follow, lower opening values indicate Q1 quarter whereas higher opening values indicate Q3 and Q2 quarters. Lower day-highs indicate Q1 quarter, higher day-highs indicate Q2 quarter. Day high value of 10498 and Q3 are in the same cluster on the other hand day-high value of 6307 and Q4 are in other cluster. If day-high is lower than 4190 then Q1 quarter. Lower closing value indicates Q1 quarter however higher closing value indicates Q2 quarter. If closing value is less than 4113 then it is Q1 quarter. If day high is greater than 9600 then Q3 quarter. Early dates refers to early quarters whereas late days refer to late quarters. 10.09.2019, 9011.06 opening value, 9264.76 day-high value, 8729.11 day-low value, 9012.70 closing value, 12203405130511520 USD Volume and Quarter Q4 are in one

cluster whereas 04.03.2019, 4150.91 opening value, 4237.53 day-high value, 4075.21 day-low value, 4167 closing value, 13894190276058918 USD Volume and Quarter Q1 are in another cluster. Closing value of 10184, 13672213744657206 USD Volume and Q3 are in one cluster on the other hand 6170 closing value, 12308265209762638 USD volume, Q4 quarter are in another cluster. Applying data mining focused time series data analysis for exploratory and confirmatory understanding may give several advantages in the problem domains of this type and provide many insights for decision makers as a decision support point or datum.

References

- [1] <https://en.wikipedia.org/wiki/Bitcoin>
- [2] "Statement of Jennifer Shasky Calvery, Director Financial Crimes Enforcement Network United States Department of the Treasury Before the United States Senate Committee on Banking, Housing, and Urban Affairs Subcommittee on National Security and International Trade and Finance Subcommittee on Economic Policy" (PDF). fincen.gov. Financial Crimes Enforcement Network. 19 November 2013. Archived (PDF) from the original on 9 October 2016. Retrieved 1 June 2014.
- [3] Böhme, Rainer, Nicolas Christin, Benjamin Edelman, and Tyler Moore. (2015). "Bitcoin: Economics, Technology, and Governance." *Journal of Economic Perspectives*, 29 (2): 213-38.
- [4] SWOT Analysis: Discover New Opportunities, Manage and Eliminate Threats". www.mindtools.com. 2016. Retrieved 24 February 2018.
- [5] https://en.wikipedia.org/wiki/SWOT_analysis#cite_note-1
- [6] <https://dictionary.cambridge.org/dictionary/english/>
- [7] Sammut-Bonnici, Tanya & Galea, David. (2015). SWOT Analysis. 10.1002/9781118785317.weom120103.
- [8] Satoshi Nakamoto, Bitcoin: A Peer-to-Peer Electronic Cash System, 2008
- [9] Águila, R.D.M., Ramírez, G.A., (2013). Series: basic statistics for busy clinicians. *Allergol Immunopathol.* 42 (5), pp. 485-492.
- [10] Blackmore, K., Bossomaier, T., (2002). Comparison of See5 and J48.PART algorithms for missing persons profiling. *International Conference on Information Technology and Applications*
- [11] Frank E. and Witten I.H. (1998). Generating Accurate Rule Sets Without Global Optimization. In Shavlik, J., ed., *Machine Learning: Proceedings of the Fifteenth International Conference*, Morgan Kaufmann Publishers.
- [12] Frank E. and Witten I.H. (2000). *Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations*. Morgan Kaufmann Publishers: San Francisco, CA.
- [13] Lemeshow S., Hosmer D.W., Klar J. & Lwanga S.K., 1990. *Adequacy of sample size in health studies*. Chichester: John Wiley and Sons.
- [14] Merriam-Webster, (2020). <https://www.merriam-webster.com> [date accessed 9 August 2020]
- [15] Ramchoun, H. r., Idrissi, M. m., Ghanou, Y. y., & Ettaouil, M. m. (2017). New Modeling of Multilayer Perceptron Architecture Optimization with Regularization: An Application to Pattern Classification. *IAENG International Journal of Computer Science*, 44(3), 261-269.
- [16] Rosenblatt, F., & Cornell Aeronautical Laboratory. (1958). *The perceptron: A theory of statistical separability in cognitive systems (Project Para)*. Buffalo, N.Y: Cornell Aeronautical Laboratory.
- [17] Shearer, C., (2000) The CRISP-DM model: the new blueprint for data mining. *Journal of Data Warehousing*, 5, 13-22.
- [18] Simoudis, E. (1996). Reality Check for Data Mining. *IEEE EXPERT*, 11(5), pp.26-33
- [19] Cohen, W. (1995). Fast effective rule induction. In A. Prieditis and S. Russell (eds.), *Proceedings of the 12th International Conference on Machine Learning*, Lake Tahoe, CA, pp.115-123.
- [20] Saravanan, N., Gayathri V., (2018). Performance and classification evaluation of J48 algorithm and Kendall's based J48 algorithm (KNJ48). *International Journal of Computer Trends and Technology*
- [21] Sasaki M., Kita K., (1998). Rule based text categorization using hierarchical categories, *IEEE*
- [22] SurveyMonkey, (2017). <https://www.surveymonkey.com/mp/sample-size-calculator/> [date accessed 28 October 2017]
- [23] Taniguchi M., Haft M., Hollm'en J., and Tresp V. (1998). Fraud detection in communications networks using neural and probabilistic methods. In *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'98)*, Volume II, pp. 1241-1244.
- [24] Venkatesan, E. V., (2015). Performance Analysis of Decision Tree Algorithms for Breast Cancer Classification. *Indian Journal of Science and Technology*.
- [25] Yavuz Ö., (2019), A data mining approach for desire and intention to participate in virtual communities. *International Journal of Electrical and Computer Engineering*, 9(5).
- [26] Ławrynowicz, A., Tresp, V., (2014). *Introducing Machine Learning. Perspectives on Ontology Learning*. AKA Heidelberg /IOS Press.
- [27] Thomas, M., (2012). Root Mean Square Error Compared to, and Contrasted with, Standard Deviation. *Surveying and Land Information Science*, 72.
- [28] Ławrynowicz, A., Tresp, V., (2014). *Introducing Machine Learning. Perspectives on Ontology Learning*. AKA Heidelberg /IOS Press.
- [29] Thomas, M., (2012). Root Mean Square Error Compared to, and Contrasted with, Standard Deviation. *Surveying and Land Information Science*, 72.
- [30] Karahoca D., Karahoca A., Yavuz Ö., (2013). An early warning system approach for the identification of currency crises with data mining techniques. *Neural Computing and Applications*, 23(7-8)
- [31] Rasmussen, C. E.; Williams, C. K. I. *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*; The MIT Press: 2005.
- [32] <http://old.opentox.org/dev/documentation/components/gaussianregressions>
- [33] https://en.wikipedia.org/wiki/Gaussian_process
- [34] Anil Rajput, (2011) J48 and JRIP Rules for E-Governance Data, *International Journal of Computer Science and Security (IJCSS)*, 5(2)
- [35] <https://plato.stanford.edu/entries/aristotle-logic/>
- [36] Dr. E. O. Aruma, (2017), Abraham Maslow's Hierarchy Of Needs And Assessment Of Needs In Community Development, 5(7), *International Journal of Development and Economic Sustainability*

- [37]Kaminsky GL, Reinhart CM (1996) The twin crises: the causes of banking and balance-of-payments problems. Board of Governors Federal Reserve System, Washington, DC 4.
- [38]Kaminsky G, Lizondo S, Reinhart CM (1998) Leading indicators of currency crisis. Staff Pap Int Monet Fund 45(1):1-48 5. Kaminsky GL,
- [39]Reinhart CM (1999) The twin crises: the causes of banking and balance-of-payments problems. Am Econ Rev 89(3):473-500
- [40]Graham, B., & Dodd, D. L. (1951). Security analysis: Principles and technique. New York: McGraw-Hill.
- [41]MLA. Kotler, Philip. Principles of Marketing. Englewood Cliffs, N.J. :Prentice Hall, 1991.
- [42]MLA. Kotler, Philip. Marketing Management. Upper Saddle River, N.J. :Prentice Hall, 2000.