TRUSTWORTHY ARTIFICIAL INTELLIGENCE FOR BLOCKCHAIN-BASED CRYPTOCURRENCY

Tiffany Zhan

USAOT, Las Vegas, Nevada, USA

ABSTRACT

Blockchain-based cryptocurrency has attracted the immersive attention of individuals and businesses. With distributed ledger technology (DLT) consisting of growing list of record blocks and securely linked together using cryptography, each block contains a cryptographic hash of the previous block, a timestamp, and transaction data. The timestamp proves that the transaction data existed when the block was created. Since each block contains information about the block previous to it, they effectively form a chain, with each additional block linking to the ones before it. Consequently, blockchain transactions are irreversible in that, once they are recorded, the data in any given block cannot be altered retroactively without altering all subsequent blocks. The blockchain-based technologies have been emerging with a fleet speed. In this paper, the trustworthy Artificial Intelligence will be explored for blockchain-based cryptocurrency where the prohibitive price leap creates a challenge for financial analysis and prediction.

KEYWORDS

Trustworthy Artificial Intelligence, Blockchain, Cryptocurrency, Financial Prediction.

1. TRUSTWORTHY ARTIFICIAL INTELLIGENCE

The flotilla development of Artificial Intelligence (AI) technology has enabled numerous applications in the world, including AI in astronomy, AI in healthcare, AI in gaming, AI in data security, AI in social media, AI in travel and transportation, AI in automotive industry, etc. [1, 2] However, many AI systems are vulnerable to indiscernible attacks which degrade people's trust in AI systems [3, 4]. One inspiring question is what does it mean to be trustworthy?

In April 2019, the European Commission's High-Level Expert Group on Artificial Intelligence (AI HLEG) published Ethics Guidelines for Trustworthy AI, stating that human beings will be able to confidently and fully reap the benefits of AI only if they have trust in it [5, 6]. There are three aspects to enforce the trustworthy AI: 1) robust and reliable technology to avoid unintentional damage due to lack of technological mastery; and 2) behave ethically and morally; 3) human in the loop. Four principles including human autonomy, prevention of harm, fairness, and explicability were set.

One definition is that to be trustworthy, an AI system should operate competently, behave ethically and morally, and interact appropriately with humans. Is this sufficient? Based on the definition provided by the Merriam-Webster dictionary, it says that firm belief in the character, strength, or truth of someone or something. Can we have a firm belief for an AI system?

David C. Wyld et al. (Eds): AI, AIMLNET, BIOS, BINLP, CSTY, MaVaS, SIGI - 2022
pp. 63-68, 2022. CS & IT - CSCP 2022 DOI: 10.5121/csit.2022.121806

Holton [7] stated that in order to trust one need not believe. He used an example of a shopkeeper who decides to trust his employee, although the latter has been convicted of petty theft. Holton argued that the shopkeeper can decide to trust the man without believing that he will not steal. He may trust him because he wants to give him moral support, a new chance to earn trust. This sort of trust has been called "therapeutic trust". However, contrary to Holton, I think that a firm belief must be implemented to be trustworthy.

Carsel [8] proposed a social-cognitive theory of trust which states that, in social contexts, trust is a fundamental element of relationships. Without trust, people may not be able to pursue valued interdependent goals [9] and meet relationship satisfaction [10]. To understand trust formally, a coherent theoretical framework is needed [11,12]. The question that typically guides psychological research and theory in trust is "Does Human A trust AI B?" An important implication is that current paradigms overlook the possibility that Human A might trust AI B differently across various contexts. For example, imagine AI B achieved a great job for Human A in Task X, Human A may trust AI in the context of Task X. However, AI performs poorly in Task Y due to data adversarial attacks. Consequently, Human A may not trust AI B in Task Y. Such possibilities raise the challenges to scientific community. One of important focus in trust community is to predict whether or not Person (Human) A will trust Person (AI) B, such as generalized anxiety [13], attachment style [14], and group membership [15]. However, such an analysis does not consider the social contexts and can only examine average levels of trust between people (human and AI).

By being able to identify when, why, and how human come to trust AI in context, practitioners need to implement policies that facilitate trust between human and AI on how to regain trust that was lost in their relationship or identify levels of (dis)trust in specific contexts that facilitate maladaptive behaviour within those contexts. An alternative question that draws attention to the potential variability in trust between human and AI across contexts is "When does Human A trust AI B?"

When discussing the trust between humans, researchers link the concept of risk [16, 17], ranging from situating trust as simply a subset of risk to locating risk as an antecedent to trust, or the current risks to the individual do indeed affect the individual's trust in others in a trust game. In the person to person trust, the risks to the trustor are not constant across contexts, even if they are similar to past interactions between the trustor and trustee, we should expect interpersonal trust to calibrate to the specific demands of the interaction. How about the trust situation between human and AI?

Instead of examining whether and to what degree an individual trusts others, the focus should be on the potentially varying levels of trust across the various contours within human and AI's relationships. In other words, the motivating question becomes "Human A trusts AI B for what?"
A careful reading of "trust" indicates that we often say "trust" when people share a goal.
I think that we should propose a novel theoretical orientation to trust between human and AI and formulate a contextual theory of trust using a new theoretical lens. Following this idea, I will apply the trustworthy AI in the blockchain-based cryptocurrency systems.

2. BLOCKCHAIN-BASED CRYPTOCURRENCY

A blockchain is a distributed ledger technology (DLT) that consists of growing list of records, called blocks, that are securely linked together using cryptography [18, 19, 20]. Each block contains a cryptographic hash of the previous block, a timestamp, and transaction data (generally represented as a Merkle tree, where data nodes are represented by leafs). The timestamp proves that the transaction data existed when the block was created. Since each block contains

information about the block before it, they effectively form a chain (compare linked list data structure), with each additional block linking to the ones before it. Consequently, blockchain transactions are irreversible in that, once they are recorded, the data in any given block cannot be altered retroactively without altering all subsequent blocks.

Cryptographer David Chaum first proposed a blockchain-like protocol in his 1982 dissertation [21] "Computer Systems Established, Maintained, and Trusted by Mutually Suspicious Groups." Further work on a cryptographically secured chain of blocks was described in 1991 by Stuart Haber and W. Scott Stornetta [22, 23]. A blockchain was created by a person (or group of people) using the name (or pseudonym) Satoshi Nakamoto in 2008 [24] to serve as the public distributed ledger for bitcoin cryptocurrency transactions, based on previous work by Stuart Haber, W. Scott Stornetta, and Dave Bayer [22, 23] The identity of Satoshi Nakamoto remains unknown to date. The implementation of the blockchain within bitcoin made it the first digital currency to solve the double-spending problem without the need of a trusted authority or central server. The bitcoin design has inspired other applications and blockchains that are readable by the public and are widely used by cryptocurrencies. The blockchain may be considered a type of payment rail [25, 26, 27].

The Bitcoin protocol is the consensus mechanism that allows users to send and receive a digital like currency called Bitcoin. Because the transfer of Bitcoin requires intensive use of cryptography, Bitcoin is referred to as a cryptocurrency. Just in this year, Bitcoin price rose to near $65k, then fell $29k the next. The current price as of today is $19,000. Additionally, the volatility of Bitcoin and other cryptocurrency is highly compared to traditional stock and indexes. Normal stock prediction is a non-trivial task, but to add extreme volatility and parameters that are internal only to Blockchain, a question of whether an AI algorithm such as a Deep Neural Network (DNN) can learn the behaviour of Bitcoin is a scorching topic in Cryptocurrency. Previous studies have shown a deep neural network is no better than traditional statistical methods [28, 29, 30].

3. CRYPTOCURRENCY PREDICTION

Stock trend prediction is challenging because there are many factors can influence the price [33]. The factors may be internal or external (or both) events to the given company. The events may not visible before it occurs. The problem of predicting Bitcoin prices is even more challenging in that Bitcoin prices may not adhere to outside business influence and government, but only on limit of coins [34, 35]. With increasing business using cryptocurrency, the United States Exchange Commission (USEC) has enforced rules to regulate cryptocurrency. To predict cryptocurrency prices, linear regression, support vector machine, logistic regression, and time series analysis have been used [36, 37, 38]. Among them, linear regression provides decent results. In this study, I would like to use deep neural network learning.

4. DATA DESCRIPTION

The data are obtained via blockchain.com (https://www.blockchain.com/charts) for blockchain data, bitcoinnity.org (https://data.bitcoinity.org/markets/price) for bitcoin data, and Yahoo Finance (www.yahoofinance.com) for indexes data. The datasets are collected in CSV format. The data range from September 2011 to September 2022.

5. ALGORITHIMC AND EXPERIMENTAL IMPLEMENTATION

In these experiments, I used the Scikit Learn Library and trained different neural network models on the datasets, including 1-hidden, 2 hidden, and 3-hidden layer deep neural network (DNN) models. I used 80% of the dataset was used for training and 20% used for testing. The DNN models are trained through a 10-fold cross-validation and 100 epochs. As discussed in class 10-fold cross validation and a high epoch were important to implement to ensure the consistency and validity of the results. I then calculated the RMSE (Root-mean-square deviation) and MAPE (Mean absolute percentage error)

$$\mathrm{RMSD} = \sqrt{\frac{\sum_{i=1}^{N} (x_i - \hat{x}_i)^2}{N}}$$

RMSD = root-mean-square deviation
i = variable i
N = number of non-missing data points
x_i = actual observations time series
$\hat{x}_i$ = estimated time series

$$M = \frac{1}{n} \sum_{t=1}^{n} \left| \frac{A_t - F_t}{A_t} \right|$$

M = mean absolute percentage error
n = number of times the summation iteration happens
A_t = actual value
F_t = forecast value

6. CONCLUSION

Trustworthy AI becomes increasingly critical for the current world applications. The scientific community should explore further about the fundamental concepts on trustworthy AI in various social contexts. This paper explores trustworthy AI for blockchain-based cryptocurrency where the prohibitive price leap creates a challenge financial analysis and prediction. Blockchain is an emerging technology which promises security and true decentralization with cryptocurrency being the first widely used application. In this paper, I used deep neural networks trained with blockchain and macroeconomic variables which provides stronger predicting power than linear regression. In the future, I would like to see how a DNN model would predict Bitcoin price with its current trend. Another idea I plan to explore is to model other cryptocurrencies such as Ethereum or Dogecoin. As each blockchain has unique parameters that could affect the coin prices differently. Such studies would contribute sanguinely to the blockchain and cryptocurrency resources.

REFERENCES

[1] Bezboruah, T., & Bora, A. (2020). Artificial intelligence: the technology, challenges and applications. Trans Mach Learn Artif Intell, 8(5), 44-51.
[2] Kaushik, M. (2022). Artificial Intelligence (Ai). In Intelligent System Algorithms and Applications in Science and Technology (pp. 119-133). Apple Academic Press.
[3] Dixon, R. B. L. (2022). Artificial Intelligence Governance: A Comparative Analysis of China, the European Union, and the United States.
[4] Petersen, E., Potdevin, Y., Mohammadi, E., Zidowitz, S., Breyer, S., Nowotka, D., ... & Herzog, C. (2022). Responsible and Regulatory Conform Machine Learning for Medicine: A Survey of Challenges and Solutions. IEEE Access.
[5] The European Commission's Artificial Intelligence Act, Available online at https://hai.stanford.edu/sites/default/files/2021-06/HAI_Issue-Brief_The-European-Commissions-Artificial-Intelligence-Act.pdf
[6] EU Commission Ethics guidelines for trustworthy AI. Available online at https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

[7] Holton, R. (1994). Deciding to trust, coming to believe. Australasian Journal of Philosophy 72, 63–76.
[8] Carsel, T. (2020) In Context We Trust: A Social-Cognitive Theory of Trust, PhD Dissertation in Psychology, University of Illinois at Chicago, 2020.
[9] Rusbult, C. E., & Van Lange, P. A. (2003). Interdependence, interaction, and relationships. Annual Review of Psychology, 54(1), 351-375.
[10] Sanderson, C. A., & Cantor, N. (1997). Creating satisfaction in steady dating relationships: The role of personal goals and situational affordances. Journal of Personality and Social Psychology, 73(6), 1424-1433.
[11] Mayer, R. C., & Davis, J. H. (1999). The effect of the performance appraisal system on trust for management: A field quasi-experiment. Journal of Applied Psychology, 84(1), 123-136. Mayer, R. C.,
[12] Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. Academy of Management Review, 20(3), 709-734.
[13] Kenworthy, J. B., & Jones, J. (2009). The roles of group importance and anxiety in predicting depersonalized ingroup trust. Group Processes & Intergroup Relations, 12(2), 227-239.
[14] Collins, N. L., & Read, S. J. (1990). Adult attachment, working models, and relationship quality in dating couples. Journal of Personality and Social Psychology, 58(4), 644-663.
[15] Brewer, M. B. (1979). In-group bias in the minimal intergroup situation: A cognitive-motivational analysis. Psychological Bulletin, 86(2), 307-324.
[16] Das, T. K., & Teng, B. S. (2004). The risk-based view of trust: A conceptual framework. Journal of Business and Psychology, 19(1), 85-116.
[17] Evans, A. M., & Krueger, J. I. (2011). Elements of trust: Risk and perspective-taking. Journal of Experimental Social Psychology, 47(1), 171-177.
[18] Patra, G. R., & Mohanty, M. N. (2022). Price Prediction of Cryptocurrency Using a Multi-Layer Gated Recurrent Unit Network with Multi Features. Computational Economics, 1-20.
[19] Shaik, A., Laxmi, P. S., Anusree, E., Abbas, S., Rajesh, S., & Teja, A. (2022). Forecast Bitcoin Price Prediction Using Time Series Analysis through Machine Learning. Journal of Algebraic Statistics, 13(3), 2113-2123.
[20] Breaves, A. and Au, B. (2015) Using the Bitcoin Transaction Graph to Predict the Price of Bitcoin, Stanford: CS224, pp. 1–6, 2015.
[21] Sherman, A., Javani, F., Zhang, H., Golaszewski, E. (2019). On the Origins and Variations of Blockchain Technologies. IEEE Security Privacy. 17 (1): 72–77.
[22] Narayanan, A., Bonneau, J., Felten, E., Miller, A. Goldfeder, S. (2016) Bitcoin and cryptocurrency technologies: a comprehensive introduction. Princeton: Princeton University Press.
[23] Haber, S., Stornetta, W. (1991). How to timestamp a digital document. Journal of Cryptology. 3 (2): 99–111.
[24] S. Nakamoto (2008) Bitcoin: A Peer-to-Peer Electronic Cash System, Tech. Rep., 2008.
[25] Bakos, Y., Halaburda, H., Mueller-Bloch, C. (2021). When Permissioned Blockchains Deliver More Decentralization Than Permissionless. Communications of the ACM. 64 (2): 20–22.
[26] Panetta, K. (2018). Digital Business: CIO Agenda 2019: Exploit Transformational Technologies.
[27] Wegner, P. (1996). Interoperability. *ACM Computing Surveys*. 28: 285–287
[28] Sun, C. (2008) Stock Market Returns Predictability: Does Volatility Matter?, Columbia University: QMSS, pp. 1-30, 2008.
[29] Ciaian, P., Rajcaniova, M. and Kancs, D. (2016) The Economics of Bitcoin Price Formation, Appl. Econ., vol. 48, no. 19, pp. 1799–1815, 2016.
[30] McNally, S., Roche, J. and Caton, S. (2018) Predicting the Price of Bitcoin Using Machine Learning, 2018 26th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP), 2018, pp. 339-343, doi: 10.1109/PDP2018.2018.00060.
[31] Gujarati, D. and Porter, D. Essentials of Econometrics. New York, NY, USA: McGraw-Hill, 1999.
[32] Unknown, S. (2015, September 15). Variance Inflation Factor. https://www.statisticshowto.datasciencecentral.com/variance-inflationfactor
[33] Berdik, D., Otoum, S., Schmidt, N., Porter, D., Jararweh, Y. A Survey on Blockchain for Information Systems Management and Security, Information Processing & Management, Volume 58, Issue 1, 2021, 102397, ISSN 0306-4573
[34] Akaike, H. (1969) Fitting autoregressive models for prediction. Ann Inst Stat Math 21:243–247

[35] Al-Khazali O, Bouri E, Roubaud D (2018) The impact of positive and negative macroeconomic news surprises: gold versus Bitcoin. Econ Bull, AccessEcon 38:373–382
[36] Almeida, J., Tata, S., Moser, A. and Smit, V. (2015) Bitcoin Prediction Using ANN, Neural Networks, Group 7, pp. 1–12, June 2015.
[37] Sun, C. (2008) Stock Market Returns Predictability: Does Volatility Matter? Columbia University: QMSS, pp. 1-30, 2008.
[38] Ciaian, P., Rajcaniova, M. and Kancs, D. (2016) The Economics of Bitcoin Price Formation," Appl. Econ., vol. 48, no. 19, pp. 1799–1815, 2016.