Smart Stock Exchange Market: A Secure Predictive decentralized Model

Gaurang Bansal[1], Vikas Hassija[1] Vinay Chamola[1], Neeraj Kumar[2], *Senior Member, IEEE*
Mohsen Guizani[3], *Fellow, IEEE*
[1]Department of Electrical and Electronics Engineering, BITS Pilani, Pilani Campus, India
[2]Department of Computer Science and Engineering, Thapar University, Patiala, Punjab, India
[3]CSE Department, Qatar University, Qatar

***Abstract*—Stock exchanges around the world are exploring the best possible solution that can improve trading efficiency, lower the risks and tighten security levels. The working and functioning of a stock exchange involves very hectic and cumbersome procedures which are time consuming, cost inefficient and can be prone to numerous risks. Machine learning and Blockchain are most popular upcoming technologies. In this paper we present a novel secure and decentralized intelligent stock market prediction model. We present a blockchain based solution for stock exchange model that uses machine learning accessible smart contracts. The machine learning model makes a prediction on the future of the stock market providing an intelligent solution for secure stock market.**

***Index Terms*—Blockchain, Commerce, Stock Exchange, Machine Learning, FinTech**

I. Introduction

The origins of stocks and stock markets date back to the 11th century, where businessmen traded debts on brokerage. It picked up momentum in the 13th century when merchants from Venice started trading government securities [1]. Since then, stock trade has evolved and changed drastically. Today, the largest of economies revolve around stock exchanges where everyday assets worth millions are traded. A stock exchange in essence is a central authority that mediates between brokers, stock traders and investors in trading securities in form of stocks, bonds and other financial instruments [2]. In addition, traditional stock exchanges act as facilities for issuing and redeeming such financial instruments including the payment of dividends [3], [4].

In the recent decades the stock market has moved away from over the counter trading(though it still exists) and has become excessively dependent on computerized systems to handle the huge number of transactions effectively and within the constraints of time. Even though these computerized systems claim to be secure and a lot of focus is on ensuring the authenticity of the transaction, these systems are still highly susceptible to tampering as they serve as central and singular points of failure, which if compromised can lead to disaster [5]. To ensure this does not happen stock exchanges spend lavishly to maintain and secure their operations, which in turn lead to high transaction fees borne by the buyers and sellers in the market [6].

The traditional stock markets are strewn with issues of centralized architectures. In such markets central components

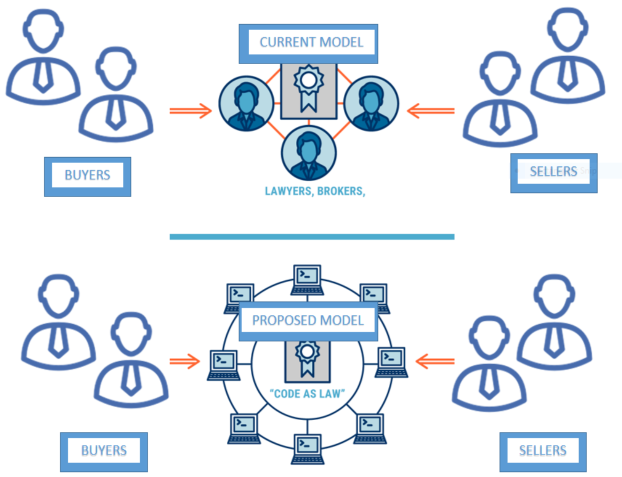

Fig. 1: Current Vs Proposed Model

gather all the data and register them at one central server. In such a scenario where the network over which transactions are happening is owned by a central third party, the network can become easy target for attackers and since the server is a single point of failure, this system is susceptible to even more damage [7]. Also, the centralized system provides for the authority to control and set the prices for processing transactions and other fees. Furthermore, the existing system is heavily dependent on intermediaries also called brokers in the market lingo who charge a fees for committing transactions for the client. This fee and unnecessary involvement can be removed through the use of our proposed system. Another major issue with the traditional system are the long processing times and long settlement delays that essentially destroy the dynamic nature of the stock markets and a completely electronic system can significantly reduce this waiting time and we may achieve the near real time stock settlement. Fig. 1 presents the difference in approach of current vs proposed model.

Stock markets across the globe are increasingly embracing blockchains native capabilities as the basis for market transactions. The likely outcome is that blockchain technology will primarily work within the existing infrastructure or ecosystem to help to restructure and simplify many existing processes

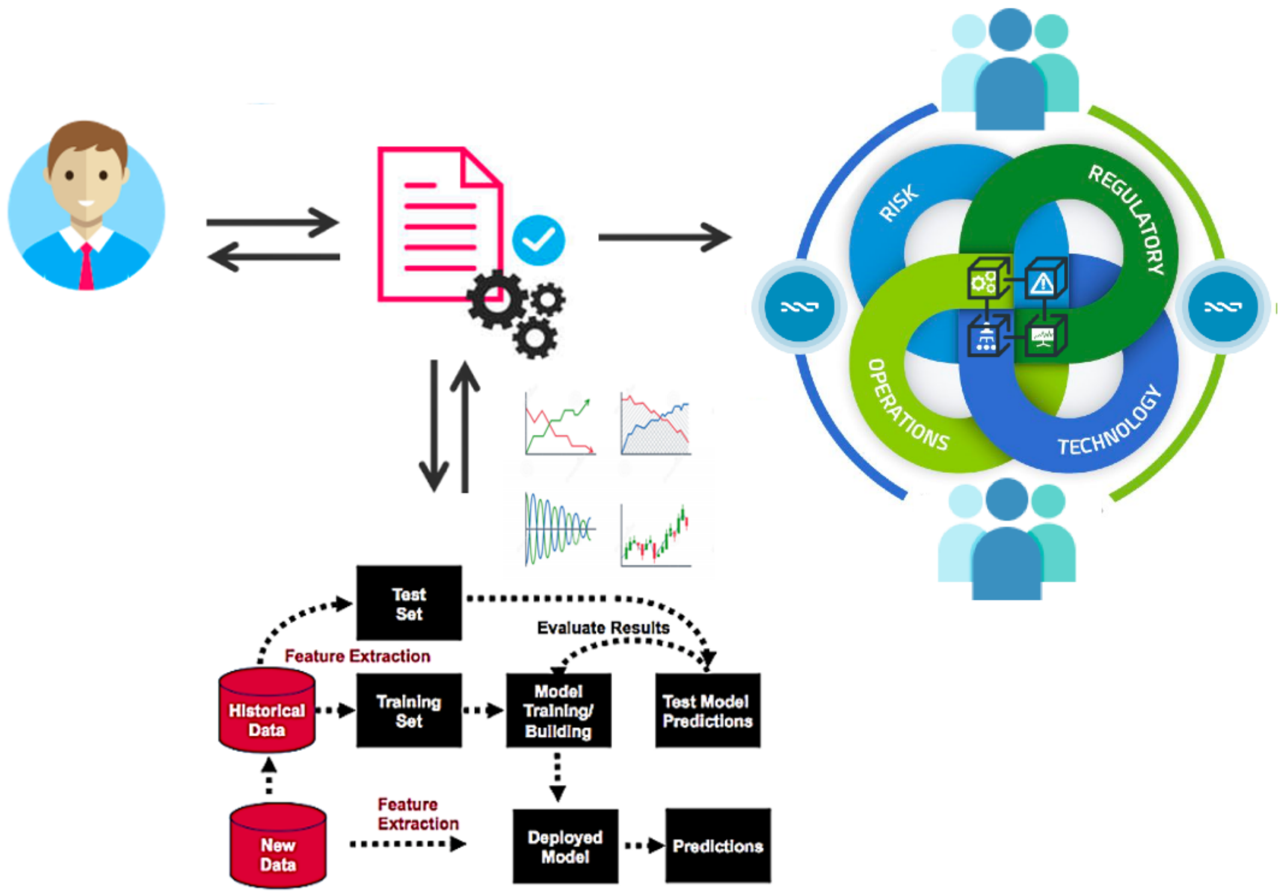

Fig. 2: When user makes a transaction, it invokes the smart contract. Smart contract takes values from user and makes use of machine learning model to predict the future of stocks in which user is investing and gives its prediction to user. User can choose to continue with transaction or change. Finally the transaction is added to blockchain

and strip out significant inefficiencies associated with reconciliation. With the advent of Blockchain and its potential to disrupt long established centralized systems, the tech industry is set for a new revolution. Blockchain has already shown a deep impact on how we transact and keep record of money and finances. The architecture of proposed model is presented in Fig. 2.

Today in the stock market prediction, models are being trained to have better and better predictive analysis of variation of stock market. Every investor is looking for more and more profit, and there is a dire need to have an intelligent model to predict the outcome of investment in a particular company. In order to make the system more future compatible, we propose a machine learning model on distributed ledger that can effectively predict the future stock market.

A. *Organization*

Rest of the paper is organized as follows. Section II discusses the related work. In Section III we give an introduction to blockchain model in stock market. Understanding of blockchain will be helpful in understanding the tangle model described in section IV. Section V discusses the machine learning modelling. Section VI presents simulation and results and finally, conclusion in presented in section VII.

II. Related Work

The potential of blockchain to simplify stock trading in a global market is the focus of today's community. The blockchain technology is an upcoming field and promises solutions that will change the way we interact with the internet and the web. Although still a topic of research and a long way from being perfect, blockchain technologies are already being employed by lots of companies and start ups for ground breaking ventures [3], [8].

More than 100 blockchain projects in 2016 alone, were identified by Moodys Investors Service [9]. From till then until now many projects have started on blockchain. Institutions dedicated to the stock market have started dabbling with the cryptocurrency markets as well. Banking heavyweight Morgan Stanley is estimated to have invested 2 billion dollars into cryptocurrencies from hedge funds over the last year. Trader View [10], a social community providing tools for stock traders, has listed many major cryptocurrency pairs on its site. And Robinhood [11], an app that introduced stock trading to a new generation of investors, announced plans to collaborate with ethereum and bitcoin on its platform.

Nasdaq Stock Market, the second-largest stock exchange in the world is also working on blockchain. Recently, it made an agreement with blockchain start ups for testing the sharing trading in its private network. The information of shares was kept secret. This was a test to see if blockchain is suitable for trading or not [12]. Blocko Inc, a blockchain startup in Korea (one of the most famous one's) has mission of making the enterprise based on blockchain. It handles all the complexities in implementation of blockchain technology in stock market companies [13]. However, many analysts believe that blockchain may provide a decentralized solution, but is it capable to handle the rapid market dynamics. With the issues of scalability, proof of work, we present a solution using Directed Acylic Graphs (DAGs). Stock market prediction tasks are difficult and cumbersome due to time series degree of variation. The variations are very subtle and can change whole dynamics of network. Jegadeesh et al. [14] studied how different predictive signals influence the market. After using handcrafted feature there is paradigm shift to neural networks. Neural networks have a great capability in learning patterns and subtle variations. Krauss et al. [15] employed deep learning and random forest for dynamic stock market prediction.

Although there have been some works on prediction of stock market, but none of the system models can be seen as a future proof solution. We in here present a decentralized distributed predictive ledger solution that is secure and is convergence of machine learning paradigm and blockchain mechanism. The major contributions of this paper are as following:

- In this paper we present an intelligent distributed and decentralized stock market based on blockchain platform that uses LSTM (Long Short-Term Memory) to predict future stocks.
- A predictive model for stock market prediction is designed achieving an accuracy of 99.71% on test dataset on New York Stock Exchange data (the world's largest stock exchange).

III. Blockchain Components

With need of decentralized consensus mechanism, there is great hype with emerging popularity of blockchain. Blockchain is an emerging completely distributed peer to peer technology for secure & consensus data sharing among network nodes without trusted third party. Blockchain can be seen as immutable ledger, which stores all the transactions that have been successfully executed and verified. It also provides a consensus mechanism where all nodes reach to same common decision. The following phases are common to any blockchain.

A. Transactions

The stock value information and digital asset records for each node that is used in trading forms a transaction. A valid transaction must have complete trade information such as amount/stocks, buyer or seller, other party information & timestamp. Each node has a valid digital signature which is private to the entity. The information is encrypted and signed with digital signatures to guarantee authenticity and accuracy. All the other nodes can verify the signature but can't forge it.

B. Validation Phase

All raw data of stock transactions are combined and shared among all authorized validating nodes. Depending on the architecture chosen for blockchain the verifying nodes may vary. In public blockchain allows any node is allowed be the mining node or verifying node. While consortium blockchain uses selected authorized nodes, having sufficient computational ability and memory resources to be the miners.

C. Proof-of-Work

A Proof-of-Work (PoW) is based on the fact that work must be feasibly hard to compute but easy to verify. It also provides protection against spam or DoS attacks where every entity is forced to do some computational task. Before a new block of transactions is inserted into block chain list, PoW is carried out for consensus mechanism. Each mining node competes to validate the block and validating node is rewarded as an incentive. If more that 51% of members agree to mining entity, the block is inserted to blockchain and considered immutable.

IV. Proposed Method

In this section we present our model using blockchain technology. The model uses smart contract. Smart contracts are a computer code running on platform containing a set of rules under which the users to that smart contract agree to work with each other. If and whenever the rules of the agreement are satisfied, the deal is processed further. The smart contract verifies and executes the negotiation of the agreement (transaction).

Whenever a buyer or seller wants to buy or sell his stocks, he initiates a transaction by calling smart contract. Company name and amount of stocks that needs to be transferred need to be specified in the transaction. Smart contract accesses Intelligent agent which is separate entity from Smart contact and can be accessed only through smart contract. Using the distributed ledger it trains the predictive model, and makes his prediction for future. The future prediction are given to the agent, where he can choose to go ahead with the transactions or choose to change. Once the transaction is made

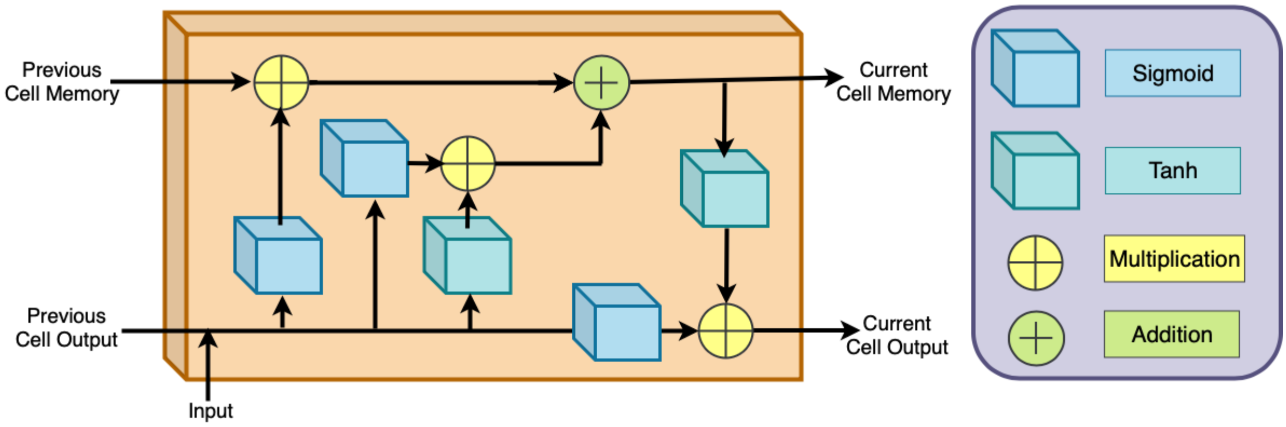

Fig. 3: LSTM architecture

it is validated and entered using the blockchain distributed consensus mechanism.

V. Intelligent Stock Market using Machine Learning

With the growing of artificial intelligence domain, systems are evolving more and more smarter. Current studies on stock exchange market use machine learning models with hand-crafted features. However, manually creating these features is complex and cumbersome task. Some of the solutions proposed such as [16], [17] have come up in predicting the stock market using neural networks, back propagation networks and convolution neural networks (CNN). CNN achieves highest prediction accuracy through Recurrent Neural Networks (RNN). RNN's are very useful in learning the short term dependencies and correlations. However, the problem with stock exchange model is that they can't take long history into account [18].

Sepp Hochreiter and Juergen Schmidhuber to enhance the RNN, proposed Long Short-Term Memory (LSTMs) to solve the vanishing gradient problem faced in RNN's [19]. LSTM's preserves the back propagation error via time and different previous layers. They have much better performance over RNN by remembering the inputs for longer duration of time. LSTM can be seen as a computer memory. Like an ordinary computer it can read, write and delete the information based on the relevance of information. The cell structure is presented in Fig. 3. It is cell which is made of gates. Gates are small decision units with binary decision capability. It decides whether the information should be stored in cell or not. It also assigns an importance factor to every piece of information it sees by giving them weights.

As shown in Fig. 3 LSTM has gates of 3 kinds: input gate (for input), forget gate (delete the information) and output gate (resultant information). These gates like traditional nodes in neural networks can decide whether to block or pass the information. The weights are adjusted similar to that in RNN. These gates are sigmoid and analog with the output range from 0 to 1. This is reason that they can make iterative process of making guesses using back-propagating error.

LSTM is an extension of RNN. Vanilla RNN is an LSTM with all input and output gates as 1, forget gate as 0. In LSTM there are 3 gates: input denoted by i, forget denoted by f and output gate represented by o. W is recurrent connection at the current and previous hidden layer. U denotes weight matrix that connects the inputs to the current hidden layer. C denotes hidden state that is computed based on the current input and the previous hidden state. It defines how previous memory is combined with new input. At gates sigmoid function is used that squashes value of vectors between 0 to 1. Then it multiplies them element by element with another vector of same size. The idea is based on this multiplication only a degree of input is “let through”. Input gate controls the degree to which the newly computed state would be ”let through”. Forget gate works similarly, how much of previous state is “let through”. Output gate defines internal state that is exposed to higher layers.

VI. Simulation & Results

For training the predicting model, we train the LSTM neural network. We use Network Stock Exchange dataset [20] for training and validating the system. The dataset is publicly available on Kaggle platform. It is datset that spans from 2010 to 2016, having the prices for different 501 companies such as Kellogs (K), Aethon Minerals Corp (AET), General Mills (GIS), Ford (F) etc. It includes 140 stock split in times for prices in companies. It also provides information regarding the general description of each company with division on sectors and the metrics extracted from annual SEC 10K fillings (2012-2016). Fig. 4 depicts the opening and closing prices for 4 companies

The LSTM model was trained with input parameters from above 4 parameters in the dataset. Number of neurons in the LSTM model were 200, with 4 neurons as number of output.

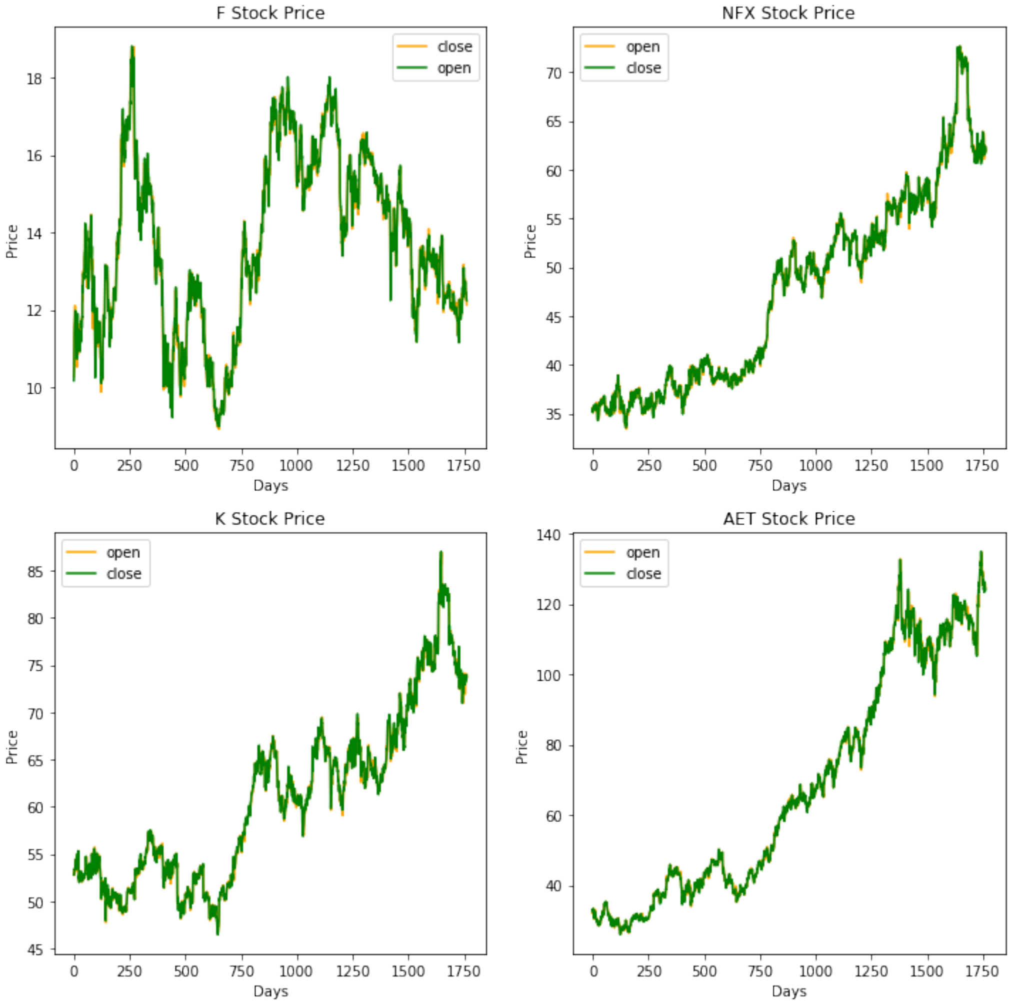

Fig. 4: Stock Price distribution for 4 companies “F”, “NFX”, “K”, and “AET”

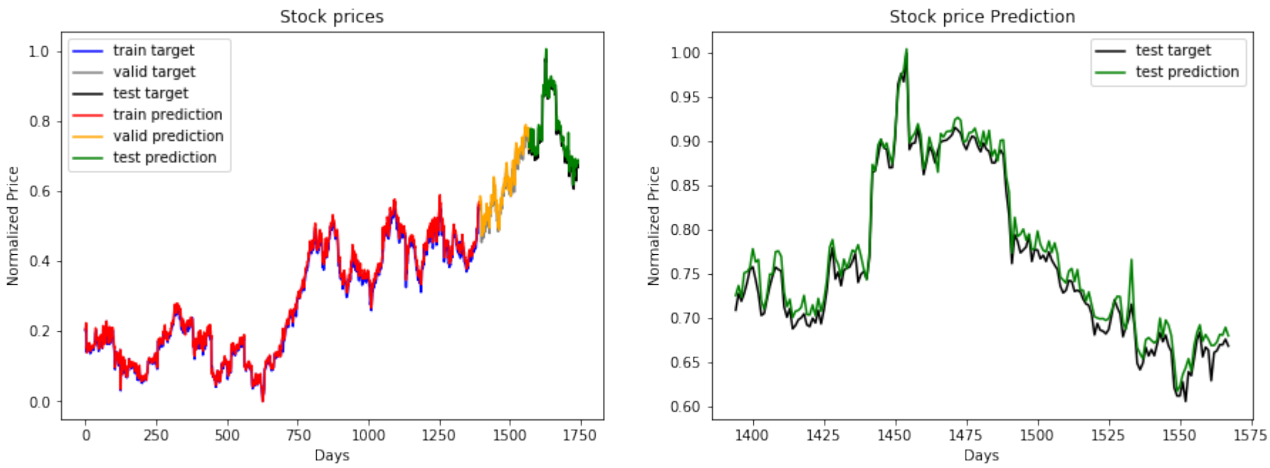

Fig. 5: Stock Market Prediction using LSTM model for stock of company Kellogs (“K”)

TABLE I: Mean square error (MSE) for the various companies

	Training MSE	Test MSE
Aethon Minerals Corp (AET)	0.00007	0.00031
Kellog (K)	0.00006	0.00040
General Mills (GIS)	0.00007	0.00030

Learning rate is set to 0.001. Batch size of 50 was used to train the model for 100 epochs. The feature vector of stock for company contains 4 parameter values i.e. 'open', 'close', 'low', 'high'. The data set for each company is distributed into 80% as training data, 10% as validation data and 10% as test data. The total instances for each company in the data set is 1742, where 1394 are training examples, 174 are validation and test samples.

Fig. 5 show the stock market prediction using LSTM model for Kellogs "K". Left figure in Fig. 5 depicts the train data set, train target, validation data set, validation target and test data and test target. A prediction of stock market is made for 200 days. Right figure in Fig. 5 depicts the actual or test target to prediction of model. We achieve an accuracy of 99.71%. Table I shows the test mean square error (MSE) is only 0.0003-0.0004 for all the companies.

VII. Conclusion

In this paper, we have proposed an Intelligent decentralized Stock market model using convergence of machine learning and DAG based cryptocurrency. The current prevalent stock exchange mechanism involves very hectic and cumbersome and susceptible to many risks. Therefore, the need of a distributed mechanism is felt. We in this paper present how blockchain can solve many problems of centralized stock market exchange. The paper also presents machine learning based model which can make the model more intelligent and provide a future proof solution for the stock market.

References

[1] S. H. Kim and S. H. Chun, "Graded forecasting using an array of bipolar predictions: application of probabilistic neural networks to a stock market index," *International Journal of Forecasting*, vol. 14, no. 3, pp. 323–337, 1998.

[2] M. Crosby, P. Pattanayak, S. Verma, V. Kalyanaraman *et al.*, "Blockchain technology: Beyond bitcoin," *Applied Innovation*, vol. 2, no. 6-10, p. 71, 2016.

[3] M. Hansson, "On stock return prediction with lstm networks," 2017.

[4] A. L. C. Lim and W. W. Wai, "Embracing blockchain applications in fundamental analysis for investment management," *Asia Proceedings of Social Sciences*, vol. 2, no. 2, pp. 111–114, 2018.

[5] H.-j. Kim and K.-s. Shin, "A hybrid approach based on neural networks and genetic algorithms for detecting temporal patterns in stock markets," *Applied Soft Computing*, vol. 7, no. 2, pp. 569–576, 2007.

[6] B. C. Florea, "Blockchain and internet of things data provider for smart applications," in *2018 7th Mediterranean Conference on Embedded Computing (MECO)*. IEEE, 2018, pp. 1–4.

[7] M. R. Islam, I. F. Al-Shaikhli, R. B. M. Nor, and V. Varadarajan, "Technical approach in text mining for stock market prediction: A systematic review," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 10, no. 2, pp. 770–777, 2018.

[8] L. Alessandretti, A. ElBahrawy, L. M. Aiello, and A. Baronchelli, "Machine learning the cryptocurrency market," *arXiv preprint arXiv:1805.08550*, 2018.

[9] R. Cantor, "Moodys investors service response to the consultative paper issued by the basel committee on bank supervision a new capital adequacy framework," *Journal of Banking & Finance*, vol. 25, no. 1, pp. 171–185, 2001.

[10] P. Egger, "An econometric view on the estimation of gravity models and the calculation of trade potentials," *World Economy*, vol. 25, no. 2, pp. 297–312, 2002.

[11] N. McAlone, "The no-fee stock trading app robinhood is now officially worth 1.3 billion dollar," *Business Insider Australia*, 2017.

[12] N. Kshetri and J. Voas, "Blockchain-enabled e-voting," *IEEE Software*, vol. 35, no. 4, pp. 95–99, 2018.

[13] K. Salah, M. H. U. Rehman, N. Nizamuddin, and A. Al-Fuqaha, "Blockchain for ai: review and open research challenges," *IEEE Access*, vol. 7, pp. 10 127–10 149, 2019.

[14] N. Jegadeesh and S. Titman, "Returns to buying winners and selling losers: Implications for stock market efficiency," *The Journal of finance*, vol. 48, no. 1, pp. 65–91, 1993.

[15] C. Krauss and J. Stübinger, "Non-linear dependence modelling with bivariate copulas: Statistical arbitrage pairs trading on the s&p 100," *Applied Economics*, vol. 49, no. 52, pp. 5352–5369, 2017.

[16] W. Chen, C. K. Yeo, C. T. Lau, and B. S. Lee, "Leveraging social media news to predict stock index movement using rnn-boost," *Data & Knowledge Engineering*, vol. 118, pp. 14–24, 2018.

[17] A. M. Rather, A. Agarwal, and V. Sastry, "Recurrent neural network and a hybrid model for prediction of stock returns," *Expert Systems with Applications*, vol. 42, no. 6, pp. 3234–3241, 2015.

[18] B.-S. Lin, W.-T. Chu, and C.-M. Wang, "Application of stock analysis using deep learning," in *2018 7th International Congress on Advanced Applied Informatics (IIAI-AAI)*. IEEE, 2019, pp. 612–617.

[19] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: Continual prediction with lstm," 1999.

[20] Kaggle, "New york stock exchange dataset," 2017. [Online]. Available: https://www.kaggle.com/dgawlik/nyse