SEMARAK ILMU PUBLISHING

Journal of Advanced Research in Applied Sciences and Engineering Technology

Journal homepage:
https://semarakilmu.com.my/journals/index.php/applied_sciences_eng_tech/index
ISSN: 2462-1943

Blockchain Based Deep Learning for Sustainable Agricultural Supply Chain Management

Kesava Rao Alla[1], Gunasekar Thangarasu[2,*]

[1] Chancellery, MAHSA University, Jalan SP 2, Bandar Saujana Putra, 42610 Jenjarom, Selangor, Malaysia
[2] Department of Professional Industry Driven Education, MAHSA University, Jalan SP 2, Bandar Saujana Putra, 42610 Jenjarom, Selangor, Malaysia

ARTICLE INFO

Article history:
Received 6 June 2023
Received in revised form 1 March 2024
Accepted 13 April 2024
Available online 22 May 2024

Keywords:
Blockchain; deep learning; sustainable; agriculture; supply chain management

ABSTRACT

Food supply chain (FSC) is an important part of the food supply chain. It is essential to establish food supply chains that are open to the public, accountable for their actions, and available in real time. The development of blockchain technology has led to an increase in the amount of data that is passed between customers and businesses, as well as the data that are passed between them. Blockchain technology is a new type of information technology that has the ability to be decentralized, safe, and trusted, which makes it an excellent option for storing sensitive data. The purpose of this study is to evaluate how blockchain with deep learning is used to find the quality evaluation of food supply chain technology. The use of BC has improved the accuracy of food traceability, while the utilization of Deep Random Forest (DRF) has boosted the efficacy of computing and shortened the reaction time. The research compares the quality evaluation system that is based on BC-DRF with some of the most common existing methods in terms of accuracy, reaction time, and sensitivity, when applied to a variety of block sizes.

1. Introduction

Instead of depending on a centralized server or any other authoritative entity, the management of the blockchain data is handled by a decentralized network of computers. Over the course of the past few months, the blockchain technology has been integrated into a variety of new applications [1]. If two parties to an agreement cannot be trusted, for instance, the conditions of the agreement can still be carried out and enforced with the use of a smart contract that is based on blockchain technology. It may reduce the amount of time and effort needed to negotiate the terms of a contract [2].

Another area where blockchain technology could potentially be applied is in the management of supply chains. Walmart has developed a blockchain system by employing Hyper Ledger Fabric. This

* *Corresponding author.*
E-mail address: gunasekar97@gmail.com (Gunasekar Thangarasu)

https://doi.org/10.37934/araset.45.2.142151

is done so that the company can evaluate the standard of the food supply that it receives. There is a significant amount of interest being shown in blockchain technology coming from other sectors [3].

In a world that is becoming increasingly dominated by digital technology, the manual processing of vast amounts of data is inefficient due to the length of time it takes. The application of machine learning as a strategy for the processing of a diverse range of data sources is gaining widespread traction. This is mostly attributable to the fact that it is able to automatically learn from its past mistakes and better itself as a result of doing so. Training supervised learning algorithms with data that has been tagged in some way is possible [4].

The fact that blockchain technology is decentralized, safe, and trusted makes it an excellent option for storing sensitive data. Large volumes of data can be processed and analysed by machine learning systems. Because of this, a substantial number of recent studies have studied the possibility of combining blockchain technology with machine learning in order to create real-time systems that are secure, efficient, and friendly to the environment. For instance, the procedures involved in accounting by employing the technologies of machine learning and blockchain. In addition, the blockchain technology that has been improved by machine learning is utilized in the medical industry [5].

The control over a series of flows, beginning with production and ending with consumption, is required for the management of the supply chain. Getting your hands on the fresh components is the initial stage in the process, and handing off the finished good to the customer is the very last thing to do.

In order for these supply chains to become more effective as a whole, new mechanisms for increased cooperation, resilience, and responsiveness will need to be built. The proliferation of blockchain technology has directly contributed to a increase in the amount of supply chain management in agricultural industry.

The advancement of digital and information technology is causing a commotion in the interactions that take place between customers and businesses, as well as the data that is passed between them. These advancements are due, in large part, to technological innovations such as smartphones, wearables, drones, computers connected to the IoT, machine learning and artificial intelligence. These cutting-edge technologies and algorithms can be helpful to the food supply chain in a variety of different ways, and the ones that have been mentioned here are only few of those ways. One such example is keeping a close eye on the weather and conducting research into the ways in which it can affect the economy of the agricultural sector [6]. It can be used in the construction of systems that track the production and distribution of food at every stage, from the farm to the consumer plate, and it can be put to use in both of those processes.

In the context of the food supply chain, the distributed ledger technology known as blockchain is regarded to be a tool that enables transactions to be honest and reliable. The costs of transactions, the governance of supply chains, and the functionality of resources are all considered to have the potential to drop as a result of blockchain technology. These benefits can be seen in the food supply chain. By utilizing blockchain technology as a data format, it is possible to realize data parity in food supply chains that are centred on blockchain technology. Smart contracts can be used to automate procedures such as the payment of vendors, and the many parties engaged in the food supply chain are able to store and share data securely. Additionally, they are able to accomplish a great deal more. In the areas of law and finance, a smart contract can also perform the functions that are traditionally associated with other types of contracts [7].

A strategic tool that has the ability to affect the reorganization of food supply chains as well as the capacities of food organizations is the distributed ledger technology, which is also known as blockchain. Partners in the supply chain were more dependent not just on one another but also on interactions that were vertically coordinated as a result of the expenses involved with installing blockchain technology at various points along the supply chain. This was a direct result of the costs associated with implementing blockchain technology at various points along the supply chain. It will also help develop alliances and collaboration inside the supply chain, which is a great benefit, so that something to look forward to. In addition to the aforementioned advantages, the use of the blockchain technology will also foster the development of skills that are beneficial to the administration of chain-wide innovation [8-10].

With the assistance of technologies such as blockchain, sensors, the Internet of Things (IoT), the cloud, and artificial intelligence (AI), it is possible to establish a food supply chain that is open to the public, accountable for its actions, and available in real time. Both blockchain technology and the Internet of Things (IoT) have the potential to have an impact on a wide range of domains, including but not limited to the management of food supply chains, automated payments, quality control, and provenance [11-14]. Blockchain technology is a distributed ledger that can be used to record transactions in a decentralized manner [15].

Processing and analysis of massive datasets made available by blockchain and AI applications are necessary for the production, distribution, and consumption of food in a sustainable manner. This is a necessary step in the process of making the entire food supply chain more sustainable. The implementation of blockchain technology within the FSC enables a transparent and trustworthy traceability system, which is vital for the upkeep of public health in the event that food, fruit, or vegetables are suspected to be traced back to their place of origin [16-20].

2. Proposed Methodology

The purpose of this study is to evaluate how many different components of blockchain technology and its potential applications in food supply chain that contributes to the accomplishment of the Sustainable Development Goals (SDGs) as illustrated in Figure 1.

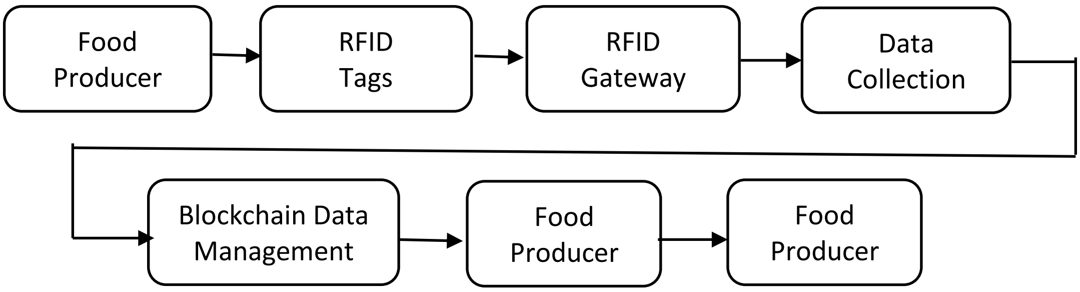


Fig. 1. Proposed BC-SCM

2.1 Data Acquisition

Let pretend you have a dataset A with n food samples, and that each sample has information about the food color, turbidity, fat, odor, taste, temperature, and pH. Let also pretend you have a dataset B with n food samples, and that dataset B also has n food samples. After being subjected to normalization, the dataset can be interpreted in the form of Eq. (1).

$$A = \{A_1, A_2, \dots A_n\} \quad (1)$$

where

A_l = l^{th} sample after the normalization process.

2.2 Data Pre-Processing

The dataset A_l goes through the pre-processing step, which is when the Z-score normalization that was utilized in order to clean the data set is applied. After that, the data set A is ready for analysis. In artificial neural networks, the Z-score is typically utilized as the first stage of the process. It has a significant computational capability and takes into account the mean as well as the variability of the raw data. Additionally, it displays the results.

The research normalizes the data by taking into account the mean as well as the standard deviation of the variables that were input, with the goal of having the mean and the standard deviation of all of the characteristics equal to zero and one, respectively. We do this so that the mean and the standard deviation of all of the characteristics are the same. The Z-score method of normalizing can be defined in Eq. (2):

$$b_l = \frac{A_l - \mu_l}{\sigma_l} \tag{2}$$

Here,

μ_l = arithmetic mean,
σ_l = variance of A_l,
b_l = normalized data.

2.3 Feature Selection

In order to obtain the most value possible from the pre-processed data, the normalized data b_l are put through a feature selection technique. This is done in order to maximize the potential for the data. In this particular scenario, features are retrieved with the assistance of the Canberra distance, which measures the amount of distance that separates a feature and a class label. The distance to Canberra is not overly susceptible to being changed by outlying data points. The following equation can be utilized in order to calculate the distance between the starting point and Canberra mentioned in Eq. (3).

$$CD(g,h) = \sum_{j=1}^{m} \left| \frac{g_j - h_j}{g_j + h_j} \right| \tag{3}$$

where,

g = candidate feature,
h = target,
m = vector dimension.

The chosen features f_s is recorded into the DRN so that they can be considered for determining the overall quality of the food Eq. (4).

$$f_s = \{f_1, f_2, \dots f_k\} \quad (4)$$

2.4 Decision Tree

The decision tree approach (DT) is a type of predictive learning that combines the opinions of subject matter experts regarding the intended outcome of an investigation with actual data pertaining to the phenomena that are the focus of the research. The goal of the inquiry is to determine whether or not a certain phenomenon exists. Learning is the method that is implemented in this technique in order to locate a DT within the dataset.

In the subject of spatial data mining, this methodology is applied rather frequently in research projects. The probable range of values for each target variable is calculated based on the values of the input variables, and the leaves of the decision tree each reflect a different range of possible outcomes for each target variable. One can unearth a DT that had been concealed by concealing it before by separating a source set according on the results of an attribute value test. It is of the utmost importance that this method be followed consistently when deriving each and every subset possible.

When it is possible to assign a single classification to each sample in the subset that was produced by the approach, the recursive operation comes to an end. This can happen either when it is no longer necessary to perform additional splitting or when it is possible to do so. These models are useful for both predictive modeling and classification, and this is due to the fact that the rules that are produced by DTs are simple enough for humans to understand.

The finished DT structure need to have an extremely restricted amount of movable parts. In order for DTs to accurately forecast new data, they need to have the ability to unearth the information that they have gained from the dataset in a manner that is consistent throughout the process. The great majority of DTs are amenable to modification or improvement when new training data are incorporated into the learning process.

There are many subgroups that fall under the umbrella of DTs. A few different DTs each produce output in the form of a separate list of numbers. CARTs allow for the development of results in both the classification and regression domains, and are therefore commonly referred to as classification categorization tree. The CART method was used throughout the duration to analyze the data and produce suggestions for selecting a path of action. The initial node, sometimes known as the root node, was the location where all of the data for the CART model was saved.

The final step involved applying the splitter variable that possesses the highest degree of purity and homogeneity across all of the branches in order to create a split in the root. This was the phase that was responsible for creating the split. At each of the tree branches, a categorization of the variables was carried out until the appropriate node included data that was enough comparable to fit into a single classification. This process would continue until the correct node had all of the data in its storage space.

The leaf nodes provide a representation of the several classes that have developed from the tree. The internal nodes of a network were situated in a location that was midway between the extremes of the network and the network central node. When trying to determine what the splitter variable is, one has access to a number of indices that they can use in their investigation.

In order to carry out the analysis on the data, the Gini index, also known as the GI, was utilized. On the other side, it was discovered that the value of the GI was at its highest when each node held data from each and every possible category. When all of the information that was connected to a specific node belonged to the same group, that node was constructed with the highest possible degree of accuracy, and the GI value that was associated with it was 0.

When all of a node variable had been determined, the GI of each one was computed, and the dataset variable that had the shortest GI was selected to serve as the splitter. This process was repeated until all of the variables of a node had been determined. The following formula was applied so that the required GI value could be determined for CART Eq. (5):

$$GINI(p) = \sum_{i=1}^{n} p_i (1 - p_i) \tag{5}$$

Where

p_i = likelihood of object placed in a category

2.5 Training of RF Using SOM

The two layers that are included in a SOM are referred to as the input layer and the output layer. There is a sizeable portion of neurons found in the output layer that have established complete connections with neurons found in the input layer. A learning algorithm that concurrently employs unsupervised and competitive learning methodologies is used to teach the SOM. This type of learning is referred to as competitive learning. The steps that are involved in SOM training are broken down into their component parts in the following condensed summary. The pattern that is used as input to the SOM can be described in mathematical terms Eq. (6):

$$x = [x_1, x_2, x_M]^T \tag{6}$$

where

x = input pattern,
M = size of x.

As a result, the SWVs of the neurons in the output layer are represented by Eq. (7):

$$w_j = [w_{j1}, w_{j2}, w_{jM}]^T \tag{7}$$

where

w_j, j = 1, 2, …, p, - j^{th} neuron SWVs,
p = Kohonen layer's.

In the initial stage of the training process, the SWVs are provided with a set of values that are relatively arbitrary to serve as a starting point. The initial values for the learning phase of the self-organizing map will be quite arbitrary, and the map will utilize these values as its point of departure. Now we will move on to the stage of the self-organizing map known as learning. Within the context of self-organizing maps, this style of education is referred to as competitive learning. This is because the neurons in the system compete with one another to see which neuron will be triggered and go on to become the winning neuron. The neuron that is stimulated will then go on to become the victorious neuron.

The SWV of each neuron is compared to the input pattern, and the winner is the neuron that exhibits the greatest degree of congruence between the two. After that, the SWV parallels are applied to arrive at a conclusion regarding which neuron proved winning. Because of this, we can determine which neuron emerged victorious by applying the following formula (8 and 9):

$$i(x)=arg\ min\ \|x-w_j\|,\ j=1,2,\ldots,p \tag{8}$$

where

$\|\cdot\|$ = Euclidean distance, and
$i(x)$ = neuron near to x.

$$\|x-wj\| = \sqrt{\sum_{i=1}^{M}\left(x_i - w_{ji}\right)^2},\ j=1,2,\ldots,p \tag{9}$$

The Euclidean distance is often used as a measuring stick by those who are comparing and contrasting two different objects. When the value of $\|x-w_j\|$ is reduced, the degree to which the input x and the SWV w_j correspond to one another increases. Following this step, the distances between the active node x in the network and each of the network output neurons are compared in order to determine which neuron in the network.

The neurons that are successful are those that are situated in the part of the brain that is now exposed to the input pattern that is considered to be the most pertinent. Using a function known as the topological neighborhood function is necessary in order to calculate how the victorious neuron will affect its neighbors. A neighborhood function can be thought of as a representation of a topological map of the region Eq. (10).

$$h_{ji(x)}(t) = \exp\left(-\frac{d_{j,i}^2}{2\sigma^2(t)}\right) \tag{10}$$

3. Results and Discussion

For evaluation, the empirical data is applied over the BC-DRF based food traceability technology. The sensitivity, response time, and accuracy of the approach that is being offered are some of the aspects that are taken into consideration in order to evaluate its effectiveness.

3.1 Experimental Setup

Python is what is utilized in the process of developing the BC-DRF food traceability system. The computer is built with a random-access memory (RAM) capacity of 8 gigabytes, an Intel i3-core processor, and the Windows 10 operating system as its foundation.

3.2 Dataset Description

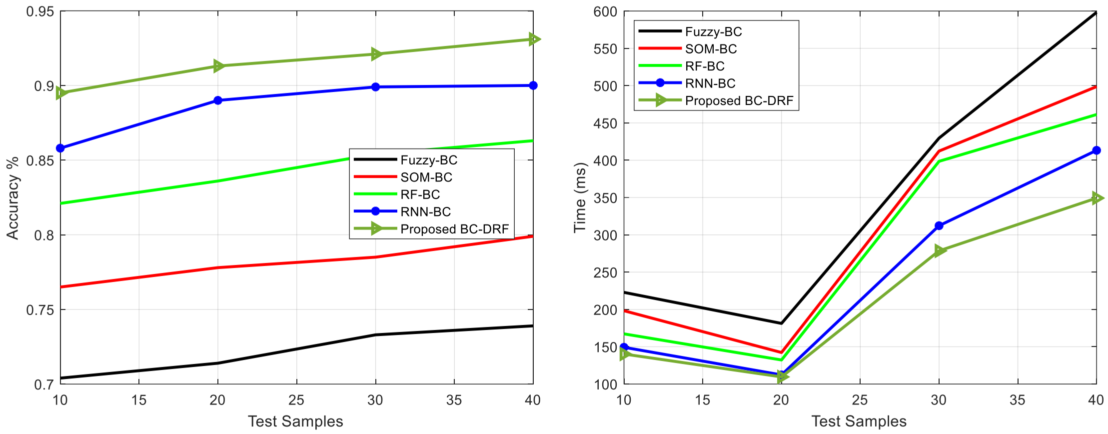
Utilizing the food grading information, the BC-DRF food traceability system is evaluated in order to determine its overall efficacy. The dataset that is used to develop machine learning models is built with the help of physical observations. These models may then be used to make predictions regarding the quality of food. The dataset contains seven separate individual metrics, including viscosity, hue, flavor, fragrance, and acidity.

The existence of these components has a considerable bearing on the accuracy with which one can anticipate the quality and grade of food. On a scale that ranges from one (good) to zero (poor), the quality of the food is graded from one (good) to half (moderate), with one being good and half indicating moderate.

Food receives a score of 0 if it does not meet the optimal standards, where the food receives a grade of 1 if it does meet these requirements. If the food does not meet the optimal requirements, then the food receives a score of 0. The temperature as well as the pH are both included in the dataset.

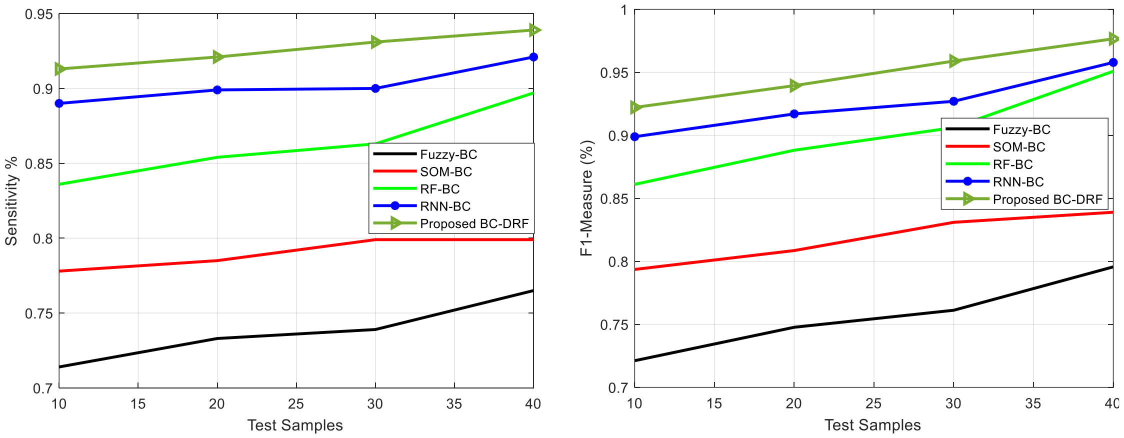
3.3 Comparative Assessment

The recently developed BC-DRF will have its utility evaluated based on how well it compares to the standard procedures used in the industry. In this section, we examined how the newly created BC-DRF stacks up against traditional methods in terms of a variety of metrics, including accuracy, reaction time, sensitivity and F1-measure when applied to a variety of block sizes in Figure 2(a) and Figure 2(b), and Figure 3(a) and Figure 3(b), respectively. The total amount of training data used allows for the examination of this comparison to be carried out.



(a) (b)

Fig. 2. (a) Testing Accuracy, (b) Response time (ms)



(a) (b)

Fig. 3. (a) Sensitivity, (b) F1-Measure

Considering diverse proportions of the training data, the BC-DRF-based food traceability system undergoes a through comparative evaluation, employing a block size of 40. The primary objective of this evaluation to make informed decisions about the efficiency of the procedure. By systematically analyzing its performance across varying data sets, the evaluation aims to provide valuable insights into the system's efficacy, facilitating informed decisions for optimizing and refining the BC-DRD-based food traceability approach.

The evaluation sought to unveil the accuracy of the newly designed BC-DRF in testing scenarios. It aimed to precisely determine the BC-DRF's efficacy. The research on response times utilized the adopted BC-DRF methodology, shedding light on the system's efficiency. These findings contribute crucial insights into the BC-DRF's performance, enriching our understanding of its accuracy and responsiveness in diverse testing conditions.

4. Conclusion

In this paper, we compare the quality evaluation system that is based on BC-DRF with some of the most common existing methods. The system that we use to evaluate the quality of food is based on BC-DRF. When analysing the differences between the various block sizes, some of the criteria that are taken into consideration include testing precision, sensitivity and reaction time. The use of BC has improved the accuracy of food traceability, while the utilization of DRF has boosted the efficacy of computing and shortened the reaction time. The utilization of DRF brings about both of these benefits. The fact that the Canberra distance was used in order to extract the features contributed to an additional increase in the system level of sensitivity.

Acknowledgment
This research was not funded by any grant.

References

[1] Dey, Somdip, Suman Saha, Amit Kumar Singh, and Klaus McDonald-Maier. "SmartNoshWaste: Using blockchain, machine learning, cloud computing and QR code to reduce food waste in decentralized web 3.0 enabled smart cities." *Smart Cities* 5, no. 1 (2022): 162-176. https://doi.org/10.3390/smartcities5010011

[2] Tsoukas, Vasileios, Anargyros Gkogkidis, Aikaterini Kampa, Georgios Spathoulas, and Athanasios Kakarountas. "Enhancing food supply chain security through the use of blockchain and TinyML." *Information* 13, no. 5 (2022): 213. https://doi.org/10.3390/info13050213

[3] Dey, Somdip, Suman Saha, Amit Kumar Singh, and Klaus McDonald-Maier. "SmartNoshWaste: Using blockchain, machine learning, cloud computing and QR code to reduce food waste in decentralized web 3.0 enabled smart cities." *Smart Cities* 5, no. 1 (2022): 162-176. https://doi.org/10.3390/smartcities5010011

[4] Sahoo, Saumyaranjan, Satish Kumar, Uthayasankar Sivarajah, Weng Marc Lim, J. Christopher Westland, and Ashwani Kumar. "Blockchain for sustainable supply chain management: trends and ways forward." *Electronic Commerce Research* (2022): 1-56. https://doi.org/10.1007/s10660-022-09569-1

[5] Chandan, Anulipt, Michele John, and Vidyasagar Potdar. "Achieving UN SDGs in Food Supply Chain Using Blockchain Technology." *Sustainability* 15, no. 3 (2023): 2109. https://doi.org/10.3390/su15032109

[6] Babu, Akhila Susan, and M. Supriya. "Blockchain Based Precision Agriculture Model Using Machine Learning Algorithms." In 2022 *International Conference on Breakthrough in Heuristics and Reciprocation of Advanced Technologies (BHARAT)*, pp. 127-132. IEEE, 2022. https://doi.org/10.1109/BHARAT53139.2022.00036

[7] Saba, Tanzila, Amjad Rehman, Khalid Haseeb, Saeed Ali Bahaj, and Jaime Lloret. "Trust-based decentralized blockchain system with machine learning using Internet of agriculture things." *Computers and Electrical Engineering* 108 (2023): 108674. https://doi.org/10.1016/j.compeleceng.2023.108674

[8] Anitha, R., and Dinesh Rai. "Internet of things with artificial intelligence detection and blockchains of crop availability for supply chain management." *International Journal of Knowledge-Based Development* 12, no. 3-4 (2022): 444-459. https://doi.org/10.1504/IJKBD.2022.128909

[9] Khan, Abdullah Ayub, Zaffar Ahmed Shaikh, Larisa Belinskaja, Laura Baitenova, Yulia Vlasova, Zhanneta Gerzelieva, Asif Ali Laghari, Abdul Ahad Abro, and Sergey Barykin. "A blockchain and metaheuristic-enabled distributed architecture for smart agricultural analysis and ledger preservation solution: A collaborative approach." *Applied Sciences* 12, no. 3 (2022): 1487. https://doi.org/10.3390/app12031487

[10] Mia, Md Uzzal, Mahfuzur Rahman, Ahmed Elbeltagi, Md Abdullah-Al-Mahbub, Gitika Sharma, HM Touhidul Islam, Subodh Chandra Pal et al. "Sustainable flood risk assessment using deep learning-based algorithms with a blockchain technology." *Geocarto International* (2022): 1-29. https://doi.org/10.1080/10106049.2022.2112982

[11] Hannah, S., A. J. Deepa, Varghese S. Chooralil, S. BrillySangeetha, N. Yuvaraj, R. Arshath Raja, C. Suresh, Rahul Vignesh, K. Srihari, and Assefa Alene. "Blockchain-based deep learning to process IoT data acquisition in cognitive data." *BioMed Research International* 2022 (2022). https://doi.org/10.1155/2022/5038851

[12] Neelakandan, S., J. Rene Beulah, L. Prathiba, G. L. N. Murthy, E. Fantin Irudaya Raj, and N. Arulkumar. "Blockchain with deep learning-enabled secure healthcare data transmission and diagnostic model." *International Journal of Modeling, Simulation, and Scientific Computing* 13, no. 04 (2022): 2241006. https://doi.org/10.1142/S1793962322410069

[13] Malik, Hassaan, Tayyaba Anees, Ahmad Naeem, Rizwan Ali Naqvi, and Woong-Kee Loh. "Blockchain-Federated and Deep-Learning-Based Ensembling of Capsule Network with Incremental Extreme Learning Machines for Classification of COVID-19 Using CT Scans." *Bioengineering* 10, no. 2 (2023): 203. https://doi.org/10.3390/bioengineering10020203

[14] Guo, Wei, and Kai Yao. "Supply chain governance of agricultural products under big data platform based on blockchain technology." *Scientific Programming* 2022 (2022): 1-16. https://doi.org/10.1155/2022/4456150

[15] Awan, Sabir, Sheeraz Ahmed, Fasee Ullah, Asif Nawaz, Atif Khan, M. Irfan Uddin, Abdullah Alharbi, Wael Alosaimi, and Hashem Alyami. "IoT with blockchain: A futuristic approach in agriculture and food supply chain." *Wireless Communications and Mobile Computing* 2021 (2021): 1-14. https://doi.org/10.1155/2021/5580179

[16] Frikha, Tarek, Jalel Ktari, Bechir Zalila, Oussama Ghorbel, and Nader Ben Amor. "Integrating blockchain and deep learning for intelligent greenhouse control and traceability." *Alexandria Engineering Journal* 79 (2023): 259-273. https://doi.org/10.1016/j.aej.2023.08.027

[17] Tkachuk, Roman-Valentyn, Dragos Ilie, Remi Robert, Victor Kebande, and Kurt Tutschku. "Towards efficient privacy and trust in decentralized blockchain-based peer-to-peer renewable energy marketplace." *Sustainable Energy, Grids and Networks* 35 (2023): 101146. https://doi.org/10.1016/j.segan.2023.101146

[18] Kazi, Sadia, and Ariyan Jahangir. "Blockchain based Agriculture Using the Application of UAV and Deep Learning Technique: Alexnet CNN." *Malaysian Journal of Science and Advanced Technology* 3, no. 2 (2023): 91-100. https://doi.org/10.56532/mjsat.v3i2.147

[19] Guo, Anting, Wenjiang Huang, Yingying Dong, Huichun Ye, Huiqin Ma, Bo Liu, Wenbin Wu, Yu Ren, Chao Ruan, and Yun Geng. "Wheat yellow rust detection using UAV-based hyperspectral technology." *Remote Sensing* 13, no. 1 (2021): 123. https://doi.org/10.3390/rs13010123

[20] Yuan, Wenan, and Daeun Choi. "UAV-based heating requirement determination for frost management in apple orchard." *Remote Sensing* 13, no. 2 (2021): 273. https://doi.org/10.3390/rs13020273