AgriTrace: Blockchain-based Agricultural Supply Chain with Price Prediction

Divyang Awasthi[1], Aadarsh Raj[2], Harshit Krishna[3], Sneha Dubey[4], Dr. Latha C A[5]

[1]Student, [2]Student, [3]Student, [4]Student, [5]Head of Department
[1]Information Science and Engineering,
[1]RV Institute of Technology and Management, Bangalore, India

Abstract – Agriculture, being a fundamental pillar of global food production and economy, plays a pivotal role in ensuring food security and sustaining livelihoods. However, the agricultural industry faces numerous challenges, including unpredictable supply chain disruptions and misinformation. These challenges often result in financial losses for farmers and hinder the growth of the agricultural sector. In this modern era, access to credible information and efficient transaction mechanisms is vital for agricultural producers to make educated decisions and overcome these challenges. Blockchain technology, with its transparent and immutable nature, has the potential to revolutionize the agricultural supply chain by ensuring data integrity, traceability, and trust among stakeholders. Furthermore, machine learning techniques can provide valuable insights, such as crop price predictions, to support farmers in making informed choices for optimizing their crop production management. The proposed system aims to explore the integration of blockchain technology and machine learning for the agricultural supply chain, providing a transparent and reliable system that empowers farmers with accurate information and tools to enhance their productivity and contribute to the holistic advancement of the agricultural sector.

Index Terms – Agricultural Supply Chain (ASC), Blockchain, Machine Learning (ML), Prediction, Regression

I. INTRODUCTION

Farmers are the backbone of the agricultural industry, responsible for cultivating crops and raising livestock to meet the global demand for food. Agriculture is a cornerstone of the economy, providing food security and livelihoods in rural communities. However, farmers face numerous challenges in today's fast-paced and dynamic agricultural landscape.

Farmers often struggle with supply chain disruptions, such as broken transportation networks or inefficient logistics, which can delay the delivery of agricultural inputs and outputs, affecting their profitability and market competitiveness. Another challenge faced by farmers is the lack of access to reliable information and credible data. Market trends, crop prices, and best practices are continually evolving, and farmers need up-to-date information to make astute judgements about crop selection, planting strategies, and marketing opportunities. However, the availability of accurate and trustworthy information is often limited, leading to misinformation and uncertainty among farmers. Furthermore, the agricultural supply chain is complex and involves multiple stakeholders, including farmers, processors, distributors, and retailers. Absence of transparency and traceability in the supply chain can result in issues such as food fraud, adulteration, and unfair pricing practices, which can impact the farmers' income and erode consumer trust in agricultural products. Blockchain-based technology and machine learning have the potential to transform the agricultural supply chain by offering transparency, traceability, and data-driven insights. The convergence of these technologies can empower farmers with credible information, optimize supply chain efficiency, and facilitate informed decision-making for sustainable agriculture and economic prosperity. Table 1 below shows that most farmers believe gaining more information about the supply chain will benefit them in increasing profits. The data is based on surveys performed by Anwar et al. [1] and Brhane et al. [2].

Table 1: Problems Faced by Farmers

Problem	Percentage of farmers affected
Poor awareness of the new crop	85
Unpredictable change in prices	83
Lack of credit facility	74
Limited availability of local information	72
Lack of transparency in the supply chain	68
Infestation of pests	65
Lack of promptly delivered information	52
Distance to information sources	25

Blockchain technology is a distributed ledger that allows for secure and tamper-proof record-keeping of digital transactions. One of its defining features is that it is a shared and decentralized system, which means that there is no single entity or node that controls the entire system. Instead, each member, also known as a stakeholder, has a copy of the ledger that contains a record of all previous transactions. By using cryptographic techniques, the ledger is made unalterable, ensuring that the data contained in it is secure and trustworthy. Unlike traditional centralized systems, blockchain technology enables transparency and accountability without relying on a central authority or intermediary. Since each activity in the system is visible and auditable by all members, this type of decentralized system creates the foundation of trust.[3] For any transaction to happen, a consensus algorithm is followed. A consensus of nodes that agree upon the issue is required which ensures its validation and authorization. As blockchain technology continues to prove its functionality and gain success in the realm of cryptocurrencies, many organizations and entities are exploring ways to leverage its transparency and fault tolerance to address challenges that arise when multiple untrusted actors are involved in the distribution of a resource. By utilizing blockchain technology, these entities aim to establish a decentralized system that enables secure and transparent transactions, without the need for a central authority to mediate the process. This has the potential to streamline workflows, reduce costs, and increase trust among participants, leading to greater efficiency and innovation in a wide range of industries and applications. Two important, highly relevant areas are agriculture and the food supply chain.[4] Figure 1 shows the fundamental working of blockchain.

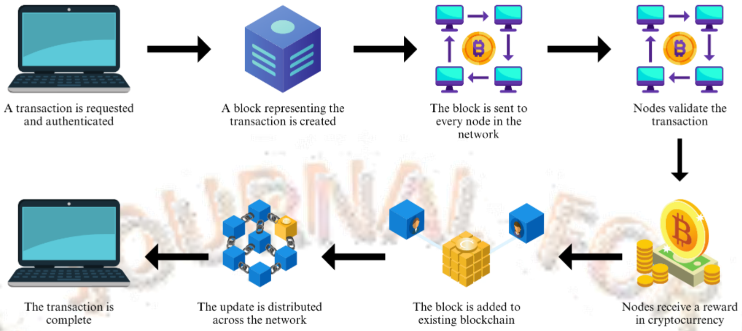

Figure 1: Working of Blockchain

The use of data mining and big data analytics tools has experienced a considerable increase in recent years. Researchers worldwide are utilizing various analysis techniques to address current challenges and forecast future trends. Data mining as a field is effective for discovering patterns in large data sets and ultimately transforming the incoherent mined data into comprehensible knowledge using a cumulation of techniques involving various fields like machine learning, statistics, and database systems.[5] Traditional crop price estimation methods are labour-intensive and lack scalability, whereas automated estimation approaches are cost-effective and highly efficient. Real-time crop production projections may assist farmers by providing high accuracy at a low cost.[6] A forward-looking model for precise crop price assessment, leveraging the advantages of automated estimation over conventional methods is essential for progress.

This pioneering system describes the amalgamation of two complex technologies – blockchain and machine learning – to establish an end-to-end system for the advancement of the agricultural supply chain.

II. LITERATURE SURVEY

Q. Wang et al. [7] mention the reduced security with an amplified rate of block generation which is a major drawback of the contemporary cryptographic algorithms used in the Ethereum blockchain. To mitigate the performance bottleneck, many approaches have been proposed: sharding technique, layer-2 protocols, sidechain technique, heterogeneous structure, hybrid consensus solutions, and assisted techniques such as modifying hard-coded parameters and cross-chain technique.

D. Y. Lin et al. [8] propose a structure used to assess the network of oil stores to determine the ideal request amount, repurchase amount and discount. It has been exhibited that production network contracts can help a food inventory network's income, client trust and sanitation.

H. Fu et al. [9] use a detailed approach based on case studies to explore new ideas in the supply chain management of agri-food industry. The article introduces concepts from new institutional economics, blockchain technology theory, and management fields and applies them to the context of the agri-food supply chain. Owing to the sensitive nature of the research topic and the fledgling status of the companies examined in the case study, the research was limited in its ability to provide company-specific data, including quotation sheets or balance sheets.

S. Khaki et al. [10] proposed the use of an ELM (Extreme Learning Machine) model based on artificial intelligence for predicting coffee yield in small farms. Many machine learning models were compared against the ELM model in their study. To predict agricultural yield, supervised machine learning techniques such as linear regression, polynomial regression, decision Tree, Random Forest, and Support Vector Machine can be used.

Jheng et al. [11] proposed a different model with an addition of the bootstrap method in the steps within the data preprocessing training data. The model being described is a hybrid SVR (Support Vector Regression) and the traditional SVR. The hybrid SVR model performs better in terms of predicting agricultural output with high reliability and high sustainability. The model was evaluated with the Root Mean Square Error (RMSE) and the Correlation Coefficient (CC).

Abdipour et al. [12] evaluated the efficiency of three different models which consisted of an Artificial Neural Network (ANN) with Radial Basis Function (RBF), component analysis and a multi-linear regression model. The aim is to differentiate the performance of yield prediction of crop seed with a substantial effect on components with error analysis consisting of coefficient of determination (R2), Root Mean Square Error (RMSE), and Mean Absolute Error (MAE).

III. Existing System

The agricultural supply chain in India is currently characterized by fragmentation, with multiple intermediaries involved in the marketing channel, often resulting in the transmission of losses to the producer while intermediaries realize profits. Despite the existing system being efficient in reference to traders making profits and customers obtaining goods, there is still a significant amount of produce wastage in India. Approximately 16% of fruits and vegetables, as well as around 10% of oilseeds, pulses, and cereals, are wasted each year [13]. This is primarily because of the lack of communication and knowledge sharing between different tiers of the supply chain. Producers usually have limited knowledge about markets, pricing, and quality control, resulting in inefficiencies. Market research conducted by higher-tier players in the supply chain, who sell directly to consumers, is not shared with lower tiers. As a result, the production and transfer of inventory are inefficient, leading to wastage.

Consider the scenario of perishable goods, where prices and demand vary greatly during sales. During the harvest season, when there is excess produce, prices tend to decline, resulting in wastage despite high demand. Conversely, during the off-season, prices tend to rise, and goods stored in cold chains fetch higher profits. However, due to uninformed decisions made by farmers or lower-tier players who did not engage in cold chain logistics, the demand may not be fully met as there is inadequate inventory available for sale. This issue arises from the lack of information sharing among different players in the supply chain, leading to suboptimal decision-making and resulting in inefficiencies.

The existing model based on Convolutional Neural Network (CNN) faced challenges with agricultural drifts, as they were not consistent with environmental factors such as temperature, weather, and soil condition. To address this, a Backpropagation Neural Network (BPNN) was employed to train the CNN model, which incorporated spatial characteristics as input for error prediction. The newly created model had the advantage of being deployed on a real-time dataset obtained from reliable geospatial resources. However, while the new model reduced relative error, it led to a decrease in forecast efficiency. Similarly, an existing model that combined a time series model with a BPNN and used a smaller dataset size had inferior performance since a smaller sample set was used for prediction. Another neural network engineering model produces good prediction results. The system handles the sensed data and applies the essential action depending on the values of temperature and soil humidity deprived of human intervention. Results obtained through experimentation indicate that random forest models yield superior outcomes comparatively.

IV. Proposed System

The proposed system delves into the utilization of blockchain technology in agricultural supply chains to enhance food traceability. An agricultural supply chain encompasses the various stakeholders engaged in farming, distribution, processing, and marketing of agricultural and horticultural products, spanning from the field to the table. This supply-chain decentralized application is an implementation of a supply-chain management system which uses blockchain to ensure a transparent and secure transfer of products from the manufacturer to the customer.

Figure 2 shows the working of the proposed blockchain mechanism of the application. The lifecycle of a product starts when the farmer makes an entry and farmer details are entered into the blockchain. The current product data is stored with the current owner – the farmer. Now this product shall be available to the wholesaler for purchase. On being purchased, the owner is set to wholesaler and the present data is pushed into the block, which helps to track the origin and handling of the product. Simultaneously, the product is shipped by the farmer and received by the wholesaler, and the wholesaler's details are entered. Each of these checkpoint's data is stored in product history with the state being updated at each step. The online purchase of the product takes place from the wholesaler. When the customer orders the product, it is shipped by the wholesaler and received by the warehouse facility. Here the customer address is stored, the owner is set to the warehouse, details of the warehouse are fed and the current data state gets pushed to the block. Finally, the product is shipped by the warehouse and received by the customer where the final state gets pushed to the block. All these juncture functions shall be executed only after complete verification of product and product history while entering a checkpoint.

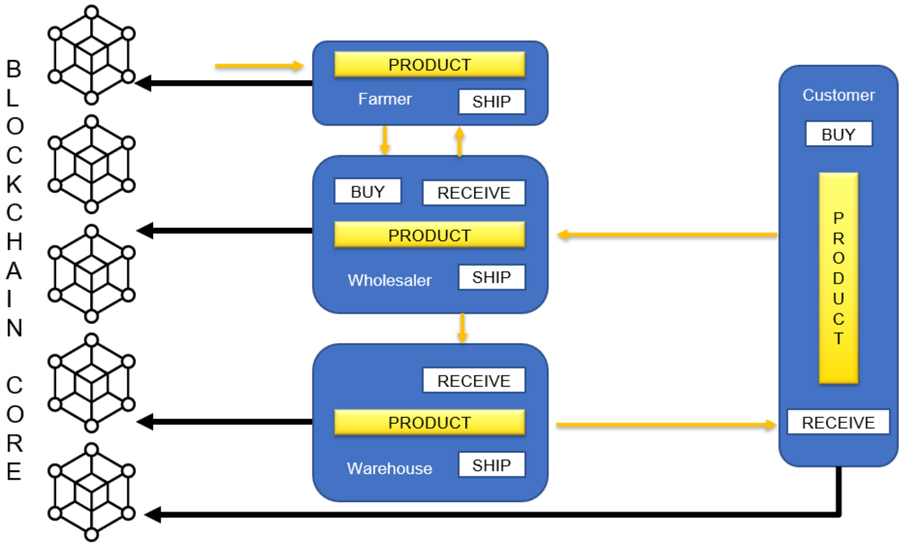

Figure 2: Blockchain Mechanism

The focus of the proposed system is on utilizing a machine learning technique, specifically the support vector regression algorithm, to predict crop prices. Regression, as a data mining technique, is used to learn patterns from the training dataset to determine the price of the crop. In this case, the regression task is treated as a classification task with defined class labels. Input values for parameters such as yield, rainfall, minimum support price, and wholesale price index are provided by the user and fed into the algorithm. The algorithm also factors in other parameters such as probability, new record input, and the number of parameters in dataset. The proposed system highlights the inclusion of various mechanisms to forecast predicted data for impoverished farmers. This involves utilizing Flask libraries from Python and employing linear regression to continuously forecast crop prices for users.

Table 2 compares the accuracy percentage of different supervised machine learning algorithms. The Random Forest (RF) algorithm showed superior accuracy comparatively. [14]

Table 2: Comparing Machine Learning Algorithms

Supervised Machine Learning Algorithm	Accuracy
Artificial Neural Network (ANN)	30%
Decision Tree (DT)	43%
K-Nearest Neighbour (KNN)	31%
Logistic Regression (LR)	25%
Naïve Bayes (NB)	30%

V. IMPLEMENTATION

The implementation involves the utilization of Smart Contracts for secure transaction execution across the supply chain. Solidity is used to code the smart contract, which is then compiled, migrated, and deployed through Truffle.js onto a local blockchain network set up via Ganache. To ensure smooth management of component and state lifecycle, the front-end is built using the React.js framework and interfaces with the smart contract and local blockchain network via Web3.js. User requests are directed to the front-end through dynamic routing using Express.js, with Nginx acting as a load balancer. Figure 3 shows the architecture of the blockchain-based application.

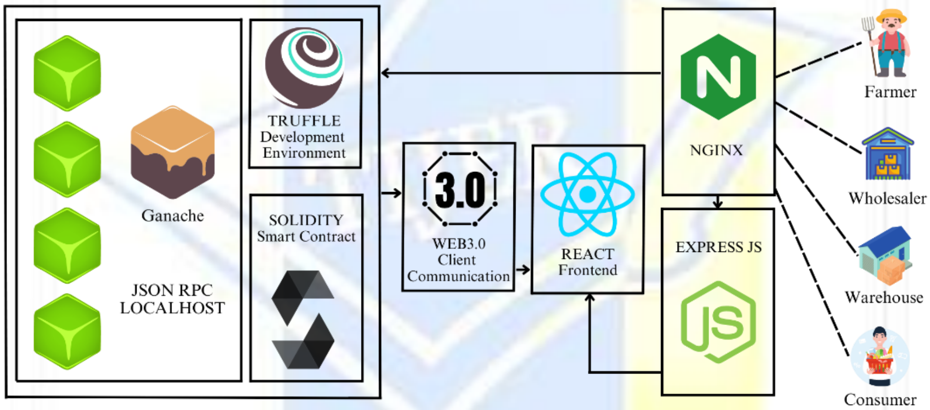

Figure 3: Architecture

Ganache is a blockchain specifically designed for Ethereum development. It is aimed at providing a personalized environment for developers to simulate the behaviour of smart contracts and decentralized applications (dApps). Created by Truffle Suite, a renowned Ethereum development framework, Ganache can be accessed through a desktop application or command-line tool and enables the creation and management of private Ethereum networks. It also includes a user-friendly graphical interface for monitoring transactions and visualizing the blockchain. Typically, Ganache is utilized by developers to test and debug smart contracts before deploying them onto the main Ethereum network.

Truffle is a framework used for Ethereum development which provides developers with a suite of tools for building, testing, and deploying dApps and smart contracts on the Ethereum network. It offers various features including an environment for development, testing, asset pipeline and deployment tools. Truffle streamlines the development process by automating contract compilation and migration and providing debugging tools.

Solidity is a language for writing smart contracts on the Ethereum blockchain. It is high-level and contract-oriented language, designed to target the Ethereum Virtual Machine (EVM). Solidity is like JavaScript with respect to syntax, and it includes features such as inheritance, libraries, and complex user-defined types. With Solidity, developers can write contracts to automate the execution of business logic and facilitate secure and transparent transactions on the Ethereum network.

Web3 (or Web3.js) is a library in JavaScript that offers developers a collection of modules for interacting with the Ethereum blockchain. It is the standard library for Ethereum development and allows developers to create and interact with decentralized applications (dApps) from within web browsers or Node.js environments. Web3.js provides a range of functions for querying and manipulating data on the blockchain, including sending and receiving transactions, querying contract data, and interacting with Decentralized Exchanges (DEXs).

Another JavaScript open-source library for web applications, React allows developers to build complex user interfaces through the use of reusable components. These components are designed in a way that they can be implemented across different parts of an application, and they can seamlessly unify with other libraries or frameworks. React uses a "virtual DOM" to improve performance by minimize the DOM updates required when the state of a component changes. This makes React a fast and efficient library for building large-scale applications.

The application uses Python (3.0 or above), Flask, Scikit-Learn Library, MaterializeCSS and Chart.js to actualize the prediction of crop prices for the farmers. Flask is a Python-based framework that offers a straightforward and user-friendly interface for building web applications. Scikit-Learn is a library that is widely used for data science tasks, providing various algorithms for classification, regression, clustering, and dimensionality reduction. A development framework based on Google's Material Design principles, MaterializeCSS comes with pre-built components and styles to create responsive user interfaces. Chart.js is a JavaScript library that enables the creation of interactive and responsive data visualizations on the web, including various chart types and customization options. These tools are popular among developers as they offer flexibility and versatility to build powerful and efficient applications that meet user requirements. By leveraging these technologies, developers can design visually-pleasing, fast, and functional applications. The process flow of the suggested model is displayed in Figure 4.

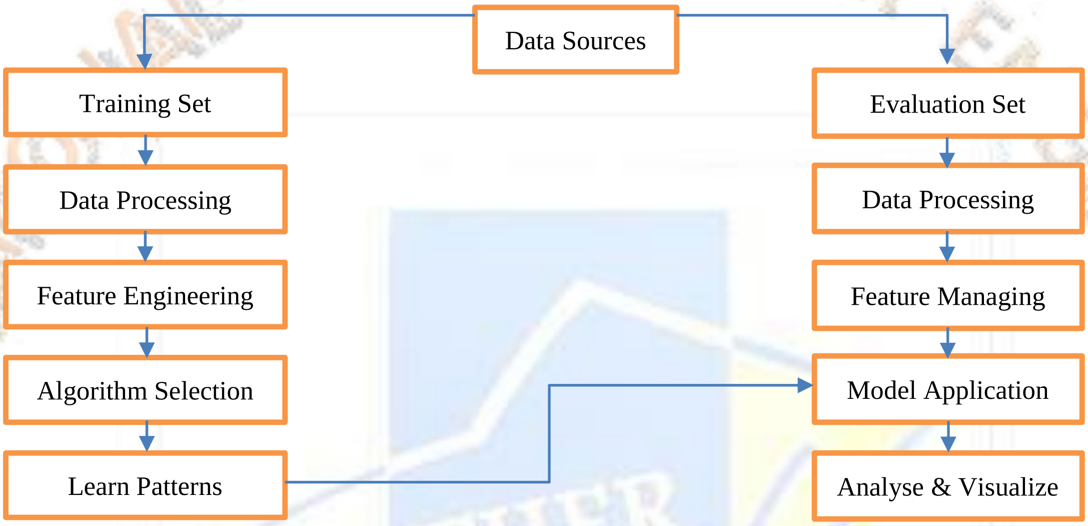

Figure 4: Workflow of Model

VI. ALGORITHMS

This agricultural supply chain tracking system utilizes Keccak-256, which is a cryptographic function belonging to the Solidity (SHA-3 Family). This function calculates the hash of an input to generate a fixed-length output, and it operates in a one-way manner, meaning it cannot be reversed. Ethereum incorporates Keccak-256 in its consensus engine, Ethash, and it follows a simple sponge construction as depicted in Figure 5, employing the innovative Keccak-f cryptographic permutation internally. Keccak has been standardized in various standards, including 3GPP TS 35.231 for mobile telephony and NIST standards FIPS 202 and SP 800-185, after being selected as the top performer of the SHA-3 competition.[15]

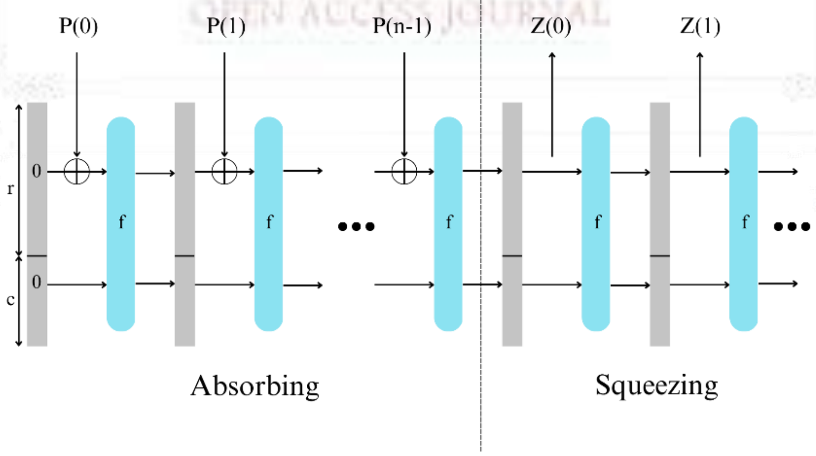

Figure 5: Construction of Keccak Algorithm

Decision Tree algorithm is used in this application to make predictions regarding crop prices based on previous years' data. Decision tree regression is a machine learning technique that examines the features of an object and constructs a tree-based model to predict future data and generate continuous output. Continuous output refers to output that is not limited to a predefined set of numbers or values.

The algorithm considers the following inputs:

- The input parameter
- The training dataset
- Formula used for prediction:

$$SSE = \sum_{i \in s_1} (y_i - y_1) + \sum_{i \in s_2} (y_i - y_2)$$ **Equation 1: Prediction Formula**

In Equation 1, the values of the dependent variable (y_1 and y_2) in groups s_1 and s_2, respectively, represent the wholesale price index parameter in the dataset. Within groups s1 and s2, which represent rainfall, the predictor values are recursively split. This process continues until the size of split group sample falls below a certain threshold.

VII. CONCLUSION

The proposed system focuses on developing a blockchain-based supply chain management application that utilizes machine learning to predict crop prices and forecast market demand. The integration of blockchain in the supply chain management application brings several benefits including increased transparency, enhanced security, improved traceability, increased efficiency, and decentralization. This system aims to help farmers in reducing their difficulties by providing them with a transparent supply chain tracking facility with reliable price and demand forecasts, ultimately improving their decision-making process in the agricultural market.

VIII. REFERENCES

[1] Anwar Mazharul, Parveen A., Ferdous Zannatul, Kafi Md. Abdulla Hel, Kabir M., "Base line survey for farmer livelihood improvement at farming system research and development", *International Journal of Business, Management and Social Research*, 2020, vol. 2, pp. 92-104, doi: 10.18801/ijbmsr.020115.10.
[2] Brhane Gebru, Mammo Yared, Negusse Gebremichael, "Sources of information and information seeking behavior of smallholder farmers of Tanqa Abergelle Wereda, central zone of Tigray, Ethiopia", *Journal of Agricultural Extension and Rural Development*, 2019, vol. 9, pp 47-52, doi: 10.5897/JAERD2016.0850.
[3] V. S. Yadav, A. R. Singh, "A Systematic Literature Review of Blockchain Technology in Agriculture", in *Proceedings of the International Conference on Industrial Engineering and Operations Management, Pilsen, Czech Republic,* July 23-26, 2019.
[4] A. Kamilaris, A. Fonts, F. X. Prenafeta-Boldύ, "The rise of blockchain technology in agriculture and food supply chains", *Trends in Food Science & Technology*, 2019, vol. 91, pp. 640-652, doi: 10.1016/j.tifs.2019.07.034.
[5] A. Saxena, M. Dhadwal, M. Kowsigan, "Indian Crop Production:Prediction And Model Deployment Using Ml And Streamlit", *Turkish Journal of Physiotherapy and Rehabilitation*, 2021, vol. 3, pp. 1872-1886.
[6] Ahmed S., "A Software Framework for Predicting the Maize Yield Using Modified Multi-Layer Perceptron" *Sustainability*, 2023, vol. 15, pp. 3017 doi: 10.3390/su15043017.
[7] Q. Wang, J. Yu, S. Chen, Y. Xiang, "SoK: DAG-based Blockchain Systems", *ACM Comput. Surv.*, Dec. 2023, vol. 55, no. 261, pp. 1-38, doi: 10.1145/3576899.
[8] D. Y. Lin, C. J. Juan, C. C. Chang, "Managing Food Safety with Pricing, Contracts and Co-ordination in Supply Chains", *IEEE Access*, 2019, vol. 7, pp. 150892-150909.
[9] H. Fu, C. Zhao, C. Cheng, H. Ma, "Blockchain-based agri-food supply chain management: case study in China", *International Food and Agribusiness Management Review*, 2020 vol. 23, no. 5, pp. 667-679, doi: 10.22434/IFAMR2019.0152.
[10] S. Khaki, L. Wang, "Crop yield prediction using deep neural networks," in Frontiers in Plant Science, 2019, vol. 10, pp. 621.
[11] T. Jheng, T. Li, C. Lee, "Using Hybrid Support Vector Regression to Predict Agriculture Output", in *27th Wireless and Optical Communications Conference (WOCC2018)*, 2018.
[12] M. Abdipour, M. Younessi-Hmazekhanlu, S. H. R. Ramazani, "Artificial Neural Networks and Multiple Linear Regression as Potential Methods for Modeling Seed Yield of Safflower", in *Elsevier*, 2018.
[13] S. Sharma, "India wastes up to 16% of its agricultural produce; fruits, vegetables squandered the most", financialexpress.com https://www.financialexpress.com/economy/india-wastesup-to-16-of-its-agricultural-produce-fruits-vegetablessquandered-the-most/1661671/ (accessed Jul. 31, 2019).
[14] S. Uddin, A. Khan, M. Hossain, "Comparing different supervised machine learning algorithms for disease prediction", *BMC Med Inform*, 2019, vol. 19, pp. 281, doi: 10.1186/s12911-019-1004-8.
[15] X. Yang, Y. Zhang, S. Wang, B. Yu, F. Li, Y. Li, W. Yan, "LedgerDB: a centralized ledger database for universal audit and verification", *Proc. VLDB Endow*, 2020, vol. 13, no. 12, pp. 3138–3151, doi: 10.14778/3415478.3415540.