Smart Agri Chain: An Integrated Machine Learning Framework for Agricultural Price Prediction and Supply Chain Transparency

Mansi Lakade, Shubham Karale, Prof Jyoti Manoorkar
Department of Computer Engineering, MIT ADT University, Pune, India

Abstract—**Agriculture continues to be one of the most important sectors in developing countries, especially in terms of food production, employment, and economic stability. However, despite its importance, farmers still face several challenges when it comes to selling their produce effectively. One of the major issues is the lack of transparency in the agricultural supply chain, which often results in limited awareness about actual market prices and increased dependence on intermediaries.**

Another important concern is the fluctuation in crop prices. These variations usually happen due to seasonal changes, demand differences, arrival quantities, and other external conditions. Because of this, farmers often find it difficult to decide the right time and place to sell their crops.

The idea behind this work is to develop a system that can assist farmers in such situations. This paper presents Smart Agri Chain, which combines machine learning techniques with a digital platform to predict crop prices and improve transparency. The system makes use of historical market data to understand pricing patterns and generate future estimates. In most cases, these predictions provide a reasonable idea of how prices might behave.

The platform is designed as a web-based application so that it remains simple and accessible. Farmers can enter crop-related information and get predicted prices in an easy-to-understand format. The system follows multiple steps including data collection, preprocessing, model training, and prediction generation.

The results show that the system is able to identify useful patterns in the data and provide meaningful predictions. While the predictions may not always be exact, they can still help farmers make better decisions and reduce uncertainty in the selling process.

I. INTRODUCTION

Agriculture plays a very important role in countries like India, where a large portion of the population depends on it for their daily income and livelihood. It supports food production, employment, and overall economic growth. Even then, farmers continue to face several real-world problems, especially related to pricing and market access.

One major issue that can be observed in the agricultural sector is the lack of transparency in the supply chain. In many cases, farmers depend on intermediaries such as traders or agents to sell their produce. Because of this, they do not always get clear or timely information about market prices. As a result, it becomes difficult for them to judge whether they are getting a fair price or not.

Another challenge is the uncertainty in crop prices. Prices do not remain constant and are influenced by many factors like seasonal production, transportation, demand, and weather conditions. In practical situations, farmers often make decisions based on guesswork or local trends, which may not always be reliable.

With the growth of technology, especially in machine learning and data analysis, it has become possible to use historical data to understand such patterns. These methods can help in estimating future crop prices to some extent. If this information is presented properly, it can support farmers in making better decisions.

Keeping this in mind, the Smart Agri Chain system is proposed. The main idea is to combine prediction with a simple digital platform so that farmers can easily access useful information without needing technical knowledge.

II. LITERATURE SURVEY

Several research efforts have explored the application of modern technologies such as machine learning, blockchain, Internet of Things (IoT), artificial intelligence, and cloud computing in agriculture. These technologies aim to improve productivity, decision-making, and supply chain management. Patel et al. proposed a machine learning-based approach for predicting agricultural commodity prices using historical data. Their study demonstrated that regression techniques can effectively capture pricing patterns and provide future estimates. However, their work mainly focused on prediction accuracy and did not address issues related to supply chain transparency.

Sharma et al. introduced a blockchain-based system de- signed to enhance transparency and traceability in agricultural supply chains. The system enabled secure recording of trans- actions among different stakeholders such as farmers, traders, and distributors. While this approach improved reliability in transactions, it did not incorporate any predictive mechanism for crop pricing.

Kumar et al. developed an IoT-based smart agriculture system for monitoring environmental factors such as soil moisture, temperature, and humidity. This system helped farmers manage crop conditions more efficiently. However, it was limited to field-level monitoring and did not provide support for market analysis or price prediction.

Chen et al. examined the role of artificial intelligence in agriculture and highlighted its potential in improving productivity and decision-making processes. Their work focused primarily on intelligent farming techniques and crop monitoring, without emphasizing market forecasting or transparency.

Singh et al. proposed a data analytics framework for studying agricultural market trends using large datasets. Their work provided insights into demand fluctuations and pricing pat- terns. However, it lacked an integrated system that combines prediction with a user-friendly platform.

Gupta et al. introduced a cloud-based agricultural information system that provided farmers with digital access to market data. While this improved accessibility, the system did not include advanced machine learning models for forecasting future prices.

From the analysis of existing literature, it can be observed that most systems focus on individual aspects such as prediction, monitoring, or supply chain management. Very few approaches combine these elements into a single unified platform. Therefore, there is a need for an integrated system that offers both predictive insights and transparency, which is addressed by the proposed Smart Agri Chain framework.

III. METHODOLOGY

The proposed Smart Agri Chain system is developed using a structured methodology that combines agricultural data analysis, machine learning techniques, and a web-based interface. The primary objective of this system is to assist farmers in predicting future crop prices while also providing a transparent and organized way to access market-related information.

The development process begins with the collection of historical agricultural market data from publicly available and reliable sources such as Agmarknet. This dataset includes important attributes such as crop name, market location, arrival quantity, date, and different price values including minimum, maximum, and modal prices. These data records are essential for understanding past market trends and serve as the foundation for building predictive models.

After data collection, the next step involves preprocessing the dataset to ensure its quality and usability. In this phase, missing values, duplicate entries, and inconsistent records are carefully handled. Data cleaning helps in improving the accuracy of the machine learning models. Additionally, the dataset is transformed into a structured format where numerical and categorical values are properly arranged. Feature selection is then performed to identify the most relevant attributes that significantly influence crop price prediction.

Once the dataset is prepared, machine learning algorithms are applied to analyze the data. In this work, models such as Linear Regression, Decision Tree, and Random Forest are considered. These models are trained using historical data so that they can learn patterns and relationships between various market factors and crop prices. Among these, Random Forest and Linear Regression have shown better performance in capturing pricing behavior and generating reliable predictions. After the training phase, the selected machine learning model is integrated into a web-based application. This allows farmers to interact with the system easily. Users can input crop-related information through the interface, and the back- end system processes this input using the trained model to generate predicted prices.

Finally, the results are displayed on the application dash- board in a clear and understandable format. This ensures that farmers can interpret the predictions without difficulty and use them to make better decisions regarding the sale of their crops. The overall system thus combines data-driven insights with an accessible platform to improve agricultural decision-making.

A. Dataset Description

The dataset used in this study is obtained from publicly available agricultural market sources such as Agmarknet. It consists of historical records collected from multiple markets across different regions. The dataset includes essential attributes such as crop name, market name, date of transaction, arrival quantity, and pricing details including minimum, maximum, and modal prices.

These attributes play a crucial role in understanding market behavior and identifying patterns in price fluctuations. The availability of diverse data from different regions helps im- prove the robustness of the model. By analyzing such data, the system can generate meaningful predictions that reflect real-world market trends.

B. System Architecture

The system architecture represents the overall structure of the Smart Agri Chain platform and illustrates how different components interact with each other. The architecture is de- signed to ensure smooth communication between the user interface, backend processing unit, machine learning model, and database.

The farmer interacts with the system through a web application interface. This interface allows users to enter crop-related information and request predictions. The web application forwards the request to the backend server, which acts as the central processing unit of the system.

The backend server communicates with the machine learning model to generate predictions and also interacts with the database to retrieve or store relevant data. Once the prediction is generated, the result is sent back to the web application and displayed to the user in an easy-to-understand format.

C. System Workflow

The workflow of the Smart Agri Chain system describes the sequence of operations performed from input to output. It explains how data flows through different stages of the system to generate predictions.

The process begins when the farmer enters crop-related details into the web application. These details are combined with the existing historical dataset stored in the system. The data then undergoes preprocessing, where it is cleaned and prepared for analysis.

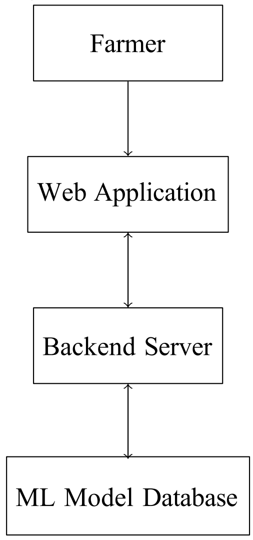

Fig. 1. System Architecture of Smart Agri Chain

After preprocessing, the machine learning model processes the data and identifies patterns based on historical trends. Using this analysis, the model generates a predicted crop price. The final output is then displayed to the user through the web interface. This structured workflow ensures that the system provides accurate and timely predictions in a systematic manner.

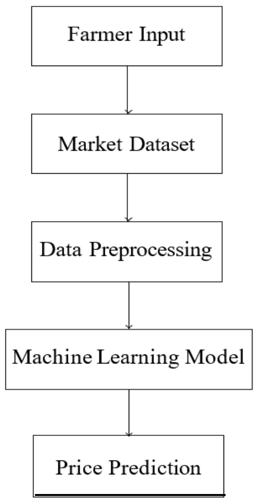

Fig. 2. Workflow of Smart Agri Chain System

D. Machine Learning Process

The machine learning process in the Smart Agri Chain system involves multiple stages, starting from raw data preparation to final prediction generation. This process ensures that the model learns meaningful patterns from historical data and produces reliable outputs.

Initially, historical crop price data is collected from various market sources. This data is then cleaned to remove missing, inconsistent, or irrelevant values. Data cleaning is an important step as it directly affects the accuracy of the model.

After cleaning, feature selection is performed to identify the most relevant variables that influence crop prices. Selecting the right features helps improve model performance and reduces unnecessary complexity.

The processed dataset is then used to train machine learning models. During training, the model learns relationships between input features and output prices. Once the training is complete, the model becomes capable of making predictions on new data.

Finally, the trained model generates predicted crop prices based on the input provided by the user. These predictions are then displayed through the web application.

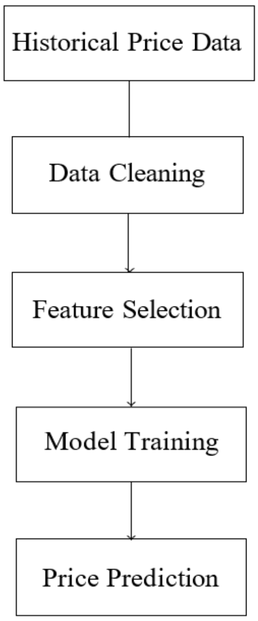

Fig. 3. Machine Learning Prediction Process

IV. RESEARCH GAP

Although significant advancements have been made in the field of smart agriculture, existing solutions still suffer from several limitations. Most of the available systems focus on addressing a single aspect of the agricultural ecosystem rather than providing a comprehensive solution that can support farmers in multiple dimensions.

A large number of studies have concentrated on crop price prediction using machine learning techniques. These approaches are effective in analyzing historical data and iden- tifying price trends. However, they are generally limited to prediction tasks and do not consider the issue of transparency in agricultural transactions or the overall supply chain.

On the other hand, systems that focus on supply chain management aim to improve traceability and record-keeping. These systems enhance transparency by maintaining proper records of transactions and product movement. However, they do not provide any predictive capability that can help farmers anticipate future price changes.

Another important limitation observed in existing systems is related to usability. Many of the proposed solutions are complex and not designed with the end user in mind. Farmers, especially those with limited technical knowledge, require systems that are simple, intuitive, and easily accessible.

Therefore, there is a clear need for an integrated platform that combines the advantages of both prediction and transparency. Such a system should not only forecast crop prices but also provide reliable and accessible market information. The Smart Agri Chain system is designed to address this gap by integrating machine learning-based prediction with a user- friendly digital platform.

V. RESULTS AND DISCUSSION

The system was tested using historical agricultural data to understand how well it can predict crop prices. During experimentation, different models were applied and their performance was compared.

It was observed that Random Forest performed better in most cases, especially when the data showed complex pat- terns. Linear Regression, on the other hand, worked well for identifying general trends but was not as effective for more complicated relationships.

When comparing the predicted values with actual prices, it can be seen that the model is able to follow the overall trend quite well. There are some differences between actual and predicted values, which is expected, but the general direction remains similar. This means the system can still provide useful guidance.

In practical terms, farmers may not require exact values but rather an idea of how prices are likely to behave. From this perspective, the system proves to be helpful. It gives an early indication of possible market trends, which can support better planning.

Another useful aspect is that all these results are available through a web interface. This makes it easier for users to understand the output without dealing with complex technical details.

A. Advantages of Proposed System

The Smart Agri Chain system offers several advantages in the agricultural domain:

- Enables crop price prediction using historical data and machine learning techniques
- Improves transparency in the agricultural supply chain
- Reduces dependency on intermediaries for market-related information
- Assists farmers in making informed decisions regarding crop sales
- Provides easy access through a user-friendly web-based platform

B. Limitations

Despite its advantages, the system has certain limitations that must be considered.

The accuracy of the predictions depends heavily on the quality and size of the dataset used for training. Incomplete or inconsistent data can affect model performance. Additionally, external factors such as sudden weather changes, government policies, transportation issues, or unexpected market disruptions may influence crop prices and reduce prediction accuracy.

C. Evaluation Metrics

The performance of the machine learning models was evaluated using standard regression metrics such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and R-squared (R^2) score.

MAE measures the average difference between actual and predicted values, providing an understanding of prediction accuracy. RMSE gives more importance to larger errors and helps identify significant deviations. The R^2 score indicates how well the model explains the variability in the dataset.

Model	Accuracy	Performance
Linear Regression	Moderate	Good
Random Forest	High	Very Good

TABLE I

COMPARISON OF PREDICTION MODELS

Based on the comparison, the Random Forest model per- forms better in capturing complex relationships within the dataset. Linear Regression, on the other hand, is simpler and useful for identifying general trends. Therefore, Random Forest is considered more suitable for this application.

VI. CONCLUSION AND FUTURE SCOPE

The Smart Agri Chain system presents a practical solution for addressing key challenges in agriculture, particularly in the areas of price prediction and supply chain transparency. By combining machine learning techniques with a web-based platform, the system provides farmers with valuable insights that can support better decision-making.

The experimental results show that machine learning models can effectively analyze historical agricultural data and generate meaningful predictions. These predictions help farmers under- stand potential market trends and make informed choices about when and where to sell their crops.

The system also highlights the importance of integrating predictive analytics with accessible digital platforms. Instead of relying solely on raw data, farmers can benefit from interpreted insights that simplify complex information.

In the future, the system can be enhanced by incorporating real-time data such as weather updates and live market prices. Advanced techniques such as deep learning can be used to further improve prediction accuracy. Additional features like mobile application support, multilingual interfaces, and region-specific recommendations can make the system more accessible and user-friendly.

With these improvements, Smart Agri Chain has the potential to become a powerful tool for improving transparency, efficiency, and profitability in the agricultural sector.

VII. ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to their project guide for continuous guidance, support, and encouragement throughout the development of this work. The authors also thank the institution for providing the necessary resources and environment to complete this research success- fully.

REFERENCES

[1] S. Patel and R. Shah, "Agricultural Commodity Price Prediction Using Machine Learning Techniques," 2021.
[2] R. Sharma and P. Gupta, "Blockchain-Based Agricul- tural Supply Chain Management System," 2020.
[3] A. Kumar et al., "IoT-Based Smart Agriculture Moni- toring System," 2019.
[4] J. Wolfert et al., "Big Data in Smart Farming," 2017.
[5] S. Kamilaris et al., "Deep Learning in Agriculture," 2018.
[6] L. Li et al., "Plant Phenotyping Techniques," 2014.
[7] K. R. Krishna, "Agricultural Supply Chain Manage- ment," 2018.
[8] S. A. Raza et al., "Crop Price Prediction Using Regres- sion," 2019.
[9] P. Singh et al., "Machine Learning for Crop Prediction," 2020.
[10] T. Liakos et al., "Machine Learning in Agriculture," 2018.
[11] A. Kamath et al., "Predictive Analytics in Agriculture," 2021.
[12] M. Caro et al., "Blockchain Traceability in Agriculture," 2018.
[13] C. Verdouw et al., "Digital Transformation in Agricul- ture," 2021.
[14] R. Ray, "Agricultural Price Forecasting," 2019.
[15] D. Mishra et al., "Smart Farming Using AI," 2020.
[16] A. Patel, "Agricultural Data Mining," 2018.
[17] N. Zhang et al., "Precision Agriculture Overview," 2002.
[18] F. Tian, "Blockchain in Food Supply Chain," 2016.
[19] J. Aker, "Mobile Phones in Agriculture," 2010.
[20] S. Shetty et al., "Neural Networks for Price Prediction," 2019.
[21] R. Jain et al., "Time Series Forecasting," 2020.
[22] P. Sharma et al., "Agricultural Data Analysis," 2021.
[23] Y. Chen, "AI in Smart Agriculture," 2020.
[24] K. Balafoutis et al., "Sustainability in Agriculture," 2017.
[25] FAO Report, "Digital Technologies in Agriculture," 2019.