



FARMER E-MARKET FOR BETTER PRODUCT DEMAND PREDICTION

Revati Vishwas Shitole¹, Dr. Shveti Chandan²

Savitribai Phule Pune University, Pune, Maharashtra, India^{1,2}

Abstract: Farmers often have to decide what to grow and how much to bring to the market without having a clear idea of future demand. When too much of a crop reaches the market, some of it may remain unsold and prices may fall. When the available quantity is too low, buyers may struggle to find the product they need. This paper proposes a Farmer E-Market system that brings farmers and market information together and uses past demand records to estimate future demand for selected crops. The proposed system allows a farmer to enter the crop name, quantity and expected harvest date. It then compares the farmer's planned supply with current supply, expected future supply and predicted demand for the selected region and period. The first version uses a Random Forest regression model trained on historical crop-demand records containing year, month, region, crop and market demand. The project uses a time-based train-test split so that earlier records are used for training and later records are used for testing. The system is intended to provide a simple, understandable view of demand and supply rather than guarantee a sale or a particular price. Its usefulness will depend on the quality, coverage and freshness of the data available.

Keywords : Farmer E-Market, Crop Demand Prediction, Random Forest, Supply-Demand Comparison, Digital Agriculture.

INTRODUCTION

Agriculture is an important source of income and food, but deciding how much of a crop to produce is not always easy. Demand can change with the season, location, consumer needs and the quantity already available in the market. Farmers may rely on experience, information from other farmers or the prices they see at local markets. These sources can be helpful, but they may not give a complete picture of future demand.

A digital marketplace can make market information easier to access. Instead of serving only as a place to list products, a Farmer E-Market can also help farmers understand the relationship between the quantity they plan to supply and the demand expected for that crop. Historical records can be used to estimate demand for a particular month, crop and region. Farmers can then compare this estimate with the supply information stored in the system.

This project focuses on a practical, small-scale system using the agricultural demand data currently available to the project. The initial version covers a limited number of regions and crops. It is therefore designed as a decision-support tool, not as a replacement for farmers' knowledge or local market advice.

PROBLEM STATEMENT

Farmers often lack accurate information about crop demand and available supply, leading to a mismatch between production and market needs. This project aims to develop a Farmer E-Market that predicts crop demand using historical data and helps farmers make better supply decisions.

OBJECTIVES

1. To develop a web-based marketplace where farmers can submit crop, quantity and expected harvest date.
2. To estimate crop demand using historical records containing year, month, region, crop and market demand.
3. To evaluate the demand prediction model using data that occurs later in time than the training data.
4. To make the results understandable to users who may not have a technical background.

SCOPE OF THE STUDY

The current study uses a demand dataset with the fields Year, Month, Region, Crop and Market_Demand. After basic text cleaning, the dataset contains 2,280 rows, with 18 crop names and 5 regions. The current coverage is limited to the regions and crops represented in the dataset. The farmer form is planned to collect the crop, quantity and expected harvest date; location will be selected from the supported regions rather than treated as unrestricted nationwide coverage.

The system's demand estimate is based on historical patterns in the available dataset. It does not automatically know every transaction happening in a market. Supply comparisons will be only as complete as the supply records entered or imported into the system.



LITERATURE REVIEW

Digital agricultural marketplaces aim to make it easier for producers and buyers to exchange product and market information. A marketplace can record what farmers intend to sell, while demand analysis can help users understand whether the expected quantity may be close to or far above the estimated requirement. These functions are related but different: listing a product connects users, whereas demand prediction estimates a future value from past records. Machine-learning methods are commonly used for prediction tasks where historical examples are available. Random Forest is one such method. It combines the outputs of multiple decision trees and can model relationships between input fields and a numeric target. For this project, the input fields are time, region and crop, while Market_Demand is the value to estimate. A model's accuracy on training records alone is not enough to show how well it may work on future months. For that reason, evaluation should preserve time order, training on earlier records and testing on later records. The project's focus is narrower than a complete national agricultural trading platform. It is a prototype built around the data currently available. A later version could include more regions, market arrivals, buyer orders, crop-season information, weather records and actual transaction outcomes. These additions should be evaluated before being used to make stronger claims about usefulness.

RESEARCH METHODOLOGY

Research approach :

The project follows a build-and-test approach. First, the available dataset is checked and cleaned. Next, a demand prediction model is trained and evaluated. The trained model is then saved so the web application can load it when a prediction is requested. After that, the application stores farmer supply records and compares them with predicted demand.

Data preparation :

The demand dataset is checked for missing values, duplicate records and inconsistent text. Spaces around crop and region names are removed so that entries such as 'Rice' and ' Rice ' are not treated as different categories. The cleaned dataset has 2,280 rows, five columns, no missing values and no exact duplicate rows. During inspection, repeated Year-Month-Region-Crop combinations were also noticed for pea/Pea variants. These records should be reviewed to confirm whether they represent different crop types or inconsistent naming before they are combined or retained.

Demand prediction model :

A Random Forest regression model is used to estimate Market_Demand. Year, Month, Region and Crop are used as input features, and Market_Demand is the target value. The text-based fields are converted into a format the model can process, and the preprocessing steps are kept together with the model in a pipeline. This makes it easier to apply the same preparation to new farmer or marketplace queries.

Model evaluation :

To test whether the model can estimate later demand, the data is divided by time rather than randomly mixing earlier and later records. The model is trained on the earlier portion and evaluated on the later portion. The reported evaluation for the current experiment is MAE = 59.13, RMSE = 101.35 and $R^2 = 0.995$, using 1,824 training rows and 456 test rows. These figures describe performance on this dataset and split only. They do not guarantee the same accuracy for new regions, new crops or future market conditions.

Mean Absolute Error (MAE) gives the average size of the prediction error. Root Mean Squared Error (RMSE) gives greater weight to larger errors. R^2 describes how much variation in the test values is accounted for by the model under the evaluation setup. All three measures should be considered together, and predictions should also be checked against actual market records as the system develops.



PROPOSED SYSTEM

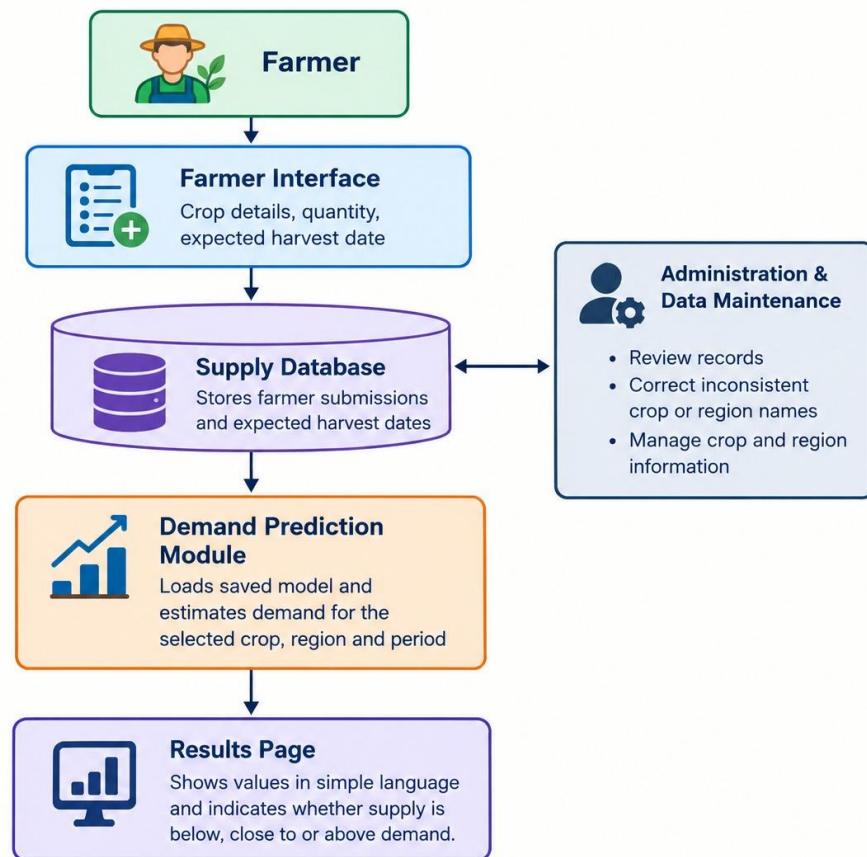


Figure 1: Proposed Farmer E-Market System

Example of system use :

A farmer selects Rice, enters a planned quantity and chooses an expected harvest date in a supported region. The application estimates demand for the matching period. It then checks supply records for rice in that region, including other farmers' expected harvests. The results page shows the predicted demand, recorded supply and the difference between them. This gives the farmer another source of information to consider before deciding how much to bring to market.

Suggested technology :

The demand model can be developed in Python using pandas and scikit-learn. The trained model can be saved as a file and loaded by a backend service. The web application can use a JavaScript-based frontend and backend, with MongoDB storing farmer supply records. The exact integration should ensure that model inputs are validated and that the same crop and region names are used across the dataset, form and database.



RESULTS AND DISCUSSION

Crop demand, forecast ahead

Know what the market will want, before the season does.

Pull up any division and crop across the region, pick a year and month, and see the demand your model expects — built on historical market data from 2015 to 2035.

Year

2026

Month

October

Region

Balod Division

Crop

Arhar

Predict demand

Your forecast will appear here.

Choose a year, month, region and crop, then run the prediction.

[Crop Demand](#) [Demand Forecast](#) [Submit Supply](#) [My Submissions](#) [Compare](#)

Crop demand, forecast ahead

Know what the market will want, before the season does.

Pull up any division and crop across the region, pick a year and month, and see the demand your model expects — built on historical market data from 2015 to 2035.

Year

2026

Month

October

Region

Balod Division

Crop

Arhar

Predict demand

Arhar - Balod Division

3,996.94

predicted market demand for October 2026



[Crop Demand](#) [Demand Forecast](#) [Submit Supply](#) [My Submissions](#) [Compare](#)

For farmers

Tell us what you're growing.

Log a crop, how much of it you expect to bring to market, and when it'll be ready. This stays on your device — it's only used to estimate supply for this session, nothing is sent to a server database.

Crop

Arhar

Region

Balod Division

Quantity

e.g. 500

Expected harvest date

04/10/2026

Add to supply list

Recently added

100 of Arhar
Balod Division · ready 2026-10-04

3,000 of Arhar
Balod Division · ready 2026-10-04

2,000 of Arhar
Balod Division · ready 2026-10-04

2,000 of Arhar
Balod Division · ready 2026-10-04

[View all submissions →](#)

[Crop Demand](#) [Demand Forecast](#) [Submit Supply](#) [My Submissions](#) [Compare](#)

This session only

My Submissions

Everything logged on this device, stored only in this browser. Close the tab or clear your browser data and it's gone.

Clear all

Crop	Region	Quantity	Expected Harvest	
Arhar	Balod Division	100	2026-10-04	Remove
Arhar	Balod Division	3,000	2026-10-04	Remove
Arhar	Balod Division	2,000	2026-10-04	Remove
Arhar	Balod Division	2,000	2026-10-04	Remove
Total		7,100		

© IJARCCE

This work is licensed under a Creative Commons Attribution 4.0 International License

168



[Crop Demand](#) [Demand Forecast](#) [Submit Supply](#) [My Submissions](#) [Compare](#)

Demand vs. supply

Will there be enough to go around?

Pick a crop, region and month. We'll weigh the model's predicted demand against what's been logged on the Submit Supply page – both what's already harvested and what's expected in time.

Crop

Region

Year

Month

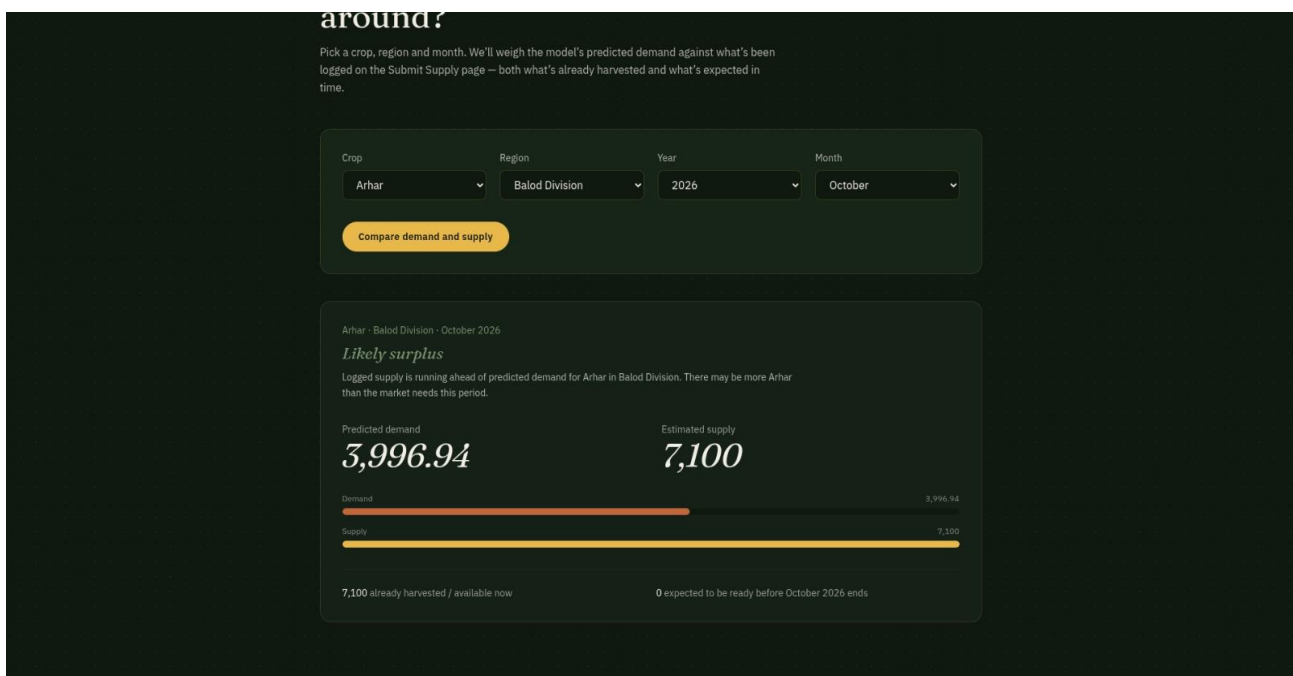
Arhar

Balod Division

2026

October

Compare demand and supply



The initial data preparation stage produced a consistent dataset of 2,280 rows after trimming extra spaces. The demand model was evaluated using a time-based split, with 1,824 rows for training and 456 rows for testing. In the current experiment, the model achieved MAE of 59.13, RMSE of 101.35 and R^2 of 0.995. These are preliminary results from the project's available dataset.

The evaluation numbers suggest that the model fitted the test records well under the chosen split. However, a high R^2 by itself does not prove that the system will make reliable decisions in every market. The dataset's coverage is limited, and patterns in historical records may change. The model should therefore be checked regularly against new observations. The next stage is to save the model as demand_model.pkl, connect it to the application, build the FarmerSupply database model and form, and implement the supply-demand comparison.

No user study or live-market trial has yet been reported in this draft. As a result, claims about reducing waste, increasing farmer income or improving sales should be treated as expected benefits to investigate, not as proven outcomes.

LIMITATIONS :



- The demand dataset covers only five regions and the crop names present in the current records.
- Historical demand values may not capture sudden changes caused by weather, transport issues, festivals, policy changes or other market events.
- The supply comparison depends on farmers entering accurate quantities and harvest dates, as well as the availability of other supply records.
- The model estimates demand; it does not guarantee a buyer, sale price or profit.
- The system needs further testing with fresh data and real users before conclusions can be made about its practical impact.

FUTURE WORK

Future work can expand the dataset to more regions and crops, add recent market arrivals and completed transactions, and introduce checks for unusual or inconsistent records. The system can also show historical demand trends and explain the difference between predicted demand and recorded supply. Model performance should be monitored as new data becomes available. A pilot involving farmers and market participants would help assess whether the results are understandable and useful in actual planning. Any later addition of price prediction should be treated as a separate task and evaluated with suitable price data.

CONCLUSION

This paper presents a proposed Farmer E-Market that combines farmer supply entry with crop demand prediction. The system is intended to help farmers compare their planned quantity with demand estimated from historical records and with supply already recorded for the same region and period. The first model uses Random Forest regression and has been evaluated using a time-based split. The current results provide a starting point, but further testing with fresh and wider market data is needed. By presenting the estimates in clear language and showing where the data is incomplete, the proposed system can offer farmers an additional source of information while leaving final decisions in their hands.

REFERENCES

- [1] L. Breiman, "Random Forests," *Machine Learning*, vol. 45, pp. 5–32, 2001.
- [2] F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [3] Crop-Demand-Data.csv, project dataset supplied for the current MCA project. Dataset fields: Year, Month, Region, Crop and Market_Demand.
- [4] M. T. Wolfert, L. Ge, C. Verdouw, and M.-J. Bogaardt, "Big Data in Smart Farming—A review," *Agricultural Systems*, vol. 153, pp. 69–80, 2017.
- [5] F. J. S. Kamilaris, A. Kartakoullis, and F. X. Prenafeta-Boldú, "A review on the practice of big data analysis in agriculture," *Computers and Electronics in Agriculture*, vol. 143, pp. 23–37, 2017.
- [6] P. K. Thornton, "Livestock production: Recent trends, future prospects," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 365, no. 1554, pp. 2853–2867, 2010.
- [7] R. Liakos, P. Busato, D. Moshou, S. Pearson, and D. Bochtis, "Machine Learning in Agriculture: A Review," *Sensors*, vol. 18, no. 8, 2674, 2018.
- [8] S. R. Elavarasan and S. D. Vincent, "Crop yield prediction using machine learning techniques," *International Journal of Engineering Research & Technology*, 2019.
- [9] J. You, X. Li, M. Low, D. Lobell, and S. Ermon, "Deep Gaussian Process for Crop Yield Prediction Based on Remote Sensing Data," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 31, no. 1, 2017.
- [10] M. A. T. Chlingaryan, S. Sukkarieh, and B. Whelan, "Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review," *Computers and Electronics in Agriculture*, vol. 151, pp. 61–69, 2018.
- [11] C. Kamilaris and F. Prenafeta-Boldú, "Deep learning in agriculture: A survey," *Computers and Electronics in Agriculture*, vol. 147, pp. 70–90, 2018.
- [12] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed. Morgan Kaufmann, 2011.
- [13] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. Springer, 2009.
- [14] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.
- [15] World Bank, *ICT in Agriculture: Connecting Smallholders to Knowledge, Networks, and Institutions*. Washington, DC: World Bank, 2017.