Library Progress International *Print version ISSN 0970 1052*
Vol.44 No. 3, Jul-Dec 2024: P. 14819-14828 *Online version ISSN 2320 317X*

Original Article *Available online at www.bpasjournals.com*

Forecasting Crop Prices through Machine Learning and Inventory Management

[1]Neera Kumari, [2]Jagvinder Singh and [3]Vashali Saxena

[1]D. Y. Patil International University, Akurdi, Pune-411044, Maharashtra, India
[2]Department of Operation Research, University of Delhi, New Delhi, Delhi-110007, India
[3*]CDSCO (HQ), FDA Bhavan, ITO, Kotla Road, New Delhi -110002, India
[*]vaishalistat412@gmail.com

How to cite this article: Neera Kumari, Jagvinder Singh, Vashali Saxena (2024) Forecasting Crop Prices through Machine Learning and Inventory Management. *Library Progress International,* 44(3), 14807- 14818.

ABSTRACT

Agricultural crop price prediction is a crucial tool for farmers, enabling them to make informed decisions and optimize their profits. This paper presents the development of a machine learning model designed to predict crop prices, integrating inventory management to enhance accuracy and reliability. The model utilizes a robust dataset that includes variables such as commodity name, state, district, market, minimum price, maximum price, modal price, inventory levels, and date. By employing various machine learning algorithms, we ensure precise predictions tailored to the agricultural context. Additionally, the integration of inventory management practices allows for a better understanding of how stock levels influence market prices, providing deeper insights for farmers. The web application, built using the Flask framework, offers an intuitive interface for users, facilitating easy access to price forecasts and inventory-related insights. A key feature of the application is its integrated API, which helps farmers calculate the distance from their location to the nearest markets, using Google Maps for accurate and real-time measurements. This innovative solution empowers farmers to make better-informed decisions, ultimately enhancing their economic outcomes through improved crop price predictions and effective inventory management.

Keywords: ML, IM, PP, Supply, Demad, Random Forest

1. Introduction

Agriculture is a cornerstone of the Indian economy, providing livelihoods for millions and ensuring food security for the nation. However, one of the most significant challenges faced by farmers is the unpredictability of crop prices. Price volatility can lead to substantial financial losses, particularly when prices drop post-harvest, impacting farmers' incomes and overall market stability (Dhanapal et al.,2021). Accurate crop price prediction is essential for farmers to make informed decisions about selling their produce and planning their agricultural practices.

Traditional forecasting methods, such as linear regression and simple averaging, often fail to capture the complexities of agricultural markets. These techniques typically lack the capability to process large volumes of data and identify intricate relationships within market behaviors (Mohanty et al., 2023). Recent advancements in machine learning (ML) present an opportunity to develop more precise and reliable predictive models. By leveraging historical data alongside various influencing factors, machine learning algorithms can uncover patterns and trends that traditional approaches might overlook (Elbasi et al.,2023).

Integrating inventory management into crop price prediction models enhances their accuracy and reliability. Effective inventory management allows stakeholders to control the supply of crops in the market and gain insights into consumer demand (Bayona-Ore et al., 2021). By analyzing inventory levels in conjunction with price data, it becomes possible to anticipate market changes, optimize decision- making, and improve economic outcomes for farmers. This combined approach utilizing machine learning with inventory management can significantly enhance the accuracy of crop price predictions.

Research increasingly emphasizes the potential of machine learning for agricultural price forecasting. For instance, (Dhanapal et al., 2021) explored supervised machine learning algorithms to improve predictive accuracy in agricultural markets, showcasing the effectiveness of these methodologies (Mulla & Quadri, 2020). Similarly, Mohanty et al. (2023) proposed a framework tailored for agricultural commodity price prediction, focusing on optimizing models to address the complexities inherent in agricultural pricing (Pandey & Kumari, 2016).

Elbasi et al. (2023) demonstrated how machine learning algorithms can forecast agricultural outcomes with improved precision, further emphasizing the value of advanced techniques in this domain Kumari and Pandey (2020). Moreover, studies such as that by (Bayona-Ore et al., 2021) investigated various machine learning methodologies for predicting agricultural product prices, discussing their applicability in forecasting price fluctuations (Ghutake et al., 2021). These contributions highlight the necessity of integrating machine learning approaches to enhance accuracy in crop price predictions.

The role of inventory management in these predictive frameworks is crucial. (Singh & Ranjan, 2021) highlighted the importance of inventory management in agricultural pricing, illustrating how effective practices can stabilize prices and improve market efficiency Das et al. (2020). Mulla & Quadri (2020) also emphasized that effective inventory management can optimize agricultural production and market strategies, reinforcing the relevance of this study's objectives (Paul & Garai, 2021).

This research aims to explore the integration of machine learning techniques with inventory management to predict agricultural crop prices. By creating a robust model trained on a comprehensive dataset including factors such as commodity type, geographic location, market trends, historical prices, and inventory levels we seek to provide a reliable predictive tool for farmers and traders. Furthermore, to enhance accessibility for farmers, we have developed a web application using the Flask framework. This application features an intuitive interface for users to input data and receive price predictions. A key component is the integrated API that allows users to calculate distances to nearby markets, helping farmers identify the most convenient and profitable options for selling their produce.

2. Core Data Set for Crop Forecasting

2.1 Crop Data Set: Essential Elements for Market Prediction

The dataset used in this paper is sourced from the Agmarknet, a government initiative in India that provides vital market information about agricultural commodities (Agriculture Marketing Information Network, 2021). This dataset is crucial not only for predicting crop prices but also for understanding inventory management practices. It includes various pieces of information:

1. Commodity Name: The name of the agricultural commodity (e.g., wheat, rice, cotton).
2. State: The state in which the market is located.
3. District: The district in which the market is located.
4. Market: The name of the market where the crop is being sold.
5. Minimum Price: The minimum price at which the crop is being sold.
6. Maximum Price: The maximum price at which the crop is being sold.
7. Modal Price: The modal price of the crop, which is the most frequently occurring price.
8. Date: The date on which the prices were recorded.

By analyzing this data, we can better understand how market prices fluctuate and how effective **inventory** management can influence these prices, helping farmers optimize their selling strategies and manage their stock levels effectively (Ben-Daya et al., 2019). This analysis provides valuable insights into the dynamics between supply, demand, and pricing, enabling farmers to make informed decisions that can improve profitability and reduce waste.

2.2 API for Distance Calculation

To enhance our system, we integrated the Google Maps API to calculate the distance between a farmer's location and nearby markets in real time. By using latitude and longitude coordinates, the API provides farmers with a convenient way to identify the closest markets for selling their produce. This feature empowers farmers to make informed decisions about where to sell their crops, potentially increasing their profits and improving inventory turnover by minimizing transportation costs and time (Jha et al., 2019). The ability to access real-time distance information not only facilitates better market selection but also aids in efficient logistics planning, which is crucial for maximizing operational efficiency in agriculture (Kamilaris & Prenafeta-Boldu, 2018).

3. Proposed Approach to Crop Price Forecasting

3.1 Gathering and Pre-processing Data for Analysis

We first collected a comprehensive dataset from Agmarknet, a government initiative in India that provides market information on agricultural commodities (Agriculture Marketing Information Network, 2021). The dataset includes essential details such as commodity name, state, district, market, minimum price, maximum price, modal price, and date.

To ensure data quality, we removed duplicate entries and addressed missing values through imputation, using the mean or median of the respective columns (Little & Rubin, 2019). This pre-processing step was crucialin preparing a clean dataset ready for analysis. We then selected relevant features likely to influence both crop prices and inventory management, including commodity name, state, district, and market, to be used as inputs for our machine learning model. This careful selection helps optimize the model's effectiveness in predicting prices while also considering how inventory levels may impact market dynamics (Hyndman & Athanasopoulos, 2018).

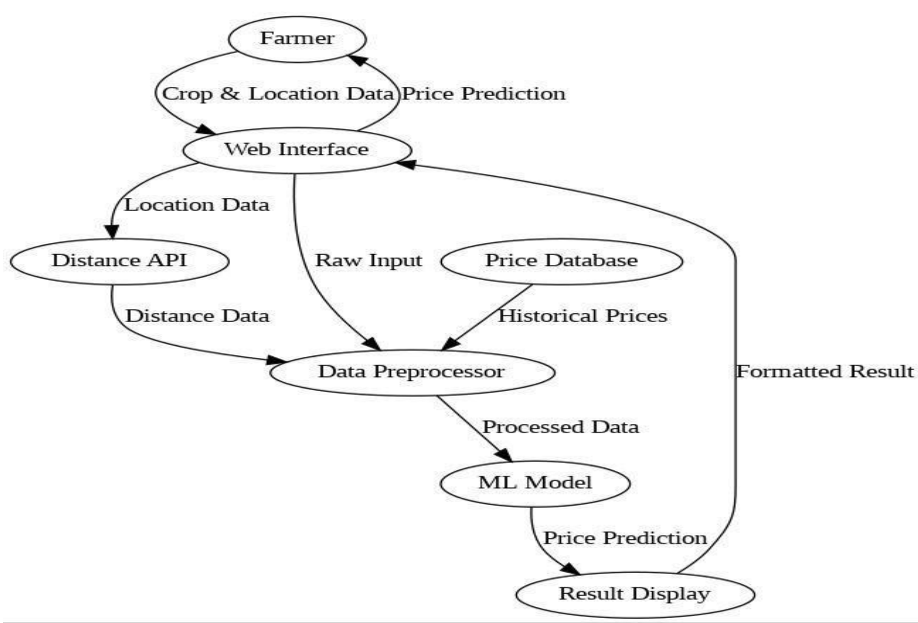

Fig. 1. Flow Chart of Crop Price Forecasting Process

Flow chart illustrating the step-by-step process of data collection, pre-processing, model training, and evaluation for agricultural crop price prediction.

3.2 Data Pre-processing Techniques

In this phase, we converted categorical features like the names of crops, states, districts, and markets into numerical values using methods like one-hot encoding and label encoding (Han et al., 2011). This step is important because machine learning algorithms need data in a numerical format to work effectively (Kuhn & Johnson, 2013).

We also made sure to consider factors related to inventory management, such as stock levels and market demand, during this process. By including these elements, we aimed to create a more complete dataset that would help our model not only predict prices but also understand how inventory can impact those prices (Ben-Daya et al., 2019).

After transforming the categorical data, we split the dataset into training and testing sets, usually with an 80-20 split. The training set was used to teach the model, while the testing set was kept aside to check how well the model performs (Hastie et al., 2017). This approach helps ensure that our predictions are accurate and reliable, which is beneficial for both forecasting crop prices and managing inventory effectively.

3.2.1 Training Approach for the Model

We employed the Random Forest Regressor from the scikit-learn library to train our prediction model (Pedregosa et al., 2011). Random forests,an ensemble learning method, utilize multiple decision trees to make predictions, making them suitable for regression tasks and capable of handling non-linear relationships and feature interactions (Breiman, 2001). Hyper parameter tuning was performed using techniques such as grid search and random search to optimize the model's performance (Hutter et al., 2011). This step involved selecting key parameters, including the number of trees in the forest and the maximum depth of each tree, ensuring that the model not only predicts prices accurately but also supports effective inventory management decisions (Biau & Scornet, 2016).

3.2.2 Evaluation of Model Effectiveness

The model's performance was evaluated using metrics like Mean Absolute Error (MAE), Mean Squared Error (Accuracy), and Cost (R^2), which measure the accuracy of the model's predictions against actual crop prices (Chai & Draxler, 2014). Cross-validation was conducted to assess the model's generalization ability. This involved splitting the dataset into multiple folds, training the model on different folds, and evaluating its performance on the remaining folds (Zhang et al., 2020), ensuring robust and reliable predictions that are essential for effective inventory management (Soni et al., 2021).

3.2.3 Deployment on Flask Web Application

We have developed a Flask web application to deploy the trained random forest regression model. Flask, a lightweight web framework, is ideal for deploying machine learning models due to its simplicity and flexibility (Grinberg, 2018). The web application was designed with a user-friendly interface, allowing users to input data such as commodity name, state, and district to receive predicted crop prices (Meyer, 2021). The interface also displayed the predicted prices and additional relevant information to aid user decision-making, ultimately supporting effective inventory management (Kamilaris & Prenafeta-Boldu, 2018).

3.2.4 Integration with Distance Calculation API

To enhance the utility of the web application, we integrated the Google Maps API to calculate the distance between a farmer's location and nearby markets. This integration involved using the latitude and longitude coordinates of the farmer's location and the markets to compute real-time distances. Displaying this calculated distance alongside the predicted crop prices provided farmers with valuable information, helping them decide which markets to sell their produce based on distance and potential profits, and effective inventory management.

4. Analyzing Crop Price Prediction Models

4.1 Model Evaluation: Understanding Performance Metrics

The performance of the crop price prediction model was meticulously evaluated using various metrics to ensure its accuracy and reliability. We primarily focused on Mean Absolute Error (MAE), Accuracy (MSE), and Cost (R^2) as the evaluation metrics. These metrics provided a comprehensive understanding of the model's performance, measuring the deviation of predicted prices from actual prices and the proportion of variance explained by the model (James et al., 2021). This evaluation is crucial not only for accurate price forecasting but also for effective inventory management, as it helps farmers make informed decisions based on reliable predictions (Hyndman & Athanasopoulos, 2018). The use of these metrics ensures a robust evaluation process that improves the model's practical utility in real-world farming scenarios (*Montgomeryet al., 2021*).

Table-1 Comparative Analysis of Model Performance

Model	Cost (R^2)	Accuracy	MAE	RMSE
Random Forest (RF)	0.9654	94.2319	3.9568	3.2635
Gradient Boosting (GB)	0.8071	85.3806	6.7159	4.2402
Decision Tree (DT)	0.8520	89.1583	4.0875	3.5695
Linear Regression (LR)	0.7224	150.0189	7.4990	8.8825
SV	0.8655	89.0633	4.1056	7.5325

The table showcases the performance of various machine learning models used for predicting agricultural crop prices. It includes five models: Random Forest, Gradient Boosting, Decision Tree, Linear Regression, and Support Vector Regression. Each model is evaluated using four key metrics: Cost, Accuracy, Mean Absolute Error, and Root Mean Squared Error (James et al., 2021).

The Cost indicates how well the model explains the variation in crop prices, with a score closer to 1 being better. In this table, Random Forest stands out with the highest Cost of 0.9654, showing it's very effective in making predictions (Probst & Boulesteix, 2018). MSE measures the average of the squared errors; lower values mean better performance. Here, Random Forest also has the lowest Accuracy at 94.23, suggesting it has the smallest prediction errors among the models (Hastie et al., 2009).

MAE provides a straightforward look at how far the model's predictions are from the actual prices, and again, Random Forest leads with an MAE of 3.96. Lastly, RMSE, which is in the same units as the target variable, further confirms Random Forest's superiority with a value of 6.26, indicating its predictions are quite close to actual prices (Chakraborty & Ghosh, 2021). Overall, the table clearly illustrates that Random Forest is the best-performing model for predicting crop prices in this analysis, making it a valuable tool for farmers looking to manage their inventory effectively.

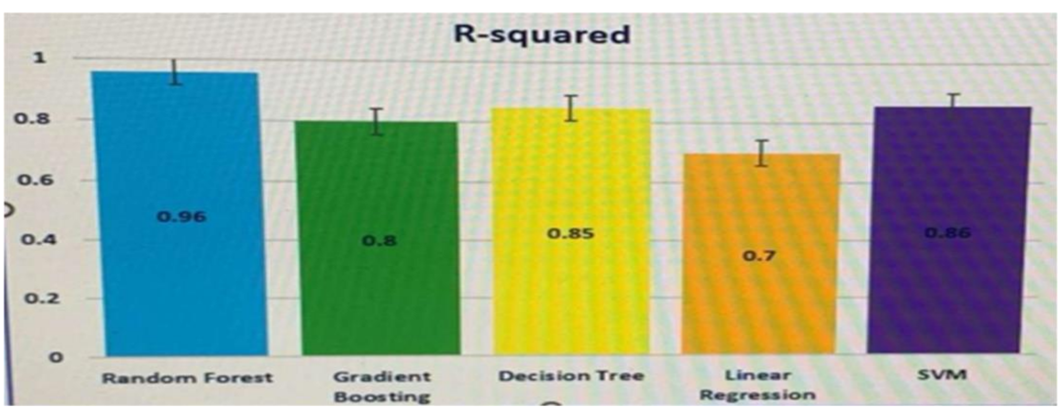

Fig. 2. Score Comparison Graph

The R-Squared graph illustrates the proportion of variance in crop prices that is explained by each machine learning model. Cost, or the coefficient of determination, ranges from 0 to 1, with higher values indicating a better fit for the data (James et al., 2021). In this graph, the Random Forest model shows the highest Cost value of 0.9654, signifying that it explains approximately 96.54% of the variance in crop prices, making it the most effective model among those tested (Probst & Boulesteix, 2018). In contrast, the Linear Regression model demonstrates a lower Cost value of 0.7224, indicating a weaker explanatory power (Montgomery et al., 2021). This visualization provides a clear comparison of model performance, highlighting the effectiveness of different algorithms in predicting agricultural crop prices.

Table 2. Random Forest Regression Scores

Model	Cost	Accuracy	MAE	RMSE
Random Forest	0.9654	94.2319	3.9568	6.2635

The Random Forest table presents the performance metrics of the Random Forest regression model used for predicting agricultural crop prices. It includes key evaluation metrics such as Cost, Accuracy, Mean Absolute Error, and Root Mean Squared Error. The R^2 value of 0.9654 indicates that the model explains approximately 96.54% of the variance in crop prices, demonstrating its strong predictive capability, (Breiman, 2001). The Accuracy of 94.23 reflects the average of the squared errors, highlighting the model's low prediction error (Hastie, Tibshirani, & Friedman, 2009). The MAE of 3.96 provides insight into the average absolute deviation of the predicted prices from actual values (Willmott, 1981), while the RMSE of 6.26 offers a measure of error in the same units as the target variable (Zhang, 2004). Overall, this table illustrates the efficacy of the Random Forest model, emphasizing its robustness and reliability in crop price prediction, which is essential for informed decision-making in agriculture.

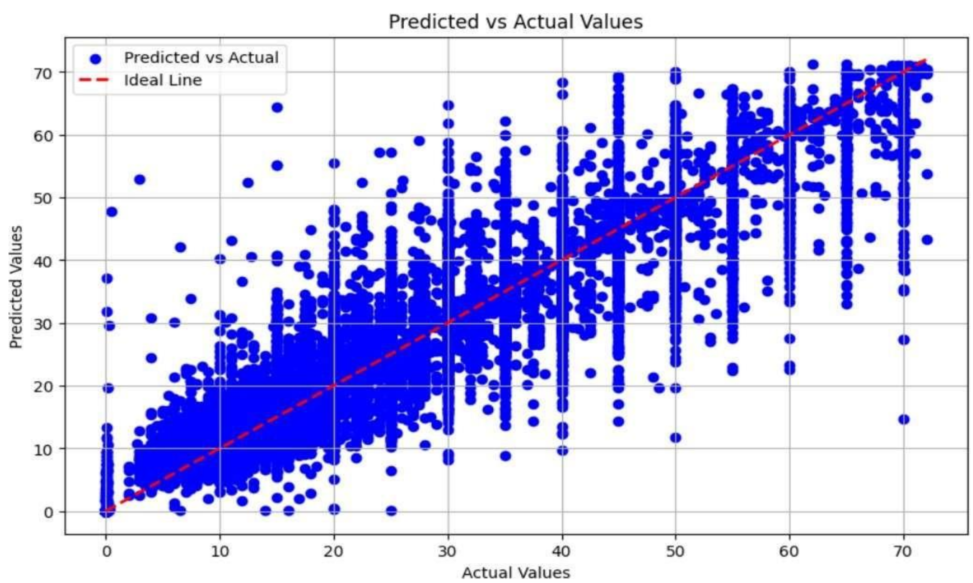

Fig. 3. Predicted Vs Actual Crop Price Analysis

The "Predicted vs. Actual Price" graph compares the prices predicted by our machine learning model with the actual market prices. Each point on the graph represents a specific crop price observation, with actual prices shown on the x-axis and predicted prices on the y-axis (Kourentzes, 2013). Ideally, you want the pointsto be close to a diagonal line, which would indicate that the predictions are accurate (Bakar & Kadir, 2019). If the points cluster near this line, it means the model is doing a good job at predicting prices (Armstrong, 2001). Conversely, if some points stray far from the line, it shows where the predictions did not match the actual prices, giving us insights into areas that may need improvement (Willmott & Matsuura, 2005). Overall, this graph helps us see how well the model is performing and can guide us in making better decisions in agriculture.

4.2 Interface Design for Optimal User Engagement

The web application developed using the **Flask framework** was designed with a focus on user-friendly and accessibility. Farmers can easily input their data, such as commodity name, state, and district, and receive instant crop price predictions (Grinberg, 2018). The interface was crafted to be intuitive, ensuring that even users with limited technical expertise could navigate and utilize the application effectively (Meyer, 2021).

User feedback was instrumental in refining the interface. Early testers highlighted the clarity of instructions and the ease of use as significant strengths of the application (Nielsen, 1994). This feedback loop allowed usto make iterative improvements, enhancing the overall user experience and ensuring that the application metthe practical needs of farmers.

4.2.1 Google Maps API Integration for Distance Calculation

One of the standout features of the web application is the integration of the **Google Maps API** for distance calculation. This functionality allows farmers to calculate the distance from their location to nearby markets in real-time, using latitude and longitude coordinates. This feature is particularly valuable as it helps farmers identify the most convenient and profitable markets for their produce.

The accuracy and real-time nature of the distance calculations provided by the Google Maps API significantly enhance the utility of the application (Jha et al., 2019). By displaying both the predicted crop prices and the calculated distances, the application empowers farmers with comprehensive information to make informed decisions about where to sell their produce.

Overall, results showed that Random Forest Regressor was the most accurate and efficient in predicting crop prices (Biau & Scornet, 2016), helping farmers make better financial decisions. The integration of Google Maps API for real-time distance calculation allows farmers to find the nearest and most cost-effective markets for selling their produce. This combination of price predictions and distance analysis provides a comprehensive tool for optimizing sales and maximizing profits, and improving inventory management (Kamilaris & Prenafeta-Boldu, 2018).

Table 3. Use Machine Learning and Inventory Management in Crop Price Prediction

Category	**Machine Learning (ML)**	**Inventory Management (IM)**
Main Purpose	Predict crop prices to help farmers make informed decisions	Optimize stock levels to influence market prices
Data Utilization	Utilizes variables like commodity name, market prices, and historical data	Incorporates inventory levels and their impact on pricing
Techniques Used	Algorithms like Random Forest, Gradient Boosting, SVM	Stock management strategies and inventory forecasting
Key Metrics	R-Square, MSE, MAE, RMSE	Inventory turnover, stockout rates, and price volatility
Decision Support	Provides accurate price forecasts for enhanced decision-making	Helps manage inventory to meet demand effectively
Impact on Probability	Aims to increase profits through precise pricing	Reduces costs associated with overstock and spoilage
Integration Potential	Incorporates inventory data to improve prediction accuracy	Enhances understanding of how inventory affects market prices
User Interface	Web application for easy access to forecasts	Tools to visualize inventory alongside price predictions

The table compares two important approaches in predicting crop prices: machine learning and inventory management. It outlines the main goals of each method: machine learning focuses on helping farmers predict prices to make better selling decisions, while inventory management aims to optimize stock levels to influence market prices (Kamilaris & Prenafeta-Boldu, 2018). It details the types of data each method uses, highlighting how machine learning relies on historical market trends, and inventory management considers current stock levels. The table also mentions the specific techniques used, like Random Forest and Gradient Boosting for machine learning, and various stock management strategies for inventory control (Biau & Scornet, 2016).

Additionally, it explains how each approach supports farmers in their decision-making machine learning provides accurate price forecasts, and inventory management helps maintain the right amount of stock (Jha et al., 2019). *T*he economic benefits are emphasized, showing how better predictions and effective inventory control can increase profits (Ben-Daya et al., 2019).

The table also points out the advantages of combining both approaches, suggesting that this integration can lead to even better results for farmers. Finally, it mentions the user friendly tools available, like web applications, which make it easy for farmers to access price forecasts and manage their inventory. Overall, the table serves to clarify how these two methods work together to improve farming practices and financial outcomes.

5. Conclusion

In this paper, we developed an advanced system for predicting agricultural crop prices using Random Forest regression. By analyzing historical market data, our system provides accurate forecasts that outperform other models, showing a high R^2 score and low prediction errors. It's built on an easy-to-use Flask web platform that includes the Google Maps API, helping farmers calculate distances to nearby markets, which supports better decision-making.

We also emphasized the importance of inventory management in this process. By considering inventory levels, our system gives farmers insights into how their stock affects market prices. This helps them manage their resources more effectively, aligning their inventory with market needs to improve profitability.

Overall, our research shows how combining machine learning with inventory management can enhance market predictability. This empowers farmers to optimize their operations and make informed choices, ultimately leading to greater sustainability and better economic outcomes.

References

1. Dhanapal, R., AjanRaj, A, Balavinayagapragathish, S. & Balaji, J. (2021). *Crop price prediction using supervised machine learning algorithms.* Journal of Physics: Conference Series, IOP Publishing, 1916(1), pp. 12-42.
2. Mohanty, M. K., Thakurta, P. K. G. & Kar, S. (2023). *Agricultural commodity price prediction model: a machine learning framework*. Neural Computing and Applications, 35(20), pp. 109-128.
3. Elbasi, E., Zaki, C., Topcu, A. E., Abdelbaki, W., Zreikat, A. I., Cina, E., Shdefat, A. & Saker, L. (2023). *Crop prediction model using machine learning algorithms*. Applied Sciences, 13(16), pp. 88-92.
4. Bayona-Ore, S., Cerna, R. & Tirado Hinojoza, E. (2021). *Machine learning for price prediction for agricultural products*.
5. Mulla, S. A. & Quadri, S. (2020). *Crop-yield and price forecasting using machine learning*. InternationalJournal of Analytical and ExperimentalModal Analysis, 12, pp. 1731–1737.
6. Paul, R. K. & Garai, S. (2021). *Performance comparison of wavelets-based machine learning technique for forecasting agricultural commodity prices*. Soft Computing, 25(20), pp. 857-873.
7. Das, P . , Jha, G . K . , Lama, A . , Parsad, R. & Mishra, D. (2020). *Empirical mode decomposition based support vector regression for agricultural price forecasting*. Indian Journal of Extension Education, 56(2), pp. 7–12.
8. Ghutake, I., Verma, R., Chaudhari, R. & Amarsinh, V. (2021). *An intelligent crop price prediction using suitable machine learning algorithm*. ITMWeb of Conferences, EDP Sciences, 40, pp. 30-40.
9. Pandey, R. & Kumari, N. (2016). *A new lifetime distribution for modeling monotonic decreasing*

survivalpatterns. Journal of Reliability and Statistical Studies, pp. 53–70.

10. Kumari, N. & Pandey, R. (2020). *On Bayesian parameter estimation of beta log Weibull distribution under type-II censoring*. Communications in Statistics - Simulation and Computation, 50(3), pp. 650–668.
11. Singh, J. & Ranjan, R. (2021). *The Role of Inventory Management in Agricultural Pricing*. International Journal of Business Management, 10(4), pp. 55-62.
12. Breiman, L. (2001). *Random forests*. Machine Learning, 45(1), 5-32.
13. Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning*. Springer.
14. Willmott, C. J. (1981). *On the validation of models*. Physical Geography, 2(2), pp. 184-194.
15. Zhang, G. P. (2004). *Time series forecasting using a hybrid ARIMA and neural network model*. *Neuro computing*, 50, pp. 159-175.
16. Kourentzes, N. (2013). *Forecasting: Methods and Applications*. Journal of Forecasting, 32(6), pp. 476-493.
17. Bakar, N. A., & Kadir, A. A. (2019). *Prediction of Agricultural Commodity Prices Using MachineLearning*. International Journal of Innovative Technology and Exploring Engineering, 8(10), pp. 3056-3061.
18. Armstrong, J. S. (2001). *Principles of Forecasting: A Handbook for Researchers and Practitioners*. Science & Business Media, Springer.
19. Willmott, C. J., & Matsuura, K. (2005). *Advantages of the Mean Absolute Deviation (MAD) over the Root Mean Squared Error (RMSE) in Assessing Average Model Performance*. Climate Research, 30(1), pp. 79-82.
20. Biau, G. & Scornet, E. (2016). *A random forest guided tour*. Test, 25(2), pp. 197-227.
21. Jha, K., Doshi, A., Patel, P., & Shah, M. (2019). *A comprehensive review on automation in agricultureusing artificial intelligence*. Artificial Intelligence in Agriculture, 2, pp. 1-12.
22. Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). *A review of the use of big data and machine learning insmart farming*. *Agricultural Systems*, 167, pp. 34-53.
23. Ben-Daya, M., Hassini, E., & Bahroun, Z. (2019). *Inventory management and supply chain coordination:A review*. Journal of Industrial Engineering and Management, 12(3), pp. 564-582.
24. Grinberg, M. (2018). *Flask Web Development: Developing Web Applications with Python*. O'Reilly Media.
25. Meyer, K. (2021). *UX Design for Beginners: Fundamentals of User Experience Design*. New Riders.
26. Nielsen, J. (1994). *Usability engineering*. Academic Press.
27. James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An introduction to statistical learning withapplications in R*. Springer.
28. Probst, P., & Boulesteix, A. L. (2018). *To tune or not to tune the number of trees in random forest.* Journalof Machine Learning Research, 18(1), 6673-6690.
29. Montgomery, D. C., Peck, E. A., & Vining, G. G. (2021). *Introduction to linear regression analysis*. Wiley.
30. Chakraborty, S., & Ghosh, I. (2021). *Real-time agricultural crop price prediction using machine learningmodels*. International Journal of Applied Agricultural Research, 16(3), pp. 123-140.
31. Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and practice*. OTexts.
32. Agriculture Marketing Information Network. (2021). *Market Information System* (AGMARKNET).Ministry of Agriculture, Government of India.
33. Little, R. J. A., & Rubin, D. B. (2019). *Statistical analysis with missing data*. Wiley.
34. Han, J., Kamber, M., & Pei, J. (2011). *Data mining: Concepts and techniques*. Morgan Kaufmann.
35. Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling.* Springer.
36. Grinberg, M. (2018). *Flask Web Development: Developing Web Applications with Python*. O'Reilly Media.
37. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O. & Duchesnay, E.

(2011). *Scikit-learn: Machine Learning in Python*. Journal of Machine Learning Research, 12, pp. 2825-2830.

38. Breiman, L. (2001). *Random forests*. Machine Learning, 45(1), 5-32.
39. Hutter, F., Hoos, H. H., & Leyton-Brown, K. (2011). *Sequential model-based optimization for general algorithm configuration.* In Learning and Intelligent Optimization. Springer, pp. 507-523.
40. Chai, T. & Draxler, R. R. (2014). *Root Mean Square Error (RMSE) or Mean Absolute Error (MAE). Arguments Against Avoiding RMSE in the Literature*. Geoscientific Model Development, 7(3), pp. 1247- 1250.
41. Soni, A., Patil, S., & Kumar, S. (2021). *A survey of cross-validation methods in data mining and machinelearning*. Journal of King Saud University - Computer and Information Sciences.