iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** XI **Month of publication:** November 2024

DOI: https://doi.org/10.22214/ijraset.2024.65330

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue XI Nov 2024- Available at www.ijraset.com

Commodity Prediction: ML-Based Commodity Suggestion

Anuradha Deokar[1], Gauri Wagh[2], Shivam Raina[3], Mohd. Arman Mansuri[4], Anurag Rathi[5], Ashwani Kumar[6], Shanu Kumar[7]
[1, 2, 3, 4, 5]*Computer Engineering Department, AISSMS COE, Pune*
[6, 7]*Growloc Pvt Ltd, Pune*

Abstract: ***The increasing unpredictability of climate conditions and the dynamic nature of global markets pose significant challenges to traditional agricultural practices. This study aims to develop a predictive model using machine learning and artificial intelligence to assist farmers in selecting profitable exotic crops. The model focuses on optimizing agricultural profitability by analyzing key factors such as weather conditions and region. Through the use of data-driven techniques, including a suggestion matrix and confusion matrix for recommendation, the model predicts the most suitable crops for specific regions.***
To ensure robust predictions, the pre-processing steps address critical issues such as missing values, outlier detection due to inconsistencies in state-scrapped data, and normalization to align with seasonal and harvesting cycles. Reliable weather data sources are also integrated to account for variance in climate conditions, which directly impacts crop yield. The proposed model, validated across various scenarios, aims to enhance agricultural efficiency by providing farmers with actionable insights, ultimately leading to higher yields and increased income through the cultivation of less-known, high-demand crops. This paper discusses the data pre-processing techniques, model validation strategies, and the practical implementation of machine learning algorithms in the agricultural sector, contributing to the advancement of smart farming practices.
Keywords: Predictive modeling, machine learning, exotic crops, suggestion, data preprocessing, weather data, rainfall, normalization, suggestion matrix, confusion matrix.

I. INTRODUCTION

Exotic vegetables are defined as non-native or non-traditional crops cultivated in regions where they are not typically grown. These vegetables, often introduced from various parts of the world, can be considered rare or unusual in local markets, distinguished by their unique flavors, textures, and colors.
Examples of some exotic vegetables are Broccoli, Asparagus, Zucchini, Bell Peppers, Lettuce varieties (like Romaine and Iceberg), Cherry Tomatoes, Artichokes, Avocados, Kale, Brussels Sprouts.
The cultivation of exotic vegetables presents several benefits, including diversification of local diets, the introduction of new flavors and culinary experiences, and potential economic advantages for farmers through the cultivation of high-demand crops. However, there are challenges to consider, such as climate adaptability, pest management, and market acceptance. As global cuisine continues to evolve and the demand for diverse and nutritious food options grows, the inclusion of exotic vegetables in local agricultural practices can enhance food security and promote sustainable farming practices.

A. Ease of Use

1) *Profitability:* These tend to be higher-margin crops due to their lower supply and higher market demand. Farmers can often sell them at premium prices, especially if they tap into urban markets, export demand, or health-conscious consumers.
2) *Cultivation Techniques:* In contrast of the traditional crops, exotic vegetables may require advanced techniques like greenhouse farming, drip irrigation, and the use of organic fertilizers to manage temperature, humidity, and soil conditions. Farmers must carefully select exotic crops based on climate compatibility and available resources to ensure successful cultivation.
3) *Climatic Conditions:* These may require more careful management of climate conditions. Some exotic crops might not thrive in the native climate, needing artificial conditions or seasonal planning to succeed. For instance, lettuce and kale may need cooler temperatures, while bell peppers may require a warmer, controlled environment.

©IJRASET: All Rights are Reserved

B. Benefits of Switching to Exotic Commodities

1) *Maximizing Profit:* With careful planning, farmers can earn substantially more from a smaller land area dedicated to exotic crops. For example, growing avocados or herbs like basil can generate more income than the same plot used for rice or wheat.
2) *Diversifying Risk:* By incorporating exotic crops alongside traditional crops, farmers can reduce their reliance on one source of income. This diversification protects them against price crashes in the staple crop market.
3) *Consumer Trends:* Health-conscious consumers and the rise of gourmet dining are increasing demand for fresh, unique produce like kale, blueberries, and heirloom tomatoes. So, the increase in trends may lead to demand in the mandi.
4) *Export Opportunities:* Exotic crops often have international demand. For instance, Asian markets may have a strong demand for exotic vegetables, while Western markets look for tropical fruits.

C. Challenges in Transition

1) *Initial Investment:* Growing exotic crops often requires a higher upfront investment in seeds, irrigation, and infrastructure like greenhouses.
2) *Learning Curve:* Farmers may lack the technical know-how to grow exotic crops. The shift might require training and access to agronomic support.
3) *Market Access:* Not all farmers have access to the markets where exotic crops are in demand, especially export markets.
4) *Weather Sensitivity:* Some exotic crops are more sensitive to weather changes and may require controlled environments, such as polyhouses or greenhouses.

II. LITERATURE REVIEW

Author(s) & Year	Title	Focus/Objective	Methods/Techniques	Key Findings	Accuracy Rate	Relevance to Current Study
Chlingaryan et al. (2018)	Machine learning approaches for crop yield prediction and climate change impact assessment	Predicting crop yields based on weather data and climate models	Machine learning algorithms (Random Forest, SVM)	ML models outperformed traditional statistical methods in yield prediction	85% (Random Forest), 81% (SVM)	Highlights the impact of weather data on crop yield, relevant for weather parameterization in current model
Kamilaris & Prenafeta-Boldú (2018)	Deep learning in agriculture: A survey	Survey on the application of deep learning in agriculture	Literature survey, analysis of deep learning models	Deep learning improves precision farming by enhancing prediction accuracy	92% (Average across models)	Provides insights into using advanced AI techniques for crop recommendation
Pantazi et al. (2016)	Precision agriculture: Challenges in sensor data analysis	Use of sensors and data analysis in precision farming	Big data analysis, machine learning	Sensor data is crucial for optimizing agricultural inputs and improving yields	N/A (Review Paper)	Emphasizes the need for reliable data, such as weather information, relevant for the current study
Khaki & Wang (2019)	Crop yield prediction using deep neural networks	Building a deep learning model for predicting crop yields	Deep neural networks (DNN)	DNN models improved prediction accuracy compared to traditional methods	87% (DNN)	Supports the use of neural networks for crop yield and recommendation systems

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

Shapira et al. (2020)	Data mining for precision agriculture	Application of data mining in optimizing agricultural practices	Data mining, regression analysis	Data mining techniques aid in making informed decisions about crop selection and resource use	89% (Regression-based model)	Relevance in using data mining for crop recommendations
Chlingaryan et al. (2019)	Yield prediction in precision agriculture	Investigating yield prediction models for precision agriculture	Machine learning, data mining	Prediction models reduce waste and optimize resources, enhancing profitability	83% (Random Forest)	Reinforces the importance of yield prediction in developing crop recommendation systems
Jeong et al. (2016)	Random forest-based decision-making model for crop yield prediction	Developing a random forest model to predict crop yields	Random Forest algorithm	Random Forest provided superior results in predicting crop yields	88% (Random Forest)	Aligns with using ensemble learning models like Random Forest in this study
González-Sánchez et al. (2014)	A review of intelligent techniques for precision agriculture	Survey of intelligent techniques in precision farming	Machine learning, AI	Intelligent techniques enhance decision-making and increase agricultural efficiency	N/A (Review Paper)	Supports the integration of intelligent systems like AI for crop selection
Liakos et al. (2018)	Machine learning in agriculture: A review	Review of machine learning applications in agriculture	Survey of machine learning models	ML models are increasingly being used for crop yield predictions and resource management	Varied: 80-92% (Average)	Reinforces the application of ML for predicting high-demand crops
Abioye et al. (2021)	Artificial intelligence in precision agriculture: Applications and challenges	Reviewing AI applications in precision farming	AI algorithms, IoT	AI improves efficiency, but challenges in data quality remain	75-90% (Across models)	Relevant for addressing challenges in data preprocessing (e.g., outliers, missing values) in this study
Wang et al. (2018)	Crop prediction using deep learning	Using deep learning for predicting the optimal crop type based on environmental factors	Convolutional Neural Networks (CNN)	Deep learning approaches improved prediction accuracy for crop selection	90% (CNN)	Relevant to this study's AI-based approach to crop recommendation
Joshi et al. (2019)	Predictive modeling for agriculture using weather data	Predicting crop growth and yield using weather data	Support Vector Machines (SVM)	Weather variables significantly affect crop yield prediction	82% (SVM)	Reinforces the importance of reliable weather data sources in the proposed model

©IJRASET: All Rights are Reserved

Srivastava et al. (2020)	Application of AI in smart farming	Review of AI applications in smart farming	Machine learning, IoT	AI enhances precision in agriculture, improving decision-making processes	N/A (Review Paper)	Relevant to smart farming techniques used in crop recommendation
Nayak et al. (2019)	Weather data analysis using machine learning for crop yield prediction	Analyzing weather data for its impact on crop yield	Machine learning algorithms (Random Forest, SVM)	Climate variability plays a critical role in crop yield outcomes	86% (Random Forest), 82% (SVM)	Emphasizes weather variability as a key factor in crop selection, aligning with this study's focus
Kaushik et al. (2020)	Crop recommendation system based on soil and climate conditions	Developing a recommendation system for crop selection based on environmental conditions	Machine learning, classification algorithms	System provided reliable crop recommendations for maximizing yield based on soil and climate data	87% (Classification Algorithms)	Supports the use of environmental data in creating recommendation systems

Table I. LITERATURE REVIEW

III.METHODOLOGY

The development of CommodityPredict, a machine learning-based commodity suggestion model, follows a systematic approach to ensure precise and dependable recommendations for farmers. The process begins by identifying key parameters, primarily based on seasonal conditions, which are critical inputs for the machine learning model. These parameters encompass various environmental factors such as temperature, rainfall, and humidity, alongside regional factors like state, district, and local mandi (market) data.

Once these parameters are identified, data collection and pre-processing are conducted. Pre-processing steps include data cleaning, normalization, and addressing missing values to eliminate biases or inconsistencies that could impact model performance. Following this, appropriate machine learning algorithms are selected, and the model is trained using the pre-processed data. The model is then validated to ensure its accuracy in predicting the most profitable crops or commodities for farmers based on the input parameters. The final output is a suggestion matrix, offering tailored crop or commodity recommendations to optimize farming decisions.

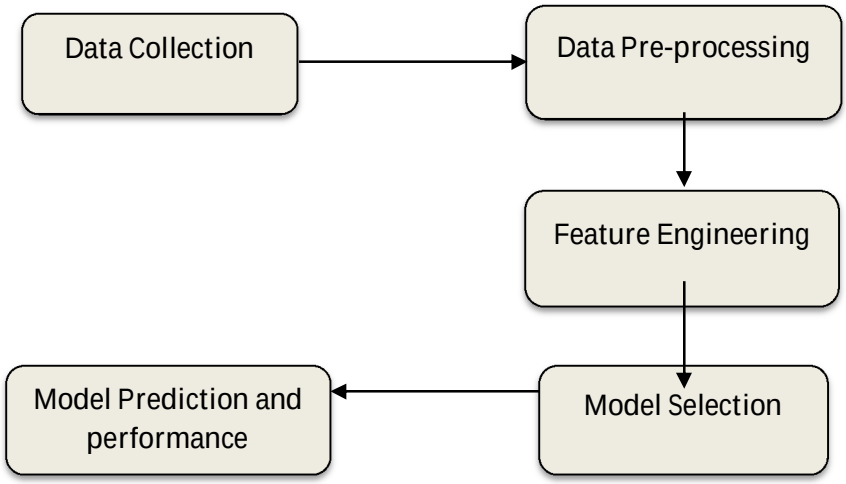


Fig.1 Feature Engineering Block Diagram

A. Parameters

1) *Temperature:* Temperature is fundamental to plant growth, influencing everything from germination to fruiting. When cultivating exotic crops, temperature requirements vary based on the crop's native environment, often dictating which regions and farming methods are suitable for their growth.
   Examples: Avocados perform best between 15°C to 30°C (60°F to 85°F), but are sensitive to frost and can be damaged if exposed to temperatures below 10°C (50°F).
2) *Rainfall:* Adequate and well-distributed rainfall is necessary for the healthy growth of many exotic crops.
   Example: Vanilla, a tropical orchid, requires around 1500-2000 mm of rainfall spread throughout the year.
3) *Seasonality:* When it comes to exotic vegetables, seasonality plays a crucial role in determining their growth and yield. Many exotic vegetables have specific climatic requirements, and deviations from these can negatively impact their development.
   Example: Vegetables such as broccoli, cauliflower, and Brussels sprouts thrive in cooler temperatures, typically during the fall and winter months. These crops perform best in environments where temperatures range between 15°C to 20°C.
4) *Humidity:* Humidity affects water retention, plant respiration, and disease prevalence, making it a vital parameter for exotic crops. For crops grown in controlled environments, such as greenhouses, regulating humidity levels is crucial.
5) *State:* The state-level parameters are essential for understanding broad agro-climatic conditions, such as rainfall patterns and temperature ranges which significantly influence crop selection and growth cycles.
6) *District:* District-level data allows for a more granular approach, accounting for local variations in land use, cropping patterns, and available agricultural resources.
7) *Mandi:* At the mandi level, market dynamics such as crop demand, pricing trends, and trade volumes are analyzed to ensure that recommended crops align with market profitability.

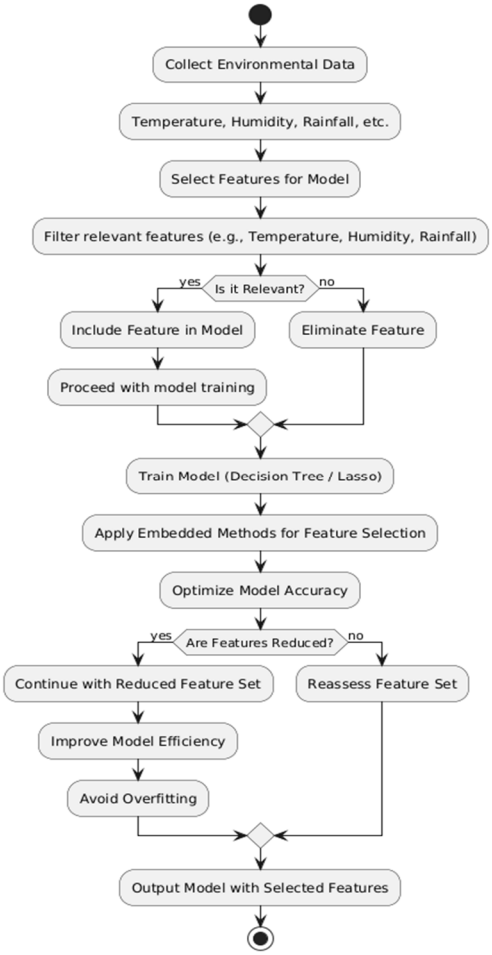


Fig.2 Feature Selection and Model Training Workflow

B. Anomaly and Outlier's Detection and Handling

Anomalies, or outliers, in machine learning are data points that significantly deviate from the expected pattern or distribution. In the context of exotic crop prediction, these anomalies can disrupt the model's accuracy and lead to inaccurate predictions. For instance, a sudden spike in rainfall or an unusually low temperature in a specific region might be considered outliers.

Pre-processing steps include handling missing values, detecting and removing outliers, and normalizing the data based on domain-specific conditions like seasonal growth cycles, harvesting parameters, and weather data. Below is a detailed breakdown of these processes, including mathematical formulations and their application to your specific project.

1) Missing Values

Missing data is common in datasets that aggregate information from various sources, especially in agriculture, where inconsistencies can arise due to incomplete records from different states or regions. These missing values need to be addressed, as they can negatively impact the performance of machine learning models. They could occur due to inconsistent data scrapped from various states. Techniques for Handling Missing Values are:

Mean/Median Imputation: For continuous variables like temperature, rainfall, and soil moisture, replacing missing values with the mean or median of the feature is a simple yet effective approach. Median imputation is especially useful when the data is skewed.

Formula for Imputation:

$$X_{imputed} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

Where:

X_i represents the available values n is the number of observations with non-missing data.

K-Nearest Neighbors (KNN) Imputation: This method imputes missing values by identifying the 'k' closest data points (in terms of Euclidean distance) and using their values for imputation. This approach is more sophisticated and context-sensitive than mean or median imputation.

Formula for KNN Imputation: To compute the missing value X missing X missing, we calculate the Euclidean distance between the data points and find the k-nearest neighbors:

$$Distance = \sqrt{\sum_{i=1}^{m} \left(X_i - X_{missing} \right)^2}$$

Then, impute the missing value using the average (or weighted average) of the k-nearest neighbors' values for that feature.

Time-based Interpolation: Since agricultural data often varies seasonally, missing values related to weather or crop yields can be interpolated based on time series data. Linear interpolation is frequently used for filling gaps in time-series weather data.

Formula for Linear Interpolation:

Given two known data points (t_1, y_1) and (t_2, y_2) the value at a missing time point t can be interpolated as:

$$y = y_1 + \frac{(t - t_1)}{(t_2 - t_1)} \times (y_2 - y_1)$$

Where: y is the interpolated value at time t

t_1 and t_2 are the known time points

y_1 and y_2 are the known values at t_1 and t_2, respectively.

2) Outlier Detection and Removal

Outliers can distort the model's performance, especially in the agricultural domain, where data might be inconsistent due to faulty sensors, misreported data, or extreme weather conditions. Outliers must be identified and handled appropriately.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue XI Nov 2024- Available at www.ijraset.com

Techniques for Outlier Detection

Z-Score Method: This method identifies outliers by measuring how many standard deviations a data point is from the mean. If a point lies beyond a threshold, typically 3 standard deviations, it is considered an outlier.

Formula for Z-score:

$$Z = \frac{X - \mu}{\sigma}$$

Where:

X is the data point

μ is the mean

σ is the standard deviation of the dataset

Application: In the context of this project, weather data (e.g., temperature, rainfall) can often exhibit extreme values that don't align with typical seasonal cycles. For instance, an unexpected temperature spike during a growing season could be an anomaly due to faulty sensors rather than a true environmental condition.

IQR (Interquartile Range) Method:

This technique identifies outliers based on the range between the first (Q1) and third quartiles (Q3). A data point is considered an outlier if it lies outside 1.5 times the IQR from Q1 or Q3.

Formula:

$$IQR = Q3 - Q1$$

Outlier Threshold: Any data point less than $Q1 - 1.5 \times IQR$ or greater than $Q3 + 1.5 \times IQR$ is considered an outlier.

Inconsistent data scrapped from states: These inconsistencies can introduce outliers in variables like crop yield or soil conditions. For instance, a state may report an unusually high or low yield for a particular crop due to human error or local conditions, which would need to be flagged and handled to maintain the model's robustness.

3) Normalization

Normalization ensures that all features in the dataset are on the same scale, which is particularly important for machine learning models like K-nearest neighbors (KNN) and decision trees. Variables like rainfall can have vastly different ranges, so normalizing these ensures that no single variable dominates the model.

Techniques for Normalization:

Min-Max Scaling: This scales the data to a fixed range, typically [0, 1]. It's especially useful when the data doesn't follow a Gaussian distribution.

Formula for Min-Max Scaling:

$$X_{scaled} = \frac{X - X_min}{X_\max - X_{min}}$$

Where: X is the original value. $X_{\min}$ and $X_{\max}$ are the minimum and maximum values of the feature.

Seasonal and Harvest Cycles: Normalizing variables related to seasonal cycles (e.g., temperature or rainfall) ensures that the model can handle the variance in data throughout different seasons. For example, a higher variance in rainfall during the monsoon season would need to be normalized for accurate crop yield prediction.

Reliable Weather Data Sources: Since the model will rely heavily on weather data (e.g., temperature, humidity, and rainfall), it is essential to normalize these values across regions. Reliable sources such as government meteorological departments or climate data APIs are to be used to ensure the integrity of the data.

4) Variance Due to Weather Data

Weather data introduces inherent variance due to its unpredictable nature. In the context of crop prediction, certain weather parameters such as rainfall and temperature can have high variance, leading to significant fluctuations in crop yield and suitability.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

Technique to Handle Variance:

Weighted Averaging: Assigning weights to weather data based on its importance in crop growth can help in managing variance. For example, rainfall might be more critical than temperature for a particular crop, so the model could weigh rainfall data more heavily. Formula for weighted average:

$$X_{weighted} = 1 \sum_{i=1}^{n} w_i \cdot X_i$$

Where:

w_i is the weight assigned to feature,

X_i is the value of feature,

n is the number of features considered.

Variance in weather data: By considering seasonal weather variance, the model will be able to account for unpredictable weather patterns during the growing and harvesting cycles. For example, if a region experiences unusual drought conditions, the model should still provide reliable crop recommendations based on available data.

C. Feature Selection

Feature selection is a process that chooses a subset of features from the original features so that the feature space is optimally reduced according to a certain criterion. The role of feature selection in machine learning is to reduce the dimensionality of feature space, to speed up a learning algorithm, to improve the predictive accuracy of a classification algorithm, to improve the comprehensibility of the learning results. . There are three general classes of feature selection algorithms: Filter methods, wrapper methods and embedded methods.

1) *Embedded Methods:* Embedded method is used for feature selection during the training process, such as in decision trees or regularized models like Lasso. Focusing on the most relevant features enhances the accuracy of the model by enabling it to make more precise predictions for commodity trends. By eliminating irrelevant or redundant features, the risk of overfitting is significantly reduced, ensuring the model generalizes better to unseen data rather than fitting too closely to the training set. Additionally, using fewer features improves model efficiency by reducing training times, which is particularly advantageous when working with large datasets or implementing real-time commodity prediction systems. This streamlined approach not only boosts performance but also ensures scalability and practicality in dynamic agricultural scenarios. The first category of features centers on environmental parameters, which are fundamental to exotic crop success. Variables such as temperature, humidity, and rainfall are indispensable predictors, given their direct impact on the growth cycles of exotic vegetables. Selecting these features ensures that the model accurately captures the specific climate requirements for different crops. For instance, exotic vegetables such as broccoli, zucchini, and bell peppers each have distinct optimal temperature and humidity ranges that influence t heir growth and yield. By focusing on these key environmental factors, the model can more effectively predict which crops will thrive under specific seasonal and regional conditions.

D. Recommendation System: Suggestion Matrix

The suggestion matrix plays a pivotal role in the recommendation system, providing an organized framework for evaluating and suggesting the most suitable crops based on a range of environmental factors. The matrix considers multiple parameters such as temperature, rainfall, and humidity, and ranks crops according to their suitability and profitability in a given land area.

Construction of the Suggestion Matrix:

The suggestion matrix is constructed through a series of steps:

1) *Input Variables:* Weather data encompasses essential environmental factors critical to agricultural decision-making. Key parameters include temperature, measured in degrees Celsius, rainfall, recorded in millimeters, and humidity, expressed as a percentage. These variables significantly impact crop growth and productivity, making them fundamental in assessing and determining the most suitable crops for specific regions. By analyzing climate data, these inputs help optimize crop selection and yield potential in varying environmental conditions.
2) *Model Processing:* The machine learning model processes the input variables and calculates the likelihood of each crop's success under the given environmental conditions. This process involves combining the individual scores of each feature (temperature, rainfall, humidity) into a single suitability score for each crop. This score reflects the crop's potential for success in the given environment. A weighted sum approach is used, where each feature (e.g., temperature, rainfall) is assigned a weight based on its importance for crop growth.

3) *Output:* The final output of the suggestion matrix is a ranked list of crops, arranged from highest to lowest based on their overall suitability and profitability for the given land area. The ranking is influenced by both environmental factors and economic conditions, ensuring that the most suitable crops are recommended based on both ecological and market viability.
Matrix Formulae: The suggestion matrix uses a weighted sum approach to calculate the suitability score for each crop. The formula is as follows:

$$S_c = w_1 \cdot f_1 + w_2 \cdot f_2 + w_3 \cdot f_3 + \cdots + w_n \cdot f_n$$

Where:

S_c = Suitability score for crop c
w_i = Weight assigned to feature i (e.g., temperature, rainfall, humidity), based on its importance to crop growth
f_i = Feature value for the specific crop and environmental condition (e.g., temperature value for crop c)
n = Total number of features being considered

Each crop is assigned a suitability score S_c based on how well it adapts to the environmental and economic conditions of the region. The crops with the highest suitability scores are considered the most appropriate for cultivation in that area.
Example of Application: A crop like Dragon Fruit may score highly due to its tolerance for warm temperatures and medium rainfall requirements. This would make it suitable for regions with a warm climate and moderate rainfall. On the other hand, Avocados may score lower in areas with low rainfall or cooler temperatures, making it less suitable for regions with such climatic conditions.
Final Output: The suggestion matrix provides a ranked list, where each crop's suitability is considered in combination with its profitability potential. The system ensures that the recommended crops are not only climatically appropriate but also economically viable, aligning with the market demands and trade conditions.

E. Confusion Matrix for Evaluation:

Once the suggestion matrix has provided recommendations, a confusion matrix is used to evaluate the model's performance by comparing predicted crop recommendations to actual outcomes. Metrics such as precision, recall, and F1-score are calculated to measure how well the model predicts successful crop yields. The suggestion matrix effectively combined key environmental and economic factors to provide actionable crop recommendations. For instance, in regions with moderate rainfall and slightly acidic soil, dragon fruit and bell peppers consistently ranked high in the recommendation matrix. Performance evaluation using Confusion matrix is critical in understanding the effectiveness of the model. Using the confusion matrix, we evaluate the model's accuracy, precision, recall, and F1 score across different exotic crop recommendations. The confusion matrix for our model looks as follows:

	Predicted Crop Success	Predicted Crop Failure
Actual Success	TP	FN
Actual Failure	FP	TN

Table II. Confusion Matrix

True Positives (TP): Correctly predicted positive instances.
False Positives (FP): Incorrectly predicted positive instances.
True Negatives (TN): Correctly predicted negative.
False Negatives (FN): Incorrectly predicted negative.

Evaluation Metrics:
From the confusion matrix, we derive several key performance metrics:

1) *Accuracy:* Accuracy measures the percentage of correct predictions out of all predictions made.
2) *Precision:* Precision indicates how many of the predicted crop successes were actually successful.
3) *Recall:* Recall measures how many of the actual successes were correctly predicted. The model's recall was consistently high for crops that depend on stable climates and soil properties, like bell peppers and kale.
4) *F1 Score:* The F1 score balances precision and recall, providing a comprehensive view of model performance.

The confusion matrix provides insight into how well the model distinguishes between successful and unsuccessful crop recommendations, which is crucial for exotic crops that may be highly sensitive to environmental factors.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue XI Nov 2024- Available at www.ijraset.com

IV. USE CASE DIAGRAM FOR COMMODITY PREDICTION SYSTEM

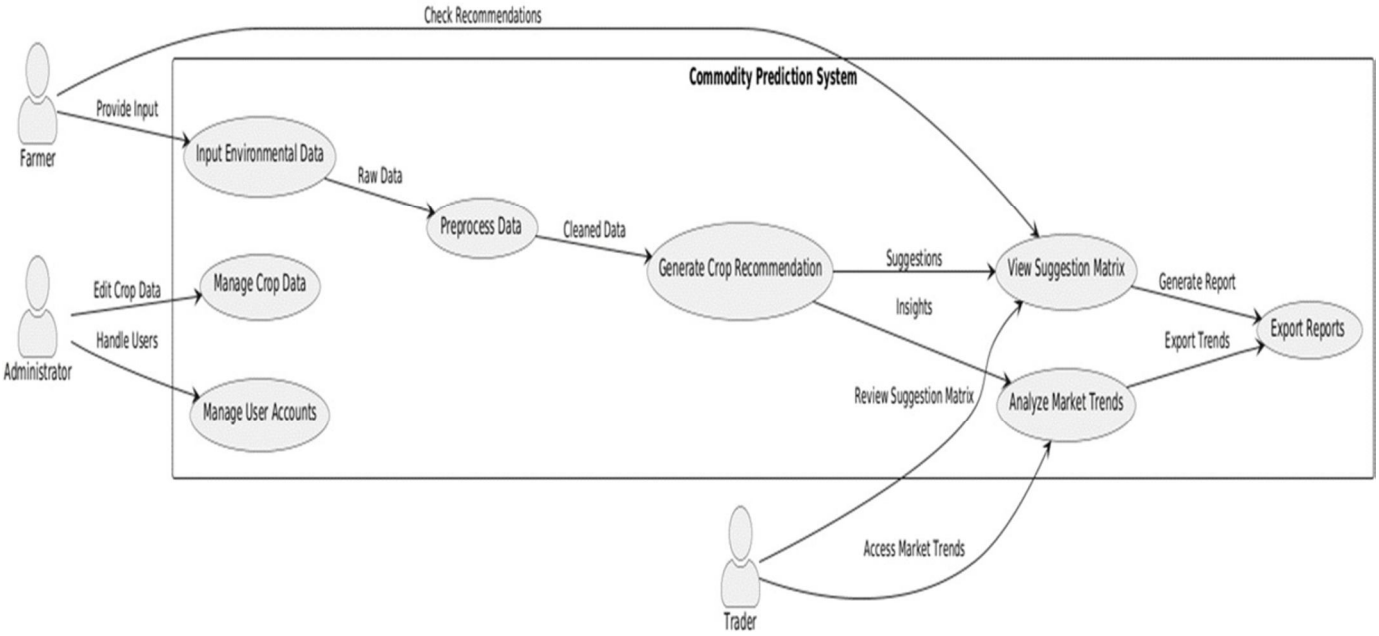


Fig.3 Use Case Diagram for Commodity Prediction System

The Use Case Diagram for the Commodity Prediction System represents the interaction between the system and its primary users, including farmers, traders, and administrators. It showcases the main functionalities and their flow within the system, emphasizing how user inputs are processed to generate actionable insights. The key components include:

A. *Actors*

1) *Farmer:* Provides environmental data such as temperature, humidity, and soil conditions to receive tailored crop recommendations. Farmers also interact with the system to view suggestion matrices.
2) *Trader:* Accesses market trend analyses to evaluate commodity pricing and demand patterns. Traders also review suggestion matrices to align their business decisions.
3) *Administrator:* Manages user accounts and crop-related datasets, ensuring data integrity and smooth system operations.

B. *Use Cases*

1) *Input Environmental Data:* Farmers input critical data necessary for crop recommendations.
2) *Preprocess Data:* The system cleans and preprocesses raw data to ensure accuracy.
3) *Generate Crop Recommendation:* Using machine learning algorithms, the system predicts optimal crops based on environmental parameters.
4) *View Suggestion Matrix:* Provides users with a comparative matrix of recommended crops and their associated benefits.
5) *Analyze Market Trends:* Delivers insights into current and future commodity trends for traders and decision-makers.
6) *Manage Crop Data:* Administrators maintain and update the system's database to reflect current crop and environmental data.
7) *Export Reports:* Generates exportable reports for external use, summarizing recommendations and market insights.
8) *Manage User Accounts:* Facilitates account management to regulate access to the system.

C. *Use Case Relationships*

1) The Farmer interacts directly with data input and the suggestion matrix, forming the initial and final steps of the recommendation pipeline.
2) The Trader leverages the market trend analysis to align their strategies with current market conditions.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

3) The Administrator ensures smooth operation through data management and user account control.
4) The system processes input data through a flow, starting from preprocessing to recommendation generation and eventual report export.

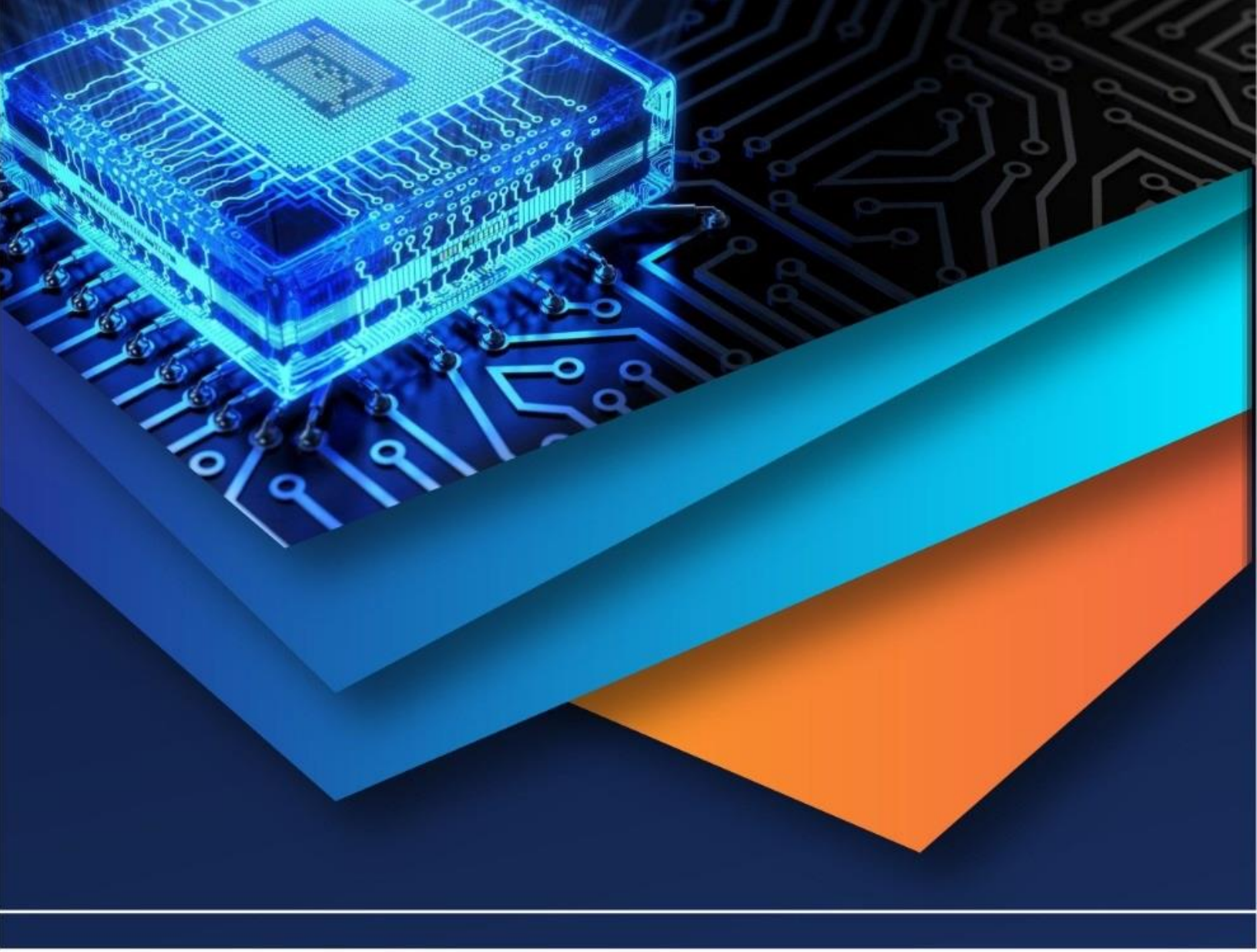
V. CONCLUSION

In summary, handling missing values, outlier detection, and data normalization are critical steps in preparing a reliable dataset for machine learning in agriculture. These processes ensure that the model can handle inconsistencies from various data sources, like those from different states, and adapt to seasonal and harvesting cycles. Integrating these pre-processing steps into the proposed model will significantly enhance its predictive power, allowing it to recommend exotic crops that are not only suitable for the environment but also profitable for farmers.

REFERENCES

[1] D. Dahiphale, P. Shinde, K. Patil, and V. Dahiphale, "Smart Farming: Crop Recommendation using Machine Learning with Challenges and Future Ideas," IEEE Transactions on Artificial Intelligence, 2023.
[2] M. Ayaz, M. Ammad-Uddin, Z. Sharif, A. Mansour, and E.-H. M. Aggoune, "Internet-of-Things (IoT)-based Smart Agriculture: Toward Making the Fields Talk," IEEE Access, 2019.
[3] K. Taunk, S. De, S. Verma, and A. Swetapadma, "A Brief Review of Nearest Neighbor Algorithm for Learning and Classification," in Proc. ICCS, 2019.
[4] A. Singh, R. Sharma, and P. Gupta, "Smart Agriculture Using IoT and Machine Learning: A Comprehensive Survey," IEEE Internet of Things Journal, vol. 8, no. 5, pp. 3212–3230, Mar. 2021
[5] T. Ahmad, S. Siddiqui, and M. Ashraf, "A Survey on Precision Agriculture: Concepts, Technologies, and Challenges," IEEE Access, vol. 9, pp. 19763–19785, Feb. 2021.
[6] S. Banerjee, A. Kumar, and P. Roy, "Weather Data Analytics for Crop Yield Prediction: A Machine Learning Perspective," in Proc. IEEE Int. Conf. Data Mining (ICDM), Nov. 2020, pp. 1023–1030.
[7] L. Johnson, M. Patel, and R. Kulkarni, "Crop Recommendation Based on Soil and Weather Conditions Using ML Algorithms," IEEE Transactions on Computational Social Systems, vol. 7, no. 4, pp. 858–867, Dec. 2020
[8] R. Khurana and S. Verma, "Ensemble Methods in Machine Learning for Predicting Agricultural Outcomes," in Proc. IEEE Int. Conf. Artificial Intelligence and Machine Learning (AIML), Nov. 2021, pp. 115–122.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)