OPEN ACCESS

International Journal of Innovations in Science & Technology

RESEARCH & INNOVATION DIVISION

IJIST

Artificial Intelligence-Based Approach for The Recommendations of Mango Supply Chain

Hamza Hussain[1], Hira Nazir[2], Muhammad Samiullah[3*], Muhammad Danial Faiz[4], Muhammad Adnan Faiz[5], Adnan Manzoor[5]

[1] Muhammad Nawaz Shareef University of Agriculture, Multan.

[2]Department of Cyber Security, Emerson University, Multan, Pakistan.

[3]Govt. Graduate College of Commerce, Qasimpur. HED, Multan, Pakistan

[4]Department of Mathematics, Institute of Southern Punjab Multan

[5]Department of Business & Public administration, Emerson University, Multan, Pakistan

*** Correspondence**: Muhammad SamiUllah, samiiub@gmail.com.

Citation | Hussain. H, Nazir. H, Samiullah. M, Faiz. M. D, Faiz. M. D, Manzoor. Adnan, “Artificial Intelligence-Based Approach for The Recommendations of Mango Supply Chain”, IJIST, Vol. 06 Issue. 04 pp 1913-1931, Nov 2024

Received | Oct 26, 2024 **Revised |** Nov 19, 2024 **Accepted |** Nov 20, 2024 **Published |** Nov 21, 2024.

This study utilizes a comprehensive dataset that encompasses variables reflecting temperature, humidity, precipitation, inventory levels, transportation modes, freshness scores, and ripeness scores. Compiled from various mango farms across different markets, this dataset provides a robust foundation for our analysis. To develop predictive models, we employed several machine learning algorithms, including Support Vector Machines (SVM), K-Nearest Neighbors (KNN), Random Forests (RF), and Decision Trees (DT). We divided the dataset into training and testing sets, using an 80-20 split for training and testing subsets, respectively. Model performance was evaluated using metrics such as accuracy, precision, recall, and F1 score. Our results indicate that Random Forests outperformed other models, achieving the highest accuracy, precision, recall, and F1 scores. A feature importance analysis revealed specific features that contributed significantly to the performance improvements of the model. These insights into feature importance can aid in refining the model's performance, making feature importance analysis a valuable component of model evaluation.

Support Vector Machines	(SVM)
K-Nearest Neighbors	(KNN)
Random Forests	(RF)
Decision Trees	(DT)
Artificial Intelligence	(AI)
Radial Basis Function	(RBF)
Euclidean Distance	(ED)
Receiver Operating Characteristic	(ROC)
Area Under the Curve	(AUC)

Keywords: Artificial Intelligence (AI), Mango, Supply Chain Optimization, Agricultural Predictive Modeling.

Introduction:

The mango supply chain is a core element of the international agriculture structure than a source of living for numerous farmers and a delight for consumers worldwide. Mango as the queen of fruit varieties is not only valued for its exquisite taste but also for its nutritional quality. Mango has a special place in the fruit market. While mango remains a favorite commodity. The mango supply chain faces a wide range of problems that are among the causes of its inefficiency and low returns on investments [1]. However, the main problems are related to the nature of mango, which is seasonal and as a result has fluctuations in the quantity and quality during the year. In addition to the inherent perishability of mangoes, the short shelf life that quickly diminishes their quality and freshness is the constant challenge that inevitably occurs throughout the supply chain. Ahead of that, external factors like unpredictable weather, period market requirements, and logistics difficulties create further barriers in the more complicated mango distribution [2]. The traditional techniques of the mango supply chain in the market are highly likely to be inadequate because they are not precise and have a low level of adaptability which is the nature of the complexities of the industry. The existing processes and human-oriented decision-making that were once efficient are now hindering progress, which leads to limited results and inefficiencies in supply chain management.

Mango is among the prominent fruits that anchor the agricultural sectors across global economies for local use and export. Nevertheless, the mango supply chain is an important subsector in the global economy, but the supply chain experiences various problems that are attributed to the fluctuations in demand, the complex logistics, and the lack of proper decision-making systems [3]. Through the use of AI-based solutions, key players will be able to receive analytics, and suggestions that are driven by intelligence that in turn enable data-based decision-making, efficient resource allocation, and reduced risks that are inherent in the operations of the mango supply chain. Besides, the advent of emerging technologies like (IoT) devices and blockchain creates a platform to trace, ensure transparency, and account for activities throughout the supply chain [4]. In the background of this, the study aims to investigate the AI-based solutions that can be applied to the mango supply chain system which involves various challenges. This is not just about searching for existing inefficiencies and cost-saving options but it is a way to build a strong foundation for future growth and to survive the changes in the mango market system. Several issues need to be addressed in the mango supply chain, beginning with the complexity of its ecosystem which consists of seasonality and perishability among many other challenges. The conventional supply chain management models fail to address the dynamic nature of mango production, distribution, and sales, which results in undesirable consequences such as overstocking, underutilization of resources, and high wastage [5].

Intensities of a mango supply chain in most cases lead to extra expenses that are incurred as a result of overstocking, unsuitable transportation routes, and poor inventory management. The research is based on the AI-driven solution approach thus it has high hopes for cost reduction. Through the improvement of routes, the erosion of inventory holding costs, and the restoration of quality-related losses, stakeholder might accomplish substantial savings as their service levels and product quality can be maintained [6]. Customer delight is an essential factor in the fast-paced mango market. Time-efficient and top-notch quality mangoes are mandatory to meet the consumers' expectations. AI can be employed to accurately forecast demand and also ensure the necessary quality controls, all of which are geared towards customer satisfaction. Stakeholders can achieve this by offering fresh mangoes and minimizing quality issues, creating a platform upon which consumers can trust and, hence, being able to compete favorably in the market [2]. In the modern agriculture landscape, sustainable production has become more and more essential. Inefficient supply chain operations not only bring economic losses but also are the cause of environmental deterioration due to the production of more waste and the use of extra resources. AI-based drive solutions can be very effective in terms of process improvement as well as a sustainable supply chain of mangoes. Stakeholders can help in the realization of long-term sustainability goals by encouraging resource conservation and decreasing the amount of pollution. This will be done without the need to go against profitability [3]. AI-powered optimization of mango, supply chain management is a new and important trend with the potential to influence the future of supply chain management practices efficiently. The research can capitalize on the latest technologies and data-centric approaches to this effect, which advances the state of the art in supply chain management approaches. The outcomes and exemplary findings of this research have the possibility of applying to different industries, facilitating the emergence of new ideas and encouraging continuous improvements in global supply chain management [4]. This study has a leading thrust to bridge the gaps and inefficiencies in the mango supply chain. However, the mango value chain has a highly significant role in the agricultural sector yet it experiences many complexities such as seasonality, perishability, and demand volatility. The aforementioned methodologies tend to fail to efficiently handle these problems, therefore leading to an increase in waste, higher costs, and poor performance. AI-based algorithms benefit the stakeholders by utilizing machine learning, predictive analysis, and business intelligence in the ability to forecast the shift in the demand and supply gap, inventory management and supply chain management more efficiently [7].

Furthermore, new recommendation systems that have been developed from the latest AI-based models are novel for the continuous development of individual and contextual management for the stakeholders of the mango chain. Therefore, it can be concluded that there is a great opportunity to transform the value creation in the mango industry by using AI technologies, including the improvement of its profitability, sustainability, and adaptability to the changes in the market environment [8]. Although this study is an effort at categorization and description of the mango production, distribution, and consumption system, it aims at providing new solutions generated in the form of algorithms and decision support systems that are useful to the players in the industry. Besides, since the study adopts mobile application technologies in the research, the study seeks to ensure that the AI information is availed to farmers, distributors, and retailers as they make decisions in the field. Therefore, this study will play a role in developing a connected and efficient supply chain for mangoes to promote the flow of socioeconomic values and to transform the sector to stand the global challenges [9]. Inventory control is also underlined in the study; attention is also paid to the identification of activities that may be pursued to increase the rate of stock turnover without a negative influence on the ability of the organization to meet client's demands. As such, by applying the AI-based approach, the stakeholders can reduce the holding cost and avoid stock-outs while identifying the right replenishment method to satisfy the customers' demand using the overall lowest cost [10].

The mango supply chain exists in a system environment where it is confronted with issues such as seasonal and perishable products. Thus, applying AI techniques to the analysis of this study will help to achieve a substantial increase in the efficiency of operations at all stages of the supply chain. By using sophisticated demand planning, route planning, inventory planning, and quality planning, the stakeholders can eliminate redundancies, minimize wastage, and enhance the utilization of resources [11]. The application of AI techniques has the potential to improve the quality of foods and fruits such as mangoes. The authors in [12] also explain how AI can enhance the quality of food through modeling. Their study supports the use of artificial intelligence solutions in the improvement of different factors affecting the food supply chain, such as food quality and product development. The research study to design recommendation systems with the help of AI technologies such as machine learning, and predictive analytics would control the flow of inventory and the routes of distribution by the trends observed in the market. This AI recommendation will help the stakeholders in the mango chain to execute their activities with low operation costs, less wastage, and high productivity [13]. By implementing the AI approach in the mango supply chain, it will be ready to meet the current demand as well as be ready to evolve in the ever-changing business environment [14]. The research conducted by [15], was intended to open new horizons in the management of the mango supply chain by incorporating AI technologies. This study therefore aimed at compiling a global database of the factors that affect mango production, distribution, and consumption using a rich variety of data sources such as satellite images, IoT sensors, and market intelligence reports. Therefore, using deep reinforcement learning and swarm intelligence optimization, the author attempted to design adaptive recommendation systems for enhancing the key supply chain processes in real-time. By applying these AI recommendations, the stakeholders in the mango chain could improve productivity, reduce losses, and thereby provide better quality fresh mangoes to the global consumers.

To assess, how the application of AI-based supply chain recommendations can revolutionize mango distribution. Thus, based on big data including factors like harvest yields, transportation logistics, market tendencies, and consumer preferences the author created complex AI algorithms to enhance distribution procedures. Predictive analysis and optimization were used to give the stakeholders suggestions on the best distribution routes, warehouse locations, and inventory stocking levels. Through implementing these recommendations, the author sought to transform mango distribution through the use of AI by increasing the efficiency of operations hence cutting costs while at the same time increasing the levels of satisfaction among customers hence a competitive edge in the market [16]. The developed AI recommendation system aimed at improving supply chain visibility, transparency, and traceability by using real-time data streams from IoT sensors and blockchain-enabled traceability systems. This intelligent approach to implementing AI can help the stakeholders in the mango supply chain to make better decisions on time concerning the quality of the products, the location of the products, and the level of compliance with the set standard [17] [18]. With the help of AI, recommendation systems for partners were given real-time information about the distribution channels, the best inventory management, and delivery schedule. According to the recommendations made by the author, he sought to transform mango distribution strategies through the simplification of distribution channels, reduction of transport costs, and improvement of the supply chain's flexibility to deliver better, fresher mangoes to consumers around the globe [19][20][21][22][23].

AI technology is becoming the engine of the mango supply chain with the potential for transforming its existing operations and addressing the numerous challenges the industry is facing. AI technologies comprised of machine learning, predictive analytics, and optimization algorithms to improve decision-making processes get the best from resource allocation, and manage the risk associated with unpredictable factors like weather conditions and market fluctuations exist [24]. The combination of IoT devices and blockchain technology offers great opportunities in the field of tracking, transparency, and building trust within the supply chain system [25]. Therefore, the research objectives are derived from the possibility of developing AI-based inventions that can respond to the mango supply chain's specific setbacks, and as a consequence, make operations more efficient, minimize waste, and improve the overall output. The mango supply chain is heavily loaded with complexities that are the result of its seasonality, the fruits' short shelf life, and the uncontrollable nature of consumer demand. Using the backbone of the traditional supply chain the management turns out to be very difficult for the challenges, which leads to the inefficiency issues like overstocking, underutilized transports, and excessive waste [26]. One of the studies recommended upgrading the existing low-cost equipment, tools, and practices by using AI-based algorithms to allow actors to maximize their incomes and achieve the objective of loss reduction [27]. An online supply chain system comprised of the blockchain network makes use of smart contracts and it is accessible to the stakeholders of the supply chain to ensure trust and authenticity. Thus, farmers will be able to reduce losses, and get profits through direct access to retailers, consumers, etc. [28]. Hence, the absence of correct forecasting tools on the one hand makes the problem more serious, and on the other hand, it results in a wrong allocation of resources and missed revenue growth opportunities. These challenges therefore make stakeholders in the mango industry struggle to find ways of dealing with these issues while at the same time trying to compete and remain profitable in the market. While recognizing the urgent call for novel techniques and approaches, the study intends to use AI as a force for change, revolutionizing the mango supply chain. Utilizing Artificial Intelligence methods like machine learning, predictive analytics, and optimization algorithms, the focus is to build a system that can generate intelligent recommendations based on various parameters of the supply chain, beginning from procurement to distribution. The AI-driven approach that combines diverse data sources including weather forecasts, market trends, and historical sales figures is intended to provide decision-makers with useful information that will lead to better decision-making. Given a set of locations representing mango distribution centers and the distances between them, the task is to find the most efficient route for delivering mangoes to customers while minimizing transportation costs.

This research study focuses on optimizing various aspects of the mango supply chain using (AI) techniques, including demand forecasting, route optimization, inventory management, and quality control. The scope encompasses the development and implementation of AI-based models and algorithms tailored specifically to address the unique challenges faced by stakeholders within the mango industry. Furthermore, the research explores the integration of emerging technologies such as IoT devices and blockchain to enhance traceability, transparency, and trustworthiness in the supply chain ecosystem.

Mathematical Formulation:

Let N be the set of locations (distribution centers and customers), and $\mathrm{c_{ij}}$ be the distance between locations i and j. The objective is to minimize the total distance traveled while visiting all locations exactly once by using a mathematical equation:

$$\text{minimize} \sum_{i=0}^{n} \sum_{j=1\, j\neq 1}^{n} c_{ij}x_{ij} \qquad (1)$$

Subject to $\sum_{j=1\, j\neq 1}^{n} x_{ij} = 1$, for all i, where x_{ij} is a binary decision variable indicating whether

there is a connection between nodes i and j.

$\sum_{i=0\, i\neq j}^{n} x_{ij} = 1$, for all j, where x_{ij} is a binary decision variable indicating whether there is a connection between nodes i and j.

Here, X_{ij} is in binary form, representing whether there is a connection between nodes i and j.

The objective function aims to minimize the total distance traveled. The expression $\sum_{i=0}^{n} \sum_{j=1\, j\neq 1}^{n} c_{ij}x_{ij}$ computes the total distance by summing the distances between all pairs of locations multiplied by binary decision variables x_{ij}, which indicates whether there is a connection between locations i and j. By minimizing this objective function subject to the constraints, we can find the optimal route for delivering mangoes, ensuring efficient distribution and minimizing transportation costs.

Objectives:

The primary objective of this study is to apply and evaluate machine learning techniques in quality control of the mango supply chain. Specifically, we aim to classify mango quality based on temperature, humidity, and transportation conditions, ultimately leading to enhanced customer satisfaction and reduced defective products.

Novelty Statement:

This study contributes to the understanding of predictive modeling in agriculture and food supply chains by evaluating different machine learning models and analyzing their performance and feature importance. Our novel approach involves applying machine learning-based algorithms for route optimization, identifying critical features, and recommending best-fit models to drive better decision-making and resource allocation in mango farming and distribution processes.

Materials and Methods:

Our dataset, obtained from Kaggle, comprises multiple attributes, including weather conditions, food inventory levels, logistics, quality control, and market demand trends. We aimed to acquire a comprehensive dataset covering various aspects of mango production and supply chain management.

Data Sources:

Weather Forecast Data: Temperature, humidity, and precipitation data were obtained from meteorological offices or online weather services. This information is crucial for determining the condition and ripeness of mangoes during transportation.

Market Demand Trends: We conducted market surveys to determine product demand data, including the number of mango units required by identified sales outlets. This data was obtained from market reports, industry databases, or past sales records analysis.

Transportation Logistics Data: Logistics-related data, such as routing (distance to market and transportation mode), were collected from transport and logistics firms or providers.

Inventory Levels Data: Logistic data on inventory levels, depicting the amount of mangoes at different stages of the supply chain, were gathered from internal sources, such as growers, distributors, and retailers.

Quality Control Metrics Data: Quality control metrics, including freshness and ripeness scores, were monitored using sensors or manual checks at various checkpoints, such as warehouses, during transportation, and at retail points.

Table 1 Feature Description Table

Feature	Description
Temperature (°C)	Ambient temperature in Celsius
Humidity (%)	Relative humidity in percentage
Precipitation (mm)	Amount of precipitation in millimeters
Market Demand	Number of mango units demanded
Distance (km)	Distance to market in kilometers
Transportation Mode	Mode of transportation (e.g., Truck, Ship, Air)
Inventory Levels	Number of mango units available in inventory
Freshness Score	Quality metric indicating the freshness of mangoes (range: 0-1)
Ripeness Score	Quality metric indicating the ripeness of mangoes (range: 0-1)
Recommendation	Binary variable indicating whether a recommendation should be made (0 or 1)

Table 1 presents the dataset features description along with the units. These key variables all together form the basis for constructing a kind of model to generate a set of recommendations for improving the entire mango supply chain. Figures 1-7 [30] represent the dataset features, their distribution, correlation, box plot etc. To this end, Figure 1 gives the distribution in terms of the histogram of each feature in the dataset, where, the subplot shows the distribution of one of the features, which helps to get an understanding of the distribution of the data and potential outliers. Figure 2 provides the correlation matrix that signifies the correlation coefficient between the features of the dataset. It gives the correlation matrix that shows how the variables are related allowing one to detect multi-collinearity and features to use in the predictive modeling. In Figure 3, the distribution of each feature was then represented using the so-called box plots. The box plots give information regarding the median of the data, variability, and symmetry of the data as well as the presence of outliers. Figure 4 builds upon the correlation map by narrowing down the analysis to numerical variables in the dataset. It offers a clear look at the association between two or more numerical variables, beneficial in establishing the right relation for prediction. Figure 5 demonstrates the distribution of the numerical features either through histograms or density plots. It is useful in determining the shape of the distribution, and the presence of any outliers as well as measuring the spread of a set of numerical data. In Figure 6, the distribution of the target variable is represented. It helped in understanding whether the classes were balanced or imbalanced in the given target variable which is important in data analysis before modeling. Figure 7 represents distribution relation plots, which are used to display the relationship between different variables of the dataset. Such plots can be scatter plots, pair plots, or any other kind of plots that would assist in finding out the patterns and relations between features.

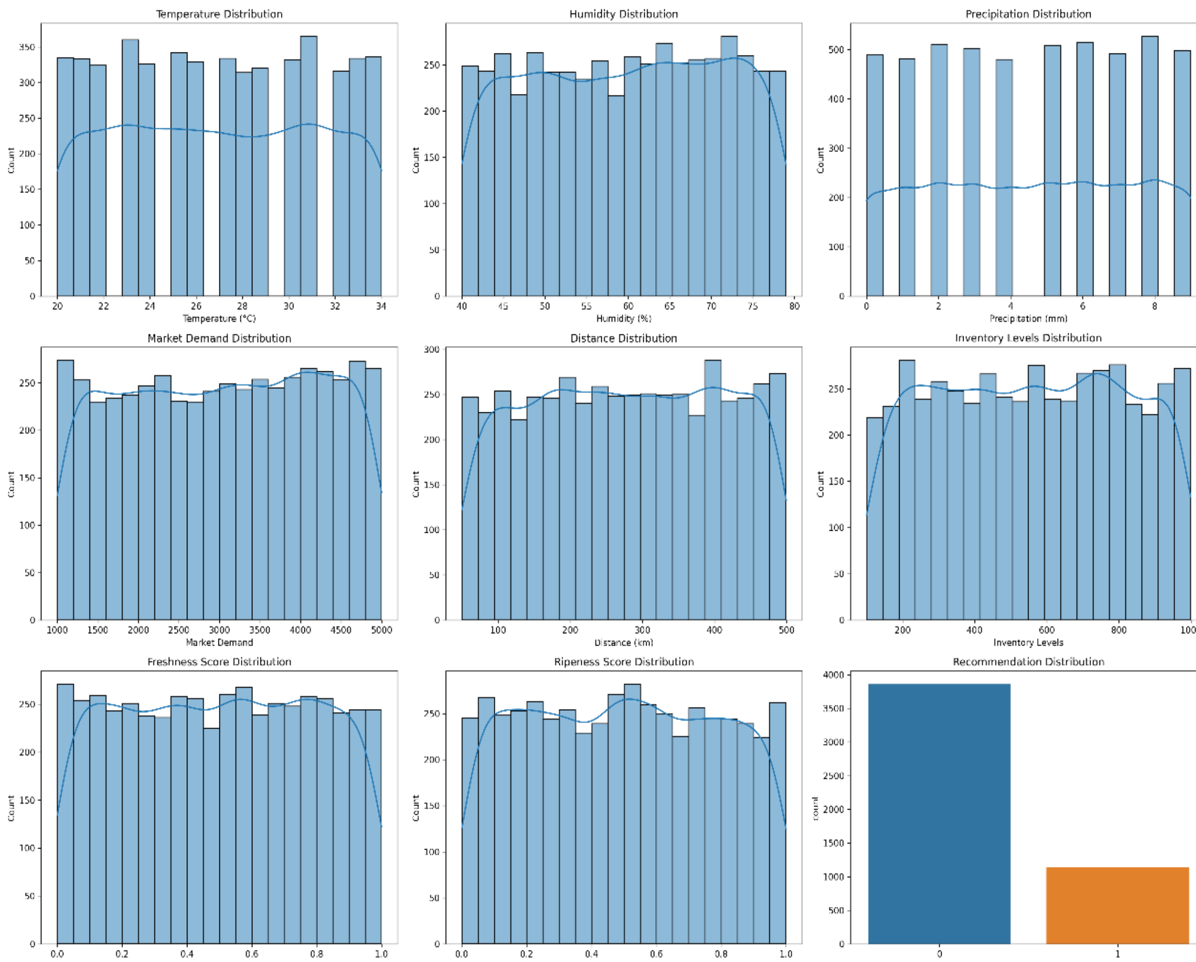

Figure 1. Distribution of features [23]

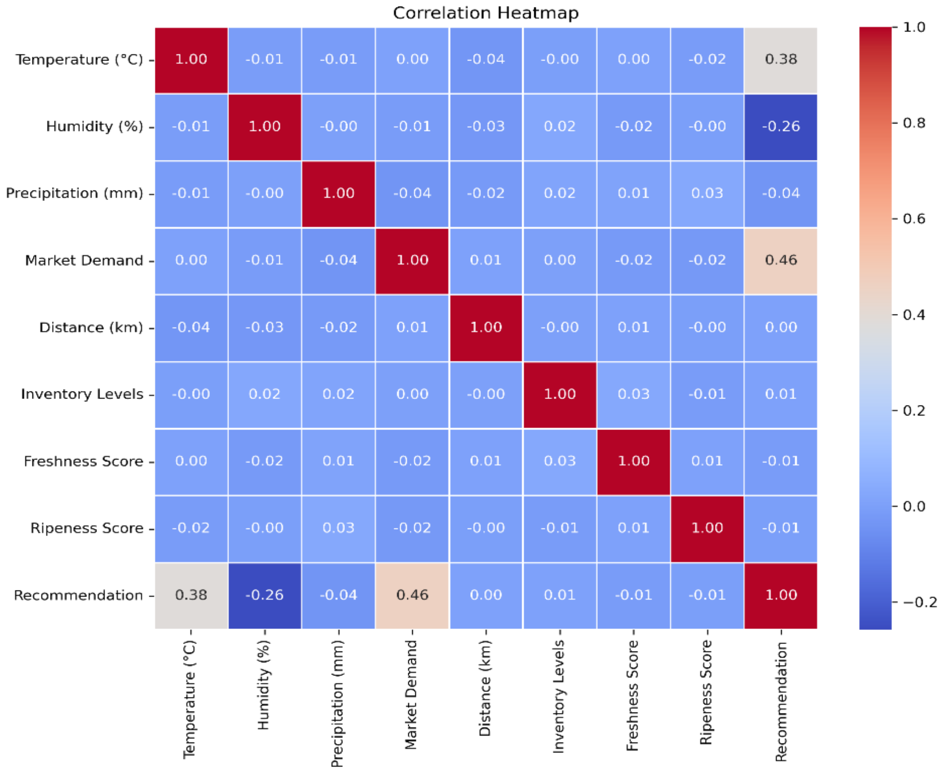

Figure 2. Correlation Map [25]

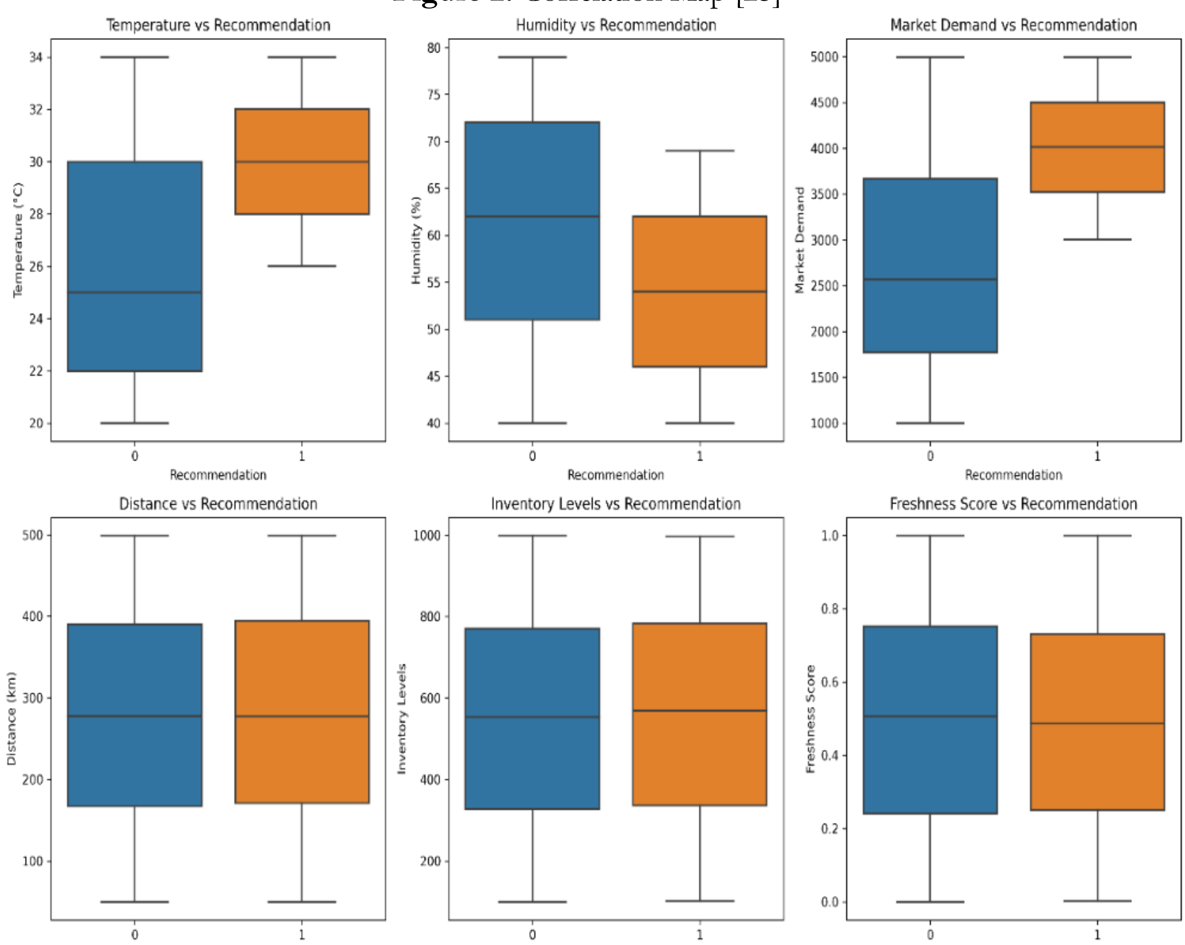

Figure 3. Box plots for each feature [24]

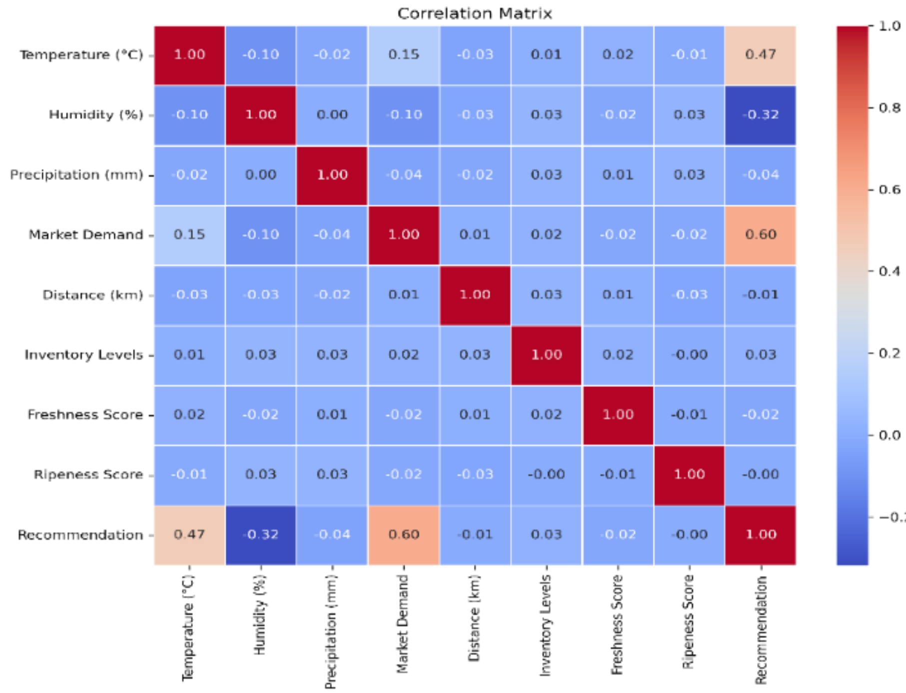

Figure 4. Correlation Map: Distribution of numerical features [24]

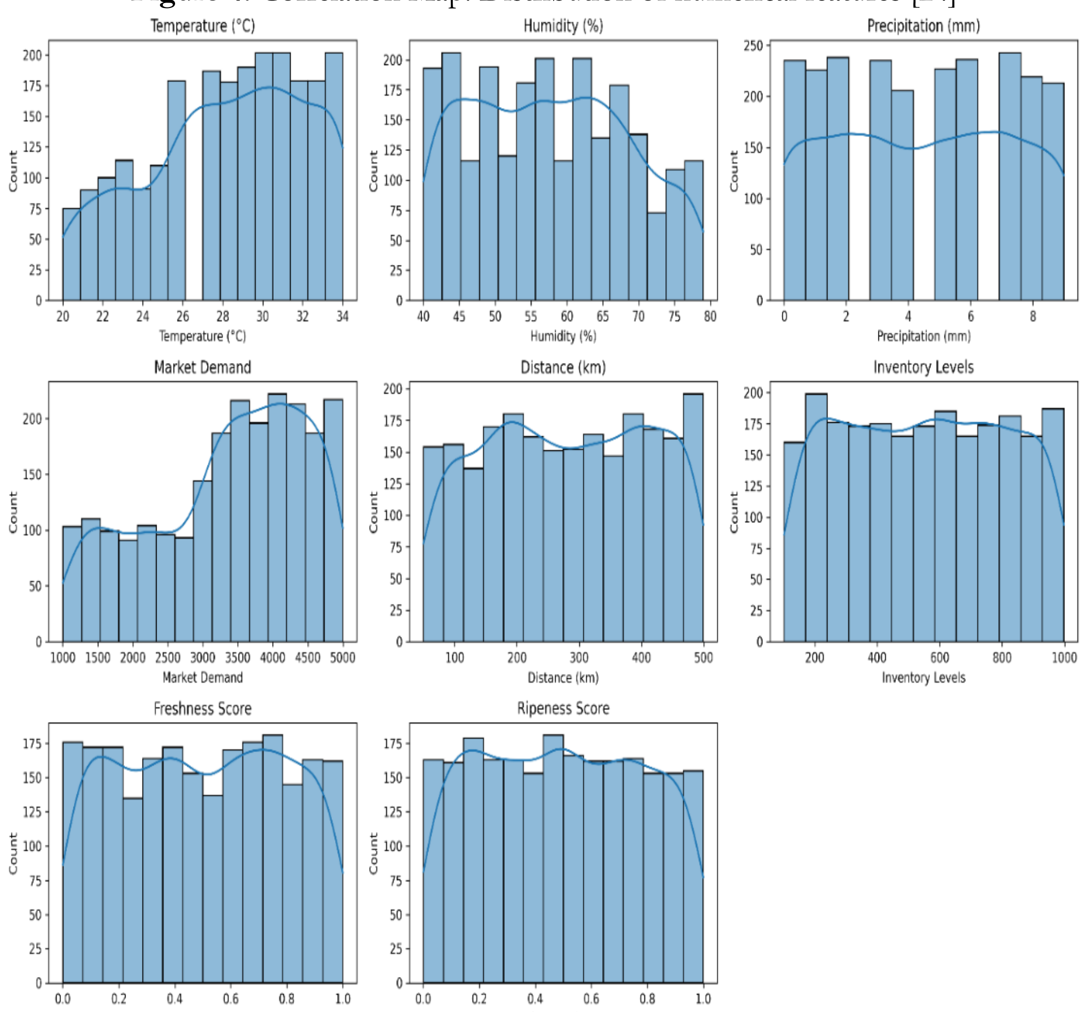

Figure 5. Distribution of numerical features [25]

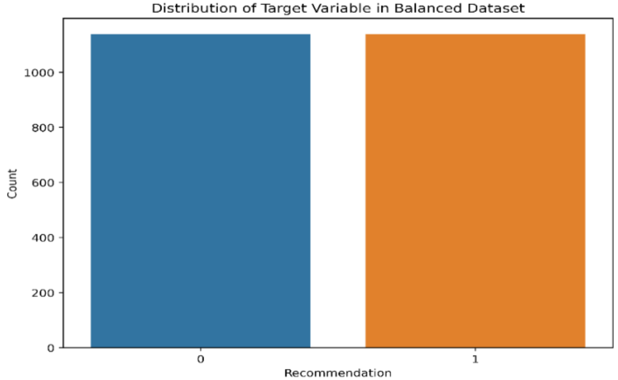

Figure 6. Distribution of the target variable [26]

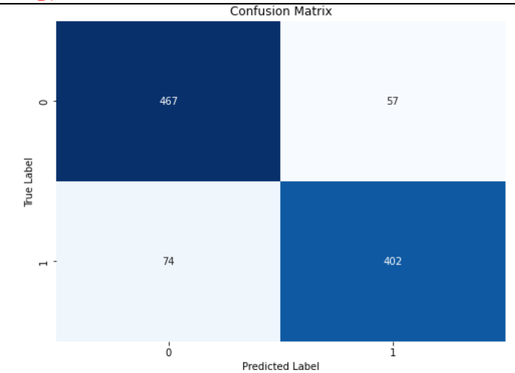

Figure 9. SVM Confusion Matrix

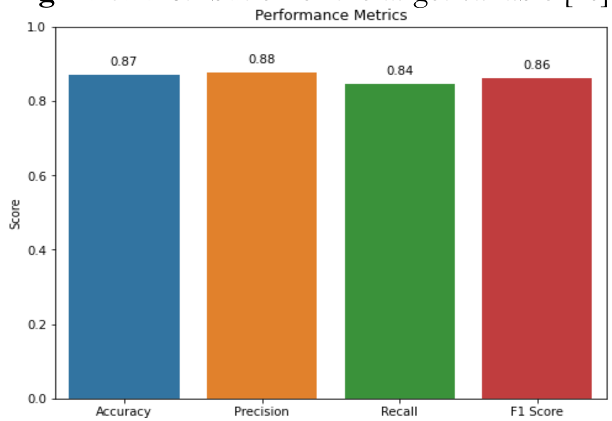

Figure 10. SVM Performance Metrics

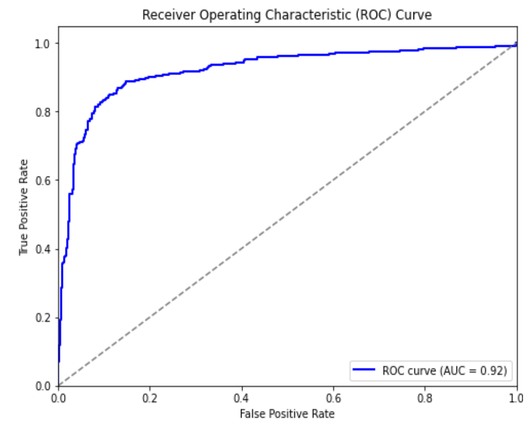

Figure 11. SVM ROC AUC

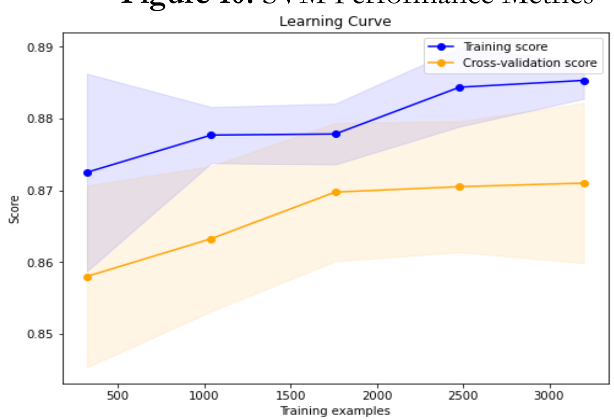

Figure 12. SVM Learning Curves

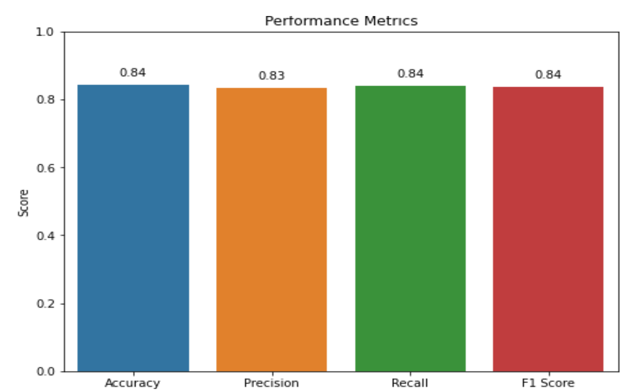

Figure 13. KNN Performance Metrics

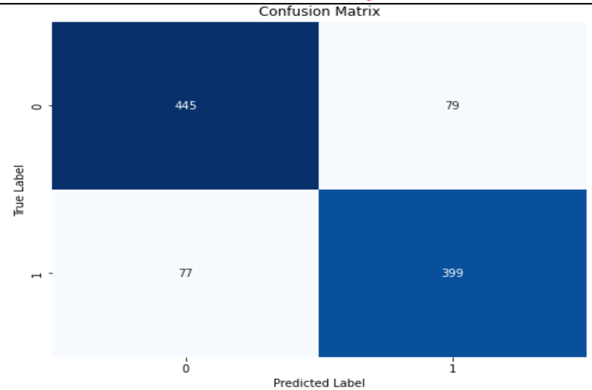

Figure 14. KNN Confusion matrix

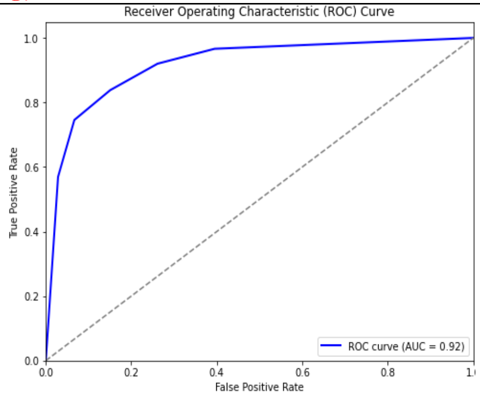

Figure 15. KNN ROC AUC

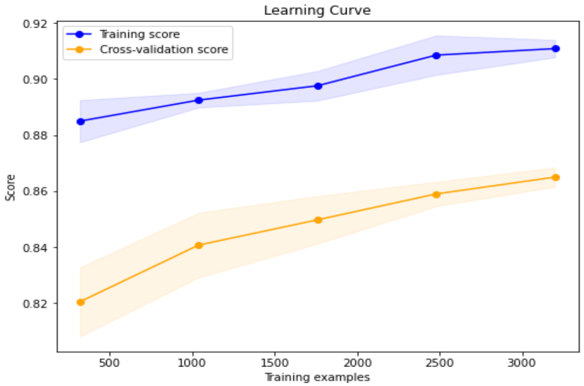

Figure 16. KNN Learning Curves

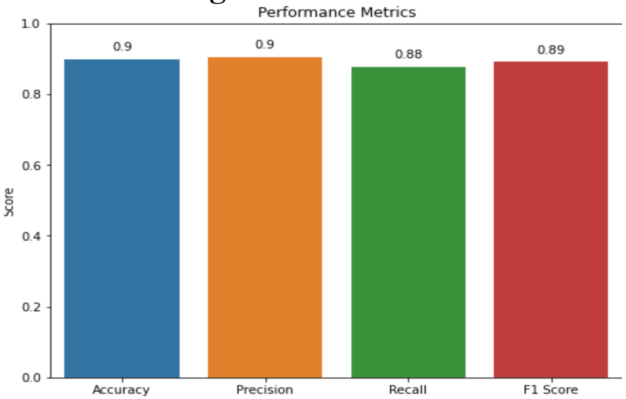

Figure 17. RF Performance Metrics

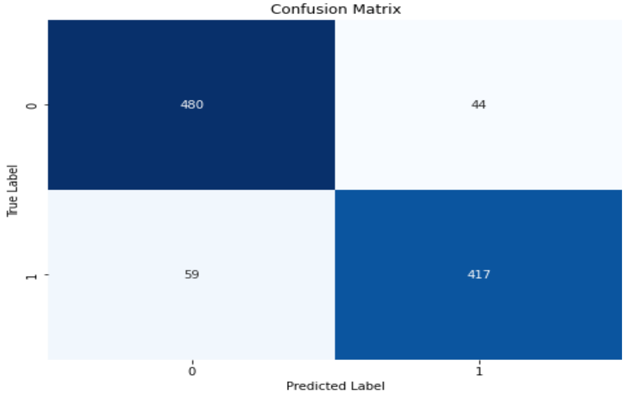

Figure 18. RF Confusion Matrix

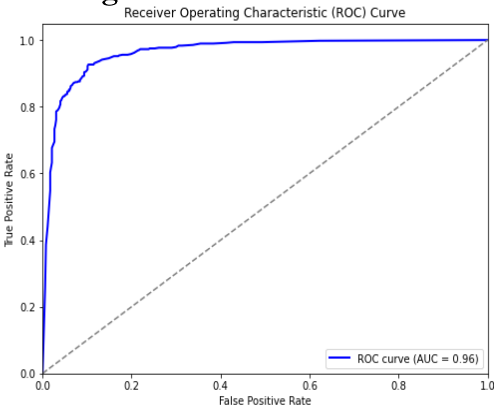

Figure 19. RF ROC AUC

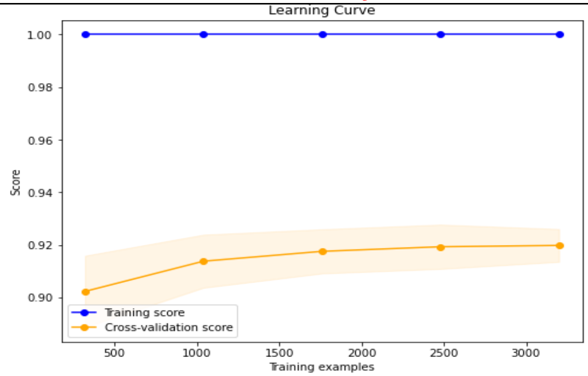

Figure 20. RF Learning Curves

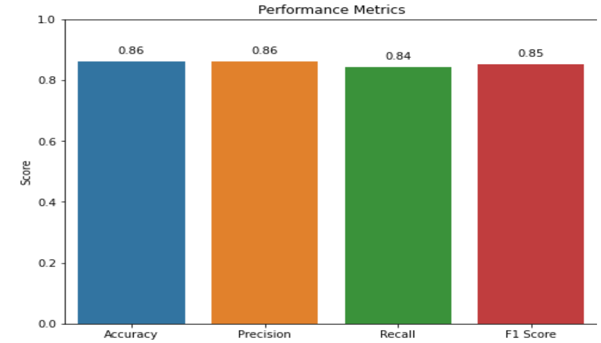

Figure 21. DT Performance Metrics

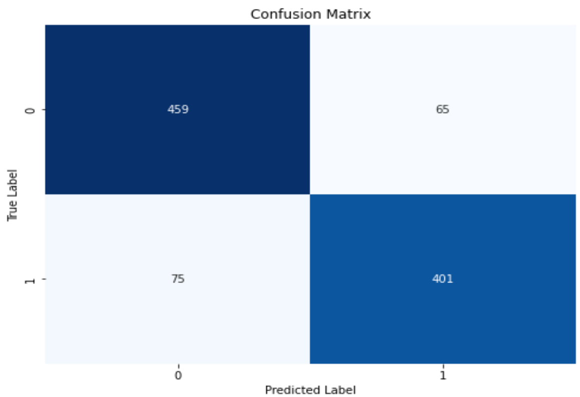

Figure 22. DT Confusion Matrix

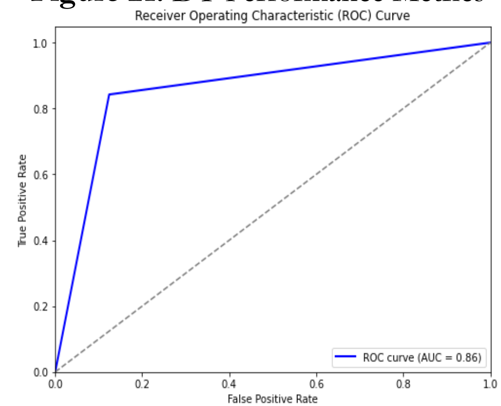

Figure 23. DT ROC AUC

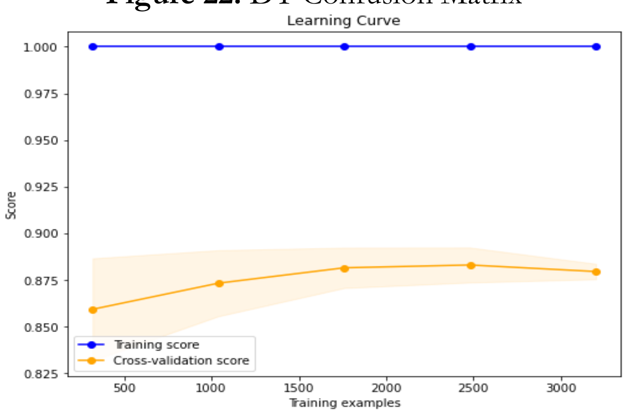

Figure 24. DT Learning Curves

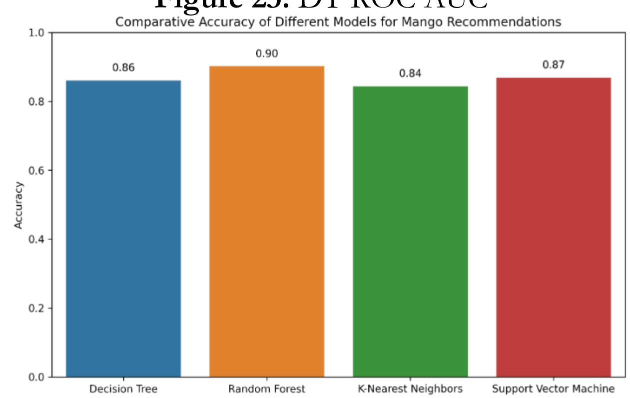

Figure 25. Comparative Accuracy

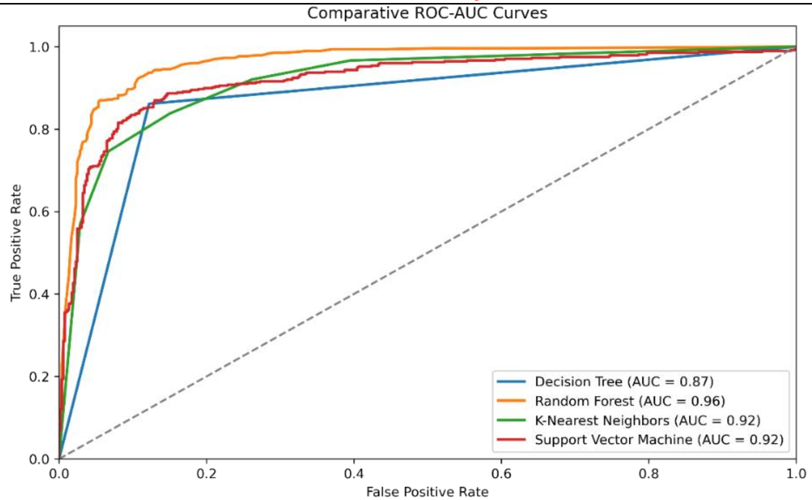

Figure 26. Comparative ROC-AUC

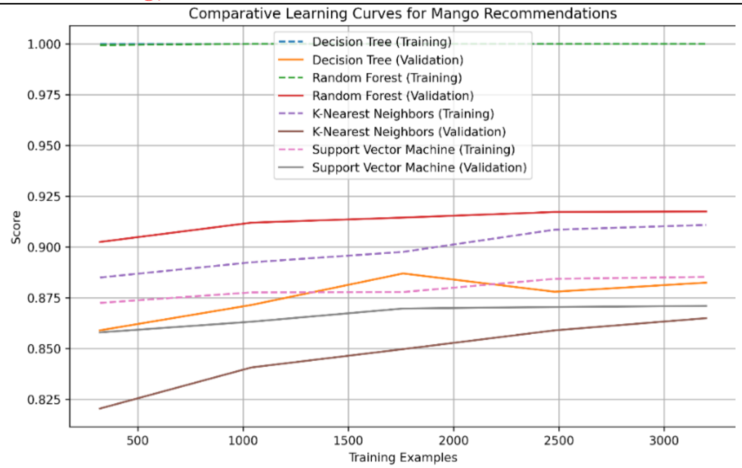

Figure 27. Comparative Learning Curves for Mango Recommendations

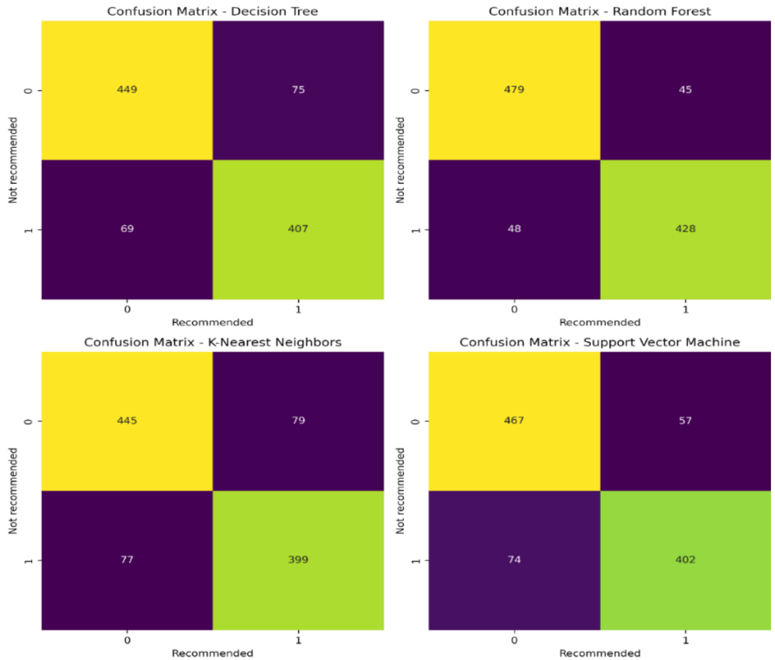

Figure 28. Comparative Confusion matrix

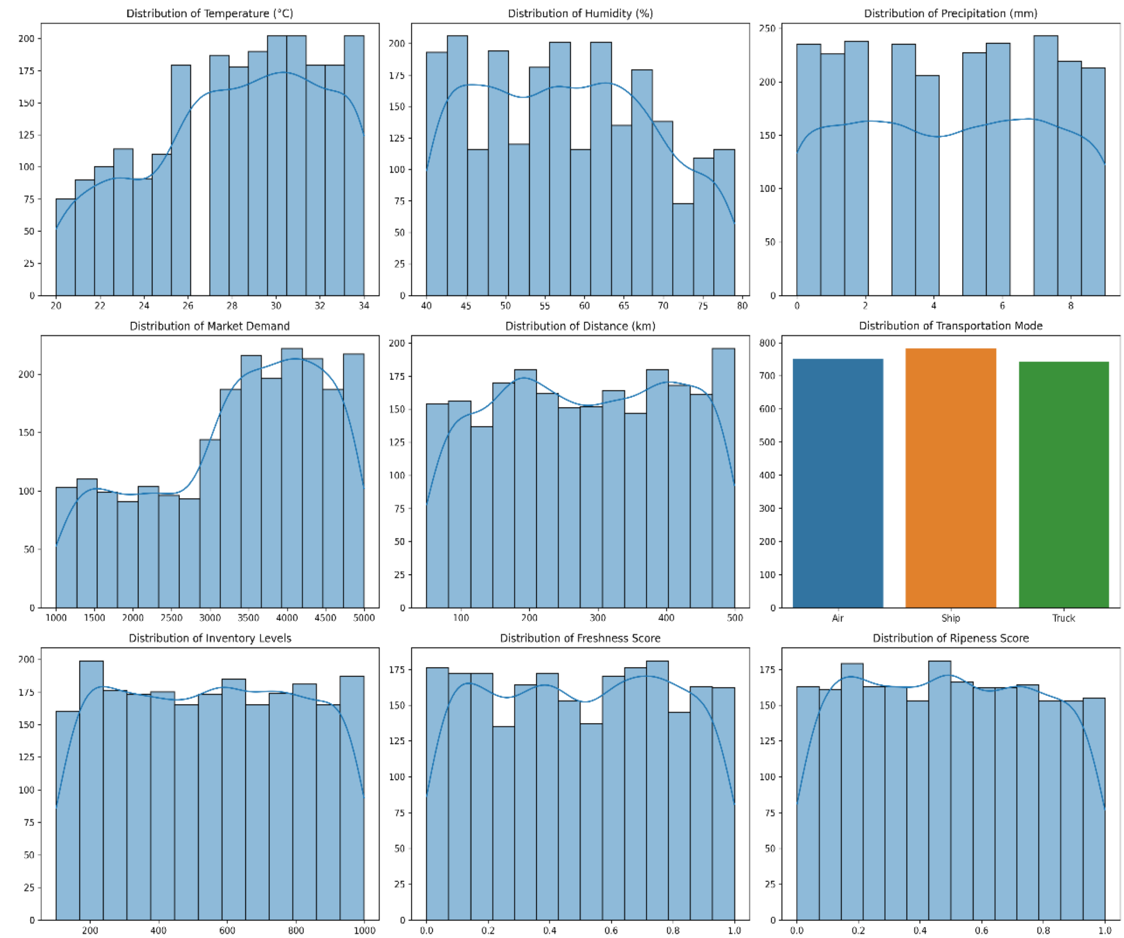

Figure 7. Distribution Relation Plots [26]

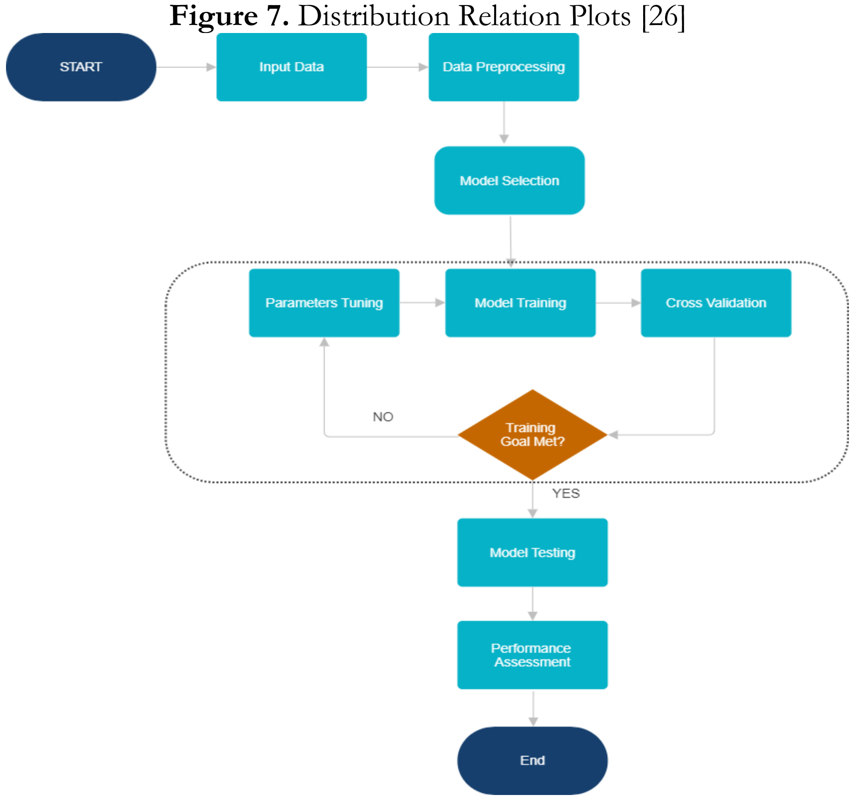

Figure 8. The flow of study diagram

Machine Learning Models:

In this section, the models of machine learning including forecasting and analysis taken from the dataset are depicted. We walk through the theory, practical details, and evaluation Figures of every model respectively. The flow diagram of the study is given in Figure 8.

Support Vector Machine (SVM):

SVM is the powerful supervised learning algorithm that conquers these classification and regression tasks of machine learning. SVM works by seeking the hyperplane that can automatically classify inputs into either two or more classes. Simultaneously, the hyperplane which is closest to the center among these classes is preferred as the classifier. The hyperplane is formed by the support vectors, which are located on the border of the support vector space, close to the hyperplane.

The SVM's purpose is to maximize the margin, which is the deliberate gap between the hyperplane and the points nearest to their respective classes. Mathematically, the objective of SVM is to solve the following optimization problem:

$$\min w, b \ \frac{1}{2} \ \|w\|^2 \text{ subject to } y_i(w.x_i + b) \geq 1, \forall i \qquad (2)$$

where:

w is the weight vector.

b is the bias term.

y_i Is the class label of the ith data point.

x_i Is the ith data point.

Support Vector Machines (SVMs) are versatile and can handle both linear and nonlinear classification tasks using various kernel functions, including linear, polynomial, Radial Basis Function (RBF), and sigmoid. The inequality constraint in SVMs ensures that data points are correctly classified with a margin of at least 1.

K-Nearest Neighbors (KNN) is a simple yet effective method for classification and regression problems. In KNN, the class label of a data point is inferred based on its nearest neighbors. The basic principle of KNN involves locating k nearest neighboring data points to the query point using a distance metric. The final result of the class label for the query point is determined by the majority class within the individuality of its k nearest neighbors. However, KNN is not scalable due to its sluggish algorithm, which requires extra memory and data storage compared to other classifiers.

Random Forest (RF) is an ensemble learning method that involves building multiple decision trees and making predictions based on the class or mean estimate of each tree output separately. The trees in the random forest are built upon randomly created bootstrap samples and feature sample bags. This randomness ensures that the model avoids overtraining, allowing for high generalization performance. The outcome of a decision tree forest model is acquired by summing up the forecasts made by individual decision forests through averaging or voting techniques.

Decision Trees (DT) are widely used supervised learning algorithms for classification and regression tasks. A decision tree recursively splits the feature space into subsets using information about the input features through impurity minimization or information maximization at each split. The decision tree learning algorithm evaluates different features and finds the best one to separate data into the most uncompromised subsets. This function is called repeatedly until the tree is either held with the maximum allowed depth or tested again with some purity criterion. A decision tree model prediction is achieved by traveling the tree from the root node to a leaf node, following the feature values of every input data point until it reaches the leaf node.

Evaluation Metrics:

To assess the performance of our models, we employed several evaluation metrics, including accuracy, precision, recall, and F1 score. These metrics provide a comprehensive view of the model's accuracy and completeness, which are crucial in situations where misclassification costs are different for different classes or when the number of objects of the given class is imbalanced in a dataset.

Accuracy:

Accuracy is the ratio of correctly classified instances to the total number of instances. It represents the proportion of correct predictions to the total number of predictions.

Precision:

Precision is the ratio of true positive predictions to the total number of positive predictions made. It measures the proportion of true positives among all positive predictions.

Recall:

Recall, also known as sensitivity, is the percentage of true-positive cases among all positive cases. It is calculated as the number of accurate positives to the sum of the number of true positives and false negatives.

F1 Score:

The F1 score is a measure of the harmonic mean of precision and recall. It balances the concepts of precision and recall and is especially useful when the dataset shows class imbalance. The F1 score can be computed as:

$$F1 = 2 \times Precision \times Recall / Precision + Recall$$

These evaluation metrics provide a comprehensive understanding of the model's performance, enabling us to identify areas for improvement and optimize the model accordingly.

Results and Discussion:

This section presents the results of experimentation and analysis for evaluating the effectiveness of different models based on their predictive capabilities for the target variable. We also provide insights into the underlying patterns and relationships within the data.

Support Vector Machine (SVM):

Figure 9 illustrates a confusion matrix that demonstrates the model's performance, categorizing true positives, true negatives, false positives, and false negatives. As shown in Figure 10, the SVM achieved an accuracy of 0.869, with a precision of 0.876, recall of 0.845, and F1 score of 0.860. The Receiver Operating Characteristic (ROC) curve in Figure 11 reveals a maximum Area Under the Curve (AUC) value of 0.92. Notably, the SVM learning curves in Figure 12 exhibit overfitting on the training data, yet still deliver better performance on the testing data.

K-Nearest Neighbors (KNN):

KNN attained an accuracy of 0.844, accompanied by a precision of 0.835, recall of 0.838, and F1 score of 0.836. Performance metrics for KNN are visualized in Figure 13. Figure 14 displays a confusion matrix illustrating the model's performance. The ROC curve in Figure 15 reveals a maximum AUC value of 0.92. The learning curves in Figure 16 demonstrate a training accuracy of 91% and a test accuracy of 86%.

Random Forests (RF):

As depicted in Figure 17, RF achieved an accuracy of 0.897, precision of 0.905, recall of 0.876, and F1 score of 0.890. The confusion matrix in Figure 17 illustrates the model's performance. The ROC curve in Figure 18 reveals a maximum AUC value of 0.96. Notably, the learning curves in Figure 19 demonstrate a training accuracy of 99% and a test accuracy of 91%.

Decision Trees (DT):

As shown in Figure 20, DT achieved an accuracy of 0.86, precision of 0.861, recall of 0.842, and F1 score of 0.851. These metrics were derived from the confusion matrix for DT displayed in Figure 21. The ROC curve in Figure 22 reveals a maximum AUC value of 0.86.

Notably, the learning curves in Figure 23 demonstrate a training accuracy of 100% and a test accuracy of 87.5%.

Comparative Performance:

The proposed models were evaluated using various metrics, including error rates, accuracy, recall, precision, F-measure, and confidence intervals. Figure 24 illustrates the comparative accuracy of the models, enabling a graphical comparison of their predictive capabilities. RF exhibits the highest accuracy of 90% among the models. Figure 25 displays the ROC curve for RF, revealing a maximum AUC value of 0.96. These curves highlight the tradeoff between true positive rates and false positive rates, providing insights into model performance.

Generalization and Confusion Matrices:

As we increase the training set size, the generalization of each model to new data is evaluated. This assessment enables us to understand how well each model generalizes. The confusion matrices for the models are illustrated in Figure 27, providing a comprehensive analysis of true positives, true negatives, false positives, and false negatives. These matrices facilitate a graphical understanding of the classes or target variables.

Numerical Evaluation Metrics:

Table 2 presents the numerical values of the evaluation metrics, including accuracy, precision, recall, and F1 score, for each model.

Table 2: Comparison of Models

Model	Accuracy	Precision	Recall	F1 Score
SVM	0.869	0.876	0.845	0.860
KNN	0.844	0.835	0.838	0.836
Random Forests	0.897	0.905	0.876	0.890
Decision Trees	0.860	0.861	0.842	0.851

Table 2 provides a quantitative comparison of the models' performance metrics, allowing for a more detailed analysis and interpretation of the results. To this end, RF exhibits a better accuracy of 89.7%, Precision=90.5%, Recall=87.6%, and F1-Score=89.0% as compared to the other models.

Conclusion:

This article provides guidelines for developing a machine learning-based mango recommendation system. By examining various models and comparison metrics, we identified the most effective prediction system for mango quality and market demand. To evaluate the performance of different models, we employed metrics such as accuracy, precision, recall, and F1 score. Our results showed that Random Forests outperformed other models, achieving the highest accuracy, precision, recall, and F1 scores. The robustness and complex data-handling capabilities of Random Forests make it an ideal choice for mango recommendation systems.

Future research directions include integrating a diverse and large dataset, and incorporating remote sensing and climate data, to enhance the precision and generalization of mango recommenders. We also plan to develop a user-friendly interface for real-time monitoring and stakeholder updates on estimated yields. Furthermore, we intend to explore the application of deep learning network topologies and adaptive learning rates to train on data clusters, rather than the complete dataset, for improved performance.

Acknowledgment: I would like to thank the Faculty of Computing Muhammad Nawaz Shareef University of Agriculture, Multan for providing the necessary resources.

Author's Contribution: Each author has made an equal contribution to this study.

Conflict of Interest: In publishing this work in IJIST, the authors state that they have no competing interests.

Project Details: Nil.

References:

[1] A. B. S. Bewketu Takele Wondifraw, Mesele Yihune Tamene, "Assessment of crop damage

by rodent pests from experimental barley crop fields in Farta District, South Gondar, Ethiopia," *PLoS One*, vol. 16, no. 8, p. 0255372., 2021, doi: 10.1371/journal.pone.0255372.
[2] E. D. Erute, Folasade O. Isinkaye, "A Smartphone-based Plant Disease Detection and Treatment Recommendation System using Machine Learning Techniques," *Trans. Mach. Learn. Artif. Intell.*, vol. 10, no. 1, pp. 1–8, 2022, doi: 10.14738/tmlai.101.11313.
[3] Y. P. Akriti Taneja, Gayathri Nair, Manisha Joshi, Somesh Sharma, Surabhi Sharma, Anet Režek Jambrak, Elena Roselló-Soto, Francisco J. Barba, Juan Manuel Castagnini, Noppol Leksawasdi, "Artificial Intelligence: Implications for the Agri-Food Sector," *Agronomy*, vol. 13, no. 5, p. 1397, 2023, doi: 10.3390/agronomy13051397.
[4] S. Bondre, D. A., & Mahagaonkar, "PREDICTION OF CROP YIELD AND FERTILIZER RECOMMENDATION USING MACHINE LEARNING ALGORITHMS," *Int. J. Eng. Appl. Sci. Technol.*, vol. 4, no. 5, pp. 371–376, 2019, [Online]. Available: https://ijeast.com/papers/371-376,Tesma405,IJEAST.pdf
[5] T. N. P. Truong Dang Le, Thanh Viet Nguyen, Nguyen Van Muoi, Ha Thanh Toan, Nguyen Mai Lan, "Supply Chain Management of Mango (Mangifera indica L.) Fruit: A Review With a Focus on Product Quality During Postharvest," *Front. Sustain. Food Syst.*, vol. 5, 2022, doi: https://doi.org/10.3389/fsufs.2021.799431.
[6] A. Gupta, A., Nagda, D., Nikhare, P., & Sandbhor, "Smart Crop Prediction Using IoT and Machine Learning," *Int. J. Eng. Res. Technol.*, vol. 9, no. 3, pp. 18–21, 2021, [Online]. Available: https://www.ijert.org/research/smart-crop-prediction-using-iot-and-machine-learning-IJERTCONV9IS03005.pdf
[7] I. B. J. Dela Cruz, "Machine Learning-Based Classification of Mango Pulp Weevil Activity Utilizing an Acoustic Sensor," *Micromachines*, vol. 14, no. 11, p. 1979, 2023, doi: 10.3390/mi14111979.
[8] T.-N. and D.-N. L.-T. Doan, "A Novel Mango Grading System Based on Image Processing and Machine Learning Methods," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 5, pp. 1118–1129, 2023, [Online]. Available: https://thesai.org/Downloads/Volume14No5/Paper_115-A_Novel_Mango_Grading_System_Based_on_Image_Processing.pdf
[9] P. D. & B. R. Jayita Dutta, "AI-based soft-sensor for shelf life prediction of 'Kesar' mango," *SN Appl. Sci.*, vol. 3, no. 6, p. 657, 2021, doi: 10.1007/s42452-021-04657-7.
[10] Y. G. Rahini Mahendran , Rongbin Xu , Shanshan Li, "Interpersonal violence associated with hot weather," *Lancet Planet Heal.*, vol. 5, no. 9, pp. 571–572, 2021, doi: 10.1016/S2542-5196(21)00210-2.
[11] S. Vishnu, "STEERING INNOVATION IN AGRICULTURE: THE DUTCH WAY," *Agric. Ext. South Asia*, 2022, [Online]. Available: https://www.researchgate.net/publication/363082449_STEERING_INNOVATION_IN_AGRICULTURE_THE_DUTCH_WAY
[12] V. K. T. and A. K. S. Saini, R. V., P. Vaid, N.K. Saini, S.S. Siwal, V.K. Gupta, "Recent Advancements in the Technologies Detecting Food Spoiling Agents," *J. Funct. Biomater.*, vol. 12, no. 4, p. 67, 2021, doi: 10.3390/jfb12040067.
[13] M. R. and A. H. G. Nithya, R., B. Santhi, R. Manikandan, "Computer Vision System for Mango Fruit Defect Detection Using Deep Convolutional Neural Network," *Foods*, vol. 11, no. 21, 2022, doi: 10.3390/foods11213483.
[14] A. D. and A. S. Rajpoot, V., R. Dubey, P.K. Mannepalli, P. Kalyani, S. Maheshwari, "Mango Plant Disease Detection System Using Hybrid BBHE and CNN Approach," *Trait. du signal*, vol. 39, no. 3, pp. 1071–1078, 2022, doi: 10.18280/ts.390334.
[15] S.-I. C. & W.-F. Chen, "The Optimal Harvest Decisions for Natural and Artificial Maturation Mangoes under Uncertain Demand, Yields and Prices," *Sustainability*, vol. 13, no. 17, pp. 1–17, 2021, [Online]. Available: https://ideas.repec.org/a/gam/jsusta/v13y2021i17p9660-d623563.html#author-abstract
[16] P. C. and R. M. R. Defraeye, T., C. Shrivastava, T. Berry, P. Verboven, D. Onwude, S. Schudel, A. Bühlmann, "Digital twins are coming: Will we need them in supply chains of

fresh horticultural produce?," *Trends Food Sci. Technol.*, vol. 109, pp. 245–258, 2021, doi: https://doi.org/10.1016/j.tifs.2021.01.025.

[17] S. J. and P. Jaidka, "Mango Leaf disease Classification using deep learning Hybrid Model," *Int. Conf. Power, Instrumentation, Energy Control (PIECON), Aligarh, India*, pp. 1–6, 2023, doi: 10.1109/PIECON56912.2023.10085869.

[18] S. S. & G. P. R. Rohan Sharma, Kartik Suvarna, "Detecting Diseases in Mango Leaves Using Convolutional Neural Networks," *Int. Conf. Innov. Comput. Commun.*, pp. 183–192, 2021, doi: https://doi.org/10.1007/978-981-16-2594-7_15.

[19] C. Q. Yingfeng Zhang, Lin Zhao, "Modeling of an IoT-enabled supply chain for perishable food with two-echelon supply hubs," *Ind. Manag. Data Syst.*, vol. 17, pp. 1890–1905, 2017, [Online]. Available: https://www.semanticscholar.org/paper/Modeling-of-an-IoT-enabled-supply-chain-for-food-Zhang-Zhao/df7dee76df5a63612c876990f28b1b9d1b47e56c

[20] A. J. M. B. and A. R. Verdouw, C.N., J. Wolfert, "Virtualization of food supply chains with the internet of things," *J. Food Eng.*, vol. 176, pp. 128–136, 2016, doi: https://doi.org/10.1016/j.jfoodeng.2015.11.009.

[21] D. . Vermani, "Farm to Fork: IOT for Food Supply Chain," *Int. J. Innov. Technol. Explor. Eng.*, vol. 8, no. 12, pp. 4915–4919, 2019, doi: 10.35940/ijitee.L3551.1081219.

[22] Y. M. Ubonrat Siripatrawan, "Hyperspectral imaging coupled with machine learning for classification of anthracnose infection on mango fruit," *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.*, vol. 309, p. 123825, 2024, doi: https://www.sciencedirect.com/science/article/abs/pii/S138614252301510X.

[23] A. K. N. Jayashri N, Veeresh Rampur, Durgaprasad Gangodkar, Abirami M, Balarengadurai C, "Improved block chain system for high secured IoT integrated supply chain," *Meas. Sensors*, vol. 25, p. 100633, 2023, doi: https://doi.org/10.1016/j.measen.2022.100633.

[24] S. D. Suveena Jantapirak, Onrawee Laguerre, Alain Denis, Yvanne Paviet Salomon, Denis Flick, Kanithaporn Vangnai, Weerachet Jittanit, "Experimental simulation of temperature non-uniformity in a loaded container along air cargo supply chain: Mango case study," *J. Food Eng.*, vol. 380, p. 112161, 2024, doi: https://doi.org/10.1016/j.jfoodeng.2024.112161.

[25] J. A. ROSE Githumbi, JOHN Mburu, ACKELLO Ogutu, "Assessment of factors influencing adoption of postharvest loss reduction practices along the mango value chain in Embu, Machakos, and Nairobi counties, Kenya," *Heliyon*, vol. 10, no. 4, p. 25146, 2024, doi: https://doi.org/10.1016/j.heliyon.2024.e25146.

[26] and I. A. B. Ayesha Hakim, Sundus Shafique, Mubashir Mehdi, "FARMER TO CONSUMER – AN ONLINE SUPPLY CHAIN PROCESS SYSTEM USING BLOCKCHAIN TECHNOLOGY," *Agric. Sci. J.,* vol. 6, no. 1, pp. 1–11, 2024.

[27] A. H. Iqbal, Hafiz Muhammad Rizwan, "Mango Variety and Grading Dataset," *Mendeley Data*, 2021, doi: 10.17632/5mc3s86982.1.

[28] S. N. & N. Anand, "Wholesalers perspectives on mango supply chain efficiency in India," *J. Agribus. Dev. Emerg. Econ.*, vol. 9, no. 2, pp. 175–200, 2019, doi: 10.1108/JADEE-02-2018-0032.

[29] A. Ballo, A.B., Diaby, M. and Coulibaly, "An Application of Machine Learning Methods to Detect Mango Varieties," *Open J. Appl. Sci.*, vol. 14, pp. 1666–1690, 2024, doi: 10.4236/ojapps.2024.147109.

[30] & S. H. W. Hardyansyah, B., Heru Sukoco, "Monitoring and Controlling System for Mango Logistics Based on Machine Learning," *Inf. Syst. Eng. Artic.*, vol. 8, no. 1, 2024, doi: https://doi.org/10.29207/resti.v8i1.5226.