IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⋇ Peer-reviewed / Refereed journal ⋇ Vol. 12, Issue 1, January 2025

DOI: 10.17148/IARJSET.2025.12106

Enhancing farming efficiency using KNN and SVM algorithms

Jemima David D[1], Mohamed Aaftaab A[2], Dr.S. Nivetha M.E, PhD[3]

B.E Computer Science and Engineering with Artificial Intelligence,

Sathyabama Institute of Science and Technology, Semmencheri, Chennai, Tamil Nadu, India-600119[1,2]

Assistant Professor, Department of Computer Science and Engineering,

Sathyabama Institute of Science and Technology, Semmancheri, Chennai, Tamil Nadu, India-600119[3]

Abstract: India being an agriculture country, its economy predominantly depends on agriculture yield growth and agro industry products. Maintaining a high yield is a very important issue in agriculture. Any farmer is interested in knowing how much yield he is about to expect. By analyzing the various related attributes like location, pH value from which alkalinity of the soil can be determined. Along with this, percentage of nutrients like Nitrogen (N), Phosphorous (P), and Potassium (K) Location can be used alongside the use of third-party apps like APIs, weather and temperature, type of soil, nutrient value of the soil in that region, quantity of rainfall, soil composition can also be determined. All these attributes of data will be analyzed, train the data with various suitable machine learning algorithms for creating a model. The system comes with a model to be precise and accurate in guiding the end user with proper recommendations about required fertilizer ratio based on atmospheric and soil parameters of the land which enhance to increase the crop yield and increase farmer revenue. It also suggests a gel, oil, or any other chemical agent that would help crops grow better. Furthermore, the system will use crop yield history data to fine-tune the predictions and recommendations. Integrating satellite imagery will enable the model to analyze vegetation health and detect issues before they occur. The inclusion of real-time sensor data from IoT devices on the farm will enhance the accuracy of the predictions. The system will also consider pest and disease predictions so that appropriate pre-emergent preventative actions and readiness notification are received in time. Moreover, it will provide advice to clients for the best times for plantings under seasonal weather regimes.

Keywords: K- Nearest Neighbor, Support Vector Machine, Machine learning, Efficient Farming, System Architecture.

INTRODUCTION

A land of agriculture and a country highly dependent on its agricultural sector is India, possessing the second highest arable area in the world, with a cultivation area above 1.6 million square kilometers. As the backbone of the Indian economy, agriculture keeps a large segment of the Indian population alive, while the entire sector's benefits are far reaching, not just for the maintenance of the country's food chain but also contributing to economic development and rural areas' development. Despite its enormous potential, India's agricultural sector has significant challenges. The farming practices of the nation remain largely traditional with little adoption of modern, technology-driven methods. These old practices, although deep in the cultural fabric of the country-they are often leading to far-from-optimal crop yields. The difference between potential and actual productivity has caused widespread disillusionment among farmers. Many farmers, who perceived agriculture as a safe source of income, now suffer from declining yields and financial strain. This disillusionment has driven a significant number of them away from farming, forcing them to seek alternative livelihoods in other sectors. The consequences of this shift are profound. As more farmers abandon agriculture, the sector’s contribution to the economy diminishes, threatening food security and rural stability. This shift away from farming has wider social implications, dislocating the traditional rural lifestyle and exacerbating issues such as urban migration and unemployment. India can become a global superpower in agriculture because of its rich natural resources and varied climatic conditions. But for this to happen, agricultural practices need urgent modernization. With advanced technologies like precision farming, data analytics, and sustainable farming methods, India can significantly boost crop yields, restore farmers' confidence in the agricultural process, and revitalize the sector as a whole. The alarming rise in attempts at suicide by farmers in India is a very serious issue that underlines the critical challenges being faced by people in the agricultural sector. The emotional and economic toll on farmers due to the failure of their livelihoods, due to agriculture, signifies that there is an urgent need for complete support systems. One of the most significant needs is for technical assistance in agriculture. Farmers need access to modern tools of knowledge and resources that can enable them to make the best decisions and significantly improve their productivity. Introduction of high-tech farming methods can go a long way to ease the pressures that force individuals to take such extreme measures. The right set of recommendations will help farmers maximize their crop yields, improve quality, and obtain a better stable income. A country like India has centuries-old agriculture experience. The data, therefore, abounds for its support. Crop performance, soil conditions, weather patterns, and other related information have been documented over the years. The history of agricultural data is thus an important asset that can be mined to analyze trends, predict outcomes, and produce practical recommendations for current farming practices. In this context, the research work proposed here will tap into this vast reservoir of historical data to build a recommendation system for farmers. By analyzing past agricultural data and combining it with modern machine learning algorithms, the system can generate a list of suitable crops tailored to the specific conditions of a farmer's land. This approach ensures that farmers are equipped with the knowledge they need to choose crops that are most likely to thrive in their particular environment, taking into account factors such as soil type, climate, and water availability.

The prime objective of the research is to empower farmers with actionable insights in order to help reduce their chances of crop failures and thereby increase their economic stability. With targeted recommendations, it is possible that such financial hardships might not be necessary and would ultimately not lead to extreme measures, like suicide. Essentially, it is not merely an improvement of agricultural productivity, but rather about helping farmers through the support of their livelihood and their well-being. The development and implementation of a robust recommendation system based on historical agricultural data and modern technology could be a game-changer for farmers, offering them a way out of the cycle of poverty and despair and leading to a more sustainable and prosperous agricultural sector. Each of these influences plays a defining role in dictating whether one crop is good for another, given land parcel. The regional location in which farming is pursued greatly impacts crop maturation. Due to the inherent differences in local soil, geography, and adjacency to water points, each locality has a respective impact on crop yield. By including geographical data into the recommendation system, farmers are likely to be provided with crop recommendations that fit their locality uniquely, thus sowing crops most likely to succeed in their respective environments. Another factor that may affect crop performance is climate. Average temperature, humidity, and seasonal weather can determine whether a crop will succeed or fail. The climatic factors are taken into consideration by the recommendation system to ensure that the crops recommended are suitable for the prevailing weather conditions in the region. For instance, crops that require cooler temperatures or specific levels of rainfall will only be recommended in regions where such conditions are consistently met. Soil quality and type are fundamental to successful farming. Different crops require different soil conditions, including pH levels, nutrient content, and texture. The recommendation system takes into account the specific soil type of a farmer's land, allowing for precise crop recommendations. By aligning the crop choice with the soil's natural characteristics, farmers can maximize their yield and ensure that their land is being used to its full potential. Water availability is a critical concern in agriculture. The system would include data on rainfall as well as groundwater levels to recommend crops best suited to the water resources that exist in the region. This is especially important when operating in regions that are arid or areas with limited irrigation resources. It suggests crops that require less water or those which are drought-resistant, helping the farmer avoid water scarcity pitfalls and manage his resources better. It also considers the temperature variations and seasonal cycles as important in fixing planting and harvesting times. There are some crops that are extremely sensitive to temperature and grow within a very small range. By evaluating the temperature data and comparing it with the needs of the crop, the system guarantees that the suggested crops will be planted at the right time, ensuring improved growth and maximum yields. By combining all of these characteristics, the recommendation system enables farmers to know exactly, through which crop types would best be suitable for cultivation on their land. Such factors eliminate the guesswork that traditionally characterized farming practices, where the farmer tried a number of crops that he or she might find not quite suitable for the conditions on the land. Such experiments can be expensive, both in terms of wasted time and money, and significant losses may result due to unfruitful yields. No universal system exists to help such farmers in their agriculture. And there is a need to reduce the suicide of farmers. The objective of the paper is to build a recommendation system to increase the crop yield. The system helps the farmers in selecting a suitable crop for their agricultural land based on the required parameters. The system is to design and develop a recommendation model to generate recommendations for crops based on geographical and climatic parameters using machine learning algorithms. This research paper proposes a hybrid crop recommendation system using machine learning classifiers.

LITERATURE REVIEW

In developing a precision agriculture crop recommendation system that could maximize crop yield and efficiency, the work of S. Pudumalar et al. (2021) takes focus. Based on analyzing the soil attributes and environmental factors, the system has been made up of machine learning techniques named K- Nearest Neighbor, and Support Vector Machine. These ensemble models with predictions improve their accuracy by combining through Majority Voting Technique. The system predicts suitable crops such as coriander, pulses, cotton, and others based on soil characteristics like pH, depth, and permeability. The ensemble model outperforms individual models in accuracy, demonstrating that combining multiple machine learning methods can significantly optimize crop recommendations and resource usage, ultimately

DOI: 10.17148/IARJSET.2025.12106

enhancing agricultural productivity and decision-making.

Rakesh Kumar et al. (2021) have reviewed crop yield improvement by using advanced crop selection methods based on machine learning techniques, including K-Nearest Neighbor (KNN) and Naive Bayes algorithms. The objective of the research is to identify the best crops for agricultural productivity through different classification rules. Algorithms applied for crop categorization and its validation according to their yield potential consisted of KNN, Naive Bayes, whose classification rules to classify the datasets have been considered and analyzed into most appropriate rules from both algorithms as their usefulness were not uniform to support crop recommendations effectively. The study focuses on the ability of machine learning to optimize crop selection but also points out the necessity of further refinement of these techniques to improve their practical application and accuracy in precision agriculture.

The Dhruvi Gosai et al. (2023) study aims at developing a crop recommendation system that maximizes crop yield through advanced machine learning techniques, such as SVM, Artificial Neural Networks (ANN), Random Tree, and Naive Bayes classifiers. This system will thus try to deliver very accurate crop recommendations by assimilating different types of data from the soil database, expert inputs, and the parameters coming out from the soil testing labs. Ensemble modeling, combined with majority voting, was applied to bring the strengths of SVM and ANN for improved accuracy in recommendations. A comprehensive range of data sources helps the model provide specific crop suggestions according to different environmental conditions. The use of ensemble techniques significantly enhances the accuracy of recommendations, and thus, leads to optimized crop yields. It underlines that the integration of various data inputs and ensemble modeling is effective for enhancing crop recommendations and agricultural productivity. However, at the same time, it does mention the fact that low fertilizer efficiency needs to be addressed for the system's full potential impact to be realized.

Meonghun Lee et al. (2021) aimed to develop an agricultural production system that would leverage the Internet of Things (IoT) for better decision-making and efficiency in production. The system combined real-time IoT data with historical agricultural data, analyzing and predicting relevant economic variables for agriculture using advanced statistical techniques. Experimental results and forecasts were presented through visualization software to better understand and plan strategically. The study showcases how IoT blended with statistical methods really processes complex data, thereby proving to be worthwhile and making way for more upgraded production strategies. The paper thus talks about the chances of upgrading agricultural systems with insights from IoT-based data, while acknowledging the challenge still faced in that regard - addressing the issues about the responsiveness of the system regarding the ever-changing scenarios of agriculture.

The paper by Dharti Vyas et al. (2022) emphasizes the development of a Smart Agriculture Monitoring and Data Acquisition System aimed at enhancing agricultural productivity through real-time data. The system presents the information from various sensors through web and mobile platforms to improve decision-making. It has a control layer, which processes the inputs through an automatic control algorithm that manages agricultural tasks. The system enables continuous remote monitoring and analysis from web and Android devices using the data collected with Tiny OS. The study further identifies how this system is effective for capturing and displaying real-time data to support decision-making in crop management. Real-time monitoring is a significant advancement in smart agriculture. However, limitations related to the Bluetooth technology and its alternative approaches such as Zigbee and Wi-Fi for good performance and reliability are suggested.

The review is done by Saeed Khaki and Lizhi Wang in the year 2023 on Developing a Crop Yield Prediction Model for Crop Yield Estimation using DNN that foresests the crop yield through evaluating environmental factors as well as the genotype data. The research aims to evaluate the performance of DNN in comparison to other methods such as Lasso, shallow neural networks (SNN), and regression trees (RT). Trained on a dataset that includes both environmental and genotype data, the DNN model demonstrated superior accuracy in predicting crop yields compared to the other techniques. The study further shows that environmental data had a much greater effect in yield predictions compared to genotype information, pointing to the influence of environmental conditions in crop productivity. Although DNN was indeed proven to be an effective predictor of higher accuracy, the "black box" character of DNN was pointed out as an even greater challenge of DNN where interpretability and transparency were required for practical agriculture applications.

To improve the crop yield prediction using advanced machine learning algorithms to suggest the best crops, the research study by Anakha Venugopal et al. (2021) was conducted. Here, there were three machine learning models applied: Logistic Regression, Naive Bayes, and Random Forest. The relevant key environmental factors of temperature, rainfall, and land area, with extensive data on these conditions and historical crop yields, were used. The objective of the study

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 1, January 2025

DOI: 10.17148/IARJSET.2025.12106

was to determine which algorithm provided the most accurate yield predictions. From the models tested, Random Forest proved to be the most effective, surpassing Logistic Regression and Naive Bayes in terms of accuracy. This shows that Random Forest is particularly adept at handling complex environmental data and predicting crop performance. The outcomes show that the use of machine learning is poised to revolutionize agricultural decision-making through providing highly accurate, data-driven information in optimizing crop choice and improving yields. It has an important role in connecting technology with conventional farming to facilitate more sustainable and efficient agriculture practices. When the algorithms evolve further, this may be able to further increase crop yields and will help benefit the larger agricultural sector.

Ramesh Medar et al. (2020) developed a crop yield prediction model using the Naïve Bayes algorithm in order to solve the problems that occur in agriculture and increase crop productivity. The study analyzed several agricultural factors, such as soil quality, weather patterns, and environmental conditions, in order to develop a predictive model. With extensive data on crop production and these variables, the Naïve Bayes algorithm was trained to identify patterns and make accurate predictions. The results indicated that Naïve Bayes improved crop yield predictions effectively, providing useful insights for optimizing crop selection and production efficiency. This method helps farmers make more informed decisions, potentially increasing crop yields and contributing to India's economic growth. This study, thus, integrates machine learning techniques like Naïve Bayes into agriculture to highlight the possibilities for enhanced productivity and economic benefits, thereby giving emphasis to some practical advantages of data-driven approaches in farming.

R. R. Rubia Gandhi et al. from 2022 highlight multiple techniques of machine learning to advance smart agriculture practices, focusing on crop monitoring, yield prediction, and resource optimization. The uses of predictive modeling in the research forecast crop yields based on historical and real-time data to assist farmers in making the decisions regarding planting and harvesting. Anomaly detection algorithms identify crop health or environmental condition problems that may otherwise lead to potential losses before they occur. Optimization algorithms optimize resource usage, such as water, fertilizers, and energy. This research aims to analyze extensive agricultural data for the development and refinement of such models, hence evaluating their effect on precision agriculture. The results show that the machine learning models improve crop monitoring, yield prediction, and resource management, resulting in increased productivity and reduced waste. These improvements indicate the promise of machine learning for transforming agriculture and making practices more efficient and sustainable. These developments have the potential to transform farming entirely toward more modernity, with chances of meeting the ever-growing global population while favoring both the farmer and the environment.

The review of Abhinav Sharma et al. (2020) covers all applications in the areas of machine learning in precision agriculture, which shows how these technologies improve agricultural practice through data-informed insights, advanced monitoring, accurate yield predictions, and resource optimization. These studies elaborate on various machine learning methods to enhance crop monitoring by achieving early detection of diseases, pests, and nutrient deficiencies, resulting in yield losses. It is also discussed whether predictive modeling has been used effectively for better estimations of crop yields, leading farmers to a proper decision over planting and harvesting while optimizing their inputs. Also, the review discusses how machine learning algorithms optimize resources to use less water, fertilizers, and energy through examining environmental conditions and crop requirements. The study highlights how useful these technologies are for improving productivity and sustainability in agriculture. This aside, the significant challenges include managing large datasets, maintaining model accuracy, and the cost involved in deploying these barriers. Even so, the review argues that the revolution by machine learning may change agriculture. Therefore, even the existence of such barriers would improve efficiency and profitability in the case of its removal.

PROPOSED MODULES

In machine learning, several key modules collaboratively form effective models and actionable insights. Data collection gathers extensive, relevant information from sources like sensors or databases that lend themselves to analysis. The dataset will organize this data, including categorizing all the attributes pertinent for building a model. Data preparation then ensues that includes cleaning (handling missing values and errors) and transformation (normalizing and encoding data). Model selection then evaluates and selects algorithms based on their suitability for tasks such as regression or classification, which is guided by performance metrics. In the analyzing phase, the selected model is applied to uncover patterns and insights, while predicting uses the trained model to forecast future outcomes. These interconnected modules ensure a structured, effective approach to transforming raw data into valuable predictions, driving informed decisions and optimizing processes.

DOI: 10.17148/IARJSET.2025.12106

1. Data Collection: Data Collection forms sound and relevant information from which valid, informed decisions and adequate analyses can be drawn. Their efficiencies bring about improved performance in both the public and the private sectors. Research, innovation, and performance evaluation find such data very important, thus finding evidence-based decisions and transparent reporting. Advanced techniques in machine learning revolutionize data collection in agriculture. Farmers use sensors, satellite imagery, and IoT devices to collect comprehensive data on soil conditions, weather, and crop health. Machine learning analyzes this data to provide insights such as optimal planting times, irrigation schedules, and pest control strategies. This approach increases crop yield, reduces resource waste, and promotes sustainability. Gathering diverse and representative datasets. Cleaning and preprocessing to eliminate noise and inconsistencies. Ensuring data accuracy and relevance to reflect real-world scenarios.
For agriculture, including data on nitrogen and potassium levels is essential for understanding soil fertility and optimizing fertilization. Models trained with this data can predict crop yields and recommend nutrient applications, leading to more effective farming strategies.

2. Dataset: Dataset provides a structured collection of data for analysis, modeling, and decision-making. It captures relevant information on specific variables, allowing the identification of patterns, testing hypotheses, and derivation of insights. An organized dataset ensures accuracy, consistency, and ease of access, supporting data-driven decisions and optimizing processes. The dataset comprises 2000 records and 14 columns, each critical to agricultural analysis. States gives the number of states in India to carry out regional comparison; Rainfall measures precipitation in millimeters, affecting crop growth; Ground Water measures the levels of underground water necessary for irrigation planning; Temperature gives climatic data in degrees Celsius, affecting crop development; Soil Type gives soil variations that affect fertility and crop suitability; Season gives optimal times to plant; Crops gives types considered for analysis; Fertilizers Required gives required fertilizers for improving soil and crop productivity; Cost of Cultivation is the total cost incurred in crop growth; Expected Revenues projects financial returns; Quantity of Seeds per Hectare details the amount of seeds needed for planting; Duration of Cultivation shows the number of days required for crop maturity; Demand of Crop classifies demand as high or low to guide market strategy; and Crops for Mixed Cropping identifies suitable crops for intercropping to boost yield.

3. Data preparation: The data preparation step prepares the raw data into a clean, organized, and structured format suitable for analysis and modeling. It includes data cleaning (addressing missing values and errors), data transformation (normalizing and encoding variables), and feature engineering, meaning creating new features or modifying existing ones. Proper data preparation ensures consistency, relevance, and alignment with analytical needs, which improves the accuracy and efficiency of subsequent analyses and machine learning models. Data preparation involves several key steps: cleaning the dataset to handle missing values and errors, transforming numerical and categorical data for consistency, and engineering features to capture underlying patterns. Data is then split into training, validation, and test sets to evaluate model performance and prevent overfitting. For integrations of multivariated input data sources to enhance predictive capability, proper pre-processing ensures great quality and data relevance that, in turn enhances better decision support and increases farmers' efficiency.

4. Model selection: Model selection is an important step in machine learning, which includes the selection of the best algorithm to solve a particular problem and affects performance, accuracy, and utility. The process starts with understanding the problem, such as predicting crop yields or optimizing irrigation, and selecting appropriate models, for example, regression for continuous variables or classification for discrete categories. Model performance can be estimated by means of accuracy and Mean Absolute Error (MAE). On the other hand, cross-validation gives a strong estimate of model performance by training and testing the model on different sub-data sets. There is also balance in terms of complexity versus interpretability. Simpler models are easier to understand, but complex models provide higher accuracy. Computational resources are also taken into account because the model has to be feasible and efficient for actual use in agriculture.

5. Analysis and Prediction: Analysis and prediction are done to extract actionable insights and predict future outcomes based on historical data and patterns. Analysis finds trends and relationships, while prediction estimates future events using statistical models. Using a dataset with seven key features—States, Rainfall, Ground Water, Temperature, Soil type, Season, and Crops—models integrate these variables to optimize farming practices. The analysis reveals patterns that guide decisions on crop types, planting schedules, and resource allocation by examining interactions between these features. Predictive models help determine the best crops for specific conditions and timing, improving yield and efficiency. This approach enhances farming strategies and supports sustainable agricultural practices by aligning planting and harvesting with optimal conditions.

DOI: 10.17148/IARJSET.2025.12106

1. User Onboarding and Profile Setup:
Users create an account, provide necessary information, and set up a profile to personalize their experience or interaction with the system.

2. Initial Data Collection:
Data relevant to the user's preferences, behavior, or requirements is gathered. This could include user inputs, system logs, or external data sources.

3. Data Preprocessing:
Collected data is cleaned, transformed, and organized in such a manner so that the analysis or modeling takes place at good quality with uniformity.

4. Training, Testing, Evaluation:
Machine learning models are trained upon the processed data; they test them to get performance measures, evaluate them, and optimize their accuracy and effectiveness in the context of the desired performance.

5. Prediction and Recommendation:
The system employs trained models for generating predictions or recommending something relevant according to the requirements of the users.

6. Real-Time Feedback and Monitoring:
The system accumulates feedback from users and tracks usage in real time to learn and improve the functionality or output of the system.

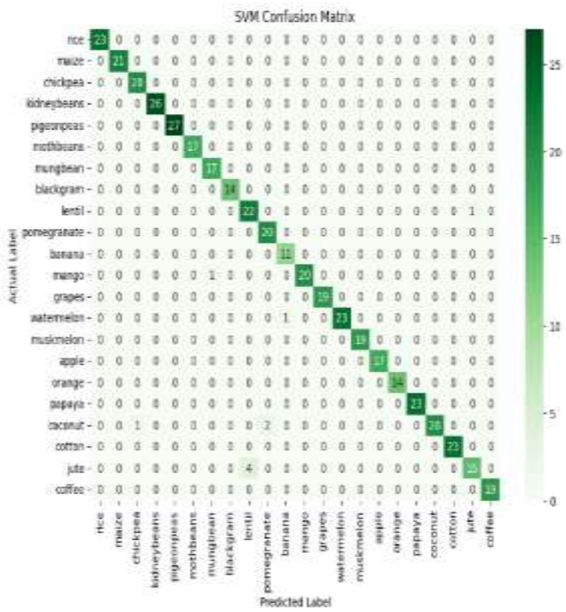
7. Progress Tracking and Reporting:
The progress of the user is tracked and reports are generated, providing insights into performance, milestones, or outcomes.

RESULT AND DISCUSSION

The results of the study show that the Support Vector Machine algorithm majorly outperformed the K-Nearest Neighbor algorithm in terms of predictive accuracy and overall classification performance. The SVM model had an impressive accuracy of 97.73%, superior to 95.68% of the KNN model, and indicates that it might better deal with complex, high-dimensional data than the KNN algorithm. In general, the confusion matrices for the two models have shown that these models are correct in predicting the crops, but for SVM, results were much more accurate and robust. This has been because of the fact that SVM is good in identifying patterns or trends in data; this is quite important in the agricultural application in which variables such as soil type, rainfall, temperature, and nutrient levels will play a major role. The system's real-time adaptability and regional customization further enhanced its utility, enabling tailored recommendations that align with specific farm conditions. These insights proved instrumental in improving crop yield predictions and

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
Impact Factor 8.066 ⁝ Peer-reviewed / Refereed journal ⁝ Vol. 12, Issue 1, January 2025
DOI: 10.17148/IARJSET.2025.12106

optimizing resource usage. By equipping farmers with actionable, data-driven guidance on crop selection, fertilization, and irrigation strategies, the system has the potential to significantly increase productivity while promoting sustainable farming practices. Ultimately, with the integration of advanced machine learning algorithms and real-time data analysis, it actually shows a promising step in making better solutions for present-day agriculture and enabling farmers to use these tools for informed decisions.



SVM Confusion Matrix

SVM Classification Report:

	precision	recall	f1-score	support
apple	1.00	1.00	1.00	23
banana	1.00	1.00	1.00	21
blackgram	0.95	1.00	0.98	20
chickpea	1.00	1.00	1.00	26
coconut	1.00	1.00	1.00	27
coffee	1.00	1.00	1.00	17
cotton	0.94	1.00	0.97	17
grapes	1.00	1.00	1.00	14
jute	0.85	0.96	0.90	23
kidneybeans	0.91	1.00	0.95	20
lentil	0.92	1.00	0.96	11
maize	1.00	0.95	0.98	21
mango	1.00	1.00	1.00	19
mothbeans	1.00	0.96	0.98	24
mungbean	1.00	1.00	1.00	19
muskmelon	1.00	1.00	1.00	17
orange	1.00	1.00	1.00	14
papaya	1.00	1.00	1.00	23
pigeonpeas	1.00	0.87	0.93	23
pomegranate	1.00	1.00	1.00	23
rice	0.94	0.79	0.86	19
watermelon	1.00	1.00	1.00	19
accuracy			0.98	440
macro avg	0.98	0.98	0.98	440
weighted avg	0.98	0.98	0.98	440

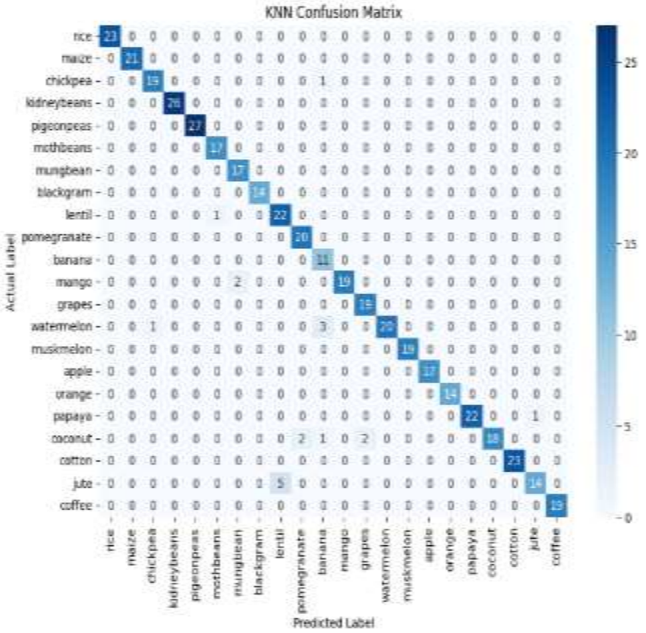
SVM Classification Report

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁐ Peer-reviewed / Refereed journal ⁐ Vol. 12, Issue 1, January 2025

DOI: 10.17148/IARJSET.2025.12106



KNN Confusion Matrix

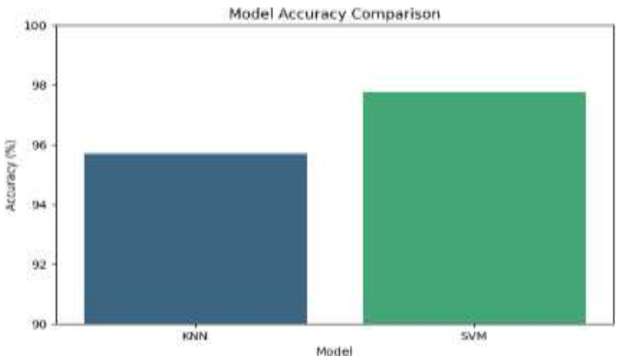
```
KNN Accuracy: 95.68%
SVM Accuracy: 97.73%

KNN Classification Report:
              precision    recall  f1-score   support

       apple       1.00      1.00      1.00        23
      banana       1.00      1.00      1.00        21
   blackgram       0.95      0.95      0.95        20
    chickpea       1.00      1.00      1.00        26
     coconut       1.00      1.00      1.00        27
      coffee       0.94      1.00      0.97        17
      cotton       0.89      1.00      0.94        17
      grapes       1.00      1.00      1.00        14
        jute       0.81      0.96      0.88        23
 kidneybeans       0.91      1.00      0.95        20
      lentil       0.69      1.00      0.81        11
       maize       1.00      0.90      0.95        21
       mango       0.90      1.00      0.95        19
   mothbeans       1.00      0.83      0.91        24
    mungbean       1.00      1.00      1.00        19
   muskmelon       1.00      1.00      1.00        17
      orange       1.00      1.00      1.00        14
      papaya       1.00      0.96      0.98        23
  pigeonpeas       1.00      0.78      0.88        23
 pomegranate       1.00      1.00      1.00        23
        rice       0.93      0.74      0.82        19
  watermelon       1.00      1.00      1.00        19

    accuracy                           0.96       440
   macro avg       0.96      0.96      0.95       440
weighted avg       0.96      0.96      0.96       440
```

KNN Classification Report



Comparison between KNN and SVM

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁒ Peer-reviewed / Refereed journal ⁒ Vol. 12, Issue 1, January 2025

DOI: 10.17148/IARJSET.2025.12106

CONCLUSION

The crop yield prediction system is a major technological leap in agricultural technology through the usage of advanced algorithms of machine learning, K-Nearest Neighbors (KNN) and Support Vector Machine (SVM). The SVM algorithm is more accurate at 97.73% as compared to KNN, which is only 95.68%. It gives more dependable and accurate results. This system stands out as it gives tailored recommendations based on individual farm conditions by analyzing data in real-time and giving feedback. It increases productivity by advising on optimal planting times, fertilization, and resource management, while also promoting sustainability by reducing resource waste. Overall, this technology enables farmers to make informed, data-driven decisions, improving agricultural efficiency and resilience.

Future enhancements for the crop yield prediction system aim to significantly improve its functionality and effectiveness. More accurate and context-specific recommendations will be achieved through the integration of satellite imagery, drone data, and real-time weather updates. Advanced algorithms, such as deep learning and ensemble methods, will be applied to complex data relationships for enhanced prediction accuracy. Dynamic adaptation and learning will enable the system to continually improve its predictions over time, with the introduction of new data and user feedback. Improvements to the user interface will include interactive dashboards and visualizations for better usability and engagement. The system will be scaled and customized to support various farm sizes and specific crops and practices. Integration with existing farm management software will streamline workflows and offer a more comprehensive farm management approach. There will be enhanced predictive features that consider future scenarios like climate change, market fluctuations, and pest threats, which will help in planning better. Also, there will be educational and support resources provided to help the users use the system's predictions and recommendations appropriately.

REFERENCES

[1]. Konstantinos G. Liakos, “ Machine Learning in Agriculture: A Review”, Sensors 2020, 18, 2674; doi:10.3390/s18082674.

[2]. A.Swarupa Rani (2021), “The Impact of Data Analytics in Crop Management based on Weather Conditions”, International Journal of Engineering Technology Science and Research, Volume 4,Issue 5,May.

[3]. Priyanka P.Chandak (2022),” Smart Farming System Using Data Mining”, International Journal of Applied Engineering Research, Volume 12, Number 11.

[4]. Ramesh Babu Palepu (2021) ” An Analysis of Agricultural Soils by using Data Mining Techniques”, International Journal of Engineering Science and Computing, Volume 7 Issue No. 10 October.

[5]. Shreya S. Bhanose, Kalyani A. Bogawar (2020) “Crop And Yield Prediction Model”, International Journal of Advance Scientific Research and Engineering Trends, Volume 1,Issue 1, April 2016.

[6]. Rajeswari and K. Arunesh (2022) “Analysing Soil Data using Data Mining Classification Techniques”, Indian Journal of Science and Technology, Volume 9, May.

[7]. Pritam Bose, Nikola K. Kasabov (2022), “Spiking Neural Networks for Crop Yield Estimation Based on Spatiotemporal Analysis of Image Time Series”, IEEE Transactions On Geoscience And Remote Sensing.

[8]. Shreya S.Bhanose (2021),”Crop and Yield Prediction Model”, International Journal of Advance Scientific Research and Engineering Trends, Volume 1,Issue 1,ISSN(online) 2456- 0774,April.

[9]. R. R. Rubia Gandhi; J Angel Ida Chellam; T. N. Prabhu;C. Kathirvel; M. Sivaramkrishnan; M. Siva Ramkumar. IEEE Xplore (2022); doi:10.1109/ICCMC53470.2022.9753841

[10]. Abhinav Sharma; Arpit Jain; Prateek Gupta; Vinay Chowdary. IEEE Xplore (2020).doi: 10.1109/ACCESS.2020.3048415