Implementation of Optimal Crop Recommendation System using Machine Learning Algorithms

Siram Sanjana
Department of ECE
V. R. Siddhartha Engineering College
Vijayawada, AP, India
sanjanasiram03@gmail.com

P. Satyanarayana
Department of ECE
V. R. Siddhartha Engineering College
Vijayawada, AP, India
satya.sp14@gmail.com

Inumula Anitha
Department of ECE
V. R. Siddhartha Engineering College
Vijayawada, AP, India
inumalanitha507@gmail.com

K Nithin Kumar
Department of ECE
V. R. Siddhartha Engineering College
Vijayawada, AP, India
knk51102@gmail.com

K. N. V. Suresh Varma
Department of ECE
S.R.K.R. Engineering College
Bhimavaram, A.P, India.
knvsureshvarma@yahoo.co.in

Dr. V. Gokula Krishnan
Department of CSE
Saveetha School of Engineering,
Saveetha Institute of Medical and Technical Sciences (SIMATS)
Chennai, Tamil Nadu, India
gokul_kris143@yahoo.com

2024 International Conference on Expert Clouds and Applications (ICOECA) | 979-8-3503-8579-3/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICOECA62351.2024.00123

***Abstract*- Agriculture consumes a significant 85% of the world's freshwater resources, a trend expected to persist due to growing population and food demand. To overcome this challenge, developing strategic solutions blending scientific principles and advanced technology are essential. Leveraging machine learning technologies, an innovative model is proposed to minimize crop loss due to adverse weather conditions. The proposed integrated system addresses three critical aspects of agriculture: crop selection, fertilizer recommendations, and automated irrigation. By analyzing various parameters such as pH levels, current weather, and soil nutrients (potassium, phosphorous, nitrogen), the proposed model adeptly selects optimal crops based on various environmental factors. A mobile application translates soil conditions into graphical data for easy interpretation. Evaluation of machine learning algorithms reveals Support Vector Machine (SVM) as the highest accuracy at 95.89%, followed closely by Naïve Bayes (NB) at 95.24%. Decision Tree also performs competitively at 95.05%, while K-nearest Neighbors (K-NN) trails slightly with 93.81% accuracy. These findings depict the efficacy of SVM in accurately classifying datasets, presenting a promising path for addressing water demand in agriculture.**

Keywords: Machine Learning, Sustainable Agriculture, Automatic Watering

I. INTRODUCTION

Crop prediction plays an essential role in optimizing agricultural productivity and resource utilization. Traditional methods [11] often rely on historical data and expert knowledge, which may not always capture the complex relationships between various environmental factors and crop outcomes. Machine learning techniques, particularly SVM, offer a data-driven approach to crop prediction, enabling more accurate and efficient forecasting. Agriculture is currently accounting for 85% of available freshwater resources globally, this substantial share in water consumption is anticipated to persist and potentially increase due to the increasing population growth and food demand. Recognizing the critical implications of this trend, there is a need to develop strategic solutions rooted in both scientific principles and cutting-edge technology. Addressing the pressing challenges in water resource management within the agricultural sector is imperative to ensure sustainable practices and meet the growing demands of an expanding global population.

In pursuit of advancing agricultural practices, this project aims to develop a comprehensive system utilizing Machine Learning algorithms [12]. This system will be designed to seamlessly integrate three key functionalities: crop selection, autonomous irrigation, and fertilizer recommendation. Here, the primary research objective is to create a solution that empowers farmers by reducing physical labor, optimizing energy usage, and ultimately enhancing productivity. The foundational aspect of the research lies in the background study, a critical component of any scientific investigation.

Precision agriculture is dominant in addressing the multifaceted challenges of agricultural production, encompassing productivity, environmental impact, food security, and sustainability. As the global population necessitates a substantial increase in food production, precision agriculture becomes instrumental in achieving this goal while ensuring the availability and high nutritional quality of food worldwide. Recent strides in sensor technology, particularly in the context of smart farming, along with the evolution of Internet of Things (IoT) technologies [13], have propelled the development of advanced irrigation systems in agriculture. IoT plays a pivotal role in the ongoing digital transformation and in the diverse applications of agriculture.

In the algorithmic aspect of the study, the focus revolves around crop recommendation utilizing key parameters such as Nitrogen (N), Phosphorous (P), and Potassium (K) values, employing Machine Learning (ML) algorithms. Specifically, a decision tree algorithm is proposed for a smart irrigation system, incorporating temperature, humidity, and moisture as critical parameters. Furthermore, an in-depth analysis of various machine learning concepts [14] is undertaken for an Internet of Things (IoT)-based smart agriculture system.

The study introduces the use of Support Vector Machine (SVM) and decision tree algorithms to determine the most suitable crop based on given soil data, aiming to enhance growth through optimized farming processes. Machine

979-8-3503-8579-3/24/$31.00 ©2024 IEEE
DOI 10.1109/ICOECA62351.2024.00123

learning, coupled with computer vision, is explored for the classification of crop images to monitor crop quality and assess yield. Transitioning to the systems and solutions proposed, a comprehensive smart farming system [15] is suggested, leveraging artificial intelligence (AI) for prediction. This system undergoes three primary phases: data collection through sensors deployed in agricultural fields, subsequent data cleaning and storage, and finally, predictive processing using various AI techniques. The system architecture is based on Wireless Sensor Network (WSN) technology.

Additionally, a novel architecture is proposed for designing Key Performance Indicators (KPIs) in precision agriculture. This entails creating a framework for assessing and monitoring performance metrics critical to precision farming practices. Both algorithmic and system-focused approaches contribute to the development of intelligent and efficient solutions for modern agriculture.

II. LITERATURE SURVEY

Rushika Ghadge, [1] Using both unsupervised and supervised learning algorithms such as Kohonen Self-Organizing Map and Back Propagation Network, the system trains on a dataset to classify soil into organic, inorganic, and real estate categories. This classification aids in predicting soil types. The system assesses the accuracy achieved through various network learning techniques, delivering the most precise result to the end user. The system not only evaluates soil quality and predicts crop yield accordingly but also offers fertilizer recommendations based on the soil's quality, if necessary.

Shridhar Mhaiskar, [2] uses the ARIMA model to estimate temperature, moisture content, and pH for crop prediction. The model predicts the value of a given parameter after a month by using the values from the database as input. Following the prediction of values, the K means technique is used to classify the crops according to pH value, resulting in the creation of k clusters of crops with comparable pH values. The top N appropriate crops are predicted using the KNN algorithm and presented to the user.

T Raghav Kumar, [3] Based on the values obtained in real time, the Machine Learning Algorithm (KNN) computes the parameter to recommend the crop that should be grown in that specific field. For crop prediction, a standardized dataset with the minimal specifications for a given crop is kept up to date. The field where the readings must be computed is expanded to include the sensors. The data is sent in real time to the cloud server by the DHT11, MQ2, Soil Moisture Sensor, and Light Intensity Sensor.

Akash Raj N, [4] The system assesses the quality of crops by considering predetermined weather and soil parameters through the application of trained datasets and Supervised and Reinforcement machine learning models. In anticipation of any adverse conditions, the system proactively employs alternative and preventive measures to protect both the planted crops and agricultural land. The precise methods used to predict the best times to sow, reap, and harvest in order to maximize overall yield are among the features of the modern agricultural revolution.

Garg G., Gupta S., Mishra P., [5] propose a sustainable real-time crop disease diagnosis and prevention system called CROPCARE is introduced in this article. By encouraging scientific methods, CROPCARE seeks to address agricultural difficulties by utilising mobile vision, IoT, and Google Cloud services. Through a mobile application, the system uses SRCNN and MobileNet-V2 for illness diagnosis. It is coupled with IoT sensors and Google Cloud for sustainability. CROPCARE meets farmer requirements by offering advice on weather and soil conditions, disease control, and help for both Hindi and English dictionaries. The Plant Village dataset is used for validation, which highlights the robust performance of the suggested system and highlights its potential for real-world application.

Jay Gholap, Anurag Ingole [6] makes it possible to anticipate soil properties like phosphorus content. In order to attain high prediction accuracy, the authors here employ a variety of classification algorithms, including Naive Bayes, C4.5, Linear Regression, and Least Median Square. Farmers may find out if the soil is suitable for a given crop by using this approach, which can be quite helpful.

T.R. Lekhaa et. al. [7] This paper addresses the challenge of farmers relying on intuition and outdated practices for crop selection due to changing soil and climate conditions. It introduces an innovative approach using Robotic Process Automation (RPA) for efficient data collection in crop recommendation models. The proposed system utilizes RPA tools like UiPath to automate data gathering, ensuring accuracy and timeliness. Unlike traditional methods, this model streamlines the multiple stages of data collection, processing, and analysis, enhancing overall efficiency. The paper highlights the application of RPA in the agricultural domain, specifically focusing on automating tasks related to meteorological and climate data for improved crop recommendations..

Punith Kumar; H Varun Prabhu [8] provides a thorough analysis on employs 50% of the country's workforce and generates a large amount of jobs. Farmers' ignorance of the elements of soil and surrounding conditions frequently results in them selecting less-than-ideal crops, which lowers yield. In order to help farmers make educated decisions, a scientific approach is suggested to help anticipate viable crops based on numerous contributing variables. For the purpose of improving agricultural techniques, reducing risks, guaranteeing food security, and promoting sustainable agriculture, accurate crop forecasting is essential. This work uses a dataset that includes meteorological and soil factors with ensemble stacking machine learning algorithms to make very accurate crop recommendations. With the help of these adaptable and effective instruments, farmers should be able to achieve better agricultural results.

R Kumar [9] This study highlights the significance of crop selection and examines the various factors that impact it, including market price, production rate, and governmental regulations. The crop selection issue is effectively addressed by the Crop Selection Method (CSM) proposed in this research. The CSM proposes a strategy that advocates for the selection of a diverse range of crops throughout the season, while considering variables such as weather conditions, soil composition, water density, and crop type. This approach results in an increased net yield rate for the chosen crops. The CSM's correctness is based on the anticipated values of significant parameters. As a result, a prediction mechanism with increased performance and accuracy must be included.

N. V Suresh Krishna; Pokala Neeraj [10] provides a foundational element of the Indian Economy, historically ensured food security and financial stability. However, unprecedented weather conditions have adversely impacted crop yields and farmers' income. This research aids novice farmers by employing machine learning algorithms to guide optimal crop selection based on soil parameters like Nitrogen, Phosphorous, Potassium, temperature, and humidity. The system, utilizing the Random Forest algorithm, assists farmers in achieving successful harvests. Additionally, a web-based application is under development, providing users with crop-specific nutrient data. Experimental results on web application data samples demonstrate a high accuracy of 91.2% in predicting crop yields, with swift response times.

III. PROPOSED METHODLOGY

Agriculture faces significant challenges in optimizing crop production among increasing water scarcity and changing climatic conditions. In this paper, a novel approach is proposed named as leveraging Support Vector Machine (SVM) to address crop loss due to adverse weather conditions. The proposed method integrates machine learning techniques to optimize crop selection, recommend suitable fertilizers, and automate irrigation, thereby enhancing agricultural productivity and sustainability.

The design and development of the proposed system with several functions, such as crop recommendation, suitable fertilizer suggestions and automatic watering is necessary to increase the agricultural output. The method recommends a crop that could grow in a given area by examining various soil and climatic parameters, which includes measuring humidity, temperature and rainfall, as well as assessing the soil's NPK and pH levels. The flow diagram of suggested method using machine learning techniques is shown in Figure 1. Compared to more recent algorithms such as neural networks, Naïve Bayes (NB), decision trees, and K-nearest neighbours (K-NN), they provide two key benefits: faster processing speed and improved results with small sample sizes (thousands). Using kernel functions to perform non-linear classification, SVM is considered as one of the finest algorithms since it can handle high-dimensional data.

Steps involved in Crop Predication using SVM algorithm:

In the training step, the crucial decision is to select the most suitable machine learning model or algorithm that delivers accurate output results. For this purpose, the Support Vector Machine (SVM) algorithm has been chosen. SVMs are highly versatile and can be applied in crop recommendation systems to classify crops based on various factors, including soil type, climate, and historical crop data.
Step 1: Data Collection and Preprocessing: Data based on the factors that influence crop choice, such as soil type, climate, historical crop yield, and other relevant features are collected and then the data will be cleaned and formatted properly by handling missing values, encoding categorical variables, and scaling numeric features.
Step 2: Data Labelling: The data will be labelled with the appropriate crop classes. These labels can represent the different types of crops to recommend.
Step 3: Data Splitting involves dividing the dataset into two distinct sets: the training set and the testing set. A common approach is to allocate 70-30 or 80-20 of the dataset, with the larger portion designated for training purposes. This ensures that the model is adequately trained on a substantial portion of the data, while also providing a separate set for evaluating the model's performance.

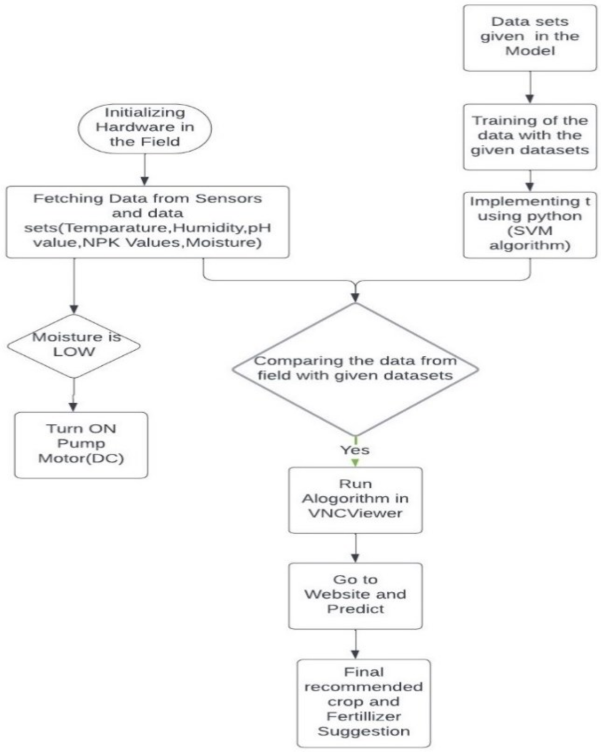

Figure.1. Flow chart of proposed method

Step 4: Model Training phase trains the SVM model on the training dataset by using the labelled data. User can use libraries like scikit-learn in Python for this purpose. The hyperparameters like the choice of kernel (linear, radial basis function, polynomial, etc.) and regularization parameters (C) can be fine-tuned.
Step 5: The trained model is then evaluated on the testing dataset using common metrics to assess its performance and identify the areas for improvement.
Step 6: Once trained SVM model, user can utilize it to recommend crops for a specific set of input features. User input these features and let the SVM model predict the most suitable crop class.

The integration of a moisture sensor and an automatic watering system adds a dynamic element to the proposed system, ensuring that the crops receive necessary hydration based on real-time moisture requirements. This promotes water efficiency while also making a significant contribution to the overall health of the crops. The decision-making process is accurate when NPK and pH sensors are included. To determine which crops would be best under these conditions, the collected data is combined with additional environmental factors and put into a machine learning method known as Support Vector Machine (SVM). By using modern technologies, farmers receive customized crop recommendation while improving the crop-selection process. The pseudocode for the proposed method is shown in figure 2.

```
1.Initialize the parameters (If Moisture is low turn on pump) and data sets
2.Import necessary libraries for implementation
3.Train and load the SVM Model as per crop requirement
4.Create a Flask web application
5.Define a crop_mapper & fertilizer_dic dictionaries
6.Identify routes for the web application
   6.1. Home route ('/'), Dashboard route ('/dashboard') shows the respective page
   6.2. Predict route ('/predict') handles prediction and shows the result page
      6.2.1 If the request method is POST:
         6.3.1.1. Get input features from the form
         6.3.1.2. Make a prediction using the SVM model
         6.3.1.3. Get the crop name from the model prediction
         6.3.1.4. Get the required nutrients from a CSV file based on the crop name
         6.3.1.5. Get the corresponding fertilizer suggestion from the fertilizer_ dic
7.Measure the accuracy of suggested SVM model
8.Compare the results of suggested method with existing methods
```

Figure 2. Pseudocode for proposed method

The SVM model's predictions are then translated into user-friendly information through a mapping process by using a dictionary containing nutrient level information for various crops. This approach ensures that farmers can easily understand and implement the recommendations provided by the proposed system. Once the predicted crop is determined, the system takes a step further by comparing recommended nutrient levels with user-input levels. This enables the identification of specific nutrient deficiencies (N, P, or K), and the system offers personalized fertilizer recommendations to address these deficiencies. This aspect of the method ensures a targeted and efficient approach to nutrient management.

The presentation of results through an HTML template enhances user accessibility and comprehension. Farmers can easily access information about the predicted crop and receive practical advice on fertilizer application. The user-friendly interface contributes to the practicality and user acceptance of the system. Moreover, the continuous monitoring of soil moisture content and the automatic irrigation system add a proactive element to the application, ensuring that crops receive consistent care. Leveraging a vast database of over 3000 datasets, the system provides farmers with not only crop recommendations but also valuable insights on rotational crops and pH level adjustments, thereby optimizing the overall yield potential.

IV. RESULTS & DISCUSSION

The proposed model leverages machine learning to prevent crop losses in adverse weather conditions by handling key agricultural challenges like crop selection, fertilizer suggestions, and automated irrigation. Through advanced algorithms, the proposed model analyzes different parameters such as pH levels, weather conditions, and nutrient concentrations like potassium, phosphorus, and nitrogen. This enables real-time selection of the best crops based on various soil and environmental conditions.

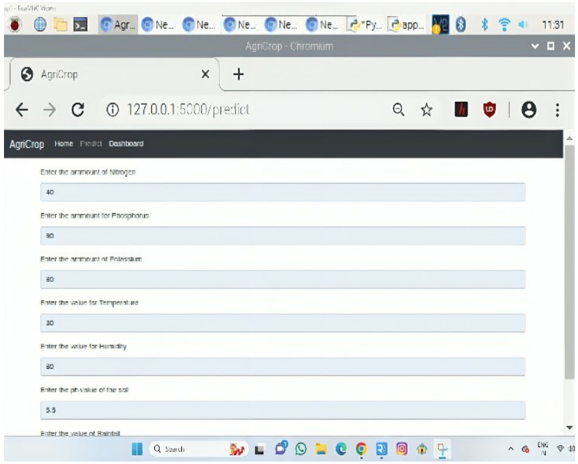

Figure 3. Input datasheet used in implementation

Figure (3) shows the datasheet used in the proposed method. In this datasheet, accurate predictions are required for essential parameters such as NPK levels, humidity, temperature, rainfall, and pH values. It is crucial to determine the appropriate values for these factors based on the information provided in the datasheet.

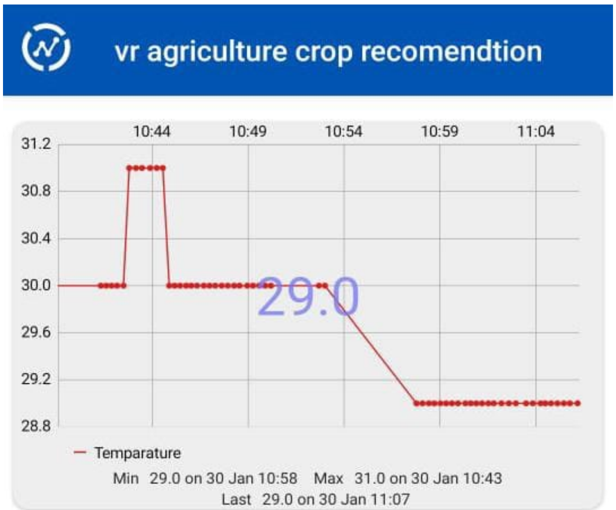

Figure 4. Graphical representation of temperature vs time at different dates

Figure (4) represents the parameter (temperature) representation of suggested method. The graph illustrates the fluctuation in temperature, with the y-axis representing temperature values and the x-axis denoting varying time in different dates. It provides a visual representation of how temperature changes over the time. From figure 3 maximum temperature range is 31 degrees, minimum temperature range is 29 degrees.

Figure (5) shows the graphical representation of humidity vs time for suggested method. The graph illustrates the fluctuation in Humidity, with the y-axis representing Humidity values and the x-axis denoting time at varying dates. It provides a visual representation of how Humidity changes over time. From figure 4, maximum humidity level is 80, minimum humidity level is 60.

Figure 5. Graphical representation of Humidity vs time at different dates

Figure 6. Graphical representation of moisture vs time at different dates

Figure (6) depicts the graphical representation of moisture vs time for suggested method. The graph illustrates the fluctuation in moisture, with the y-axis representing moisture values and the x-axis denoting time at varying dates. It provides a visual representation of how moisture changes over time. From figure 5, maximum moisture level is and minimum is 0.

Figure (7) illustrates the results of suggested machine learning algorithm with various parameters (Temperature, humidity, moisture, status of pump and light intensity). The results presented in this section provide a complete overview of the system's performance, showcasing key parameters such as temperature values, light intensity levels, motor status, and the operational state (ON/OFF) of the pump. These results are derived from the data collected by the employed sensors, offering a detailed understanding into the real-time conditions and functioning of the monitored environment. The accurate and reliable information obtained from these sensors serves as a foundation for assessing the effectiveness and efficiency of the implemented system in managing and controlling the various aspects of the monitored system.

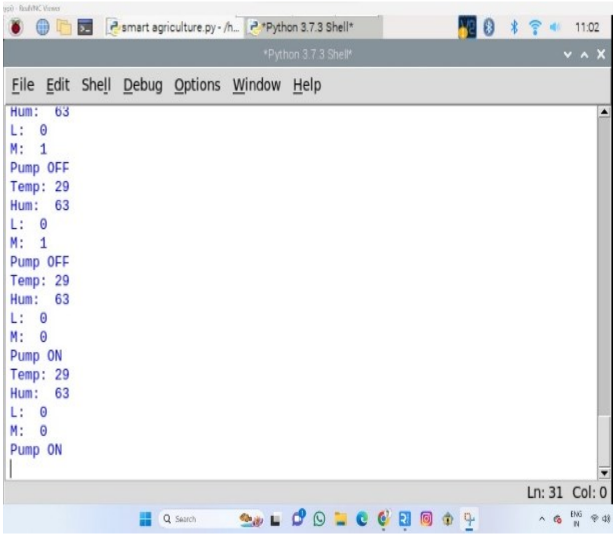

Figure 7. Results of suggested machine learning algorithm with various parameters

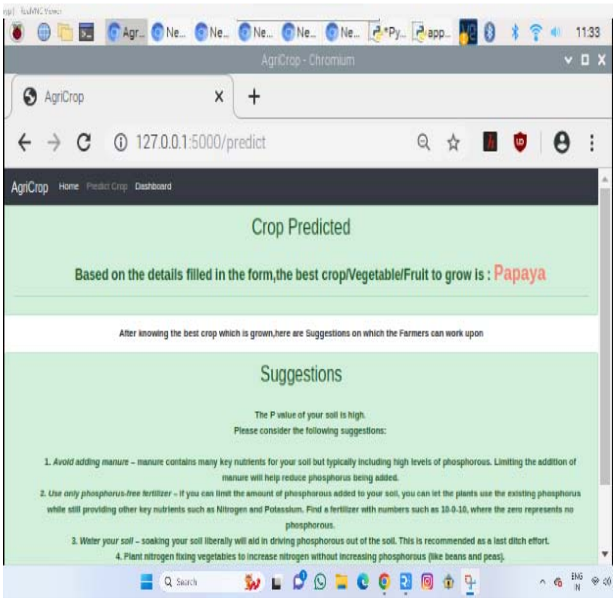

Figure 8. Results of crop prediction and fertilizer suggestion of proposed method using python

Figure (8) depicts the Results of crop prediction and fertilizer suggestion of proposed method using python. The crop recommendation is visualized in the output display, highlighting crucial details for enhanced visibility. Subsequent to proposing the optimal crop, the system goes a step further by furnishing detailed suggestions regarding the specific fertilizers for the recommended crop, along with guidance on necessary pH value adjustments. This comprehensive approach ensures that users receive not only the most suitable crop recommendation but also valuable insights into the associated fertilization and pH management strategies for optimal agricultural results.

Type of algorithms	**Naïve Bayes (NB)**	**Decision tree**	**K-nearest Neighbors (K-NN)**	**Support Vector Machine (SVM)**
Accuracy	95.24%	95.05%	93.81%	95.89%
Size of the sample	Small size	Small size	Small size	Large size
Learning Speed	Highest speed	High speed	Highest speed	Low speed

Table1: Comparing SVM with other algorithms

In this section, the performance evaluation of the Naïve Bayes (NB), Decision Tree, K-nearest Neighbors (K-NN), and Support Vector Machine (SVM) algorithms based on their accuracy metrics is shown in Table (1). The Support Vector Machine (SVM) algorithm exhibited the highest accuracy of 95.89%, closely followed by Naïve Bayes (NB) with an accuracy of 95.24%. Decision Tree also demonstrated competitive performance with an accuracy of 95.05%. However, K-nearest Neighbors (K-NN) showed slightly lower accuracy compared to the other algorithms, achieving 93.81%. These results indicate that the selected algorithms perform effectively in classifying the dataset, with SVM showing the highest classification accuracy among them.

V. CONCLUSION

This study presented an innovative approach to address the emerging challenge of optimizing agricultural practices amidst increasing water scarcity. The proposed solution integrates cutting-edge machine learning techniques with agricultural science to mitigate crop losses resulting from adverse weather conditions. By harnessing the power of data analytics, the proposed model offers tailored recommendations for crop selection, fertilizer application, and automated irrigation systems. The resultant findings highlight the potential of selecting SVM as a robust tool for optimizing agricultural decision-making processes, particularly in the context of water resource management. Further research could focus on enhancing the proposed model by integrating additional data sources such as satellite imagery, soil quality metrics, and real-time weather data. Incorporating these factors could improve the accuracy and robustness of the crop recommendations provided by the proposed system.

REFERENCES

[1] Rushika Ghadge, Juilee Kulkarni, Pooja More, Sachee Nene, Priya R L, “Prediction of Crop Yield using Machine Learning,” International Research Journal of Engineering and Technology (IRJET) Feb 2018, pp. 2237-2239

[2] Shridhar Mhaiskar, Chinmay Patil, Piyush Wadhai, Aniket Patil, Vaishali Deshmukh, “A Survey on Predicting Suitable Crops for Cultivation Using IoT,” International Journal of Innovative Research in Computer and Communication Engineering January 2017, pp. 318-323

[3] T Raghav Kumar, Bhagavatula Aiswarya, Aashish Suresh, Drishti Jain, Natesh Balaji, Varshini Sankaran, “Smart Management of Crop Cultivation using IOT and Machine Learning,” International Research Journal of Engineering and Technology (IRJET) Nov 2018, pp. 845-850

[4] Akash Raj N, Balaji Srinivasan, Deepit Abhishek D, Sarath Jeyavanth J, Vinith Kannan A, “IoT based Agro Automation System using Machine Learning Algorithms”, International Journal of Innovative Research in Science, Engineering and Technology November 2016, pp. 19938-19342

[5] Garg, G., Gupta, S., Mishra, P., Vidyarthi, A., Singh, A., Ali, A.: CROPCARE: an intelligent real-time sustainable IOT system for crop disease detection using mobile vision. IEEE Internet Things J. 10(4), 2840–2851 (2023)

[6] Jay Gholap, Anurag Ingole, Jayesh Gohil, Shailesh Gargade and Vahida Attar, “Soil Data Analysis Using Classification Techniques and Soil Attribute Prediction”, International Journal of Computer Science Issues, Volume 9, Issue 3

[7] T.R. Lekhaa, “Efficient Crop Yield and Pesticide Prediction for Improving Agricultural Economy using Data Mining Techniques”, International Journal of Modern Trends in Engineering and Science (IJMTES), 2016, Volume 03, Issue 10

[8] A Punith Kumar; H Varun Prabhu; H N Champa:Crop Recommendation using Ensemble Stacking Machine Learning approach , IEEE 3rd Mysore Sub Section International Conference ,2023.

[9] R. Kumar, M. P. Singh, P. Kumar and J. P. Singh, "Crop Selection Method to maximize crop yield rate using machine learning technique," 2015 International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM), Avadi, India, 2015, pp. 138-145.

[10] N. V Suresh Krishna; Pokala Neeraj; Bathina Surya; B.V. Prudhvi; V. Hima Deepthi:Machine Learning Algorithms for Crop Yield Prediction in Real-Time Scenarios, 2023 4th International Conference on Signal Processing and Communication (ICSPC).

[11] Haedong Lee and Aekyung Moon, “Development of Yield Prediction System Based on Real-time Agricultural Meteorological Information”, 16th International Conference on Advanced Communication Technology, 2014

[12] Anshal Savla, Parul Dhawan, Himtanaya Bhadada, Nivedita Israni, Alisha Mandholia , Sanya Bhardwaj “Survey of classification algorithms for formulating yield prediction accuracy in precision agriculture”, Innovations in Information, Embedded and Communication systems (ICIIECS).

[13] Aymen E Khedr, Mona Kadry, Ghada Walid (2015), Proposed Framework for implementing Data Mining Techniques to Enhance Decisions in Agriculture Sector Applied Case on Food Security Information Center Ministry of Agriculture, Egypt, December 2015, Procedia Computer Science 65:633-642

[14] Devi, K & Kumar, Josephine. (2023). An efficient data collection tool for crop recommendations model using Robotic Process Automation. 1-5. 10.1109/ICCCNT56998.2023.10308274.

[15] Janmejay Pant, R.P. Pant, Manoj Kumar Singh, Devesh Pratap Singh, Himanshu Pant, Analysis of agricultural crop yield prediction using statistical techniques of machine learning, Materials Today: Proceedings,Volume 46, Part 20,2021,Pages 10922-10926.