

Research

Revolutionizing Somali agriculture: harnessing machine learning and IoT for optimal crop recommendations

Mohamed Omar Abdullahi¹ · Abdulkadir Dahir Jimale¹ · Yahye Abukar Ahmed¹ · Abdulaziz Yasin Nageye¹

Received: 3 September 2023 / Accepted: 15 February 2024

Published online: 19 February 2024

© The Author(s) 2024 [OPEN](#)

Abstract

Agriculture drives Somalia's economy, but challenges like unpredictable weather, limited resources, and poor infrastructure hamper productivity and economic progress. Emerging technologies like machine learning and IoT offer transformative solutions, optimizing crop yield and resource use. This research demonstrates the substantial impact of integrating Machine Learning (ML) and Internet of Things (IoT) technologies to improve agricultural decisions in Somalia. The study conducts a comprehensive comparison of Decision Trees (DT), K-nearest Neighbor (KNN), and Random Forest algorithms within a Crop Recommendation System. The Decision Tree algorithm emerges as the standout performer, boasting an impressive accuracy of 99.2% and achieving well-balanced precision, recall, and F1-score metrics. Its transparency and interpretability render it an optimal choice for guiding agricultural choices. Despite slightly trailing in performance, KNN and Random Forest algorithms achieve accuracies of 97.2% and 99.0% respectively, presenting valuable alternatives for various contexts. The successful implementation of the Crop Recommendation System, particularly in Somalia's Balcad District, underscores the tangible advantages of real-time IoT data and the Decision Tree model. This system enables farmers to optimize crop selection, thereby enhancing sustainability and yield potential. In a broader context, this research underscores the capacity of data-driven agriculture to tackle food security challenges and drive economic advancement. The transparent and accurate attributes of the Decision Tree algorithm, coupled with IoT capabilities, establish a framework for modernizing traditional farming practices and shaping a more productive future for Somalia and the global agriculture. The continuous evolution of crop recommendation systems holds promise for further transformative opportunities within the agricultural sector.

Article Highlights

- One algorithm, called Decision Trees, was good at predicting crops, even better than two others, KNN and Random Forest.
- Using Decision Trees with smart tech in Somalia helped farmers choose better crops and grow more food.
- This technology could change farming in tough places, making it better, helping everyone get more food, and making farming smarter for the future.

Keywords Agriculture · Data driven farming · Somalia · Internet of Things · Machine learning

✉ Mohamed Omar Abdullahi, momar@simad.edu.so | ¹Faculty of Computing, SIMAD University, Mogadishu, Somalia.



Discover Applied Sciences

(2024) 6:77

| <https://doi.org/10.1007/s42452-024-05739-y>

1 Introduction

Agriculture plays a significant role in Somalia's economy, contributing to its economic growth. Studies have shown a positive relationship between the agriculture sector and the country's Gross Domestic Product (GDP) [1, 2]. In Somalia, most of the population relies on farming for their livelihoods, and it is essential to ensure the optimal utilization of available resources and maximize agricultural productivity [3]. However, the country faces several challenges in agriculture, including unpredictable weather patterns, limited access to resources, and low crop yields [4]. The agriculture sector needs innovative solutions to address these challenges to help farmers make informed decisions about their crops [5]. A study on agricultural productivity growth and incidence of poverty in Africa, including Somalia, found that the continent experienced improved technology with a 2.1 percent upward shift in the production frontier and a 1.8 percent decline in efficiency [6]. East, South, and North Africa experienced growth of 3.3, 2.6, and 3.6 percent, respectively [6].

Livestock and crops remain the primary sources of economic activity, employment, and exports in Somalia, with agriculture's share of GDP at approximately 75%, representing 93% of total exports, mostly linked to robust livestock exports in the recent pre-drought years [7]. However, the agriculture sector in Somalia faces several challenges, including environmental degradation, political instability, and climate change. Climate change has significantly impacted production in Somalia, with temperature changes leading to reduced soil moisture, evaporation, drier conditions, and rain failures [8, 9]. In addition, farming practices tend to be constrained by skill level, a lack of government extension services, few protected storage facilities, a lack of technology, and poor infrastructure [10, 11]. Moreover, food security is a significant issue in Somalia, with acute food insecurity in remote and politically sensitive regions close to Somalia. The area suffers from acute food insecurity, and the dependency on food programs is very high [12].

Somalia has a vast potential for agriculture, and geospatial techniques have been used to assess the country's agricultural potential. The terrain of Somalia is primarily flat, with river basins that can be used as arable lands for agriculture. Groundwater points can sustainably irrigate large parcels for crop production [13]. Despite challenges such as weak institutions, insecurity, and environmental degradation, agriculture has been identified as a critical sector for Somalia's economic development [14]. Farmers in Afgoye District, Lower Shabelle, face unstable weather, water scarcity, pests damaging crops, poor transportation, and lack of access to seeds and fertilizers [15]. However, there are potential solutions. Implementing simple agricultural best management practices can increase agriculture production in Somalia [16].

Machine learning (ML) and Internet of Things (IoT) technology have the potential to revolutionize the agriculture industry. With the help of these technologies, it is possible to optimize crop yield, reduce resource consumption, and improve farm management [17]. ML algorithms and IoT devices can analyze data collected from agricultural fields to monitor crop growth, detect diseases, and manage water resources [18]. Integrating IoT networks with sensors can create a "smart agriculture" system to analyze all information generated in agricultural operations [19]. ML algorithms can assist in analyzing agricultural sensor data and improve forecasting, decision-making, and sensor dependency management [20]. Applying these technologies can help farmers make informed decisions and improve the productivity and quality of agricultural products [21]. Integrating ML and IoT technology can help create a more sustainable and productive agriculture industry.

This paper will discuss designing and implementing a crop recommendation system for Somalia farms. The system will use IoT sensors to collect soil conditions and weather data. ML algorithms will then use this data to train and test machine models. The best model will be used to build a system recommending crops most likely successful in a given area. By leveraging ML algorithms, the system can also provide personalized recommendations based on the specific needs of each farm. This research paper explores the potential of such a system for Somalia farms, highlighting its benefits, challenges, and prospects. A crop recommendation system using IoT and ML techniques could significantly impact food security in Somalia.

After the introduction, the study analyzes existing literature on IoT and ML technologies in agriculture, specifically focusing on their application in developing regions. The methodology section explains the steps undertaken to create and execute the crop recommendation system. This includes gathering soil and weather data, choosing and training ML algorithms, and constructing the recommendation model. The ensuing section presents the results and discussions, which thoroughly examine the performance of the developed system, explain its implications for agriculture in Somalia, and address any issues encountered during implementation. The report finishes by

summarizing the main findings and emphasizing the potential influence of the crop recommendation system on food security in Somalia. It also proposes future research directions to improve the system's effectiveness and scalability.

2 Related work

Crop recommendation systems are crucial for optimizing agricultural practices by providing farmers with informed decisions on suitable crop selections based on environmental factors and historical data [22]. This literature review presents an overview of existing research in the field, highlighting methodologies, algorithms, and data sources used. By analyzing related work, this review identifies the current state of the art, challenges, and future directions in crop recommendation systems, contributing to sustainable and productive agriculture.

Krishna and Amrutha propose an ensemble model for crop recommendation in agriculture using ML algorithms. Their proposed model considers soil type, location, and economic factors and has been tested against previous research. The authors suggest that their model can help farmers increase productivity, reduce soil degradation, reduce chemical usage, and efficiently use water resources. The authors compare the performance of various machine learning algorithms and find that an Artificial Neural Networks (ANN) with a rectified linear unit (ReLU) activation function gives the best performance. They also compare their proposed model with existing ensemble techniques and find it achieves higher accuracy [23].

Dhabarde et al. propose an IoT and ML based Agriculture system that can assist farmers or agriculturists in crop prediction based on Metrological Agriculture theory by getting live Metrological data from the crop field using IoT technology and ML for prediction, which will enable smart farming and increase their overall yield and quality of products [24].

Madhuri and Indiramma propose a system that uses ANN to recommend suitable crops based on soil properties, crop characteristics, and climate parameters. The user interface developed takes the location-specific soil properties as real-time input and recommends the suitable crop considering the input and climate parameters. The model was tested with ANN and a Decision Tree (DT), and the overall accuracy value of ANN is 96%, whereas the accuracy value of the DT is 91.5% [25].

Ali et al. propose a ML based Crop Recommendation System for Local Farmers of Pakistan that suggests the appropriate crops based on the temperature. This system reduces the monetary losses faced by the farmers caused by establishing unfavorable harvests and provides information on the seasonal characterization of yields. The proposed algorithm has an average accuracy of 90% on the given dataset [26].

The convergence of ML and IoT for crop recommendation systems presents a promising avenue for bolstering agricultural productivity. Through empirical experiments comparing various algorithms, particularly the DT, this study underscores its exceptional accuracy and interpretability, positioning it as a leading contender for precise crop recommendations. However, existing research lacks tailored, adaptable systems attuned to regional nuances and specific environmental factors. Conducting these experiments aimed to fill this void by providing empirical evidence in a real-world setting, emphasizing the need for robust, region-specific solutions that integrate ML and IoT technologies. The study intended to bridge the gap between theoretical insights and practical implementation, offering a user-friendly, real-time web application that caters to localized agricultural needs. This approach sought to address the deficiency in comprehensive, region-specific systems, particularly in Somalia's agricultural landscape, aiming to enhance agricultural decision-making and productivity in a targeted, adaptable manner.

3 Methodology

This research introduces a system designed to improve crop production in Somalia's Agricultural sector by utilizing ML and IoT technologies. The core idea is to offer recommendations for suitable crops to cultivate. The system is divided into two main components: a Hardware part and a Software part.

The hardware configuration comprises essential components for efficient crop monitoring. The Arduino Uno microcontroller is the control center at its core, orchestrating the entire system. Accompanying this is a suite of sensors: The Nitrogen, Phosphorus, and Potassium sensor, which gauges soil nutrient levels crucial for optimal plant growth; the pH sensor, assessing soil acidity and alkalinity to maintain the ideal pH range; and the Humidity and Temperature sensor, providing insights into atmospheric conditions that influence plant health and development; the rainfall Sensor

Fig. 1 Proposed system architecture

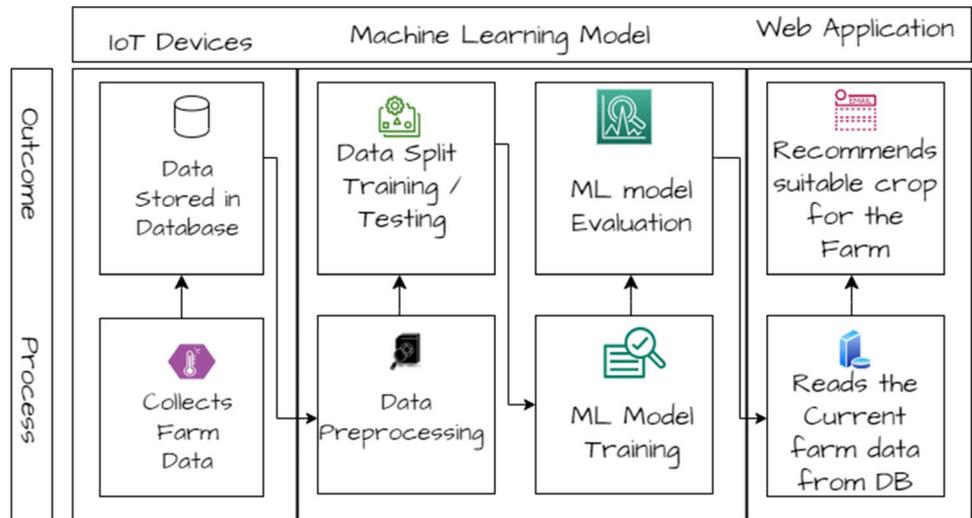


Table 1 Summary of dataset

No.	Crop type (labels)	Crop quantity
1	Banana	100
2	Coconut	100
3	Kidney beans	100
4	Maize	100
5	Mango	100
6	Lemon	100
7	Papaya	100
8	Rice	100
9	Watermelon	100
10	Pumpkin	100
Total		1000

measuring the amount of rain received by crop fields. This integrated setup is used to access real-time data on soil nutrient content, pH balance, humidity, temperature, and rainfall, enabling informed decisions for fertilization and cultivation practices. Ultimately, this holistic approach collects crop data to enhance yields, resource efficiency, and thriving crops.

The Software part is responsible for processing the data the hardware devices collect. This project component is a crucial aspect that encompasses the training and testing various ML algorithms including DT, Random Forests (RF), Support Vector Machines (SVM), and K-Nearest Neighbor (KNN). The primary objective is to identify, amongst these algorithms, the model with the highest accuracy rate, which will then be employed to provide recommendations about crop suitability. Moreover, the software segment encompasses creating a web application using the Django framework for the backend, accompanied by MySQL as the database management system. The application’s front end uses HTML, CSS, and JavaScript. This amalgamation of technologies ensures a seamless and user-friendly interface. The web application aims to display the farm’s fields and respective recommended crops dynamically in real-time. This functionality will empower users, such as farmers and agricultural stakeholders, to make informed decisions regarding crop selection based on the ML model’s predictions. The application’s interactivity and real-time updates enhance the precision and timeliness of decision-making in agricultural practices. Below Fig. 1 shows the proposed system architecture.

3.1 Dataset description

This dataset contains data from various farms in the Afgooye, Balcad, and Jowhar Districts of Somalia. The data was collected using IoT devices to gather essential parameters such as soil nitrogen, potassium, phosphorus, pH, air

temperature, humidity, and rainfall, influencing crop growth and yield. The data was collected based on Somalia's seasons from May 2022 until April 2023. The collected data and their respective labels are transmitted to a central database.

The dataset consists of ten crop labels that Somali farmers cultivate and their quantity summing the overall dataset, as Table 1 below summarizes.

The data in this dataset will be used to train ML models to recommend a crop and optimize crop production. The data will be used to develop decision support tools for farmers to help them make better-informed decisions about their crops.

3.2 Data preprocessing

The dataset originates from farms in the Afgooye, Balcad, and Jowhar Districts of Somalia, collected using IoT devices to capture crucial parameters of crop cultivation. With a commitment to data integrity, efforts were made to guarantee the completeness of information stored in the database, minimizing the presence of missing values. The dataset's structure was inspected for proper organization, and any missing values were effectively addressed through imputation or removal. Duplicate data was meticulously handled to mitigate redundancy and potential biases. The overarching goal was to ensure data consistency by rectifying any inconsistencies arising from duplicates or missing values, ultimately priming the dataset for training a reliable ML model.

3.3 Machine learning models

After data preprocessing, we split the data into a 75:25 ratio for training and testing models. Three ML Models were employed (RF, DT, and KNN) to Select the best model for the crop recommendation System. We discussed the algorithms below.

RF is a potent and widely used algorithm for classification tasks. As an ensemble learning method, it harnesses multiple DT to achieve accurate and robust classifications. By training DT on different subsets of data, a process known as "bagging," RF diminishes overfitting and enhances generalization to new data. This technique is further enhanced by introducing feature randomness during training, where subsets of features are considered for each tree split, bolstering the model's resilience. For classification, RF combines individual tree predictions through majority voting, resulting in a final predicted class label. This approach mitigates noise and outlier effects. Another advantage is its Out-of-Bag (OOB) evaluation, utilizing instances not part of specific trees' training sets to estimate accuracy. The algorithm's strengths encompass achieving high accuracy, minimizing overfitting, revealing feature importance insights, accommodating various data types, and addressing imbalanced data. Its versatility and effectiveness make it applicable in finance, healthcare, marketing, and more [27].

A DT Classifier is a ML algorithm for classification and regression tasks. It constructs a tree-like structure where each internal node represents a feature, branches depict decision rules and leaf nodes correspond to class labels (for classification) or predicted values (for regression). The algorithm iteratively selects the most significant feature to split the data based on criteria like Gini impurity or Information Gain. This splitting process continues recursively, forming branches and nodes until predefined stopping conditions are met, such as reaching a maximum depth or achieving pure subsets. DT are easy to interpret and can handle various data types. However, they can overfit, especially with complex trees or noisy data [28].

The KNN algorithm is a widely-used supervised ML method for classification and regression tasks. This technique relies on the concept that data points sharing similar features tend to exhibit similar outcomes. In practice, KNN predicts the label or value for a new instance based on the majority class or average value of its KNN within the training dataset. This involves calculating distances between the new instance and all instances in the training set, such as Euclidean distance, then selecting the k closest neighbors to make predictions. While KNN is straightforward to comprehend and implement and suits moderate-sized datasets, it can be computationally intensive during prediction, particularly for larger datasets. The parameter k's choice significantly impacts its performance, requiring careful consideration. Moreover, the algorithm's effectiveness is influenced by the choice of distance metric and data distribution [29].

The dataset comprises measured parameters from IoT devices on Somali farms, encompassing soil attributes (nitrogen, potassium, phosphorus, pH) and environmental factors (temperature, humidity, rainfall), with ten distinct crops cultivated. The ML models aim to establish connections between these input variables and the crop type, predicting the

most suitable crop based on these parameters. They analyze how variations in soil attributes and environmental factors correspond to successful crop cultivation, seeking patterns or combinations of these variables that optimize growth and yield for specific crops. Ultimately, these models aim to recommend the most appropriate crop based on the measured parameters, facilitating informed decision-making for farmers in Somalia's agricultural landscape.

3.4 Deployment on the web

ML algorithms in the system's web architecture entail a network of interconnected components. The hardware arrangement consists of sensors coupled to an Arduino Uno microcontroller. The sensors collect up-to-the-minute information on soil nutrients, pH levels, humidity, temperature, and rainfall. The data is sent to the backend for processing.

Multiple sequential steps are executed within the data processing backend. The backend server receives and processes the data collected by the Arduino Uno microcontroller. This processing includes data cleansing, preprocessing, and feature extraction. Afterward, different ML techniques like DT, RF, SVM, and KNN are trained using historical and real-time data to forecast the appropriateness of crops. The trained models are subsequently assessed according to their accuracy rates to identify the model with the highest performance.

The user interface for farmers and agricultural stakeholders uses HTML, CSS, and JavaScript in the front-end development. The backend, constructed with the Django framework, manages user requests, establishes communication with the database, and combines with the machine learning model for providing crop suggestions. MySQL database management system stores sensor data, model parameters, and user information. Real-time updates are essential to the web application, providing up-to-date information on farm fields and suggesting crops based on the most recent sensor data and machine learning forecasts.

After training and evaluating a machine learning model, it is deployed in the web application's backend. Upon a user's request for crop suggestions, the web application activates the installed machine learning model to produce predictions using the current environmental conditions acquired from the hardware sensors. The anticipated agricultural suggestions are exhibited to the user via the online interface.

4 Results and discussion

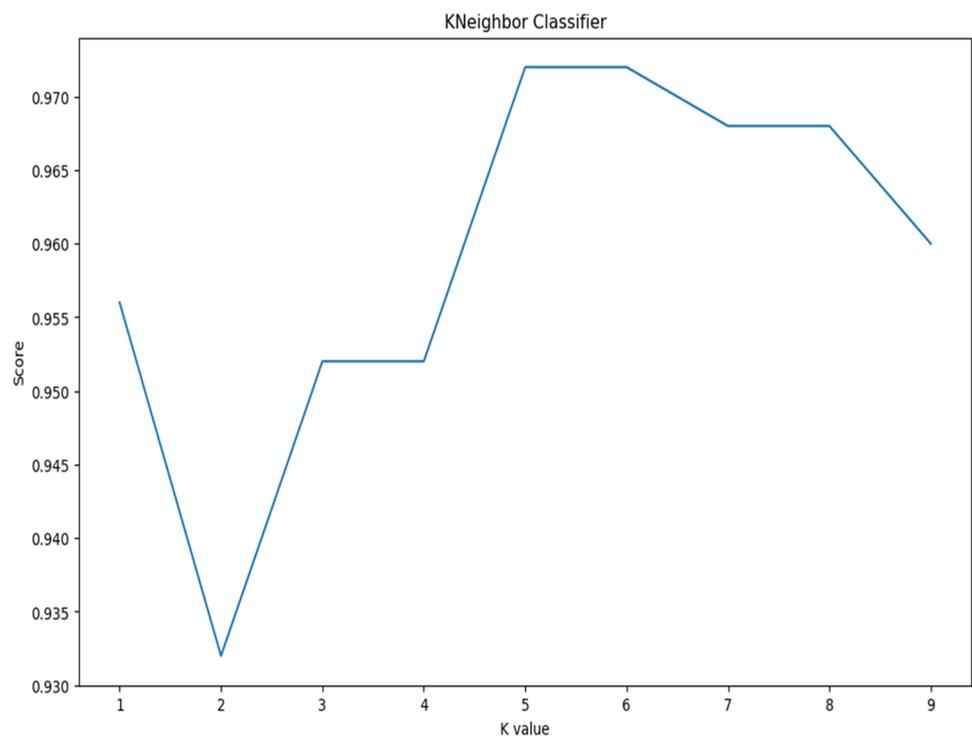
The study aimed to enhance crop production in Somalia's agricultural sector by leveraging ML and IoT technologies. In this paper, we present the evaluation of a Crop Recommendation System by utilizing three prominent ML algorithms: DT, KNN, and RF. This investigation aims to compare and contrast the performance of these algorithms, considering a range of vital metrics such as accuracy, precision, recall, and F1-score. Then, develop a web application platform that the farmers support in decision making on which crop to cultivate in which field of the farm. The ultimate goal is to ascertain the algorithm that offers the most effective and reliable crop recommendations, aiding Somalia's agricultural decision-making processes to enhance Somalia's agricultural production and tackle food insecurity. Following data preprocessing and allocating 75% of the data for training and 25% for testing the various ML models, we evaluated the three models, RF, DT, and KNN, to identify the best-performing model for the crop recommendation system.

The DT algorithm, renowned for its simplicity and interpretability, remarkably outperformed the other models in this study. Achieving an extraordinary accuracy rate of 99.2% showcased unparalleled precision in predicting the most suitable crops for specific conditions. The precision, recall, and F1-score, each at 99.0%, underscore the algorithm's ability to strike a harmonious balance between minimizing false positives and false negatives. These results show the DT's capability to dissect complex decision boundaries within the dataset and provide valuable recommendations for crop selection, thus reaffirming its efficacy as an agricultural decision support tool. Table 2. Below are the precision, recall, f1-score, and support of each crop trained with the DT model.

The KNN algorithm, a classic choice for pattern recognition, exhibited commendable performance when configured with $k=5$. With an accuracy of 97.2%, it offered slightly less accurate predictions than the DT. The precision, recall, and F1-score metrics, all recording values of 96.0%, indicated consistent performance in trade-offs between accuracy and misclassification. While KNN's results are marginally lower than the DT, its effectiveness in capturing localized patterns in the data showcases its potential for accommodating specific regional variations in crop suitability. Below is Fig. 2. It shows different K values and their accuracy, and when $K=5$ and 6 produces the highest accuracy of the KNN model.

Table 2 DT performance evaluation

Crop type	Precision	Recall	F1-score	Support
Lemon	1.00	1.00	1.00	27
Pumpkin	1.00	0.97	0.98	30
Banana	1.00	1.00	1.00	26
Coconut	1.00	1.00	1.00	20
Kidney beans	1.00	1.00	1.00	30
Maize	1.00	1.00	1.00	16
Mango	1.00	1.00	1.00	28
Papaya	0.96	1.00	0.98	26
Rice	1.00	0.95	0.98	21
Watermelon	0.96	1.00	0.98	26
Macro Avg	0.99	0.99	0.99	250
Weighted Avg	0.99	0.99	0.99	250

Fig. 2 K values and their accuracy

The Table 3. Below are the precision, recall, f1-score, and support of each crop trained with the KNN model.

The RF algorithm, a robust ensemble method, displayed an impressive accuracy level of 99.0%. The ensemble of DT working together contributed to its success in delivering accurate crop recommendations. Similarly, the precision, recall, and F1-score values of 99.0% emphasized the algorithm's balanced and dependable performance across multiple evaluation metrics. RF's unique ability to alleviate overfitting and encapsulate intricate relationships within the dataset is central to its capacity to generate reliable predictions for crop recommendations. Table 4. Below are the precision, recall, f1-score, and support of each crop trained with the RF model.

In the comparative analysis of the three algorithms, the DT emerges as the frontrunner, exhibiting the highest accuracy, precision, recall, and f1-score. These results collectively indicate its suitability for accurate and reliable crop recommendations. The KNN and RF algorithms also demonstrate strong performance, albeit slightly trailing behind the DT. While the other algorithms have merits, the DT's consistently superior performance suggests its potential as a primary choice for crop recommendation systems, as Table 5 shows.

Table 3 KNN performance evaluation

Crop type	Precision	Recall	F1-score	Support
Lemon	1.00	1.00	1.00	27
Pumpkin	0.86	0.83	0.85	30
Banana	1.00	1.00	1.00	26
Coconut	1.00	1.00	1.00	20
Kidney beans	1.00	1.00	1.00	30
Maize	1.00	1.00	1.00	16
Mango	1.00	1.00	1.00	28
Papaya	1.00	0.96	0.98	26
Rice	0.95	1.00	0.98	21
Watermelon	0.81	0.85	0.83	26
Macro Avg	0.96	0.96	0.96	250
Weighted Avg	0.96	0.96	0.96	250

Table 4 RF performance evaluation

Crop type	Precision	Recall	F1-score	Support
Lemon	1.00	1.00	1.00	25
Pumpkin	0.95	0.95	0.95	21
Banana	1.00	1.00	1.00	21
Coconut	1.00	1.00	1.00	23
Kidney beans	1.00	1.00	1.00	27
Maize	1.00	1.00	1.00	13
Mango	1.00	1.00	1.00	15
Papaya	1.00	1.00	1.00	13
Rice	1.00	1.00	1.00	20
Watermelon	0.95	0.95	0.95	22
Macro Avg	0.99	0.99	0.99	200
Weighted Avg	0.99	0.99	0.99	200

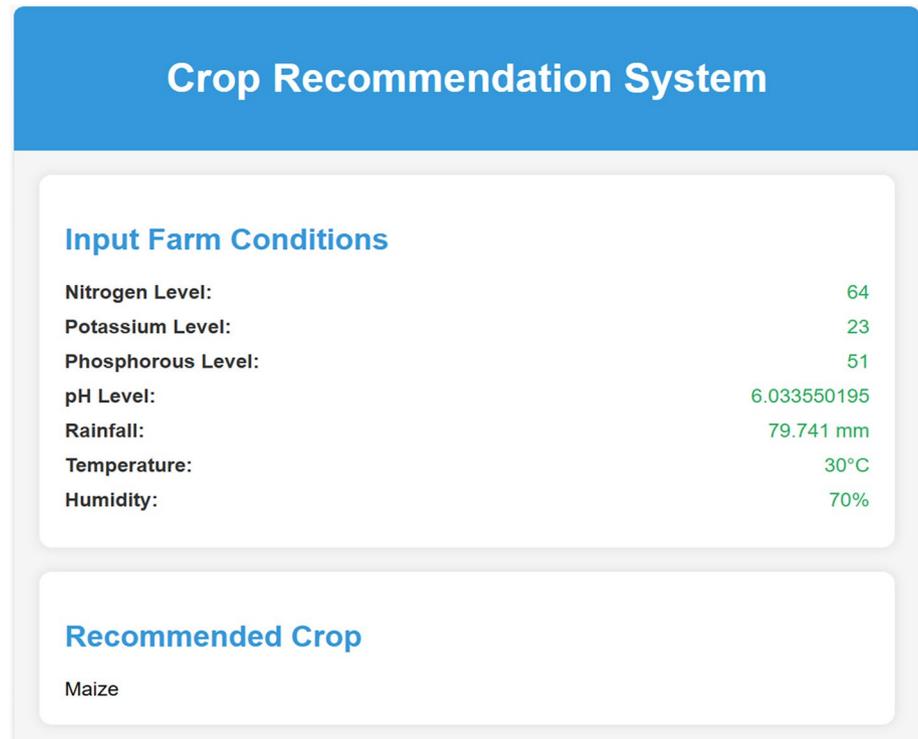
Table 5 Performance matrix of the models

Model	Accuracy	Precision	Recall	F1-score
DT	0.992	0.99	0.99	0.99
KNN	0.972	0.96	0.96	0.96
RF	0.99	0.99	0.99	0.99

We have successfully constructed a real-time and cost-effective web application tailored for a crop recommendation system by utilizing the DT model for crop recommendation in conjunction with IoT devices. This innovative system can potentially revolutionize farming practices by providing insightful recommendations to farmers. To manifest the practicality of our approach, we executed the developed system in Balcad District, situated in the Middle Shabelle Region of Hirshabelle State in Somalia. This specific locale was chosen due to its representative nature of regional agricultural practices. By applying it in this context, we aimed to demonstrate its efficacy in a real-world setting. Below Fig. 3. Shows the web interface of the proposed system.

The system effectively analyzed the collected data from IoT devices, processed it through the DT model, and generated accurate crop recommendations. This can substantially improve the yield and sustainability of agricultural efforts in the region. Furthermore, our study opens avenues for enhancing traditional farming methodologies through integrating modern technology. By offering real-time insights into optimal crop choices based on environmental and soil conditions, our system empowers farmers to make informed decisions and optimize their productivity.

Fig. 3 Web interface



The implications of this study are substantial for the agricultural domain. The DT algorithm's combination of high accuracy and interpretability positions it as an invaluable tool for guiding farmers and agricultural experts of Somalia in making informed decisions regarding crop selection. The algorithm's transparent decision-making process can provide insights into why specific recommendations are made, fostering greater trust and adoption by end-users. The amalgamation of IoT and ML, as demonstrated through our DT based crop recommendation system, showcases a significant leap toward data-driven agriculture. Our system's successful implementation and functionality in the agricultural setting of Balcad District underscore its potential to revolutionize farming practices and contribute to food security and economic growth in Somalia and similar regions worldwide.

Furthermore, the Crop Recommendation System evaluated in this research, featuring the DT, KNN, and RF algorithms, underscores the potential of ML in augmenting agricultural decision-making. The DT's outstanding accuracy and balanced performance position it as a prime contender for generating accurate and comprehensible crop recommendations. However, ongoing exploration and refinement of crop recommendation systems are encouraged, aiming to advance the agricultural industry through innovative data-driven solutions.

5 Conclusion

In this study, we have explored the integration of ML and IoT technologies to develop a Crop Recommendation System tailored for Somali agriculture. We rigorously evaluated three ML algorithms, DT, KNN, and RF, to identify the most effective approach for guiding agricultural decisions in Somalia. The DT algorithm emerged as the standout performer, with an impressive accuracy rate of 99.2% and well-balanced precision, recall, and F1-score metrics. Its transparency and interpretability make it an optimal choice for guiding agricultural choices. KNN and RF Show Promise, although slightly trailing in performance, both KNN and RF algorithms achieved accuracies of 97.2% and 99.0%, respectively, presenting valuable alternatives for various contexts. The successful implementation of the Crop Recommendation System, particularly in Somalia's Balcad District, underscored the tangible advantages of real-time IoT data and the DT model. This system enables farmers to optimize crop selection, enhancing sustainability and yield potential. The implications of our findings are substantial for Somalia's agricultural sector, showcasing the potential of data-driven agriculture in addressing food security and economic growth. The DT's accuracy and interpretability position it as a valuable tool for guiding

farmers and agricultural experts in making informed decisions regarding crop selection. It also modernizes traditional Farming by offering real-time insights into optimal crop choices based on environmental and soil conditions; our system empowers farmers to modernize traditional farming practices, improving productivity and sustainability.

While our research demonstrates promising results, it is not without limitations as we used a dataset primarily sourced from a specific region, specifically Balcad. To enhance the robustness of the system, there is a need to expand the dataset to encompass a broader geographical area, considering the diverse agricultural settings found in various regions of Somalia. Also, we evaluated three ML algorithms, but there is room for further exploring additional algorithms and ensemble techniques to improve recommendation accuracy.

5.1 Future research directions

Future research directions encompass various critical aspects. First, there is a need to enhance data collection efforts, focusing on gathering comprehensive and real-time IoT data. This should include incorporating additional parameters pertinent to crop health and soil conditions, enriching the dataset. Second, algorithmic optimization remains a crucial avenue, with the potential exploration of advanced optimization techniques and ensemble methods to enhance crop recommendations' accuracy further. Third, user interface enhancements are essential for facilitating farmer accessibility and usability, making the Crop Recommendation System more user-friendly. Finally, scaling the system for regional impact is imperative, involving expanding the system's implementation to encompass a more comprehensive array of regions within Somalia and potentially extending its application to other countries facing similar agricultural challenges. These research directions collectively contribute to the continued improvement of data-driven agriculture and its potential for transformative impact.

Author contributions M.O. and A.D. wrote the main manuscript text, Analysis, and interpretations. Y. A. and A. Y. N. Algorithms comparison. All authors reviewed the manuscript.

Funding The authors have not disclosed any funding.

Data availability This study leveraged data collected from IoT sensors. The data used in this research is currently in the author's possession and has not been deposited into an online repository due to confidentiality and privacy concerns. However, we are committed to ensuring data accessibility for the purpose of scientific rigor and reproducibility. Therefore, interested parties may request access to the dataset by contacting the corresponding author directly. We will make every effort to provide reasonable access to the data while safeguarding sensitive information as required. We believe that this approach strikes a balance between preserving data integrity and promoting research transparency.

Declarations

Competing interests The authors have not disclosed any competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Sheikdon AA. An analysis of the contribution of agriculture sector to the Somali economy. *Int J Econ Manag Sci.* 2020;9(1):1–3.
2. Osundina KC. Government expenditure on agriculture and economic growth in Nigeria. 2014. www.ijsr.net.
3. Warsame AA, Sheik-Ali IA, Ali AO, Sarkodie SA. Climate change and crop production nexus in Somalia: an empirical evidence from ARDL technique. *Environ Sci Pollut Res.* 2021;28(16):19838–50. <https://doi.org/10.1007/s11356-020-11739-3>.
4. Bouët et al. Africa Agriculture Trade Monitor. AKADEMIYA2063, Kigali and Washington, DC, 2021. <https://doi.org/10.54067/9781737916406>.
5. Mathenge M, Sonneveld BGJS, Broerse JEW. Application of GIS in agriculture in promoting evidence-informed decision making for improving agriculture sustainability: a systematic review. *Sustainability (Switzerland).* 2022. <https://doi.org/10.3390/su14169974>.

6. Ajao AO, Ogunniyi LT, Oyedele GA. agricultural productivity growth and incidence of poverty: an experience from Africa. 2013. www.iiste.org.
7. Zanini G et al. Rebuilding Resilient and Sustainable Agriculture in Somalia Volume I SOMALIA. 2018. <https://documents.worldbank.org/en/publication/documents-reports/documentdetail/781281522164647812/volume-1-main-report>. Accessed 10 Jul 2023.
8. Warsame AA, Mohamed J, Mohamed AA. The relationship between environmental degradation, agricultural crops, and livestock production in Somalia. *Environ Sci Pollut Res*. 2023;30(3):7825–35. <https://doi.org/10.1007/s11356-022-22595-8>.
9. Ali Warsame A, Abdulkadir Sheik-Ali I, Mohamed Barre G, Ahmed A. Examining the effects of climate change and political instability on maize production in Somalia examining the effects of climate change and political instability on maize production in Somalia 2. 2021. <https://doi.org/10.21203/rs.3.rs-709389/v1>.
10. Michael YG. Vulnerability and local innovation in adaptation to climate change among the pastoralists: Harshin District, Somali Region, Ethiopia. *Environ Manage Sustain Dev*. 2017;6(2):65. <https://doi.org/10.5296/emsd.v6i2.11211>.
11. Ryan G, Hussein H, Nicolas J, Allison H, Paul P. On-farm irrigated maize production in the Somali Guu season. *Afr J Agric Res*. 2018;13(19):969–77. <https://doi.org/10.5897/ajar2018.13111>.
12. Dube DK, Reddy RU, Girmay A. Tribes perception for agriculture and rural house hold food in-security in Kelafo Woreda, Gode Zone in Somali Region. 2017. www.iiste.org.
13. Boitt MK, Langat FC, Kapoi JK. Geospatial agro-climatic characterization for assessment of potential agricultural areas in Somalia, Africa. *J Agric Inform*. 2018. <https://doi.org/10.17700/jai.2018.9.3.479>.
14. The World Bank and the Food and Agriculture Organization of the United Nations. Rebuilding Resilient and Sustainable Agriculture in Somalia Volume I. 2018.
15. Abdi-Soojeede MI. Crop production challenges faced by farmers in Somalia: a case study of Afgoye district farmers. *Agric Sci*. 2018;09(08):1032–46. <https://doi.org/10.4236/as.2018.98071>.
16. Ryan G, Hussein H, Nicolas J, Paul P. An on-farm comparison of the agronomics and economics of irrigated maize production systems in the Somali Deyr season. *Afr J Agric Res*. 2018;13(23):1177–85. <https://doi.org/10.5897/ajar2018.13177>.
17. Dewangan AK. Application of IoT and machine learning in agriculture. *Int J Eng Res Technol*. 2020;9(7):110–4.
18. Akhter R, Sofi SA. Precision agriculture using IoT data analytics and machine learning. *J King Saud Univ Comput Inf Sci*. 2022;34(8):5602–18. <https://doi.org/10.1016/j.jksuci.2021.05.013>.
19. Mekonnen Y, Namuduri S, Burton L, Sarwat A, Bhansali S. Review—machine learning techniques in wireless sensor network based precision agriculture. *J Electrochem Soc*. 2020;167(3): 037522. <https://doi.org/10.1149/2.0222003jes>.
20. Araby AA et al. Smart IoT Monitoring System for Agriculture with Predictive Analysis. In 2019 8th International Conference on Modern Circuits and Systems Technologies (MOCAST), IEEE, May 2019, pp. 1–4.
21. Reddy KSP, Roopa YM, Rajeev K, LN, Nandan NS. IoT based smart agriculture using machine learning. In Proceedings of the Second International Conference on Inventive Research in Computing Applications (ICIRCA-2020), IEEE, Jul. 2020, pp. 130–134.
22. Bodake K, Ghate R, Doshi H, Jadhav P, Tarle B. Soil based fertilizer recommendation system using Internet of Things. *MVP J Eng Sci*. 2018. <https://doi.org/10.18311/mvpjes/2018/v1i1/18273>.
23. Rama Krishna S, Amrutha K. A novel ensembling technique for crop recommendation system using machine learning. *Int J Adv Sci Technol*. 2020;29(5):874–82.
24. Dhabarde S, Gupta A, Pote D, Yadav A, Bisane S. Agricultural Crop Recommendation System Using IoT and ML. *Int J Adv Res Sci Commun Technol*. 2022. <https://doi.org/10.48175/ijarsct-3376>.
25. Madhuri J, Indiramma M. Artificial neural networks based integrated crop recommendation system using soil and climatic parameters. In 5th International Conference on Computing Methodologies and Communication (ICCMC), 2021, pp. 843–848. doi: <https://doi.org/10.17485/IJST/v14i19.64>.
26. Ali SM, Das B, Kumar D. Machine learning based crop recommendation system for local farmers of Pakistan. *Rev Gestão Inovação Tecnol*. 2021. <https://doi.org/10.47059/revistageintec.v11i4.2613>.
27. Svetnik V, Liaw A, Tong C, Christopher Culbertson J, Sheridan RP, Feuston BP. Random forest: a classification and regression tool for compound classification and QSAR modeling. *J Chem Inf Comput Sci*. 2003;43(6):1947–58. <https://doi.org/10.1021/ci034160g>.
28. Safavian SR, Landgrebe D. A survey of decision tree classifier methodology. *IEEE Trans Syst Man Cybern*. 1991;21(3):660–74.
29. Liu ZG, Pan Q, Dezert J. A new belief-based K-nearest neighbor classification method. *Pattern Recognit*. 2013;46(3):834–44. <https://doi.org/10.1016/j.patcog.2012.10.001>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.