

Intelligent Soil Classification for Precision Agriculture in Algeria:

Integrating Machine Learning and Image Processing for Optimized Crop Yield and Recommendation

Souheila BOUDOUDA⁽¹⁾, Mountaha BOUKALLEL⁽²⁾

(1) (2) LIRE laboratory, Department of Software Technologies and Information Systems, University of Constantine 2-Abdelhamid Mehri, Algeria

Abstract: Agriculture is a vital pillar of Algeria's economic development and a fundamental contributor to the well-being of its population. The productivity of this sector largely depends on soil characteristics, which play a crucial role in determining suitable crop varieties. This paper presents a methodology for soil classification that considers both nutrient composition and physical attributes such as color and texture. By integrating data mining techniques with image classification algorithms, our approach aims to enhance the accuracy of soil type categorization. Image classification, in particular, enables a detailed analysis of soil texture and structure, improving classification precision. Various machine learning algorithms, including K-Nearest Neighbor (KNN) and data mining methods, are employed to construct the proposed classification framework. By leveraging these technologies, our methodology provides farmers with actionable insights into soil properties, enabling them to make informed crop selection decisions and optimize agricultural productivity.

Keywords: Agriculture, Soil Classification, Crop Recommendation, Machine Learning

1. INTRODUCTION

The Agriculture is crucial for the sustainability of any economy and plays a significant role in long-term economic growth. It has evolved from a primary source of livelihood to a vital source of national trade, as it's considered as a primary source of income and very helpful for the development of the economic condition of the country [1].

Technological advancements emerge as a catalyst for agricultural development, ushering in an era of innovation and transformation within the sector [2], [3]. Concepts such as precision agriculture and smart farming, coupled with cutting-edge technologies like image processing, Geographic Information Systems (GIS), and Unmanned Aerial Vehicles (UAVs), hold promise for revolutionizing agricultural production. Digital agriculture, facilitated by information and communication technologies, furnishes farmers with vital crop and market information, thereby enhancing decision-making processes [4].

The rapid development of technologies for collecting, storing, and processing Big Data has prompted a transition from a classical approach to a more measured and intelligent one. Despite this advancement, in fields such as Artificial Intelligent (AI), robotics, and Internet Of Things sensors (IoT), which are being utilized to enhance the accuracy of

agri-business and enable farmers to monitor crop fertility, soil conditions, temperature, and durability. Algeria has yet to fully leverage these technologies to drive growth in the agricultural sector, as the country struggles with lack of market awareness, basic farmer to farmer communication and traditional ways to crop selection. This created a need for farmers and individuals to innovate new ways to manage farms to achieve food security.

Traditional farming methods often rely on experience and intuition, but these practices may only sometimes guarantee optimal results when it comes to soil management due to this, countries are producing yields much below their potential [4]. Therefore, to address this knowledge gap, numerous researchers (like the ones will be mentioned in the next sections), have explored the potential of technology to provide accessible tools for soil classification and analysis. These tools have the potential to empower farmers in Algeria and worldwide, enabling informed decision-making about crop selection, soil health, and sustainable farming practices.

Soil is the most important part of the agricultural process as it contains the crop. Farmers in Algeria and rural areas lack the information or knowledge about the texture, or type of the soil. They are currently farming without proper testing and are oblivious to

choosing the proper crop for their land type, which directly affects the yield and crop growth. Thus, farmers need to be equipped with solutions that provide soil classification.

This paper proposes a technical approach to address the challenges within the farming community in Algeria, this includes giving farmers the ability to connect, and better exploit their land for maximum yielding by establishing a farmer-centric approach by engaging farmers with others near them.

Transforming the Agri-sector by building solutions based on the available technologies such as AI, GIS, Big Data and various practices. Minimize food waste by providing more organized market practices. Upgrade planting techniques by recommending the most suitable crop for each soil type. Additionally, the study explores computer-based soil classification techniques. To classify soil, there are 3 main factors to discuss, fertility, type and elevation. Each one of these plays a huge role in crop recommendation and needs to be studied to find the effect each one poses on soil productivity and crop yielding.

This paper presents a comprehensive approach to soil classification utilizing machine learning and image analysis techniques. By harnessing the power of data-driven methodologies, we aim to contribute to the advancement of agricultural practices in Algeria, facilitating sustainable crop production and bolstering the nation's agricultural sector.

The rest of this paper is structured as follows, Section 2 presents a brief overview of the initiatives taken by other researchers and countries. Section 3 introduces the main concepts and rationale based on which the system's architecture has been designed, while Section 4 represent the methodology and the architecture of our system. Finally the paper's conclusions are presented in section 5.

2. RELATED WORK

The literature of this research is going to look at the applications of machine learning in precision agriculture. Firstly, it is necessary to fully understand machine learning and predictive agriculture. With the blooming of Artificial Neural Networks (ANN), which are biologically inspired computational models that can exceed the work of previous types of AI, one of the most impressive forms of ANN is CNNs. The main objective of CNN is to solve difficult image-driven recognition tasks.

In the past few years, many computer-based methods and techniques have been developed, such as Support Vector Machines (SVM), ANN, and the K-Nearest Neighbors' Algorithm (k-NN). It has been observed that SVM is the most developed classifier for the soil in which it can work efficiently with a high level of accuracy. Therefore, Precision Agriculture includes all those agricultural production practices that use information technology either to tailor input use in order to achieve desired outcomes, such as Variable Rate Application (VRA), yield monitoring and remote sensing [5].

Berman and Dev's paper [6] shows that the Support Vector Machine classifier can be used to classify soil images using kernels which returned an average accuracy of 91.37% and a result that nearly matches the United States Department of Agriculture soil classification.

Srivastava et al. [7] presented databases created under different environmental conditions for soil image processing and aimed to develop a low-cost digital soil classification system for rural farmers, focusing on soil texture important in crop selection and water transmission.

In a research carried out by Saranya [8], it was proven that the use of data mining techniques to analyze soil's chemical attributes to propose a suitable crop, and the accuracy of this model is 96% which is higher than other methods proposed by other researchers.

In a paper presented by Raju et al. [9], authors evaluated HybridTransferNet on a soil classification dataset and reported multiple performance metrics, including accuracy, precision and recall. Their experimental results showed that HybridTransferNet outperforms traditional transfer learning approaches and achieves state-of-the-art performance on the soil classification task.

Chiranjeevi [10] carried out research for classifying soil types so farmers can analyze it and predict the best crop to grow based on their soil. They have considered the data mining algorithms for classifying the soil. They used algorithms such as J48 decision tree classifier and Naïve Bayes classifier among these two algorithms Naïve Bayes has obtained the maximum accuracy of 98% [11].

In a research carried out by Armstrong [12], a comparative study of data mining algorithms was presented. They have used a large dataset extracted from the Australian Department of Agriculture and Food (AGRIC) to conduct the research. The study evaluated

five machine learning models using Kaggle dataset to recommend crops and determine nutrient needs based on environmental factors. Separate evaluation for agricultural and horticultural crops improved prediction accuracy. XGBoost model outperformed others, achieving high precision rates for crop recommendations. The findings suggest potential for developing a user-friendly AI interface aiding crop selection and fertilizer application decisions in diverse environments [13].

In a research carried by Aboubakar et al. [14], an IoT based smart agriculture monitoring system was presented. It aimed at increasing agricultural productivity and efficiency by automating and optimizing crop management.

Barkataki et al. [15] proposed an IoT system comprising various sensors for measuring soil characteristics, plus utilized SVM and Decision Tree algorithms for recommending suitable crops based on soil data.

We can note that Remote Sensing is one of the important data sources used in GIS for accessing data acquired through satellites. Remote Sensing relies heavily on the reflection of visible light energy from an external source. The sun serves as the external source of energy for passive systems. Image sensors now enhance satellite data collection. [16] Satellite photos can be studied digitally, thanks to advancements in image processing techniques such as enhancement, restoration and analysis. These advancements have enabled remote sensing to progress independently of GIS. Remote sensing monitors the Earth's surface, measuring geographical, biological, and physical factors to detect and analyze land cover materials [17]. One study presented by Pandey [18] was intended to apply wireless sensors and IoT in the agriculture domain. The development of the ventilate, wireless sensor, VAVS, cloud computer and a new method for improving crop yield and handling were also discussed. IoT-based sensors and communication are mandatory in agriculture to maximize crop production .

Our approach emphasizes finding the right dataset with the right attributes is vital for crop prediction/recommendation, thus due to the lack of data about Algerian soil types and chemical attribute, an IoT system needs to be build that allows farmers to analyze their soil's chemical attributes, that's also connected to a database that allows us to collect the necessary data to classify soil hens recommend crops. Classification is an

important part of this system as it allows us to determine soil type, and also used to predict the most suitable crop for that particular field. Classification is used to predict the class for each record in a dataset. This study uses supervised learning techniques for prediction that handle high-dimensional data, like the k-Nearest Neighbour (kNN) [19] , Naïve Bayes (NB) [20], SVM [21], and Random Forest (RF) [22]. The KNN makes no assumptions about the data, though data scaling is a must. The NB is scalable with large datasets but does not work well if the training data is not representative of the population. The SVM, which is the best algorithm for separable classes, needs a lot of processing time for large datasets. The RF handles huge volumes of data with high-dimensional attributes, though the attributes only work if they have predictive power. Each algorithm has it's own way for the prediction process, so it's crucial to test them out before deciding the best feature selection technique with the classifier for crop prediction process.

In the following sections, we will provide a detailed explanation of our proposed approach, outlining how it addresses key challenges in Algerian agriculture through machine learning, IoT-based data collection and predictive analytics.

3. PROPOSED APPROACH

A system's architecture forms the foundation for its design, functioning and performance. In the context of a smart farming system, this architecture plays an important role, as it allows for comprehensive data collecting and analysis. The Figure 1 explains the architecture of our proposed system, including an overview of the components used in each tier.

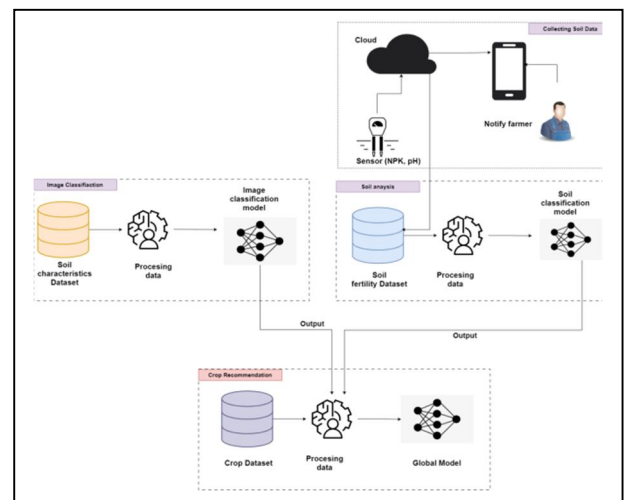


Fig. 1 Proposed architecture

The complete process is defined in the methodology section, where each aspect of the architecture is further explained.

4. METHODOLOGY

Soil classification and regional topography play a huge role in crop selection, as there are specific needs for each crop type, to anticipate crops more accurately, a real-time soil analysis system needs to be developed.

To create such a system, there are two factors to take into consideration, soil type, which can be determined using several machine learning techniques, for this research we will be using Convolutional Neural Networks (CNNs) and its chemical attributes.

K-nearest neighbour Classifier

According to Balducci et al. [23], the kNN is a non-parametric supervised learning technique that employs training sets to categorize data points into distinct groups. Based on a new instance, data is gathered from all educational cases and the correlations in fundamental classifications. The largest number of prior cases (neighbors) k is found in the training dataset, and new instances (x) are calculated by adding up the output attributes for the k cases.

Naive bayes Classifier

The Naive Bayes classifier [24] is one of a family of elegant probabilistic categorization algorithms in machine learning that is based on the Bayes' theorem's freedom properties. Every class label is projected based on the likelihood of a specific instance. Naive Bayes classifiers infer that, given a class attribute, the value of one quality is independent of the value of all other qualities.

Support Vector Machine classifier

The SVM divides data into decision surfaces, which are then divided into two hyper-plane groups [25]. Training points identify the vector that supports the hyper-plane. Presumably, due of larger margins, a hyper-plane with the largest distance. The nearest learning data point typically has greater margins, less mistakes, and excellent classifier generalization.

Random Forest classifier

The Random Forest classifier blends multiple decision trees to get the most accurate forecast [26]. The RF divides each node and

finds the most significant parameter before selecting the best from a subset of random attributes.

Soil classification

In this section, we will discuss the procedure of soil classification using CNN. The CNNs are a specialized class of neural networks designed to effectively process grid-like data, such as images, with their ability to automatically learn features from raw pixel data, it achieved a remarkable success in the market of object recognition.

CNNs consist of multiple layers, each with unique functions, the key components are input layer convolutional layers, pooling layers, activation functions, fully connected layers and finally an output layer. The proposed system inputs soil images from the user and states the type of soil as an output. SVM and CNN architectures are used for soil image classification and evaluated accuracy of each classifier.

Crop recommendation and classification

For this section, a machine learning-based approach for recommending suitable crops to farmers based on various environmental factors is presented. The basic crop recommendation is given below. Soil parameters are given as an input and the type is provided as an output.

- **Step 01:** Crop dataset is given as an input, and the set of data imported.
- **Step 02:** The attributes used in the set of data are transformed into a particular range, bringing the set of data into a consistent state, thus avoiding anomalies. Any missing values are removed and normalization used to standardize the data. Redundancy is minimized once, the dataset is structured and also it helps to make the efficient data for the prediction processing.
- **Step 03:** Splitting the dataset into training and testing sets
- **Step04:** The classification algorithm such as KNN or SVM is applied to the training sample.
- **Step 05:** The trained classifier is applied to the testing samples to predict the most suitable crop for cultivation in a particular piece of land.
- **Step 06:** The target label for new instances is found by the trained classifier so as to identify a suitable crop.
- **Step 07:** Finally, a suitable crop for cultivation is recommended by the results.

The pipeline involves data pre-processing, feature engineering, model training using classifiers such as K-Nearest Neighbors, Support Vector Machines, Decision Trees, and Random Forests and evaluation of model performance through various metrics including accuracy, confusion matrices and feature importance analysis [27].

Soil parameter collection

Algeria's vast agricultural sector holds immense potential for growth. However, a crucial obstacle exists (the lack of comprehensive soil data). This data is essential for developing a robust crop recommendation system based on soil fertility. To bridge this gap, we are turning towards the Internet of Things (IoT) to empower Algerian farmers with the tools they need for success. By providing a network of sensors that farmers can deploy across their fields, they can collect real-time data on crucial soil parameters that will be later used to recommend the best crop based on the percentage of each chemical factor such as:

- **Soil Moisture:** Gaining insights into water content allows for precise irrigation, preventing over watering and ensuring optimal nutrient uptake for plants.
- **Nutrient Levels:** Sensors can estimate the levels of Nitrogen (N), Phosphorus (P), and Potassium (K), essential for healthy plant growth.
- **Soil pH:** Knowing the acidity or alkalinity of the soil helps farmers understand nutrient availability for their crops.
- **Temperature:** Soil temperature plays a vital role, influencing the rate of nutrient breakdown and plant growth.

With a robust data foundation in place, the next step is harnessing the power of analytics. Using machine learning algorithms, now our Crop recommendation model can analyze the collected soil data alongside historical crop yields and regional climate information. This empowers the prediction of crops specifically tailored to Algerian agricultural conditions. This engine will suggest the most suitable crops for a particular field, considering soil fertility, soil type and optimal climatic conditions.

5. IMPLEMENTATION AND EXPERIMENTATION

The dataset utilized in this research was sourced from Kaggle repositories due to the lack of Algerian data, which was collected over a period of time by Jaya Prakashpondy and cleaned by Jhislaine Matchouathé. The

dataset is consistent of 1244 images belonging to 4 soil classes "Alluvial Soil", "Red Soil", "Clay Soil", "Black Soil" (See Figure 2).

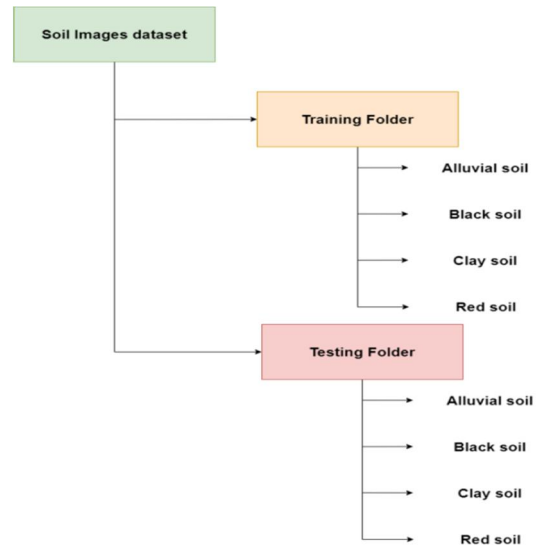


Fig. 2 Dataset organization

A sample of the soil images provided to the model are depicted in Figure 3 below.

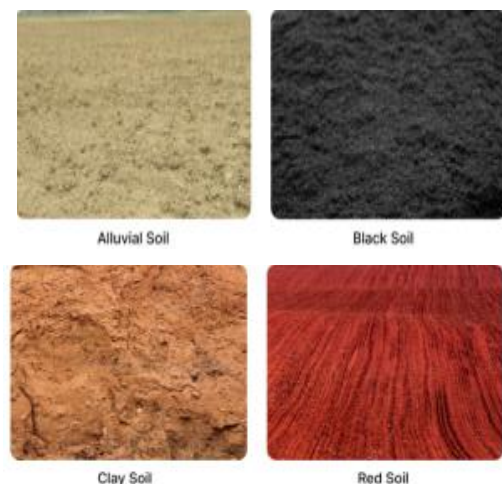


Fig. 3 Sample images from soil image classification dataset

Data pre-processing

The dataset included photos of various sizes and resolutions. To guarantee uniformity, all photographs were under 'jpeg', 'jpg', 'bmp', 'png' formats. Data augmentation techniques were used to resize the photos and make them uniform such as flipping, zooming, rotation, sharing, and shifting. Finally, the pre-processed data is loaded into Python memory using `flow_from_directory`, configuring parameters such as image dimensions, batch size, and label type. This comprehensive pre-processing pipeline ensures the dataset's

integrity and diversity, facilitating robust model training and evaluation for accurate image classification. (See table 1)

Table 1 Summary of Soil Image Classification dataset

Category	Training images	Testing images
Alluvial Soil	469	39
Black Soil	187	94
Clay Soil	179	56
Red Soil	158	62

Algorithm 1: Training and evaluation pipeline for soil classification

Algorithm 1 CNN Algorithm: Training and evaluation pipeline

Require: Dataset
 Ensure: Trained model
 Split the dataset into training and testing sets
 Define data transforms and pre-process datasets.
 Define the CNN model
 Compile the model using an optimizer (e.g., Adam)
 Define learning rate scheduler
 While Not converged do
 Train the model on the training set
 Calculate validation loss
 Update learning rate scheduler
 End while
 Evaluate the best model on the test set
 Calculate performance metric
 Plot training loss curve
 Save model
 Define prediction function

Mathematical model

The mathematical model for the Convolutional Neural Network (CNN) used in this paper involves a series of mathematical operations performed on the input data to learn and extract features, followed by classification using a softmax activation function. Here's an overview of the mathematical operations involved:

1) *Convolution Operation:*

The convolutional layers in the CNN apply convolution operations to the input images using learnable filters (kernels). This operation involves element-wise multiplication of the input image and the filter, followed by summation, to produce a feature map. Mathematically, the convolution operation can be represented as:

2) *Activation Function:*

After convolution, an activation function (typically ReLU) is applied element-wise to introduce non-linearity into the model.

Mathematically, ReLU activation can be represented as:

3) *Pooling Operation:*

Max-pooling layers down sample the feature maps by taking the maximum value within each window of the feature map. Mathematically, max-pooling can be represented as:

4) *Fully Connected Layers:*

The flattened features are passed through fully connected (dense) layers, which perform matrix multiplication with weight matrices and add bias terms. Mathematically, the fully connected layer can be represented as:

5) *Softmax Activation:*

The output layer typically uses the softmax activation function for multi-class classification tasks. It converts the raw scores (logits) into class probabilities. Mathematically, softmax can be represented as:

6) *Loss Function:*

Categorical cross-entropy loss is commonly used for multi-class classification tasks. It measures the dissimilarity between the predicted probability distribution and the true distribution of the labels.

These mathematical operations together form the backbone of the CNN model used for image classification.

6. RESULTS AND DISCUSSION

The results of the soil image classification experiments using the CNN architecture are presented below, showing the model's accuracy. Overall, the results exemplify the efficiency of the evaluated CNN architectures in soil classification, highlighting their potential for practical applications in agriculture, land management and environmental monitoring. Furthermore, the choice of optimizer plays a significant part in determining the effectiveness of the models, influencing their convergence speed and overall accuracy.

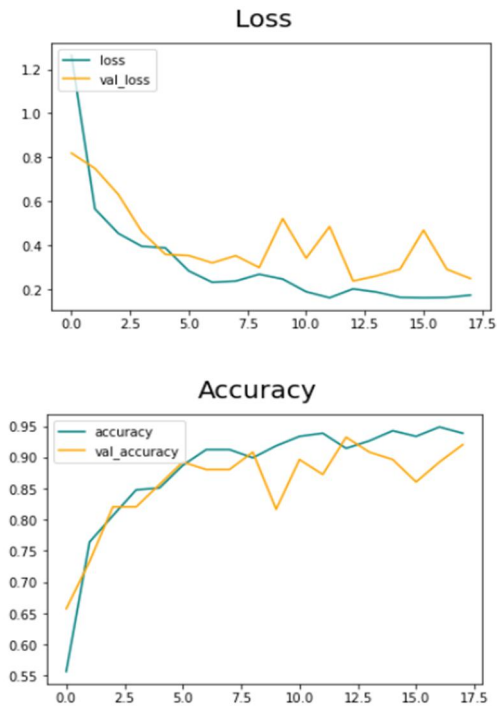


Fig. 4 Training Accuracy and loss of CNN Adam optimizer

As seen in the figure 4, the model's accuracy is around 95%, which is a high value for an early model, with a relatively small dataset.

Crop recommendation

The dataset was sourced from Kaggle repositories [28] due to the lack of Algerian data used consists of parameters such as temperature, humidity, pH level, and soil nutrients, which are crucial for crop growth. The goal is to develop predictive models capable of accurately recommending crops based on these environmental variables .

Table 2 Chemical attributes (Crop Recommendation Dataset, 2020)

Attribute	Details
Label	The name of the crops
N	Ratio of Nitrogen content in soil
P	Ratio of Phosphorous content in soil
K	Ratio of Potassium content in soil
Temperature	Temperature in degree Celsius
Humidity	Percentage
pH	pH value of the soil
Rainfall	millimetre
Soil type	A set of predefined soil types mentioned in the previous section(Soil Classification) "Alluvial Soil", "Red Soil", "Clay Soil", "Black Soil"

Table 3 presents a summary of the crop recommendation dataset.

Table 3 Summary of the crop recommendation dataset

Crop	Count
chickpea	200
lentil	200
Potato	109
Paddy	100
Tomato	100
pomegranate	100
watermelon	100
muskmelon	100
maize	100
kidneybeans	100
grapes	100
apple	100
orange	100
Eggplant	96
Onion	96

Data pre-processing

The initial stages involve data preparation as the dataset is collected from different Kaggle repositories using techniques to manage missing values and identifies potential inconsistencies.

- **Exploratory Data Analysis:** A comprehensive exploration of the data is conducted to understand the distribution of features and potential relationships between them. This includes visualizations like distribution plots, pair plots and joint plots.
- **Feature Engineering:** Feature engineering techniques are applied to potentially improve the model's performance. This may involve data filtering, transformation, or creating new features based on existing ones. In our case a categorical encoding is applied to the target variable, and features are normalized using MinMaxScaler to ensure uniformity in scale. The visualization of the dataset has returned the results in figure 5, it helps understating how different variables correlate with each other and how they contribute to distinguishing different categories or classes represented by 'CROP'.

	N	P	K	temperature	humidity	ph	rainfall	label
0	71	54	16	22.613600	63.690706	5.749914	87.759539	maize
1	61	44	17	26.100184	71.574769	6.931757	102.266244	maize
2	80	43	16	23.558821	71.593514	6.657965	66.719955	maize
3	73	58	21	19.972160	57.682729	6.596061	60.651715	maize
4	61	38	20	18.478913	62.695039	5.970458	65.438354	maize

Fig. 5 Data loading and exploration into Pandas

The experimental results indicate that the Random Forest classifier achieves the highest accuracy among the models evaluated, with an accuracy of 100% on the test set (See table 4). Feature importance analysis reveals that soil nutrients (Nitrogen, Phosphorus, Potassium) play a significant role in crop recommendation, followed by environmental factors like temperature and humidity.

Table 4 Result comparison

Method	Kernel	Accuracy
K-Nearest Neighbor		99.6%
Support Vector Machines	Linear	100%
	RBF	100%
	Poly	100%
Decision Trees		99.6%
Random Forests		100%

Figure 6 represents a study limited to certain types of crops. To obtain more data on other crops, a new study needs to be conducted using IoT. In this study, a sample field will be established and equipped with various monitoring and control systems to experiment with the aspects mentioned above, aiming to improve yield and data collection. Different sensors, including temperature, soil moisture, pH, and NPK attributes, will be used. The collected data will first be stored in the Arduino's memory and then transmitted to a database, which will later assist in determining the optimal conditions for growing specific plants.

Farmers are increasingly adopting monitoring and control systems to optimize key agricultural parameters such as temperature, humidity, soil moisture, and pH. Automating these parameters helps farmers enhance crop yields by ensuring optimal growing conditions. Additionally, a user-friendly interface will be integrated into the

communication platform, allowing users to remotely monitor their farms

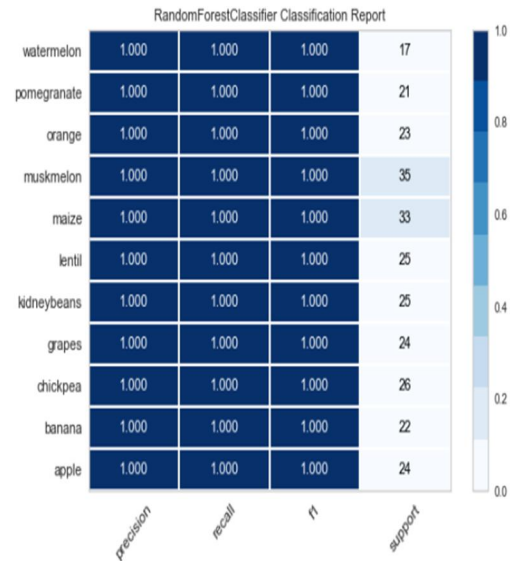


Fig. 6 Random Forest classifier report

The required components needed to build such system are:

- **Soil Moisture Sensor:** is used for measuring the moisture in soil and similar materials.
- **NPK Sensors:** These emerging sensors estimate levels of Nitrogen, Phosphorus, and Potassium, key nutrients for plants.
- **pH Sensor:** Measures acidity/alkalinity, affecting nutrient availability.
- **Temperature Sensor:** Soil temperature influences nutrient breakdown and plant growth.

7. CONCLUSION

This work introduces a data-driven, farmer-centric approach to addressing the key challenges faced by Algerian farmers, including crop selection, soil management, and limited access to agricultural information. By leveraging machine learning and an Internet of Things (IoT)-based data collection system, the proposed solution enables precise soil analysis and intelligent crop recommendations, empowering farmers to make informed decisions. This approach not only enhances agricultural productivity but also promotes sustainable farm management practices.

The integration of predictive analytics, combining real-time and historical data, further strengthens decision-making

processes, leading to increased yields, reduced environmental impact, and improved profitability. As the system evolves, continuous retraining of machine learning models can refine its accuracy and adaptability, ensuring more precise and context-aware crop recommendations.

Looking ahead, several avenues for future development can be explored. First, integrating remote sensing technologies and satellite imagery could enhance soil and crop monitoring capabilities, providing more comprehensive insights. Additionally, expanding the system to incorporate climate forecasting and pest detection could further support farmers in mitigating agricultural risks. Finally, fostering collaboration between researchers, policymakers, and agricultural stakeholders will be essential for scaling up this innovation and ensuring its widespread adoption.

Ultimately, the successful implementation of this system has the potential to revolutionize agricultural practices in Algeria, fostering a more resilient, efficient, and sustainable farming sector.

References

- [1] R. Kapur. Usage of technology in the agricultural sector. *Acta Scientific Agriculture*, pp. 78–84. 2018.
- [2] A. Rehman & I. Hussain. Modern agricultural technology adoption its importance, role and usage for the improvement of agriculture. *American-Eurasian Journal of Agricultural & Environmental Sciences*. 2016.
- [3] S. A. Kumar & P. Ilango. The impact of wireless sensor network in the field of precision agriculture: A review. *Wireless Personal Communications*, pp. 685–698. 2018.
- [4] C. Costa, F. Antonucci, F. Pallottino, J. Aguzzi, D. W. Sun & P. Menesatti. Shape analysis of agricultural products: A review of recent research advances and potential application to computer vision. *Food and Bioprocess Technology*, Vol. 4, pp. 673–692. 2011.
- [5] D. Jansirani, R. N. Karthick, R. J. Hariprasanth, S. Sweetin Preethi & R. S. A. Sorna Kumar. Synthesis of colloidal starched silver nanoparticles by sonochemical method and evaluation of its antibacterial activity. *Journal of Chemical and Pharmaceutical Sciences*, Vol. 9, pp. 177–179. 2016.
- [6] Barman & R. Dev. Soil texture classification using multi-class support vector machine. *Information Processing in Agriculture*, pp. 1–15. 2019.
- [7] P. Srivastava, A. Shukla & A. Bansal. A comprehensive review of soil classification using deep learning and computer vision techniques. *Multimed Tools Appl*. 14887–14914, 2021.
- [8] N. Saranya, & A. Mythili. Classification of Soil and Crop Suggestion using Machine Learning Techniques. *International Journal of Engineering Research and Technology*. 2020.
- [9] C. Raju, A. Ashoka & B. Prakash. Advancing Soil Image Classification through Comprehensive Evaluation for Crop Suggestion. [10.21203/rs.3.rs-3181331/v1](https://doi.org/10.21203/rs.3.rs-3181331/v1). 2023.
- [10] M. N. Chiranjeevi & B. Ranajana. Analysis of Soil Nutrients using Data Mining Techniques
- [11] R. S., Parmar, V., Mehra, & G. J. Kamani. Analyzing soil fertility using data mining techniques. *Gujarat Journal of Extension Education*. 2022.
- [12] L. Armstrong, D. Diepeveen & R. Maddern. The Application of Data Mining Techniques to categorize agricultural soil profiles. 2004.
- [13] U. Ahmed, J. C. Lin, G. Srivastava & Y. Djenouri. A nutrient recommendation system for soil fertilization based on evolutionary computation. *Computers and Electronics in Agriculture*. <https://doi.org/10.1016/j.compag.2021.106407>. 2021.
- [14] B. M. Aboubakar & A. Hassan & L. Rizwan & A. Muhammad. IOT Based Smart Agriculture Monitoring System. 2023.
- [15] N. Barkataki, S. Mazumdar, P. B. D. Singha, J. Kumari, B. Tiru and U. Sarma. Classification of soil types from GPR B Scans using deep learning techniques, 2021 International Conference on Recent Trends on Electronics, Information, Communication & Technology (RTEICT), 2021.
- [16] A. B. Muhammad. Systematic Review of the Literature on Various Soil Classification Methods, *International Journal on Recent Technologies in Mechanical and Electrical Engineering*, page 4, 2022.
- [17] G. Yanxia, W. Baozhong, & R. Zhenhui, Research of precision farming expert system based on GIS. In *The eighth international conference on electronic measurement and instruments ICEMI*. pp. 858–861. 2007.
- [18] K. S. Pandey, V. Santi, & A. Tande. Smart agriculture: Technological advancements on agriculture— A systematical review. In *Deep Learning for Sustainable Agriculture*. 2022.
- [19] M. Lango & J. Stefanowski, Multi-class and feature selection extensions of roughly balanced bagging for imbalanced data, *J Intell Inf Syst*; Vol. 50, pp. 97–127. 2018.
- [20] Z. Karimi, M.M.R. Kashani & A. Harounabadi, Feature ranking in intrusion detection dataset using combination of filtering methods, *International Journal of Computer Applications* Vol. 78, pp. 21–27. 2013.
- [21] V. N. Vapnik, *Statistical Learning Theory*. New York, NY, USA: Wiley, first edition. 1998.
- [22] P.S. Maya Gopal & R. Bhargavi. Performance evaluation of best feature subsets for crop yield prediction using machine learning algorithms, *Appl. Artif. Intell.* 33, pp. 621–642. 2019.
- [23] F. Balducci, D. Impedovo & G. Pirlo, Machine learning applications on agricultural datasets for smart farm enhancement, *Machines* 6, pp. 38–59. 2018.
- [24] R. Jahan. Applying Naive Bayes classification technique for classification of improved agricultural land soils, *International Journal for Research in Applied Science & Engineering Technology (IJRASET)* VI. 6, pp. 189–193. 2018.

- [25] J.A.K. Suykens & J. Vandewalle, Least squares support vector machine classifiers, *Neural Processing Letters*. Vol. 9 , pp. 293–300. 1999.
- [26] D. Anantha Reddy, B. Dadore & A. Watekar, Crop recommendation system to maximize crop yield in Ramtek region using machine learning, *International Journal of Scientific Research in Science and Technology*. Vol. 6 , pp. 485–489. 2019.
- [27] J. Li. Machine Learning Approaches for Crop Yield Prediction and Nitrogen Recommendation in Precision Agriculture, *Computers and Electronics in Agriculture*, vol. 176, 2020.
- [28] S. Khan. A Review on Crop Recommendation Systems, *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 3, 2020.