Journal of Artificial Intelligence and Computational Technology

Journal homepage: https://ojs.omgfzc.com/index.php/JECT

Predictive Modelling of Crop Rotation Using Data Mining Approaches

Ibrahim Abdallah Hageltoum , Esraa Kamal Abdallah, Mahmoud I Alfeel ,& Salwa Awad Abbas.

[1]Department of Information Technology, Faculty of Computer Science and Information Technology, Kassala University. Sudan.
[2] *Graduate College, University of Medical Sciences and Technology. Sudan.*
[3]Department of Information Technology, Merowe University of Technology. Sudan.
[4]Department of Information Technology, Gulf Colleges, Hafr Al-Batin, 2600, Saudi Arabia.

**Correspondence: E-mail: Hijo0101@hotmail.com,*

Article Info

Article History:
Submitted/Received 15 July 2024
Revised in revised format 20 August 2024
Accepted 20 September 2024
Available 15 October 2024
Publication Date 15 October 2024

Keyword:
Crop Rotation, Random Forest Algorithm, Decision Tree Algorithm, ***K-Neighbors*** *Classifier.*

Cite this article: *Hageltoum, I. A., Abdallah, E. K., Alfeel, M. I., & Abbas, S. A. (2024). Predictive modelling of crop rotation using data mining approaches. Journal of Artificial Intelligence and Computational Technology, 1(1).*

ABSTRACT

Agriculture is crucial for economic growth and food security, particularly in agro-based countries. As the global population grows, the demand for food increases, necessitating improvements in agricultural productivity. Traditional methods have often fallen short, and innovative approaches such as data mining and machine learning are needed. This research aims to develop a predictive model for crop rotation using machine learning techniques. A comprehensive dataset was collected and preprocessed to train various algorithms. The proposed model demonstrated that machine learning could effectively predict suitable crops for cultivation, thereby enhancing crop yield and sustainability. The evaluation results were promising, with the Random Forest model achieving a precision of 0.67 to 1.00, recall of 0.43 to 1.00, and F1-score of 0.60 to 1.00; the Decision Tree model had a precision of 0.50 to 1.00, recall of 0.43 to 1.00, and F1-score of 0.50 to 1.00; and the K-Neighbors Classifier model showed precision of 0.40 to 1.00, recall of 0.43 to 1.00, and F1-score of 0.50 to 1.00.

1. Introduction

Agriculture is known as a most valuable active sector for most countries. The increase of the world population is expected to attain 9.8 billion in 2050 and 11.2 billion in 2100 [1], which automatically increase the food demand; as a consequence, Agriculture proved to be the main source of food and clothes for the population of the world. This initiates the non-stop need to improve the food production to satisfy the food demand of the society. Yield of the crop depends on various meteorological and agronomical factors such as seed quality, rainfall, soil fertility, fertilizers, crop rotation and husbandry. In order to assess the relationship between these factors, crop yield and to identify the input variables in enhancing the productivity of crop, a real data set should be collected from farmers [2]. With food security challenges facing the nations globally, agriculture sustainability has been a significant consideration for the international concerned agencies like the Food and Agriculture Organization (FAO) and United Nations (UN) [3]. Agriculture sustainability rely on reasonable use of natural resources, reducing environmental degradation as well as adapting to climate change. Hence the concept of crop rotation must be handled with care. Crop Rotation is one of the solutions to deal with the growing demand for food while meeting sustainability. Crop Rotation is the practice of growing a series of different types of crops in the same area across sequence of growing seasons. Moreover, crop rotation is very important in optimizing land and labor productivities, enhancing higher cropping intensities, and producing better crop yield [4]. It achieves land productivity, ensuring fertility across the whole year, and best used for the land. However, crop rotation supports the integrated land use assessment and soil nutrients. It affects the economic and environmental performance of cropping systems and is important for the design and realization of sustainable agricultural systems. Therefore, there is a need for scientific analysis and appropriate techniques. For that, a model will be developed based on data mining techniques. Machine learning methods are viewed as an appropriate consideration to the model. This is because machine learning methods have shown benefits in many industries and have also been successfully applied to predict crop yield in other studies. [5]

2. Related Works

Kumar et al. they presented a technique named Crop Selection Method (CSM) to select sequence of crops to be sowing over season. This method takes crop, their sowing time, plantation days and predicted yield rate for the season as input and finds a sequence of crops whose production per day are maximum over season. The crop sowing table data are gathered from farmer(s) of Patna District, Bihar (India). Performance and accuracy of this method depends on predicted value of influenced parameters so there is a need to adopt a prediction method with more accuracy and high performance. [23]

This is a research project on the management of crops, Karthikeyan et al. explored the efficiency and usefulness of the crop deployment methods. They used Random Forest technique, which showed it can make an efficient processes and the accuracy of the prediction is high. [24] The work done by Waikar et al. [9], it built a system that suggest crop based on soil classification with assembling classifiers system has been created. The system combined Artificial Neural Network (ANN), Bagged Tree, Naive Bayes, Adaboost, and Support Vector Machine (SVM) algorithms to improve the accuracy of the selection which gives the list appropriate crop according to the soil type. In order to anticipate the crop selection for an increase in crop yield rate and to provide more profit to the farmers, this study used the Crop Variety Selection Method, or CVSM, which used machine learning techniques and artificial learning algorithms in agriculture. Waikar et al. suggest that the more input parameters, such as micronutrients, fertilizer needs, and disease susceptibilities, the more precise and trustworthy the findings of production rate prediction. [14] This work introduced by Majumdar et al. [10], which

DOI: https://doi.org/10.70274/jaict.2024.1.1.18.
p- ISSN 34562-3478 e- ISSN 5443-1243

focuses on the analysis of the agriculture data and finding optimal parameters to maximize the crop production using data mining techniques, based on historical data of crop yield. It covered the Partition around medoids (PAM), clustering large applications (CLARA), Modified DBSCAN clustering methods and multiple linear regression method. Using these methods crop data set is analyzed and determined the optimal parameters for the wheat crop production. They used Multiple linear regression to find the significant attributes and form the equation for the yield prediction. This work found out that DBSCAN gives the better clustering quality than PAM and CLARA, CLARA gives the better clustering quality than the PAM. Becker et al., work on agriculture planning, to analyze and predict the crop using soil properties parameters. There are 1600 datasets used in this application, support vector machine model used to train and test this application. The accuracy achieved to predict the suitable crops and fertilizers for the field using the SVM model is 100%. [25]

3. Problem Statement

Sudan was classified in the 90s as World Basket Food, because of its huge resources (water, good fertile lands and good soils) which can comprise sustainable economic and crop production. Food security has become a real challenge for organizations in charge of the food program and for the majority of countries, especially African countries [6]. Sustainable agriculture involves the selection of crop suited to the location, conditions of the farm, crops diversity, proper soil management, and efficient use of farm resources. Most of our practices in agriculture production are totally traditional, that means agriculture harvests cannot attend the maximum benefits. The crop diversity in any cultivation plan must be based on scientific approaches so as not to drop the fertility of the soil, which leads to poor production. This research proposes stable prediction model for a stable crop rotation to guarantee the sustainability of land using data mining techniques.

4. Proposed Model

The current crop selection process used by farmers in Sudan heavily relies on past experiences, often resulting in unsatisfactory yields despite significant effort. This highlights the need for a data-driven approach to enhance agricultural outcomes. The proposed model, illustrated in Figure 1, begins with the collection of a comprehensive dataset. This dataset undergoes preprocessing to address missing values and normalize the data. Data preparation involves scaling and normalizing features to ensure consistency. The dataset is divided into 80% for training and 20% for testing to maintain representative samples in both subsets. Various data mining algorithms are applied to train and validate the models, with their performance rigorously evaluated. The parameters listed below will help farmers recommend optimum crops to cultivate, yielding better results. The major parameters considered are: 1) Crop name, 2) Sowing time (month/season), 3) Region, 4) Irrigation type, and 5) Historical data.

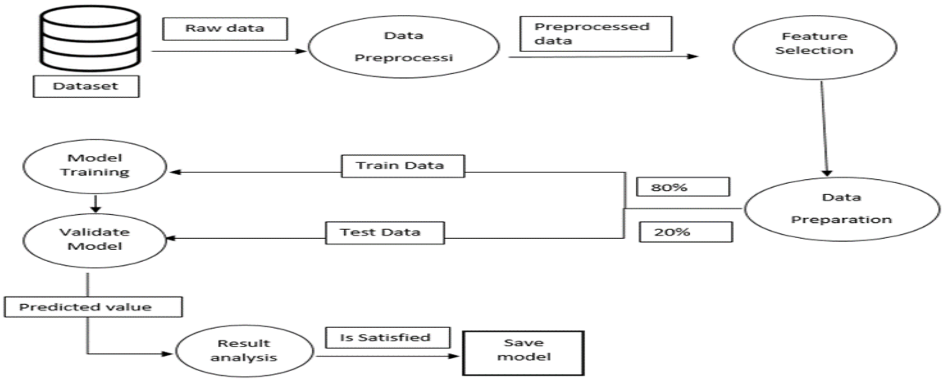

Figure 1. Proposed Model

DOI: https://doi.org/10.70274/jaict.2024.1.1.18.
p- ISSN 34562-3478 e- ISSN 5443-1243

4.1. Data Collection

This step is the most important and crucial during the project development. For this research, the researcher collected a dataset used to create a comprehensive dataset for modeling the crop prediction algorithm. The sources of our dataset include the Arab Organization for Agricultural Development (Arab Agricultural Statistics Annual Book) and the Department of Planning & Agricultural Economics (Ministry of Agriculture).

Due to the lack of uniformity in how the data was presented, the researcher constructed a modified dataset by selecting appropriate parameters. The modified dataset includes the following columns: Irrigation Type, Area, Year, Planted, Harvested, Total Yield, and Yield per Feddan.

Interviews were conducted as a method of data collection through direct communication between the researcher and the respondents. Both structured and unstructured questions were utilized in the interviews because of their flexibility, adaptability, and the depth of insight they provide on the topic. The researcher interviewed an expert in agriculture multiple times to understand the philosophy of crop rotation practices and the general rules governing production activities. The expert highlighted several points:

1. The same crop should not be planted in the same piece of land for two consecutive seasons to avoid crop infections with pests and diseases specific to that crop.
2. Replanting the same crop in the same piece of land twice can create high competition for resources, exhausting the soil and reducing the chances of success for the repeated crop.
3. Crops from the same family (e.g., cotton, karkady, and okra) should not be planted in the same area because they host the same pests and diseases.

4.2. Data Integration & Preparation

The data also had several inconsistencies since it was obtained from various sources. The collected data was first filtered based on the following criteria:

1. **(Missing Values:** The missing values were removed from the dataset as they could lead to inconsistencies and ultimately result in incorrect predictions.
2. **Redundant Data:** Redundant data were discarded because they did not add any additional significance to the process.
3. **Spelling Aberrations:** Inconsistent spellings throughout the files were normalized to a single value to improve the model's accuracy.

The data was eventually consolidated into a single dataset by flattening it, which facilitated the model building and computation process.

DOI: https://doi.org/10.70274/jaict.2024.1.1.18.
p- ISSN 34562-3478 e- ISSN 5443-1243

4.3. Predictive Modeling

We employed various data mining techniques to develop predictive models for crop rotation. The primary methods used include the Random Forest Classifier, Decision Tree Classifier, and K-Neighbors Classifier. The dataset has been loaded into Jupyter and divided into 80% for training and 20% for testing to ensure representative samples in both subsets.

The steps are as follows:

1. Data Splitting: The dataset is split into 80% for training and 20% for testing.
2. Model Training: The training data is used to train the Random Forest Classifier, Decision Tree Classifier, and K-Neighbors Classifier algorithms.
3. Model Testing: The trained models are then tested on the test data to predict the crops.

The model is trained using the training data, where each algorithm learns the patterns and relationships within the dataset. After training, the model is given the test data to predict the crops.

By comparing these algorithms, we aim to identify the most effective model for predicting crop selection, thereby providing farmers with data-driven recommendations for optimal crop cultivation.

4.4. Evaluation

Positive (1)

To measure the performance of the predictive models, we employed a confusion matrix. The confusion matrix is a fundamental tool in machine learning used to evaluate the accuracy of a model by comparing predicted values with actual values. It provides a detailed breakdown of the model's performance by categorizing predictions into four possible outcomes:

True Positive (TP): The model correctly predicts the positive class.

True Negative (TN): The model correctly predicts the negative class.

False Positive (FP): The model incorrectly predicts the positive class.

False Negative (FN): The model incorrectly predicts the negative class.

These outcomes are represented in Figure 2. The confusion matrix helps to identify not only the accuracy of the model but also the types of errors it makes. By analyzing the confusion matrix, we can determine the precision, recall, and overall accuracy of the model. This analysis provides valuable insights into the model's performance and helps in identifying areas for improvement. For instance, a high number of false positives may indicate that the model is overly optimistic in its predictions, while a high number of false negatives may suggest that the model is missing positive cases. In this study, the confusion matrix was used to evaluate the performance of the Random Forest Classifier, Decision Tree Classifier, and K-Neighbors Classifier. By comparing the confusion matrices of these models, we were able to select the one with the highest accuracy and the most balanced performance across all four categories. The detailed evaluation using the confusion matrix ensures that the selected model provides reliable recommendations for crop rotation, ultimately aiding farmers in making data-driven decisions for optimal crop cultivation [31].

DOI: https://doi.org/10.70274/jaict.2024.1.1.18.
p- ISSN 34562-3478 e- ISSN 5443-1243

Actual Values

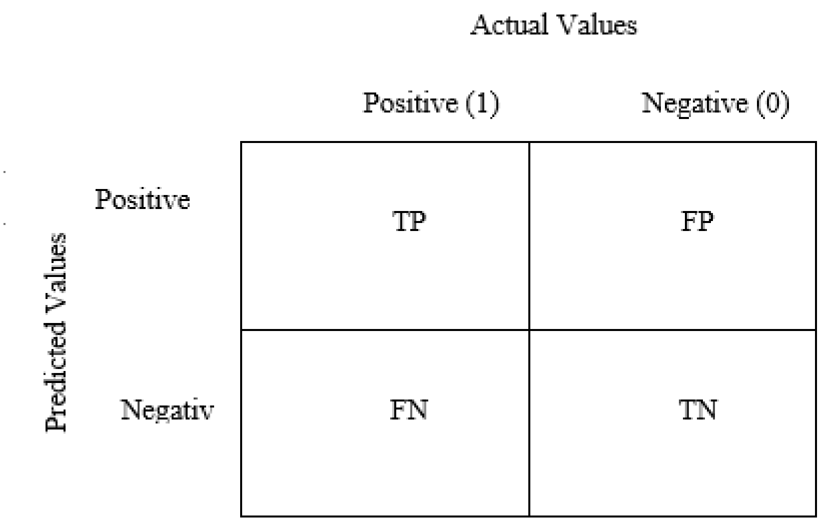

Figure 2. Confusion Metrix

Precision: also called Positive predictive value. The ratio of correct positive predictions to the total predicted positives

$$P = \frac{TP}{TP+FP} \qquad (1)$$

Recall: also called Sensitivity, Probability of Detection, True Positive Rate. The ratio of correct positive predictions to the total positives.

F1 score: is the measure of a test's accuracy. It considers both the precision p and recall r of the test to compute the score.

$$F1\ score = 2 \times \frac{p \times r}{p+r} \qquad (2)$$

Accuracy: Accuracy is defined as the ratio of correctly predicted examples by the total.

$$Accuracy = \frac{TP+TN}{TP+FN+FP+FN} \qquad (3)$$

5. Experimental Results and dissection

Model is designed by using Anaconda3, Anaconda is an open-source program that contains numerous Python packages which can be used during programming to make implementation easier for the developer through the use of predetermined functions. [32]

Dataset

Figure 3 shows the first draft of the data (raw data), it contains five columns (Area, Planted, Harvested, Yield, and Production) and the type of irrigation been used. Two columns with string data type, and

DOI: https://doi.org/10.70274/jaict.2024.1.1.18.
p- ISSN 34562-3478 e- ISSN 5443-1243

four columns with integer data type. The size of data differs from any type of crop. The data composed of 36 rows of groundnuts, 74 rows of sorghum, 42 rows of millet, 40 rows of sesame, 20 rows of wheat, 32 rows of sunflower, and 26 rows of cotton

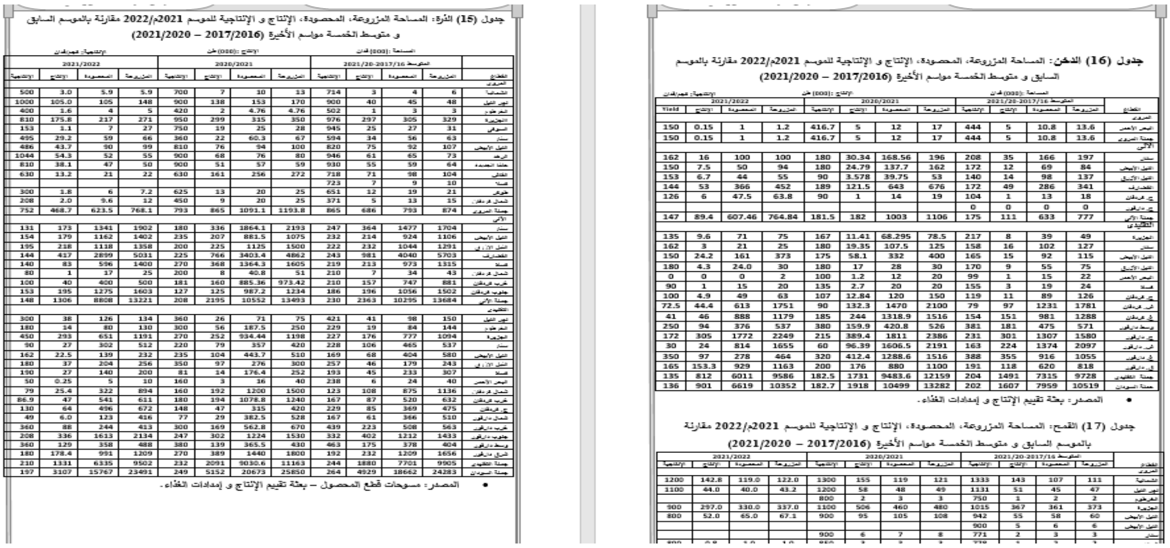

Figure 3 Raw Data (Level Zero)

Data been inserted manually into excel sheet and then a lot of filtering and editing added as showed in figure 4.

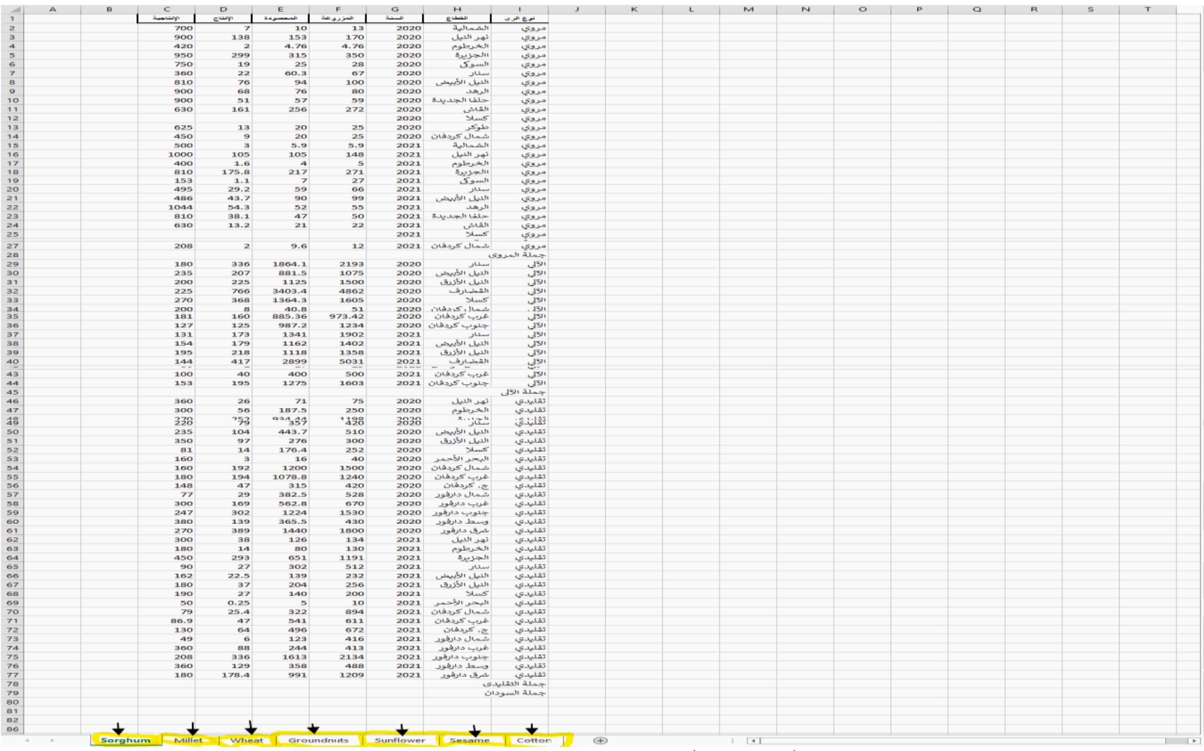

Figure 4 Raw Data (Level 1)

Note that data has been added spirited manually to different sheets as been highlighted in the figure above.

DOI: https://doi.org/10.70274/jaict.2024.1.1.18.
p- ISSN 34562-3478 e- ISSN 5443-1243

Then it's been combined in one sheet to ease its uploading in Juypter, Figure 5 shows the final data, with 276 rows and 6 columns.

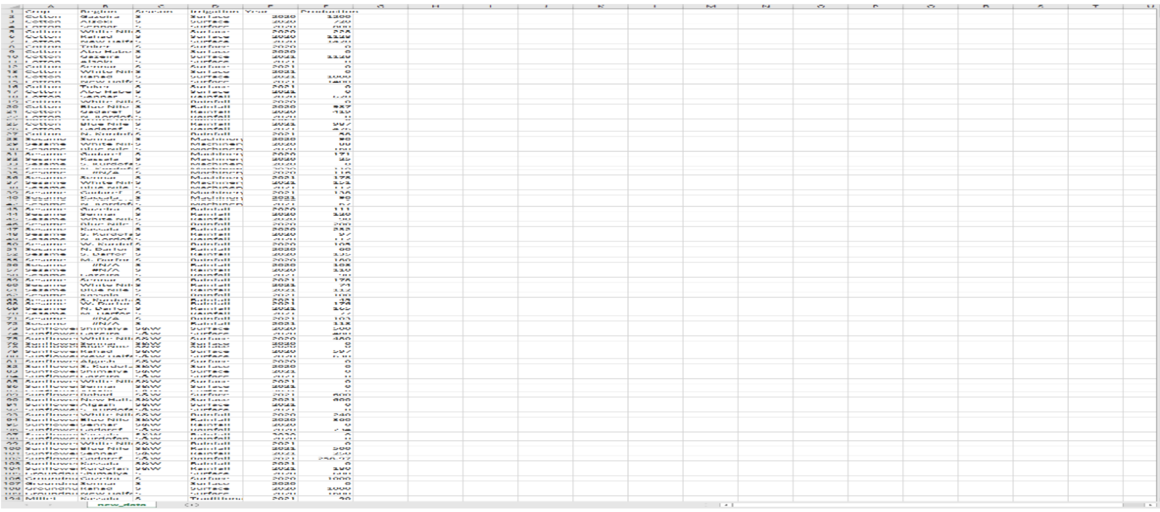

Figure 5 Final Dataset

Figure 6 shows the production of each crop.

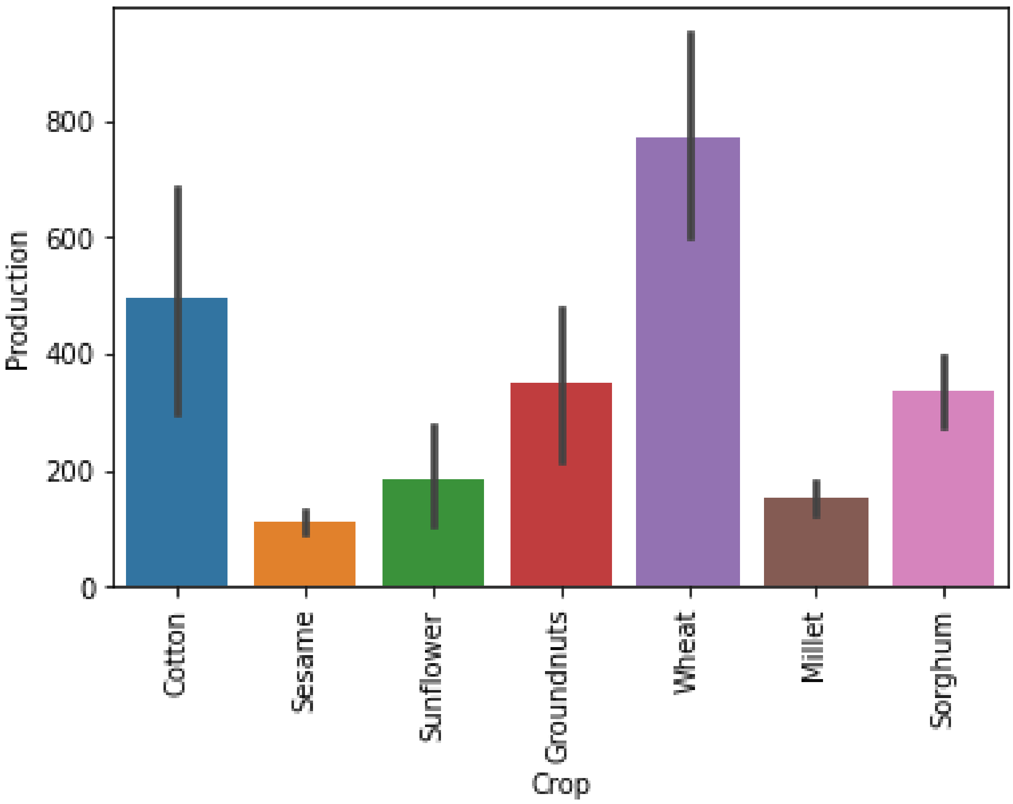

Figure 6 Production for Crops

DOI: https://doi.org/10.70274/jaict.2024.1.1.18.
p- ISSN 34562-3478 e- ISSN 5443-1243

6. Modelling Approach

Firstly, data has been shuffled. Secondly, the data was split into a training and test set. Thirdly, a variety of machine learning methods were applied to build a suitable prediction model using the training dataset. Lastly, the final models for each of the methods were evaluated using the test set.

1. **Phase One:** Shuffle data During the first phase, and due to the fact that data will be split into training and testing datasets; data has been shuffled randomly. This is to ensure that the model is not overfitting to certain pattern duo sort order and to avoid any element of bias/patterns in the split datasets before training the model.

2. **Phase Two:** Data Split: Throughout the second phase, data has been spilt into 80% training and 20% testing set. Training set is the set of data used for training a model and this will also be the largest set of data, where the model will use and learns the behavior from and understand it's patterns. Test set is used to evaluate the performance of a model built using a training dataset. The approach used here is HOLD-OUT method; since this approach is often used when the data set is small and there is not enough data to split into three sets (training, validation, and testing) and it is simple to implement as shown in figure 7.

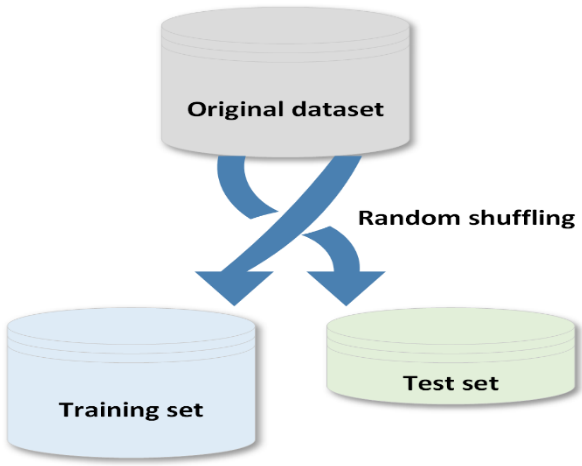

Figure 7 Random Shuffle[34]

3. **Phase Three:** Build Baseline Model Baseline model acts as a reference in a machine learning project. It serves as a benchmark, which enable more informative evaluation of a trained model. [35]. The accuracy of baseline model was: 31.48 %

DOI: https://doi.org/10.70274/jaict.2024.1.1.18.
p- ISSN 34562-3478 e- ISSN 5443-1243

4. **Phase Four:** Methods used

This phase includes three machine learning classifications methods that been used: decision Tree, Random Forest, and K-Nearest Neighbors.

Evaluation of Decision Tree Model

The precision, recall, and F1-score using Decision Tree are listed in Table 1. The results demonstrate that Cotton and Groundnuts have the lowest measurements among others. Cotton has a precision of 0.50, recall of 0.50, and F1-score of 0.50, and Groundnuts has a precision of 0.60, recall of 0.43, and F1-score of 0.50. Where Sorghum, Sunflower and Wheat get the highest score of precision, recall and F1-score.

Table 1 Precision, Recall and f1-score values for Decision Tree Model

Crop	Precision	Recall	F1-score
Cotton	0.50	0.50	0.50
Groundnuts	0.60	0.43	0.50
Millet	0.67	0.60	0.63
Sesame	0.50	0.67	0.57
Sorghum	1.00	1.00	1.00
Sunflower	1.00	1.00	1.00
Wheat	1.00	1.00	1.00

6.1. Evaluation of Random Forest Model

The precision, recall, and F1-score using Decision Tree are listed in Table 2. The results signify that Sesame and Millet have the lowest precision measurements among the other. The Sesame has a precision of 0.46 and Millet has a precision of 0.62. Where Sorghum, Sunflower and Wheat get the highest score of precision, recall and F1-score.

Table 2 Precision, Recall and f1-score values for Random Forest Model

Crop	Precision	Recall	F1-score
Cotton	0.67	1.00	0.80
Groundnuts	1.00	0.43	0.60
Millet	0.62	0.50	0.56
Sesame	0.46	0.67	0.55
Sorghum	1.00	1.00	1.00
Sunflower	1.00	1.00	1.00
Wheat	1.00	1.00	1.00

DOI: https://doi.org/10.70274/jaict.2024.1.1.18.
p- ISSN 34562-3478 e- ISSN 5443-1243

6.2. Evaluation of K- Nearest Neighbor

The precision, recall, and F1-score using K-Nearest Neighbor are listed in Table 3. The results show Groundnuts has the lowest measurements among others. Groundnuts and Sesame the only one with precision lower than 0.50. Where Sorghum, Sunflower and Wheat get the highest score of precision, recall and F1-score. 33

Table 3 Precision, Recall and f1-score values for K- Nearest Neighbor

Crop	Precision	Recall	F1-score
Cotton	0.50	0.75	0.60
Groundnuts	0.40	0.29	0.33
Millet	0.75	0.60	0.67
Sesame	0.45	0.56	0.50
Sorghum	1.00	1.00	1.00
Sunflower	1.00	1.00	1.00
Wheat	1.00	1.00	1.00

Figure 8 summary the performance measurements for Decision Tree, Random Forest and K-Nearest Nighbour models

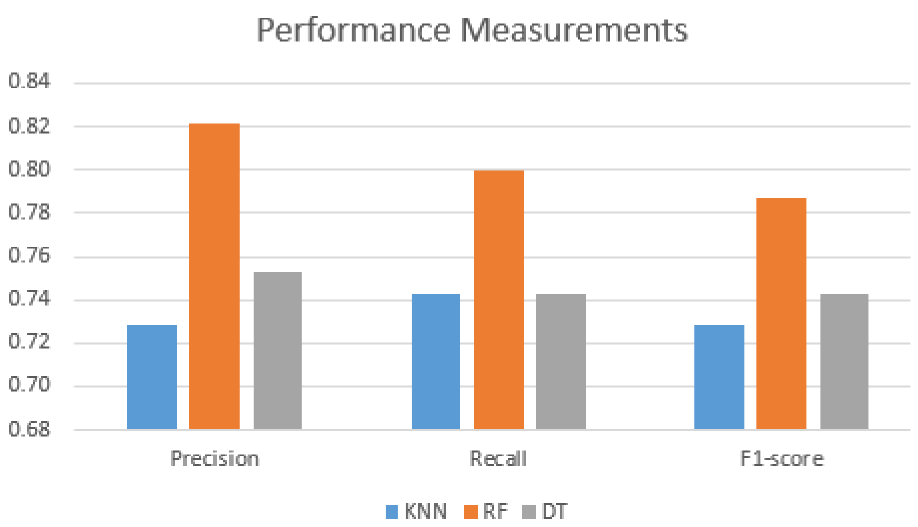

Figure 8 summary the performance measurements

Figure 9 F1-score, precision, and recall performance measurements of machine learning algorithms.

In order to provide a brief overview, table 4 and figure 10 show an aggregation of the presented results and illustrates the general performance of the used algorithms.

DOI: https://doi.org/10.70274/jaict.2024.1.1.18.
p- ISSN 34562-3478 e- ISSN 5443-1243

Table 4 Accuracy Results

Algorithm	Accuracy
Random Forest	**77.78**
Decision Tree	75.93
K-Nearest Neighbors	70.37

Figure 10 Machine Learning Techniques' Results

Results showed that random forest performed much better than decision tree and KNN. This lead to build the model with random forest.

7. Conclusion

The study highlights the potential of data mining and machine learning techniques in improving agricultural practices, particularly in crop rotation. By employing algorithms such as Random Forest, Decision Tree, and K-Neighbors Classifier, the model was able to provide accurate crop recommendations. The evaluation, based on precision, recall, and F1-score, confirmed the model's effectiveness. Specifically, the Random Forest model achieved a precision of 0.67 to 1.00, recall of 0.43 to 1.00, and F1-score of 0.60 to 1.00; the Decision Tree model had a precision of 0.50 to 1.00, recall of 0.43 to 1.00, and F1-score of 0.50 to 1.00; and the K-Neighbors Classifier model showed precision of 0.40 to 1.00, recall of 0.43 to 1.00, and F1-score of 0.50 to 1.00. The results indicate that data-driven approaches can significantly aid farmers in optimizing crop selection, leading to better yield and resource management. This research underscores the importance of integrating advanced technologies in agriculture to meet the increasing food demands and ensure sustainable farming practices.

DOI: https://doi.org/10.70274/jaict.2024.1.1.18.
p- ISSN 34562-3478 e- ISSN 5443-1243

References

UN, "World population projected to reach 9.8 billion in 2050, and 11.2 billion in 2100," 2017. https://www.un.org/development/desa/en/news/population/world-population-prospects-2017.html (accessed Aug. 10, 2022).

[A. Arumugam, "A predictive modeling approach for improving paddy crop productivity using data mining techniques," Turkish J. Electr. Eng. Comput. Sci., vol. 25, no. 6, pp. 4777–4787, 2017, doi: 10.3906/elk-1612-361.

S. F. Nisperos and F. D. McKenzie, "Assessing crop rotation sustainability using analytical hierarchy process," Proc. - Eur. Counc. Model. Simulation, ECMS, pp. 336–340, 2018, doi: 10.7148/2018-0336.

C. Singha, K. C. Swain, and S. K. Swain, "Agriculture-10-00213.Pdf," 2020.

S. Mishra, D. Mishra, and G. H. Santra, "Applications of machine learning techniques in agricultural crop production: A review paper," Indian J. Sci. Technol., vol. 9, no. 38, 2016, doi: 10.17485/ijst/2016/v9i38/95032.

C. Lontsi Saadio, W. Y. H. Adoni, R. Aworka, J. T. Zoueu, F. Kalala Mutombo, and C. L. M. Kimpolo, "Crops Yield Prediction Based on Machine Learning Models: Case of West African Countries," SSRN Electron. J., vol. 2, no. December 2021, 2022, doi: 10.2139/ssrn.4003105.

H. Ait Issad, R. Aoudjit, and J. J. P. C. Rodrigues, "A comprehensive review of Data Mining techniques in smart agriculture," Eng. Agric. Environ. Food, vol. 12, no. 4, pp. 511–525, 2019, doi: 10.1016/j.eaef.2019.11.003.

S. Mishra, P. Paygude, S. Chaudhary, and S. Idate, "Use of data mining in crop yield prediction," Proc. 2nd Int. Conf. Inven. Syst. Control. ICISC 2018, no. Icisc, pp. 796–802, 2018, doi: 10.1109/ICISC.2018.8398908.

V. C. Waikar, S. Y. Thorat, A. A. Ghute, P. P. Rajput, and M. S. Shinde, "Crop Prediction based on Soil Classification using Machine Learning with Classifier Ensembling," Int. Res. J. Eng. Technol., no. May, pp. 4857–4861, 2020, [Online]. Available: www.irjet.net.

J. Majumdar, S. Naraseeyappa, and S. Ankalaki, "Analysis of agriculture data using data mining techniques: application of big data," J. Big Data, vol. 4, no. 1, 2017, doi: 10.1186/s40537-017-0077-4.

J. W. Grzymala-Busse and W. J. Grzymala-Busse, Handling Missing Attribute Values. 2009.

S. Devadhe, A. Kausadikar, P. Daphal, and A. Joshi, "Expert System for Crop selection," vol. 3, no. 3, pp. 436–438, 2017.

"Types of Machine Learning Algorithms," 2020. https://7wdata.be/visualization/types-of-machine-learning-algorithms-2/ (accessed Aug. 31, 2022).

G. Vishwa, J. Venkatesh, and C. Geetha, "Crop Variety Selection Method using Machine Learning," vol. 12, no. 4, pp. 35–38, 2019.

B. T. Jijo and A. M. Abdulazeez, "Classification Based on Decision Tree Algorithm for Machine Learning," vol. 02, no. 01, pp. 20–28, 2021, doi: 10.38094/jastt20165.

P. Patil, V. Panpatil, and P. S. Kokate, "Crop Prediction System using Machine Learning Algorithms," pp. 748–753, 2020.

A. O. Ok, O. Akar, and O. Gungor, "Evaluation of random forest method for agricultural crop classification," Eur. J. Remote Sens., vol. 45, no. 1, pp. 421–432, 2012, doi: 10.5721/EuJRS20124535.

A. B. Shaik and S. Srinivasan, A brief survey on random forest ensembles in classification model, vol. 56. Springer Singapore, 2019.

K. Taunk, "A Brief Review of Nearest Neighbor Algorithm for Learning and Classification," 2019 Int. Conf. Intell. Comput. Control Syst., no. Iciccs, pp. 1255–1260, 2019.

S. Pudumalar, E. Ramanujam, R. H. Rajashree, C. Kavya, T. Kiruthika, and J. Nisha, "Crop recommendation system for precision agriculture," 2016 8th Int. Conf. Adv. Comput. ICoAC 2016, pp. 32–36, 2017, doi: 10.1109/ICoAC.2017.7951740.

R. Priya, D. Ramesh, and E. Khosla, "Naïve Bayes MapReduce Precision Agricultural Model," 2018 Int. Conf. Adv. Comput. Commun. Informatics, pp. 99–104, 2018.

DOI: https://doi.org/10.70274/jaict.2024.1.1.18.
p- ISSN 34562-3478 e- ISSN 5443-1243

A. Tharwat, "Parameter investigation of support vector machine classifier with kernel functions," Knowl. Inf. Syst., vol. 61, no. 3, pp. 1269–1302, 2019, doi: 10.1007/s10115-019-01335-4.

R. Kumar, M. P. Singh, P. Kumar, and J. P. Singh, "(Paper 7 Tafur)," no. May, pp. 138–145, 2015.

R. Karthikeyan, M. Gowthami, A. Abhishhek, and P. Karthikeyan, "Implementation of effective crop selection by using the Random Forest algorithm," Int. J. Eng. Technol., vol. 7, no. 3.34 Special Issue 34, pp. 287–290, 2018, doi: 10.14419/ijet.v7i3.34.19209.

F. G. Becker et al., "No 主観的健康感を中心とした在宅高齢者における 健康関連指標に関する共分散構造分析Title," Syria Stud., vol. 7, no. 1, pp. 37–72, 2015, [Online]. Available: https://www.researchgate.net/publication/269107473_What_is_governance/link/548173090cf22525dcb61443/download%0Ahttp://www.econ.upf.edu/~reynal/Civil wars_12December2010.pdf%0Ahttps://think-asia.org/handle/11540/8282%0Ahttps://www.jstor.org/stable/41857625.

S. Veeragandham and H. Santhi, "A review on the role of machine learning in agriculture," Scalable Comput., vol. 21, no. 4, pp. 583–589, 2020, doi: 10.12694:/scpe.v21i4.1699.

T. Setiadi, F. Noviyanto, H. Hardianto, A. Tarmuji, A. Fadlil, and M. Wibowo, "Implementation of naïve bayes method in food crops planting recommendation," Int. J. Sci. Technol. Res., vol. 9, no. 2, pp. 4750–4755, 2020.

S. K. S. Raja, R. Rishi, E. Sundaresan, and V. Srijit, "Demand based crop recommender system for farmers," Proc. - 2017 IEEE Technol. Innov. ICT Agric. Rural Dev. TIAR 2017, vol. 2018-Janua, no. April 2017, pp. 194–199, 2018, doi: 10.1109/TIAR.2017.8273714.

G. Zeng, "On the confusion matrix in credit scoring and its analytical properties," Commun. Stat. - Theory Methods, vol. 49, no. 9, pp. 2080–2093, 2020, doi: 10.1080/03610926.2019.1568485.

A. M. Jain, "Confusion Matrix," 2020. https://medium.com/@anishajain2910/confusion-matrix-30249214041d (accessed Aug. 24, 2022).

Z. Fayyaz, M. Ebrahimian, D. Nawara, A. Ibrahim, and R. Kashef, "Recommendation systems: Algorithms, challenges, metrics, and business opportunities," Appl. Sci., vol. 10, no. 21, pp. 1–20, 2020, doi: 10.3390/app10217748.

B. O. F. Technology and C. Engineering, "Data Processing Unit for Crop Monitoring and Advisory System," pp. 2019–2020.

"Libraries in Python." https://www.geeksforgeeks.org/libraries-in-python/ (accessed Sep. 02, 2022).

"Creating training and test sets," 2022. https://www.oreilly.com/library/view/machine-learning-algorithms/9781789347999/f394f9d6-19dc-4103-8715-8895ef4d10b8.xhtml (accessed Sep. 02, 2022).

A. Nair, "Baseline Models: Your Guide For Model Building." https://towardsdatascience.com/baseline-models-your-guide-for-model-building-1ec3aa244b8d (accessed Sep. 05, 2022).

DOI: https://doi.org/10.70274/jaict.2024.1.1.18.
p- ISSN 34562-3478 e- ISSN 5443-1243