Crop Recommendation using Machine Learning Algorithms and Soil Attributes Data

Takalani Orifha Mufamadi
School of Computer Science and Applied Mathematics
The University of the Witwatersrand
Johannesburg, South Africa
1100162@students.wits.ac.za

Ritesh Ajoodha
School of Computer Science and Applied Mathematics
The University of the Witwatersrand
Johannesburg, South Africa
ritesh.ajoodha@wits.ac.za

***Abstract*—Yield optimisation in farming is very crucial for countries whose agricultural sector plays an important and significant role in the economy. Better yields could be achieved if farmers knew what to grow based on their location. In this study we implement precision agriculture to develop a crop recommendation system that can help farmers increase crop yield. We use various machine learning algorithms for the classification of soil which allows our system to be able to recommend crops to grow based on the attributes of the particular soil. The only soil attributes we consider in this study are: three main soil nutrients, namely, nitrogen, potassium, and phosphorus; as well as soil pH level. This means that our solution is more accessible to most farmers and thus would have a greater impact on the agricultural sector of South Africa, because this type of soil data is more accessible. Our contribution to the literature is that we are able to develop a highly accurate crop recommendation system using a soil attribute data that is more accessible. We also managed to establish which algorithm is best suited for the task of soil classification.**

***Index Terms*—Agriculture, Precision agriculture, Crop recommendation, Machine learning, Kernel-based support vector machines, Random forests, South Africa**

I. Introduction

Agricultural activity plays an important and significant role in the economies of most countries, especially in less developed countries where the contribution of agricultural activity toward the economy is substantial [1]. Hence, high levels of production in the agricultural sector is of utmost importance for the economic well-being of most countries and the welfare of their people. [2] [3] [1]. This is indeed the case for most countries in Africa including South Africa and it's neighbouring countries Besides the need to increase (maximize) productivity, there is also a variety of factors such as climate change [2], reduction of farmland due to population growth [4], and globalization [3] [4] that lead to the need for farmers to adapt by changing the crops they choose to grow. The greatest challenge for most farmers and in particular subsistence farmers whose contribution in the agricultural sector is a significant one, is being able to decide what to grow and where [5] [4] [6].

In other words, farmers would be more productive if they knew what types of crops or livestock to grow depending on their location and its attributes, as well as other environmental factors that affect farm productivity. Which would result in the improvement of the food security of a country and it's overall economic well-being.

Our main aim is to come up with a crop recommendation (CR) system that will help farmers select crops that will produce the greatest yield for that particular soil type. This CR system should be able to recommend a crop based only on the attributes of the soil, which would make it more accessible to most farmers because of ease of access to such data. An inaccurate CR system would result in great losses for farmers, the agricultural sector, and the economy as a whole if the CR system is adopted by a significant number of farmers. Therefore, it is crucial that the CR system is highly accurate to avoid such losses.

Other authors have incorporated more information about the soil (or land) and other environmental factors that affect crop growth the development of their CR systems, which yields higher accuracy than the model we develop. However, this also makes their models more complicated which makes it difficult for most farmers to adopt (use) these models. Therefore, our research aims to develop a CR system that can suggest crops to farmers using only the soil attributes.

Guided by the literature, we use machine learning algorithms, including SVM and RF, to develop a CR system with high accuracy by just considering the three main soil nutrients, namely, nitrogen, potassium, and phosphorus; as well as soil pH level. This would mean that farmers would be able to use the CR system with minimum resources required because this information is more readily available.

The main contributions of our study to the body of literature are: a) we develop a high accuracy CR system that is more accessible to farmers using only soil attributes such as the three main soil nutrients, namely, nitrogen, potassium, and phosphorus; as well as soil pH level. b) We compare various machine learning algorithms, including SVM and RF algorithms, to see which one is best suited for developing a CR system using only soil attribute data. c) We compare the power of the features we use in this study with that of other previously used features to establish which features will yield higher accuracy.

This research report is structured as follows: we begin by introducing the topic in Section I. In section II we provide an overview of the literature related to our study. Following this, we highlight and discuss the methodology and data used to implement the our research in section III. Then in section IV we highlight and discuss the results obtained using the methodology highlighted in III. Section V is the final chapter, wherein, we summarise the main points and findings discussed of this study as well as potential future work.

II. Related Work

A. Precision Agriculture (PA)

Researchers have over the years attempted to come up with ways that can assist farmers make better crop choices to maximize their yield using precision agriculture (PA) [1] [3]. The maximization of farm productivity can be achieved by implementing PA, which could have a great impact on the well-being of farmers, the agricultural sector, and the economy as a whole. PA is a modern farming technique which uses research data and algorithms to suggest the most appropriate crop to grow and/or livestock to raise based on parameters that are specific to the farmer's location [3] [7] [8].

Many authors have developed various PA models that attempt to assist farmers in choosing the best crop for their soil based on the attributes of the soil [3] [7] [8]. The majority of these PA models that have been developed over the years recommend crops for particular land based on the attributes of the soil, and can be referred to as crop recommendation (CR) systems [9].

B. Recommendation Systems (RSs)

CR systems are part of a broader category of systems called recommendation systems (RSs), that are used to recommend items to users. RSs are programs that are able to recommend or suggest the best items to specific users based on the history of the user, information regarding the items, and the interaction between the user and items [9]. Such systems use a variety of data processing methods to analyze the user's history and similar users' behaviour to be able to predict the interest of the user.

There are three categories of RSs, namely, content-based, collaborative filtering, and hybrid; CR systems belong to the content-based category of recommendation systems. RS that are content based suffer from the problem of new user cold-start. What this entails for CR systems is that they are unable to recommend a crop if the crop characteristics are not available [9]. Various authors have attempted to remedy this problem in order to yield better accuracy [10] [11] [12] [13]. They use opinion mining methods as well as neural networks to circumvent the cold-start problem associated with content-based RS. For our future work we will consider using similar methods for our CR system to circumvent the cold-start problem.

RSs play a major role in multiple areas of application such as e-commerce wherein items that could be interesting to the target user are recommended. Perhaps one of the most important areas of application of RSs is in PA for the development of CR systems [9].

C. Crop Recommendation (CR) Systems

The development of a crop recommendation system at any level requires information and knowledge about the crops and the soil in which crops are grown as well as other environmental factors which affect crop growth such as temperature, precipitation and humidity [1] [9] [7] [3]. The soil attributes that can be considered are the soil nutrients, soil pH level and soil colour, to name a few [7] [3] [14] [8] [6]. When it comes to the sample of crops to be used, we can choose the crops that are most dominant in the location being studied [7] [3] [14].

Some CR systems are machine learning based models [3] [15] [16] [7], some are probabilistic models [2], others are based on artificial intelligence [8], and so on. The most prevalent models amongst these are machine learning based models, which use machine learning algorithms to solve the classification task at hand, namely, soil classification.

The scientific field that gives learning ability to machines without the need for strict programming is referred to as machine learning [17]. Machine learning based models are preferred for the development of CR systems because they provide the best (most accurate) soil classification regardless of the performance measures that are used to evaluate the performance of algorithms [14].

[16] use naive Bayes (NB) and k-nearest neighbours (KNN) for soil classification which allows them to predict crop yield using the data that is available. They also suggest that classification algorithms like support vector machines can be used to create models that are more efficient. This was confirmed by [15] who use kernel based support vector machines (SVM) to develop their CR system and report an accuracy of 94.95%. However, [3] found that random forests (RF) based models performed better than their SVM counterparts when it comes to soil classification. They report an accuracy of 75.73% and 86.35% for SVM and RF, respectively.

D. Evaluation Criteria

Various evaluation parameters are used to rank the accuracy of the models and how effective they are. These performance measures can be categorized into two broad categories, namely, offline and online assessment [9]. Offline assessment uses offline data to assess the recommendation method's efficiency and reliability whereas online assessment uses real-time data for this purpose [9].

Examples of offline performance measures for the evaluation of CR systems include: standard deviation, standard error, root mean square error (RMSE), mean square error (MSE), recall and precision, cross-validation tests, and train-test analysis [9]. Conversion rate, A/B testing, and click-through rate (CTR) are examples of online performance measures [9].

For the purposes of this study, we will focus only on offline methods of model evaluation since offline data is more readily available than real-time data. Using this approach instead of the alternative should not have any significant impact on the results we obtain, and thus, our conclusions are not affected by the evaluation criteria we choose.

In this chapter we have gathered that we can be able to help farmers increase their productivity by implementing PA [1] [15]. The development of a CR system is one way of implementing PA that could help crop farmers increase their productivity [9] [8]. What we also learned from the above discussion is that information about the soil is the most important information we will need to develop our CR system. [18] [19] [20].

We also gathered from the chapter that machine learning algorithms were better suited for development of a CR system [21] [22] [9] and that, in particular, the RF algorithm is best suited for soil classification [3]. This insight was gained by ranking the models and algorithms based on the various evaluation parameters that are commonly used in the literature as discussed above. Another insight we obtained, is that, CR systems suffer from the cold-start problem of content-based RSs which can be solved using neural networks and opinion mining [9].

III. Methodology

The methodology used in this study is similar to that of [3] and [6]

who use machine learning algorithms for soil classification, which is the basis of our CR system. [3] suggest that the RF machine learning algorithm, in particular, is best suited for soil classification. Following this, we develop our CR system using RF algorithm as well as other machine learning algorithms including SVM.

These algorithms are implemented on soil data that includes information about the nutrients in the soil and the pH level of the soil. The reason for this is that the development of a CR system requires information about factors that affect crop growth, for example, soil attributes. The various offline measures that we use to assess the accuracy and the performance of our CR system are borrowed from [9]

The rest of the section is structured as follows: in subsection III-A we discuss the data and features that are required for the implementation of crop recommendation (CR) systems. In subsection III-B we discuss the models that are used to develop the CR system. The performance measures that are used to evaluate the performance of the developed CR systems are discussed in subsection III-C.

A. Data and Features

To develop our crop recommendation system we use a cross-sectional data-set that was sourced from Kaggle and consists of soil attributes as well as other environmental factors that affect crop growth, namely, temperature, humidity, and precipitation. There are no ethical considerations that we had to undertake when obtaining the crop and soil data used in this study.

The following soil attributes are included in the dataset: soil pH, phosphorus, potassium, nitrogen, and magnesium. Soil pH (level of acidity or alkalinity) is also included in the data-set because it affects the availability of nutrients in the soil. It is a master variable which can affect the level of exchangeable aluminium in the soil and also activity of micro- organisms present in the soil.

Temperature, humidity, and precipitation were found to have the greatest power for crop recommendation and models that include these features yield higher accuracy. However, this also makes the model more complicated and less accessible to farmers. As a results we do not use these features in this research which means that our model can be adopted most farmers. We also exclude magnesium since it does not have a great impact on plant growth for plants grown here in South Africa. Thus, for our final dataset we only remain with the soil pH as well as the three main soil nutrients.

In Table I below, we provide the definition or descriptions of the features (soil attributes) we use for our proposed research. Most of these attributes are described in terms of the role they play in overall crop growth in order to highlight their importance in our research.

Variable	Definition or descriptions
1. Nitrogen (N)	The atomic element that is responsible for photosynthesis in the plant.
2. Phosphorus (P)	The atomic element that plays a major role in a crop to store and transfer energy for growth of the crop.
3. Potassium (K)	The atomic element that is required for reproduction of crops.
4. pH level (pH)	The level of alkalinity or acidity.

Table I: Definition or descriptions of the features

Our crop selection criteria is based on crops that are important in the location of study, namely, South Africa. The crops included in the original crop data are maize, rice, banana, mango, grapes, watermelon, apple, orange, papaya, coconut, cotton, jute, coffee, muskmelon, lentil, black-gram, kidney beans, pigeon beans, mung beans, moth beans, and pomegranate. Therefore, we select the following crops from the original crop data: maize, kidney beans, banana, mango, grapes, watermelon, apple, orange, papaya, and cotton. These crops are the ones that are commonly grown in South African and have the greatest contribution to the agricultural sector.

B. Machine Learning Algorithms

In the subsections below, we provide a brief overview of SVM and RF. We focus mainly on SVM and RF because the literature suggests that they are the best performing algorithms when it comes to soil classification [3] [15]. From our results we see that between these two algorithms, RF outperforms SVM when it comes to soil classification. Other algorithms that have been used in the literature include K-NN, K-Star, and NB to name a few. In this study we will compare the performances of some of these algorithms to establish which ones are best suited for soil classification. We also can see from our results that there exist other machine learning algorithms that perform better than those suggested by the literature which contradicts the finding suggested by the literature mentioned above.

a) SVM: SVM uses a set of hyper-planes or a single hyper-plane for different errands, relapse, or characterisation.

These hyper-planes are developed in boundless or high dimension space. SVM calculation stays away from over-fitting by having a regularisation parameter. It also uses portion traps which allows us to develop master learning regarding the issue [3].

b) RF: The RF algorithm belongs in the category of supervised machine learning algorithms and can be used for classification and regression problems. This algorithm is based on ensemble learning and one of its advantages is that it has multiple trees and a subset of the data is used for each tree which makes it unbiased. The advantage of RF algorithm is that the introduction of a new data point in the data-set does not affect the overall algorithm which makes it stable [3].

C. Evaluation Parameters

In the subsections below, we provide a quick overview of some of the online evaluation parameters that were mentioned earlier (section II). See [9] for the definitions of RMSE, MAE, and RAE.

a) Train-test analysis: When it comes to train-test analysis the data is split into data for training and data for testing which are used separately. The training data (data used for training the model) will usually consist of between 70% and 80% of the original data. The testing data is used for model evaluation and would consist of the remaining 20% to 30% of the original data [9]. For this study we use 70% of the original data as our training data and use the remaining 30% for testing purposes.

b) Cross-validation tests: Cross validation tests can be thought of as an n-folds repetition of train-test analysis because the same process of splitting data into train and test data that is used in train-test analysis is repeated n-times. At each fold the accuracy of the is assessed for all the folding models and these accuracies are ultimately used to determine the final accuracy of the model [9]. In this study we use 10-fold cross validation.

c) Recall and precision: Precision and recall are the other measures of accuracy that are widely used to evaluate the accuracy of crop recommendation systems by identifying the number of labels that are classified correctly compared to the actual class. The proportion of important labels that are in the top-k set is known as precision at k (nearest neighbours). The percentage of important labels that are in the list of recommendations out of total important labels is the recall at k [9].

IV. Results and Discussion

In this section we highlight analyse discuss the results we obtained using the methodology and data discussed in the preceding section. Table II shows the results of RF and SVM based CR systems using the 10-fold cross-validation test. We have included the results of other machine learning algorithms for comparison purposes, and we can see that neither the SVM algorithm nor the RF algorithm are the best algorithm for soil classification. The best performing machine learning algorithm is the K-Star algorithm whose accuracy is 91.3%,

Variable	Precision	Recall	RMSE	MAE	RAE
1. RF model	0.911	0.911	0.1044	0.0197	10.9522 %
2. SVM model	0.886	0.884	0.2731	0.1605	89.1778 %
3. K-NN model	0.910	0.910	0.1335	0.0198	10.989 %
4. K-Star model	0.913	0.913	0.1037	0.024	13.309 %
5. NB model	0.906	0.906	0.1028	0.0213	11.8521 %

Table II: Evaluation of the RF and SVM models

which is significantly greater than the 88.6% accuracy of SVM algorithm.

However, it is very close to the 91.1% accuracy of RF algorithm which means that we can conclude that the RF algorithm performs just as well as the best performing algorithm. Therefore, RF algorithm is well suited for soil classification as suggested by the literature, and as such we can use it for the development of our CR system. Using precision and recall as our evaluation criteria we can see that the RF algorithm has a recall of 0.911 and a precision value of 0.911. On the other hand, SVM has a recall value of 0.884 and a precision value of 0.886.

The general theme across all evaluation parameters including RMSE and MAE is that RF based models yield more favourable results compared to their SVM counterparts. In particular, when using RAE as the evaluation parameter, we see that the RF model (RAE value of 10.9522%) outperforms the SVM model (RAE value of 89.1778%) by a significant margin.

From the two confusion matrices above (i.e Figure 2 and Figure 1) we can see that the RF based model yields an accuracy of 91.1% whereas the SVM based model yields an accuracy of 88.4%. Our results are in agreement with those of [3] who found that RF models are better than SVM models when it comes to soil classification.

Figure 1: Support vector machines (SVM) - 88.4% Accuracy, Correctly identified 884/1000 labels

Predicted

Actual	a = Maize	b = Kidney beans	c = Banana	d = Mango	e = Grapes	f = Watermelon	g = Apple	h = Orange	i = Papaya	j = Cotton
a	98	0	0	0	0	0	0	0	0	2
b	0	100	0	0	0	0	0	0	0	0
c	0	0	100	0	0	0	0	0	0	0
d	0	0	0	93	0	0	0	7	0	0
e	0	0	0	0	78	0	22	0	0	0
f	0	0	0	0	0	100	0	0	0	0
g	0	0	0	0	73	0	27	0	0	0
h	0	0	0	8	0	0	0	92	0	0
i	0	0	0	0	0	0	0	0	100	0
j	4	0	0	0	0	0	0	0	0	96

Figure 2: Random forest (RF) - 91.1% Accuracy, Correctly identified 911/1000 labels

Predicted

Actual	a = Maize	b = Kidney beans	c = Banana	d = Mango	e = Grapes	f = Watermelon	g = Apple	h = Orange	i = Papaya	j = Cotton
a	99	0	0	0	0	0	0	0	0	1
b	0	100	0	0	0	0	0	0	0	0
c	0	0	100	0	0	0	0	0	0	0
d	0	0	0	100	0	0	0	0	0	0
e	0	0	0	0	58	0	42	0	0	0
f	0	0	0	0	0	100	0	0	0	0
g	0	0	0	0	46	0	54	0	0	0
h	0	0	0	0	0	0	0	100	0	0
i	0	0	0	0	0	0	0	0	100	0
j	0	0	0	0	0	0	0	0	0	100

The main purpose of our CR system is to classify soil based on its attributes and determine which crops are suitable for that particular soil. By using the RF algorithm, we were able to develop a model with an accuracy of 91.1% that only considers the three main soil nutrients, namely, nitrogen, phosphorus, and potassium; as well as the pH level of the soil.

The RF algorithm-based CR system we have developed is of high accuracy relative to previous implementations of such a CR system, for example [3]. Especially considering that we only considered the three main soil nutrients of the soil as well as the pH level in the development of our CR system. Incorporating more information about the soil or other environmental factors that affect the soil could increase the accuracy of our system, but it would also make our system more complicated and less accessible.

An example of a model that uses more features and yields better accuracy is that of [15] which considers all the features used in our model as well the following chemical attributes: zinc, sulphur, boron, calcium, magnesium, iron, copper, manganese, and organic matter %. This model yields a higher accuracy of 94.95% when comparing it to our model whose accuracy is 91.1%. However, we can conclude that our model is more efficient because it yields a comparable accuracy regardless of having used fewer features. This means that the features used in this study are better suited for the development of a CR system. We can also conclude that our model is more accessible to farmers because the data for the features we use for our model is more readily available.

V. Conclusion

In this study we were able to develop a crop recommendation system that can determine which crop to grow based on the attributes of the soil. We employed RF algorithm and SVM for the development of our CR system and found that RF algorithm works best for soil classification which is the basis of our CR system. Our final system uses RF algorithm and yields an accuracy of 91.1 % as shown in the discussion of our results.

The soil attributes that we considered were the three main soil nutrients, namely, nitrogen, phosphorus, and potassium, as well as the soil pH level. Choosing to only consider these four soil attributes whose data is more readily available and inexpensive to attain, allows our system to be more accessible to most farmers. Therefore, our developed CR system would have a greater impact on the farming industry, the agricultural sector, and hence, the economy as a whole. This is especially the case for South Africa and other countries similar to it, whose agricultural sector plays a major role in the economy.

Acknowledgment

This work is based on the research supported in part by the National Research Foundation of South Africa (Grant numbers: 121835). The crop recommendation system developed in this study can be adopted by farms of all sizes across South Africa because it is easily accessible and not resource intensive, and if widely adopted, this system would have a great impact on the economic output of the agricultural sector.

References

[1] S. Babu, "A software model for precision agriculture for small and marginal farmers," in *2013 IEEE Global Humanitarian Technology Conference: South Asia Satellite (GHTC-SAS)*. IEEE, 2013, pp. 352–355.

[2] S. N. Seo and R. Mendelsohn, "An analysis of crop choice: Adapting to climate change in south american farms," *Ecological economics*, vol. 67, no. 1, pp. 109–116, 2008.

[3] D. A. Bondre and S. Mahagaonkar, "Prediction of crop yield and fertilizer recommendation using machine learning algorithms," *International Journal of Engineering Applied Sciences and Technology*, vol. 4, no. 5, pp. 371–376, 2019.

[4] P. Bandara, T. Weerasooriya, T. Ruchirawya, W. Nanayakkara, M. Dimantha, and M. Pabasara, "Crop recommendation system," *International Journal of Computer Applications*, vol. 975, p. 8887.

[5] M. Shinde, K. Ekbote, S. Ghorpade, S. Pawar, and S. Mone, "Crop recommendation and fertilizer purchase system," *International Journal of Computer Science and Information Technologies*, vol. 7, no. 2, pp. 665–667, 2016.

[6] A. Kumar, S. Sarkar, and C. Pradhan, "Recommendation system for crop identification and pest control technique in agriculture," in *2019 International Conference on Communication and Signal Processing (ICCSP)*. IEEE, 2019, pp. 0185–0189.

[7] R. K. Rajak, A. Pawar, M. Pendke, P. Shinde, S. Rathod, and A. Devare, "Crop recommendation system to maximize crop yield using machine learning technique," *International Research Journal of Engineering and Technology*, vol. 4, no. 12, pp. 950–953, 2017.

[8] G. Suresh, A. S. Kumar, S. Lekashri, and R. Manikandan, "Efficient crop yield recommendation system using machine learning for digital farming," *International Journal of Modern Agriculture*, vol. 10, no. 1, pp. 906–914, 2021.

[9] K. Patel and H. B. Patel, "A state-of-the-art survey on recommendation system and prospective extensions," *Computers and Electronics in Agriculture*, vol. 178, p. 105779, 2020.

[10] K. K. Fletcher, "A method for dealing with data sparsity and cold-start limitations in service recommendation using personalized preferences," in *2017 IEEE international conference on cognitive computing (ICCC)*. IEEE, 2017, pp. 72–79.

[11] A. Sang and S. K. Vishwakarma, "A ranking based recommender system for cold start & data sparsity problem," in *2017 Tenth International Conference on Contemporary Computing (IC3)*. IEEE, 2017, pp. 1–3.

[12] N. Kumbhar and K. Belerao, "Microblogging reviews based cross-lingual sentimental classification for cold-start product recommendation," in *2017 International Conference on Computing, Communication, Control and Automation (ICCUBEA)*. IEEE, 2017, pp. 1–4.

[13] J. Wei, J. He, K. Chen, Y. Zhou, and Z. Tang, "Collaborative filtering and deep learning based recommendation system for cold start items," *Expert Systems with Applications*, vol. 69, pp. 29–39, 2017.

[14] K. G. Liakos, P. Busato, D. Moshou, S. Pearson, and D. Bochtis, "Machine learning in agriculture: A review," *Sensors*, vol. 18, no. 8, p. 2674, 2018.

[15] S. A. Z. Rahman, K. C. Mitra, and S. M. Islam, "Soil classification using machine learning methods and crop suggestion based on soil series," in *2018 21st International Conference of Computer and Information Technology (ICCIT)*. IEEE, 2018, pp. 1–4.

[16] M. Paul, S. K. Vishwakarma, and A. Verma, "Analysis of soil behaviour and prediction of crop yield using data mining approach," in *2015 International Conference on Computational Intelligence and Communication Networks (CICN)*. IEEE, 2015, pp. 766–771.

[17] A. L. Samuel, "Some studies in machine learning using the game of checkers," *IBM Journal of research and development*, vol. 3, no. 3, pp. 210–229, 1959.

[18] S. Pudumalar, E. Ramanujam, R. H. Rajashree, C. Kavya, T. Kiruthika, and J. Nisha, "Crop recommendation system for precision agriculture," in *2016 Eighth International Conference on Advanced Computing (ICoAC)*. IEEE, 2017, pp. 32–36.

[19] J. Lacasta, F. J. Lopez-Pellicer, B. Espejo-García, J. Nogueras-Iso, and F. J. Zarazaga-Soria, "Agricultural recommendation system for crop protection," *Computers and Electronics in Agriculture*, vol. 152, pp. 82–89, 2018.

[20] K. Akshatha and K. Shreedhara, "Implementation of machine learning algorithms for crop recommendation using precision agriculture," *International Journal of Research in Engineering, Science and Management (IJRESM)*, vol. 1, no. 6, pp. 58–60, 2018.

[21] R. Ghadge, J. Kulkarni, P. More, S. Nene, and R. Priya, "Prediction of crop yield using machine learning," *Int. Res. J. Eng. Technol.(IRJET)*, vol. 5, 2018.

[22] A. Chougule, V. K. Jha, and D. Mukhopadhyay, "Crop suitability and fertilizers recommendation using data mining techniques," in *Progress in Advanced Computing and Intelligent Engineering*. Springer, 2019, pp. 205–213.