Automation in Agriculture: A Systematic Survey of Research Activities in Agriculture Decision Support Systems Using Machine Learning

Sushma Vispute and Madan Lal Saini

Abstract In this age of automation, Machine learning (ML) plays the main role in agriculture sector to suggest suitable advice, crop advice, which includes decisions of growing crops, and advice related to growing season for precision farming. This systematic literature review performs a review of 103 documents of different ML approaches to analyze the performance of algorithms and used features in the work of prediction of crop yield and decision support systems to solve agriculture problems. These 103 documents are retrieved from different electronic databases, for analysis. The paperwork presents methods, accuracy measures, and used agriculture parameters, to understand the existing work done by authors. According to analysis, most of the authors used N, P, and K values and type of soil, and most of the authors used classification techniques such as Support Vector Machine, Decision Trees, Regression techniques, Random Forest, and Naive Bayes algorithm; the most applied clustering algorithm in the existing work is K-means. As per the additional survey, the Convolution Neural Network (CNN) algorithm is used by most of the authors for image processing in their work. Also, survey shows that very few authors used associative classifiers and association rule mining techniques to solve the agriculture problems.

Keywords Associative classifier · Agriculture · Classification · Association rule mining · Clustering · Neural network · Decision support system · Machine learning · Convolution neural network

Madan Lal Saini: This author contributed equally to this work.

S. Vispute (✉) · M. L. Saini
Department of Computer Science and Engineering, Poornima University, Jaipur 303905, Rajasthan, India
e-mail: sushma.vispute@pccoepune.org

M. L. Saini
e-mail: madan.saini@poornima.edu.in

1   Introduction

Machine Learning techniques are used in many sectors, such as healthcare to predict the suitable treatment, supermarkets, manufacturing companies to analyze the customers' behavior like products used by customers. From several years Artificial Intelligence and ML techniques are also being applied in the agriculture-Farming sector to solve farmers' problems. Crop production is based on several parameters like seed type, climate, fertilizer used, weather and soil type, etc.

A problem with most of Indian farmers is, Lack of knowledge and Lack of proper assistance for precision Farming and so the objective of this Literature Review is to study and analyze existing Indian agriculture problems and solutions provided to these problems using Machine Learning techniques, to study and analyze different soil parameters which affect the agriculture production, to find the novel approach for proposed work.

The beauty of Machine Learning algorithm is to train the model using a training dataset and predict the class of new samples even though the new example is not completely matching with training samples. For example, where the training dataset contains CAT and DOG faces and predicted class for Tiger face as CAT.

There are three main categories of Machine Learning approaches. First, Supervised Learning includes learning from experience data, i.e., empirical data and its examples includes Classification, Regression (KNN, Decision Tree, and Linear Regression). Second Unsupervised Learning, i.e., Learning from observations given in the dataset, i.e., patterns in the dataset, its examples include, Clustering Techniques such as K-means, DBSCAN, third, Reinforcement Learning, i.e., learning from environment feedback in the form of penalty and rewards, for example, Deep Q Networks. Nowadays, deep learning algorithms are used for optimization of models because they attempt to learn by using a hierarchy of multiple layers [1].

2   Related Work

This work of literature review includes a survey of existing Indian agriculture problems and solutions provided to these problems using Machine Learning techniques, survey of different soil parameters which affect the agriculture production, and survey of different ML techniques to find the novel approach for proposed work.

Identified Research questions are:

Q1. Identify machine learning algorithm used for the Agriculture Support System.

Q2. Identify features used to design Agriculture Support System.

Q3. Identify model evaluation parameters and evaluation approaches used for the agriculture Support System.

Q4. Identify the Gaps in the field of Agriculture Support System.

2.1   Bibliography Analysis

Figure 1 shows bibliography analysis for distribution of papers used for the literature review.

Table 1 gives the count of documents referred for the survey on the basis of type of document and Table 2 gives the count of documents referred for the survey on the basis of publication year (Fig. 2).
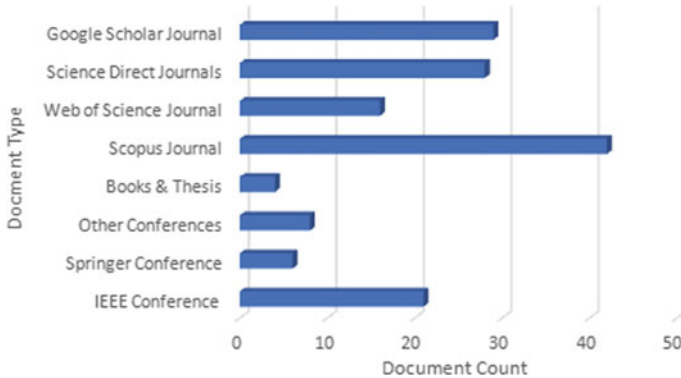

Fig. 1   Bibliography Analysis

Table 1 Number of papers referred on the basis of type	Document type	# Documents
	IEEE conference	21
	Springer Conference	06
	Web of science	16
	Books and thesis	04
	Science direct	28
	Scopus	42
	Google scholar	29
	Others	08

Table 2 Distribution of documents based on the publication year	Publication year	#Documents
	2020	20
	2019	21
	2018	19
	2017	11
	2016	09
	<2016	22

Fig. 2 Number of documents referred on the basis of publication year

Sirsat et al. developed 20 different classification models for Classifying Indian agricultural soil parameters. They developed Soil nutrients N, P, K Classification model, Soil pH Classification model, model for Classification of Crop, model for Soil classification by type. These Classification problems are studied and implemented for the Marathwada dataset using Bagging, Boosting, Decision Tree (DT), K-Nearest Neighbor (KNN), Rule-Based (RB), Neural Network (NN), Random Forest (RF), and Support Vector Machine (SVM) models. Cohen kappa (k) in % is used by authors for measuring accuracy of these models. Model results are discussed below [2].

- Best k for Decision Tree Soil classifier using Weka tool is 97.82%,
- Best k for Random Forest Crop classifier using R language is 88.13%,
- Best k for Random Forest pH classifier using R language is 47.32%,
- Best k for Random Forest NPK classifier using R language is 33.6%,
- Best k for Random Forest OC-F classifier using Weka tool is 90.65%.

Sirsat et al. proposed 76 different models to predict soil fertility based on nutrients values of organic carbon (OC), phosphorus pentoxide (P_2O_5), Zn-Zinc, Fe-iron, and manganese (Mn) using different Regression Techniques such as Linear Regression (LR), Generalized Linear Regression (GLR), Least Square (LS), Partial Least Square (PLS), LASSO, Ridge, Neural Network, Deep Learning, SVM, Random Tree [3].

R2 accuracy measure is used by authors to find the best Regressor. Authors concluded following results of proposed models [3] (Table 3, Fig. 3).

ZhaoyuZhai et al. represented a survey and challenges of agriculture (4.0) decision support systems (DSS). They did systematic survey of 13 representative DSS including their applications for planning missions, management of water resources, for controlling food waste, etc. [4].

Alexandre Barbosa et al. proposed a model for optimizing nutrient management for predicting crop yield response using CNN. Authors developed CNN with Early Fusion (EF), CNN with Late Fusion (LF), and 3D CNN and compared results with Multiple Linear Regression (MLR), Full Connected Network, Support Vector Network, and Random Forest models. Root Mean Square Error (RMSE) measure is used by authors to measure accuracy of CNN model. Results shows, CNN-LF with lowest error for nine tested fields (0.66), and CNN-RF with second best result (0.76) [5].

Suchithra et al. proposed a model for proper fertilizer utilization, to reduce the analysis time experts, and to improve quality of soil. In this work accuracy measures

Table 3 Comparative analysis of Indian agricultural soil parameters models accuracy

Model	Accuracy measure	Accuracy in %
Random forest pH classifier	Cohen Kappa	47.32
Random forest NPK classifier	Cohen Kappa	33.6
Random forest Crop classifier	Cohen Kappa	88.13
Random forest OC-F classifier	Cohen Kappa	90.65
Decision tree soil classifier	Cohen Kappa	97.82
Random forest—Boruta FS regressor	R^2	69.8
Organic carbon (OC) extra trees regressor	R^2	69
Phosphorus pentoxide (P$_2$O$_5$) extra trees regressor	R^2	60.3
Iron (Fe) extra trees regressor	R^2	66.6
Manganese (Mn) extra trees regressor	R^2	57.5
Zinc (Zn) extra trees	R^2	70.7
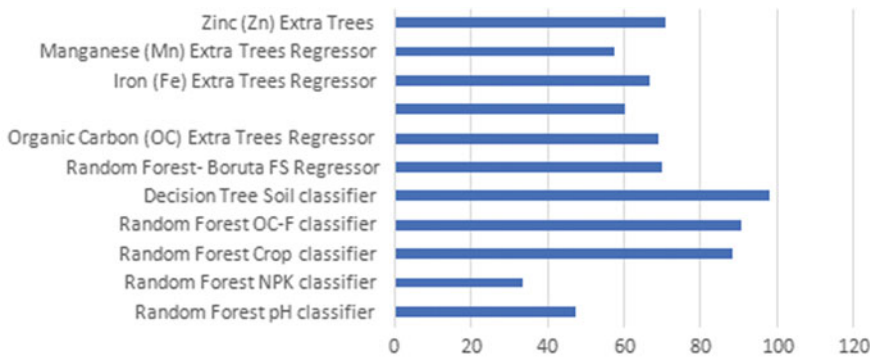
	Random Forest pH classifier	Random Forest NPK classifier	Random Forest Crop classifier	Random Forest OC-F classifier	Decision Tree Soil classifier	Random Forest-Boruta FS Regressor	Organic Carbon (OC) Extra Trees Regressor	Phosphorus pentoxide (P2O5) Extra Trees Regressor	Iron (Fe) Extra Trees Regressor	Manganese (Mn) Extra Trees Regressor	Zinc (Zn) Extra Trees
■	47.32	33.6	88.13	90.65	97.82	69.8	69	60.3	66.6	57.5	70.7

Fig. 3 Indian agricultural soil parameters models accuracy with data chart

used are Accuracy, Kappa, Precision, Recall, FScore, and results given by models are as follows [6].

- Soil Nutrient Classification for Gaussian radial basis function: 80% (Optimal neurons 50),
- pH classification for hyperbolic tangent function: 90% (Optimal neurons 150).

Himanshu Pant et al. proposed a model to enhance the precision of crop-fertility prediction using different supervised ML techniques. K-Means is used to identify quality and fertility of the Soil with levels 1, 2, and 3 for Nainital District dataset. Accuracy measures used for Classification problems are Precision, Recall, F1 Score, Support, Accuracy, and results are as below [7].

- SVM with 96.62% Accuracy (Best classifier among all),
- KNN with 91.01% Accuracy
- LR with 89.88% Accuracy
- LDA with 91.01% Accuracy.

Santhi et al. proposed a model to compare the categories of Farming and types of crops using crop and fertilizer recommendation methods based on soil test reports [8].

Manpriya et al. proposed a model for effective crop prediction technique for better crop production with more crop datasets. Deep NN with two hidden layers is used to predict appropriate crops for every district of India. 124 crops are included in the work. Performance parameters used by authors are Accuracy, Mean Absolute Error (MAE), and MSE. Sigmoid as activation function (SGD optimizer) is used for updating parameters and weights to reduce the loss function. Values of performance parameters are Accuracy with 99.19%, MAE with 0.0157, and MSE with 0.0078 [9].

Deshmukh et al. proposed a model for Soil Health Analysis and Soil quality prediction with N, P, and K Soil parameters. Results for soil quality prediction models and crop prediction models are shown in figure. CN2 Rule Inducer with accuracy of 0.94 declared as Best Classifier. Figure 4 shows performance comparison of Soil Quality and Crop Advice Prediction using different classifiers [10].

Labhade et al. developed a model to predict the outcomes based on the selected data and business requirements. Predictive Analytics is done using KNIME Tool and its results are as follows. Figure 5 shows Accuracy and Error rate for different classifiers using KNIME tool. As per analysis Logistic Regression method gives best accuracy for student datasets [11].

Viviliya et al. developed Hybrid model of J48 and Naive Bayes classifiers for recommending crops using ML techniques, to increase crop yield. Models are developed using dataset of parameters State, District, Crop year, Area, etc. and yield info from 1997 to 2015, Season, Temperature, Rainfall, Water requirement, and type of soil. J48 has given best accuracy 95.53% [12].

Devdatta et al. implemented a model of crop yield prediction using historical data by using machine learning on agriculture dataset and recommending fertilizers suitable for crop. Classification models using SVM and RF are built and authors used
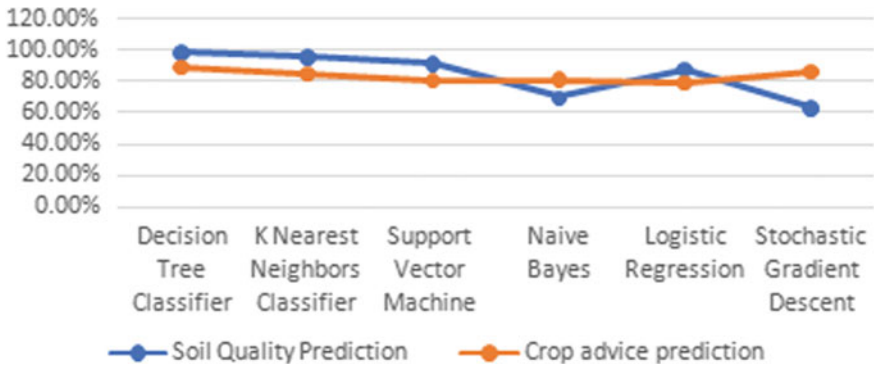
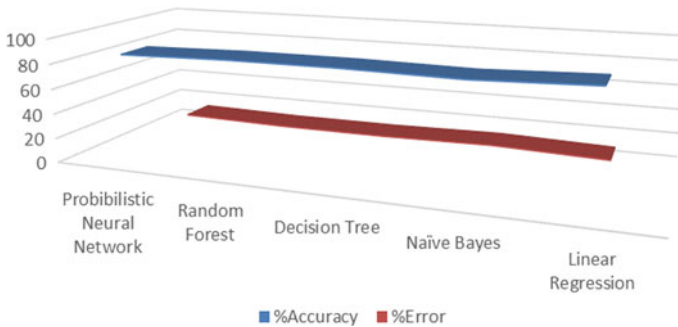**Fig. 4** Soil quality and crop advice prediction

Fig. 5 Accuracy and error rate for different classifiers using KNIME tool

Precision, Recall, f1-score, and accuracy in % performance measuring parameters and discussed the results are as below [13].

- Soil Classification model using RF with accuracy of 86.35% and SVM with 73.75%.
- Crop Yield Prediction model using SVM with 99.47% accuracy and RF with 97.48% accuracy.

Rafael Hernández Moreno et al. presented a Multi-Layer Perceptron (MLP) ANN model with an input layer formed by soil parameters, an output layer with fertilizers and amendments. A GridSearchCV is used to test and optimize the model [14].

Archana et al. proposed a DSS model using Voting Based Ensemble Classifier. Voting based ensemble classifier for Crop recommendation (Random Forest Classifier, Naive Bayes Classifier, and CHAID Classifier) with input parameters, N, P, K, Temperature, and other soil parameter is built and got the 92% accuracy [15].

Rajak et al. developed a model for crop prediction using Ensemble technique (Majority Voting technique). In Ensemble technique different selected algorithms are SVM, Random Forest, NAÏVE Bayes, ANN- Multi-layer Perceptron [16].

Devotha et al. presented a review for survey of use of Characterization techniques in agriculture sector. They applied probabilistic and deterministic approaches, where the supervised algorithms are used in deterministic approaches, while the unsupervised algorithms are used in probabilistic approaches [17].

Srivastava et al. presented survey paper to electorate on different Clustering Techniques such as DBSCAN, Agglomerative, K-means, EM algorithms for Agriculture applications to bring a good advancement in the agricultural area for Forecasting Pollution, Combined Classification of Soil with GPS [18].

Bouighoulouden et al. proposed a model using PCA for reduction of the features and K-means implemented on Rstudio, Orange DM tools to identify groups of productive and non-productive yield [19].

Dr. Madhavi Gudavalli et al. applied Clustering on Wheat seed dataset using different clustering techniques. 3 clusters are formed Kama, Rosa, Canadian with pair of attributes using R tool, authors reported that k-mean is good for large datasets and Hierarchical is good for small datasets [20]

Priya et al. built a model for depiction of management zones and soil dataset analysis using K-means, GK clustering, and Farthest First (obtained Best-faster) Algorithms [21].

Utkarsha et al. developed Modified K-Means Algorithm and used it for crop prediction. District, zone, and selection of seasons, max temperature, min temperature, soil type, and average rainfall are considered for training the model. Work shows comparison of k-Means++ and k-Means with modified k-Means on Crop data. Modified k-Means gave the maximum quality clusters, maximum accuracy count, and correct prediction of crop [22].

Silas et al. used Association Rule Mining and Clustering Techniques for Tea Production prediction in Kenya country. Dataset contains 156 tea production records from year 2003 to 2015. Clustering techniques are used to form the groups of similar productions using (SPSS) K-Means [23].

Majumdar et al. presented analysis using different ML techniques such as Multiple LR, CLARA, PAM, and Modified DBSCAN to identify optimal parameters to maximize crop production. Modified DBSCAN was declared as a Best to cluster the data having similar rainfall, temperature, and soil type [24].

Vandana et al. proposed model for crop production and US arrest dataset analysis. Techniques used are Hybrid K- means which declared as a Best. Elbow, Gap Statistic, Silhouette Methods are used to select optimal "K" value [25].

Aurelia-Vasilicalana et al. used clustering methods for Organic farming patterns analysis. Work identified three possible clusters using clustering methods [26].

Chunjiang et al. built a model using Frequent Pattern Tree for mining association rules with multiple inputs of minimum supports (MSDMFIA). It overcomes the problem of single minimum support used in tradition method [27].

Geetha et al. used Apriori algorithm to assess different association algorithms and used them into a soil science database to identify meaningful relationships [28].

Kane et al. proposed model for Classification of home loan sales in an Irish retail banking using Association Rules. Associative classifier models used are CMAR, Classification Based Association (CBA), and SPARCCC [29].

Vasoya et al. proposed distributed model based on distributed and parallel computing for large dataset association rule mining to find frequent patterns in less time. Clustering process is used to divide large data into number of clusters and these clustered data are used for mining process [30].

Thakkar et al. used Association rule mining algorithms like Apriori and classification techniques like ID3 and C4.5, to solve agriculture crops problems [31]. Khan and Singh [32], presented survey of Association Rule Mining methods for agriculture problems. Survey represents techniques used to solve problems using the Partition Algorithm, Apriori, Pincer search Algorithm, FP-Tree Growth Algorithm, Dynamic Itemset Counting Algorithm [32].

Mishra et al. presented survey of Associative Classifiers (CBA, CMAR, MCAR, and GARC) used on Soil dataset of Bhopal M.P District [33]. Sun et al. [34] presented an Overview of Associative Classifiers. The conventional classification system such as C4.5 is compared with associative classifier. Total 27 UCI datasets are used for comparison [34].

Prachitee et al. proposed a model of Classification Technique using Associative Classifier based on the Neural Network system (NNAC) to improve its accuracy. NNAC system performance is compared with the Classification Based Association on four different datasets from UCI repository [35].

Soni et al. proposed solution for Health Care domain using Associative Classifiers to predict the disease with some suitable treatments. Authors used class rule mining—Associative Classification (AC), classification Association rule (CAR) techniques [36]. Classifier to assist the physician to find association among patient parameters (e.g., personal data, medical tests,) have also been developed, and advanced association rule mining with classifiers are used to develop models of an AC based on positive and negative rules, Temporal AC, AC using Fuzzy Association Rule, Weighted AC [36].

Jinubala et al. proposed a model to classify Pest Level based on whether data using Constraint-based AC (Accuracy 92%) and Traditional method accuracy of 59% [37]. Mattieva and Kavšeka [38], proposed a model using associative classification techniques such as AC based on strong association rules. Average accuracy given by model is 91.3%. Experiments are done on 15 UCI ML D/B repositories [38].

Li Yu Hu et al. work presented Novel CBA-based method: MMSCBA, (multiple minimum supports (MMSs)) [39]. Dalvi et al. [40], proposed a Ontology-based model for agricultural (IR) using NLP, to extract knowledge in Marathi language [40].

Pai et al. presented ML models for Identification of Kannada Farmer's Query using a speech recognition system for agricultural dataset in Kannada language. The dataset consists of the name of the crops and name of the districts of Karnataka state. MFCC is the most prominent feature extraction method used in speech recognition. MFCC for CROP, District Data [41].

Savant et al. presented survey of existing system of Maharashtra Government, Survey of clustering Techniques, and Classification of farmer's feedback [42]. Vispute et al. [43] proposed a model for automatic personalized Marathi content generation in Marathi language using LINGO algorithm. Work has experimented on

five different datasets and personalization is done using "Time Session", "Number of hits" and Bookmark methods [43].

Vispute et al. extended previous work using HADOOP parallel system platform for Marathi dataset [44]. Vispute et al. [44], developed a model for categorizing Marathi text documents automatically for dataset of three categories- Health Programs, Tourism, and Maharashtra festivals using Lingo Clustering algorithm. Dataset contains 107 Marathi documents [45].

Sonigara et al. built a model for effective information retrieval system to input the data in heterogeneous forms and represent it into a common format, i.e., a text file, and categorizing Marathi data automatically using LINGO algorithm [46].

Tayal and Meena developed parallel system solution using the MapReduce approach on HADOOP platform for associative classification and experimented on six datasets available on UCI repositories. To provide solution to problems they developed two algorithms MRMCAR-F and MRMCAR-L [47]. Figure 6 shows accuracy comparison of proposed association classification techniques by Devendra et al. for six different datasets of UCI data repository.

Figure 7 shows comparison of time required to execute different associative classification techniques proposed by Tayal and Meena [47].

Dang Nguyen et al. proposed an efficient constraint-based CARs model with the item set constraint. To test the performance of novel model authors used 14 different datasets like Adult, Breast, German, Chess, Connect4, etc. Figure 8 shows proposed models for adult dataset [48].

Wang et al. proposed an improved model using dynamic property in the associative classification [49]. Villuendas-Rey et al. [50] used and evaluated the Naïve Associative Classifier on financial dataset for simple, transparent, and accurate classification [50].
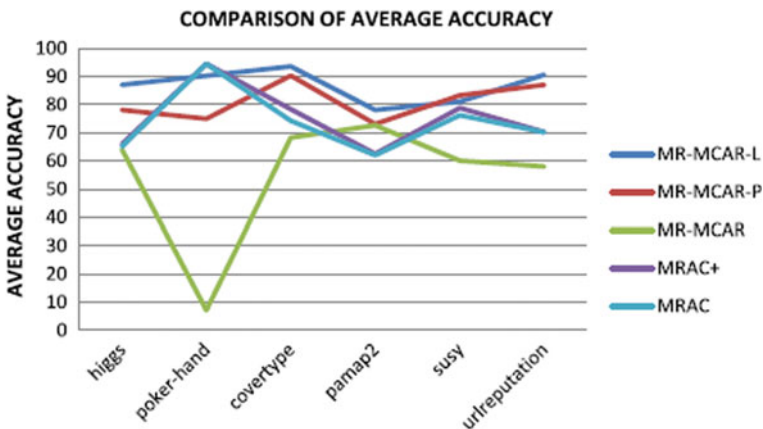

Fig. 6 Comparison of proposed association classification techniques for six different datasets of UCI data repository [47]
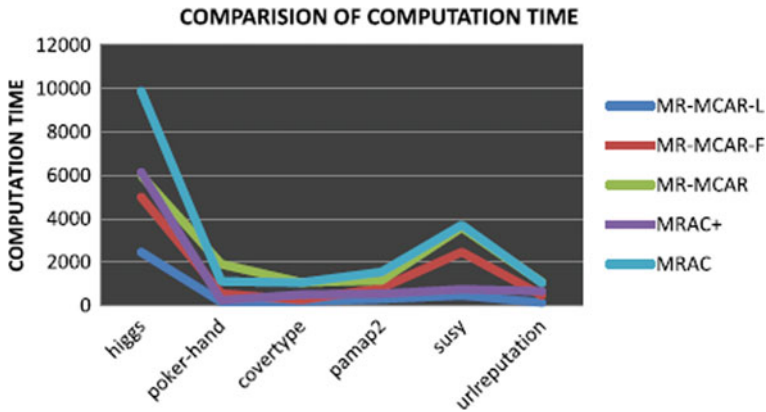
Fig. 7 Comparison of execution time with the proposed associative classification techniques [47]
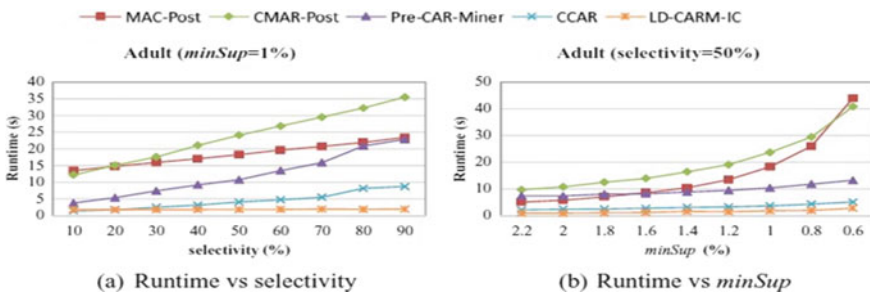

Fig. 8 Comparison of execution time required for different proposed models for adult dataset [48]

Figure 9 shows the overall AUC results of NAC, compared with other classifiers. It shows that NAC outperforms as compared to other algorithms.

Chen et al. proposed an efficient classification approach, Principal Association Mining to design a compact classifier for generating reduced association rules [51].
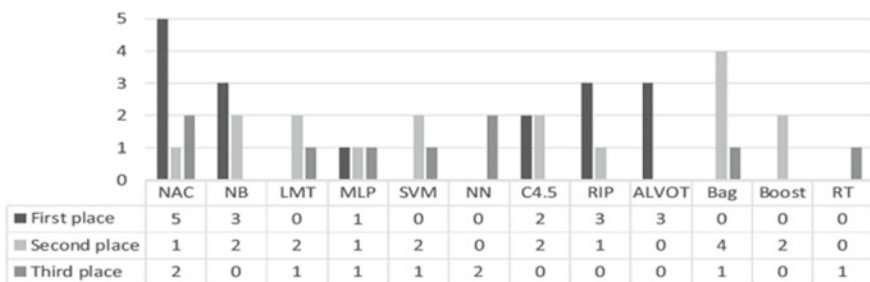

Fig. 9 AUC results of NAC, compared with other classifiers [50]

Padillo et al. [52] introduced a new Library of JAVA language for Associative classification, i.e., LAC. This library package includes the full taxonomy of associative classification paradigm [52]. Loan et al. [53], developed a new model for extracting class-association rules [53]. Antonell et al. [54] used fuzzy-frequent pattern mining algorithm to proposed a novel classification model. Authors tested the new approach on 17 datasets and represented comparative analysis. New model gave better results than existing [54]. Hadi et al. [55] proposed efficient model for text classification which combines features of Naïve Bayes and associative classifiers [55]. Thasleena et al. [56] developed an efficient classifier for XML documents using associative classifier to overcome the drawback of the existing technology [56]. Mattieva et al. [57], proposed simple classification with "strong" class-association rules to improve the classifier performance with good accuracy [57].

Nguyena et al. and Wang et al. proposed hybrid and an efficient method to solve problem using associative classification techniques [58, 59]. Villuendas-Reya et al. proposed new model NAC, based on Associative classifier, and tested and evaluated model on financial dataset [60].

In the next literature survey of agricultural decision support systems for precision farming we compared different ML and Deep Learning algorithms and explored possible uses of these algorithms to solve multiple problems related to farming.

Many algorithms like SVM, Random Forest, and CNN were used to detect plant diseases. The result shows that CNN detects a greater number of diseases of plants with high accuracy [61–65].

In scenarios where there is huge difference between size or color difference between crop and weed, image processing-based algorithm works well. Survey tells that CNN performs better than the SVM and ANN due of its ability of learning in depth to learn related features from the image dataset. ANN is very accurate but requires huge amounts of training data and is slower [66–70].

For weather forecasting research shows that different models such as ANN, CNN, and Recurrent NN can be used. Out of these models, Long Short-Term Memory LSTM (type of RNN) works exceptionally well for sequential data of weather prediction [71–75].

Many algorithms like ARIMA, SARIMA, and RNN algorithms such as LSTM and Gated Recurrent Units (GRU) can be used to predict agricultural prices. The results show that in general LSTM models perform better than others with higher data while ARIMA and SARIMA can perform reasonably well even with less data [76–80].

The next survey of work shows, solution to a variety of problems like prediction of soil fertility level, disease detection, prediction of yield based on weather conditions, identifying correct action during farming in different situations, etc. [81–85].

3   Common Findings from Literature Review

3.1   Results and Common Methodology

Most commonly used methodology by authors is given in the below Fig. 10. It includes following basic steps to develop a model for solving problems.

Table 4 shows most used machine learning algorithms in the existing work with efficient model details. (Answer of question1).

In the review of existing work, it is found that Agriculture decision support systems are developed to provide decisions about single areas of farming such as recommendations for Crop yield prediction, recommendation for fertilizer, etc. by using different machine learning techniques mentioned in Table 4.

In the survey, it is found that very little work is done to provide solutions to the agriculture problems using associative classifiers, only three paperwork shows solutions to agriculture problems using associative classifiers. This existing work only provides a single decision to farmers at a time by considering different agriculture parameters such as N, P, K, Ph, Crop year, Rainfall, etc. So, the more effective agriculture decision support system needs for precision farming.
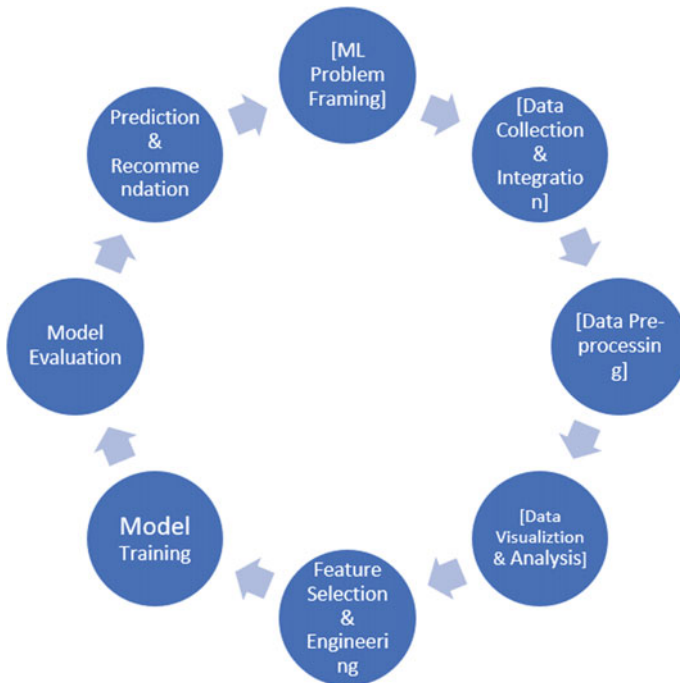

Fig. 10   Common flow diagram for Developing ML Model

Table 4 Algorithms used by most of the existing work with best model details

Algorithms	Frequency count	Efficient model details
Classification: RT, RF, NB, DT, SVM	29	Random forest is better in most of the work
Neural network: CNN, DNN	27	CNN with Relu and optimized parameters
Clustering: K-means	07	K-Means by most of the work for efficient analysis
Clustering: agglomerative, DBSCAN	02	
ARM: Apriori	04	Apriori for appropriate results
ARM: FP-growth	01	
Associative classifier	17 (only 2 for Agri problems)	Hybrid associative classifier

Features used in most of the work (Answer to question 2) are Sulfur, Magnesium, potassium, zinc, nitrogen, calcium, boron, and Phosphorus, pH-value State, District, Crop year, Season, Area, Production and yield details, Rainfall details, Temperature details, Groundwater level, Water availability, type of soil, Organic carbon (OC).

Most used evaluation parameters (Answer of question 3) are Accuracy (36 times), Kappa (8), Precision (27), Recall (27), FScore (24), RMSE(6), R^2(5), WCSS (9), Support and confidence(5).

4 Conclusion

This systematic literature review showed that the work in the referred documents those used a several features, depending on the research type and requirements with the selected dataset. Most of the work is done for prediction of yield and applied machine learning algorithms but on different features. Also, work is done for plant disease prediction, weed detection. Selected features are dependent on the objectives of the research. The best model can be identified by testing models with more features and fewer features and also models with different ML techniques. According to survey study and analysis, most of the authors used rainfall, temperature, and type of soil, and most preferred classification algorithms are Neural Networks, Regression techniques, SVM, Random Forest, andRandom Forest worked better in most of the work. The most applied clustering algorithm is K-means for finding efficient solution to problem.

As per the additional survey, Convolution Neural Networks (CNN) with optimized parameters is used by most of the authors for image processing and classification and then another widely used DL algorithm is Deep Neural Networks (DNN). Also, survey shows that very few authors (only 2) used associative classifiers and association rule mining techniques to solve the agriculture problems.

References

1. Han J, Kamber M (2011) Data mining: concepts and techniques. Morgan Kaufmann, 3rd edn. A volume in The Morgan Kaufmann Series in Data Management Systems

2. Sirsat MS, Cernadas E, Fern´andez-Delgado M, Khan R (2017) Classification of agricultural soil parameters in India. Comp Elect Agri 135:269–279

3. Sirsat EC, Fern´andez-Delgado S, Barro S (2018) Automatic prediction of village-wise soil fertility for several nutrients in India using a wide range of regression methods. Comp Elect Agri 154:120–133

4. ZhaoyuZhai J, Martínez J-F, Beltran V, Martínez NL (2020) Decision support systems for agriculture 4.0: Survey and challenges. Comp Elect Agri Science Direct 170:105256

5. Barbosa A, Trevisan R, Hovakimyan N, Martin NF (2020) Modeling yield response to crop management using convolutional neural networks. Comp Elect Agri Sci Direct 170:105197

6. Suchithra MS, Pai ML (2019) Improving the prediction accuracy of soil nutrient classification by optimizing extreme learning machine parameters. Sci Direct, Info Process Agri 7(1):72–82

7. Pant H, Lohani MC, Bhatt A, Pant J, Joshi A (2020) Soil quality analysis and fertility assessment to improve the prediction accuracy using machine learning approach. Int J Adv Sci Tech 29(3):10032–10043

8. Santhi P, Priyanka T (2020) Smart India agricultural Info retrieval system. Int J Adv. Sci Tech 29(7):1169–1175

9. Manpriya D, Jindal V (2020) Crop prediction using deep neural network. Int J Mech Product Eng Res Develop 3:2249–6890

10. Deshmukh S, Dhannawat D, Dalvi M, Gawali P, Vispute SR, Kekane S (2019) Application of data analytics in agriculture sector for soil health analysis: Literature review. In: 5th International Conference on Computing, Communication, Control and Automation (ICCUBEA-2019), pp 1–4. https://doi.org/10.1109/ICCUBEA47591.2019.9129104

11. Labhade D, Lakare N, Mohite A, Bhavsar S, Vispute S, Mahajan G (2019) An overview of machine learning techniques and tools for predictive analytics. Asian J Conv Tech 5(3):63–66

12. Viviliya B, Vaidhehi V (2019) The design of hybrid crop recommendation system using machine learning algorithms. Int J Innov Tech Exploring Eng 9(2):4305–4311

13. Devdatta AB, Mahagaonkar S (2019) Prediction of crop yield and fertilizer recommendation using machine learning algorithms. Int J Eng Appl Sci Tech 4(5):371–376

14. Hernández Moreno R, Garcia O, Luis Alejandro Arias R (2018) Model of neural networks for fertilizer recommendation and amendments in pasture crops. In: 2018 IEEE, 978-1-5386-9459-6/18/$31.00.

15. Archana K, Saranya KG (2020) Crop yield prediction, forecasting and fertilizer recommendation using voting based ensemble classifier. Int J Comp Sci Eng 7(5):1–4

16. Rajak RK, Pawar A, Pendke M, Shinde P, Rathod S, Devare A (2017) Crop recommendation system to maximize crop yield using machine learning technique. Int Res J Eng Tech 4(12):950–953

17. Devotha G. Nyambo ETL, Yonah ZQ (2019) A review of characterization approaches for smallholder farmers: towards predictive farm typologies. Hindawi: Scient World J Wiley, 1–10

18. Srivastava V, Aggarwal KK, Srivastava AK (2019) A revisit to clustering techniques with its application in agriculture sector. HEB 3(1):1–1

19. Bouighoulouden A, Kissani I (2020) Crop yield prediction using K-means clustering, school of science and engineering. Al Akhawayn University, Spring

20. Gudavalli M, Vidyasree P, Viswanadha Raju S (2017) Clustering analysis for appropriate crop prediction using hierarchical, fuzzy C-means, k-means and model based techniques. Int J Adv Eng Res Develop 4(11):1233–1242

21. Krishna Priya CB, Venkateswari S (2018) Delineation of management zones in precision agriculture using different clustering algorithms. Int J Appl Eng Res 13(22):15951–15955

22. Utkarsha PN, Adhiya KP (2016) Evaluation of modified k-means clustering algorithm in crop prediction. Int J Adv Comp Res 4(16):799–807

23. Silas NM, Nderu L (2017) Prediction of tea production in Kenya using clustering and association rule mining techniques. American J Comp Sci Info Tech 5:1–7
24. Majumdar J, Naraseeyappa S, Ankalaki S (2017) Analysis of agriculture data using data mining techniques: application of big data. J Big Data 4(20):1–15
25. Vandana B, Sathish Kumar K (2019) Hybrid k mean clustering algorithm for crop production analysis in agriculture. Int J Inno Techn Explo Eng 9(2S):9–12
26. Aurelia-Vasilicalana ET, Dobrea C, Soarea E (2015) Organic farming patterns analysis based on clustering methods, Science Direct. Agri Agricultural Sci Procedia 6:639–646
27. Zhao Chunjiang W, Huarui SX, Baozhu Y (2010) Algorithm for mining association rules with multiple minimum supports based on FP-Tree. N Z J Agric Res 50(5):1375–1381. https://doi.org/10.1080/00288230709510425
28. Geetha MCS (2015) Implementation of association rule Mining for different soil types in agriculture. Int J Adv Res Comp Comm Eng 4(4):520–523
29. Kane C (2018) Classification using association rules. Technological University Dublin, ARROW@TU Dublin
30. Vasoya A, Koli N (2016) Mining of association rules on large database using distributed and parallel computing. In: 7th International Conference on Communication, Computing and Virtualization.Proccedia Computer Science, 79, 221–230
31. Rahul GT, Kayasth M, Desai H (2014) Rule based and association rule mining on agriculture dataset. IJRICCE
32. Khan F, Singh D (2014) Association rule mining in the field of agriculture: a survey. Int J Inno Res Comp Comm Eng 2(11):6381–6384
33. Mishra AK, Sharma P (2014) A review on associative classification data mining approach in agricultural soil land. Int J Modern Trends in Eng Res 1(4):65–69
34. Sun Y, Andrew KC, Wong F, Wang Y, Member IEEE (2006) An overview of associative classifiers. In: International Conference on Data Mining, DMIN, Las Vegas, Nevada, USA, 1–7.
35. Prachitee B. Sheetal S, Dhande S (2011) A classification technique using associative classification. Int J Comp Appl 20(5):20–28
36. Soni S, Vyas OP (2010) Using Associative Classifiers for Predictive Analysis in Health Care Data Mining. International Journal of Computer Applications, 2010, 4(5), 33–37.
37. Jinubala V, Raj L (2018) Mining pest level based on weather using associate classification. Pestology Wiley XLII(3):1–8
38. Mattieva J, Kavšeka B (2020) A compact and understandable associative classifier based on overall coverage. Procedia Computer Science170:1161–1167
39. Hu L-Y, Hu Y-H, Tsai C-F, Wang J-S, Huang M-W (2016) Building an associative classifier with multiple minimum supports. Springer Plus 5(528):1–19
40. Dalvi P, Mandave V, Gothkhindi M, Patil A, Kadam S, Pawar S (2016) Ontology extraction for agriculture domain in Marathi language using NLP techniques, ICTACT-J Soft Comp 7(1):1359–1365
41. Pai A, Hegde S (2019) Study on machine learning for identification of farmer's query in Kannada language. Int J Comp Appl (0975–8887) 178:40–47
42. Savant V, Shinde A, Yedle B, Pantawane S, Vispute SR, Pede SV (2015) A survey on farmer's need and feedback, IJIRCCE
43. Vispute SR, Kanthekar S, Kadam A, Kunte C, Kadam P (2014) Automatic personalized Marathi content generation. In: International Conference on Circuits, Systems, Communication and Information Technology Applications (CSCITA), pp 294–299. https://doi.org/10.1109/CSCITA.2014.6839275
44. Vispute SR, Patil S, Sangale S, Padwal A, Ukarde A (2015) Parallel processing system for Marathi content generation. International Conference on Computing Communication Control and Automation 2015:575–579. https://doi.org/10.1109/ICCUBEA.2015.118
45. Vispute SR, Potey MA (2015) Automatic text categorization of marathi documents using clustering technique. In: 15th International Conference on Advanced Computing Technologies (ICACT), pp 1–5, https://doi.org/10.1109/ICACT.2013.6710543

46. Prachi S, Kirti P, Pooja N, Alisha S, Sushma V (2018) Automatic integration and clustering of Marathi documents in different formats for effective information retrieval. In: Proceedings of International Conference on Recent Advancement on Computer and Communication, Lecture Notes in Networks and Systems. https://doi.org/10.1007/978-981-10-8198-9_36

47. Tayal DK, Meena K (2020) A new MapReduce solution for associative classification to handle scalability and skewness in vertical data structure. Future Generation Comp Syst 103:44–57

48. Nguyen D, Loan TT, Nguyen BV, Pedry W (2016) Efficient mining of class association rules with the itemset constraint. J Know Based Syst 103:73–88

49. Wang X, Yue K, JiaNiu W, Shi Z (2011) An approach for adaptive associative classification. Expert Syst Appl ScienceDirect 38:11873–11883

50. YennyVilluendas-Rey CF, Rey-Benguría Á-S, Camacho-Nieto O, Yáñez-Márquez C (2017) The Naïve Associative Classifier (NAC): a novel, simple, transparent, and accurate classification model evaluated on financial data. Neurocomputing J 265:105–115

51. Chen F, Wang Y, Li M, Harris W, Tian J (2014) Principal association mining: an efficient classification approach. Knowledge-Based Syst J 67:6–25

52. Padillo F, Luna JM, Ventura S (2019) LAC: library for associative classification. Knowledge-Based Systems J 193:105432

53. Loan TT, Nguyen B, Vo B, Hong T-P, Thanh HC (2012) Classification based on association rules: A lattice-based approach. J Expert Syst Appl 39:11357–11366

54. Antonelli M, Ducange P, Marcelloni F, Segatori A (2015) A novel associative classification model based on a fuzzy frequent pattern mining algorithm. J Expert Syst Appl 42:2086–2097

55. Hadi W, Qasem A, Al-Radaideh SA (2018) Integrating associative rule-based classification with Naïve Bayes for text classification. J Appl Soft Comp 69:344–356

56. Thasleena NT, Varghese SC (2014) Enhanced associative classification of XML documents supported by semantic concepts. Int Conf Information Comm Tech 46:194–201

57. Mattiev J, Kavšeka B (2020) A compact and understandable associative classifier based on overall coverage. In: International Workshop on Statistical Methods and Artificial Intelligence (IWSMAI), April 6–9, Warsaw, Poland, 170, 1161–1167

58. Dang Nguyena, Loan T.T. Nguyenb, Bay Vo, Witold Pedryczf. Efficient mining of class association rules with the itemset constraint, Knowledge- Based Systems journal, 2016, doi:https://doi.org/10.1016/j.knosys.2016.03.025

59. Wang X, Yue K, JiaNiu W, Shi Z (2011) An approach for adaptive associative classification. Expert Syst Appl ScienceDirect J. https://doi.org/10.1016/j.knosys.2016.03.025.

60. YennyVilluendas-Reya CF, Rey-Benguría Á, Santiagoc O-N, Yáñez-Márquez C (2017) The Naïve Associative Classifier (NAC): a novel, simple, transparent, and accurate classification model evaluated on financial data. Neurocomputing 000:1–11

61. Sharma P, Berwal YPS, Ghai W (2018) KrishiMitr (farmer's friend): using machine learning to identify diseases in plants. In: IEEE International Conference on Internet of Things and Intelligence System (IOTAIS), pp 29–34. https://doi.org/10.1109/IOTAIS.2018.8600898

62. Ramesh S, et al (2018) Plant disease detection using machine learning. In: International Conference on Design Innovations for 3Cs Compute Communicate Control (ICDI3C), pp. 41–45. https://doi.org/10.1109/ICDI3C.2018.00017

63. Ferentinos K (2018) Deep learning models for plant disease detection and diagnosis. Comput Electron Agric 145:311–318

64. Gaikwad VP, Musande V (2017) Wheat disease detection using image processing. In: 1st International Conference on Intelligent Systems and Information Management (ICISIM), pp 110–112. https://doi.org/10.1109/ICISIM.2017.8122158

65. Barure S, Mahadik B, Thorat M, Kalal A (2020) Disease detection in plant using machine learning. IRJET 7(3), Mar

66. Irías Tejeda J, Castro R (2019) Algorithm of weed detection in crops by computational vision. In: 2019 International Conference on Electronics, Communications and Computers (CONI-ELECOMP), Cholula, Mexico, pp 124–128. https://doi.org/10.1109/CONIELECOMP.2019.8673182

67. Kumaraswamy R, et al (2019) Performance comparison of weed detection algorithms. In: 2019 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, pp 0843–0847. https://doi.org/10.1109/ICCSP.2019.8698094

68. Umamaheswari S, Arjun R, Meganathan D (2018) Weed detection in farm crops using parallel image processing. In: IEEE 2018 Conference on Information and Communication Technology (CICT), Jabalpur, India, pp 1–4. https://doi.org/10.1109/INFOCOMTECH.2018.8722369

69. Hameed S, Amin I (2018) Detection of weed and wheat using image processing. In: 5th International Conference on Engineering Technologies and Applied Sciences (ICETAS), Bangkok, Thailand, pp 1–5. https://doi.org/10.1109/ICETAS.2018.8629137

70. Barrero O, Rojas D, Gonzalez C, Perdomo S (2016) Weed detection in rice fields using aerial images and neural networks. In: 2016 XXI Symposium on Signal Processing, Images and Artificial Vision (STSIVA), Bucaramanga, pp 1-4. https://doi.org/10.1109/STSIVA.2016.7743317

71. Mishra M, Srivastava M (2014) A view of artificial neural network. In; International Conference on Advances in Engineering & Technology Research (ICAETR—2014), pp.1–3. https://doi.org/10.1109/ICAETR.2014.7012785

72. Albawi S, Mohammed TA, Al-Zawi S (2017) Understanding of a convolutional neural network. In: International Conference on Engineering and Technology (ICET), pp 1–6. https://doi.org/10.1109/ICEngTechnol.2017.8308186

73. Denny Prabowo Y, Warnars HLHS, Budiharto W, Kistijantoro AI, Heryadi Y, Lukas (2018) LSTM and simple RNN comparison in the problem of sequence to sequence on conversation data using Bahasa Indonesia. In: Indonesian Association for Pattern Recognition International Conference (INAPR-2018), pp 51–56. https://doi.org/10.1109/INAPR.2018.8627029

74. Salman AG, Kanigoro B, Heryadi Y (2015) Weather forecasting using deep learning techniques. In: International Conference on Advanced Computer Science and Information Systems (ICACSIS), pp 281–285. https://doi.org/10.1109/ICACSIS.2015.7415154

75. Naveen L, Mohan HS (2019) Analyzing impact of weather forecasting through deep learning in agricultural crop model predictions. Int J Appl Eng Res 14:4379–4386

76. Selvanayagam T, Suganya S, Palendrarajah P (2019) Agro-genius: crop prediction using machine learning. Int J Inno Sci Res Tech 4(10):243–249

77. Sabu KM, Manoj Kumar TK (2020) Predictive analytics in Agriculture: Forecasting prices of Arecanuts in Kerala. Procedia Comp Sci 171:699–708

78. Peng Y, Hsu C, Huang P (2015) Developing crop price forecasting service using open data from Taiwan markets. In: Conference on Technologies and Applications of Artificial Intelligence (TAAI), Tainan, Taiwan, pp 172–175. https://doi.org/10.1109/TAAI.2015.7407108. ©2015 IEEE

79. Vohra NP, Khatri SK (2019) Decision making support system for prediction of prices in agricultural commodity. In: Amity International Conference on Artificial Intelligence (AICAI), Dubai, United Arab Emirates, pp 345–348. https://doi.org/10.1109/AICAI.2019.8701273 ©2019 IEEE

80. Kurumatani K (2018) Time series prediction of agricultural products price based on time alignment of recurrent neural networks. In: 17th IEEE International Conference on Machine Learning and Applications (ICMLA), Orlando, FL, USA, pp 81–88. https://doi.org/10.1109/ICMLA.2018.00020

81. Mhudchuay T, Kasetkasem T, Attavanich W, Kumazawa I, Chanwimaluang T (2019). Rice cultivation planning using a deep learning neural network. In: 16th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), pp 822–825. https://doi.org/10.1109/ECTI-CON47248.2019.8955227

82. Rajesh R, Saradhambal D, Latha S (2018) Plant disease detection and its solution using image classification. Int J Pure Appl Math 119:879–884

83. Mishra DK, Veenadhari S, Singh CD (2011) Soybean productivity modelling using decision tree algorithms. Int J Comp Appl

84. Sehgal A, Mathur S (2019) Plant disease classification using soft computing supervised machine learning. In: 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA), pp 75–80. https://doi.org/10.1109/ICECA.2019.8822213

85. Mehta P, Shah H, Kori V, Vikani S, Shukla, Shenoy M (2015) Survey of unsupervised machine learning algorithms on precision agricultural data. In; International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS-2015). pp 1–8. https://doi.org/10.1109/ICIIECS.2015.7193070

86. van Klompenburga T, Kassahuna A, Catal C (2020) Crop yield prediction using machine learning: A systematic literature review. Comp Elect Agri, Sci Direct 177:105709

87. Bacco M, Barsocchi P, Ferro E, Gotta A, Ruggeri M (2019) The digitisation of agriculture: a survey of research activities on smart farming. Array 3–4:100009

88. Jha K, Doshi A, Patel P, Shah M (2019) A comprehensive review on automation in agriculture using artificial intelligence. Artificial Intell Agri Sci Dir 2:1–12

89. Yuzhen L, Young S (2020) A survey of public datasets for computer vision tasks in precision agriculture. Comput Electron Agric 178:105760

90. Padarian J, Minasny B, McBratney AB (2019) Machine learning and soil sciences: a review aided by machine learning tools. Soil, EGU, 6(1):35–52

91. Bachhav NB (2012) Information needs of the rural farmers: a study from Maharashtra, India: a survey. Digital Commons @University of Nebraska. Library Phil Pract (e-journal) 866:1–13

92. Young A, Mahan J, Dodge W, Payton P (2020) BLOB-based AOMs: a method for the extraction of crop data from aerial images of cotton. MDPI-Agriculture. https://doi.org/10.3390/agriculture10010019

93. Jain L et al (2017) Cloud-based system for supervised classification of plant diseases using convolutional neural networks. In: IEEE International Conference on Cloud Computing in Emerging Markets (CCEM), pp 63–68. https://doi.org/10.1109/CCEM.2017.22

94. Kokane P, Vispute S, Aarti Kalekar KB, Kamble M (2018) Automated generation, calculation of the village soil fertility index and analysis of soil health card. IJNTSE-ISSN5

95. Medar R, Rajpurohit VS, Shweta S (2019) Crop yield prediction using machine learning techniques. In: IEEE 5th International Conference for Convergence in Technology (I2CT-2019), pp. 1–5. https://doi.org/10.1109/I2CT45611.2019.9033611

96. Deshmukh PR, Badnuke MR (2012) Infected leaf analysis and comparison by Otsu threshold and k-means clustering. Int J Adv Res Comp Sci Soft Eng 2(3):449–452

97. Sun L, Yang Y, Hu J, Porter D, Marek T, Hillyer C (2017) Reinforcement learning control for water-efficient agricultural irrigation. In: IEEE International Symposium on Parallel and Distributed Processing with Applications and 2017 IEEE International Conference on Ubiquitous Computing and Communications (ISPA/IUCC), pp 1334–1341. https://doi.org/10.1109/ISPA/IUCC.2017.00203

98. Vispute SR, Kolekar V, Bhujbal H, Gudle M, Kadam A, Kadu K (2018) An application for e-marketing of agricultural commodities and analysis of marketable surplus. Int J Adv Res Comp Comm Eng 7(4):77–79

99. Sutton R, Barto A (2018) Introduction to reinforcement learning, 2nd edn. The MIT Press, Cambridge MA

100. Gawali P, Dalvi M, Dhannawat D, Deshmukh S, Vispute SR (2020) An application of data analytics in agriculture sector for multi-advice generator in native language. J Critical Rev 7(19):2389–2394

101. Waghmare H, Kokare R, Dandawate Y (2016) Detection and classification of diseases of Grape plant using opposite colour Local Binary Pattern feature and machine learning for automated decision support system. In: 3rd International Conference on Signal Processing and Integrated Networks (SPIN), pp 513–518. https://doi.org/10.1109/SPIN.2016.7566749

102. Elavarasan D, Vincent PMD (2020) Crop yield prediction using deep reinforcement learning model for sustainable agrarian applications. IEEE Access 8:86886–86901. https://doi.org/10.1109/ACCESS.2020.2992480