**Journal of Energy and Natural Resource Management (JENRM), Vol. 8, No. 2, 51-62, 2022**
**DOI: 10.26796/jenrm.v8i2.209 (Research Article)**

# Application of Machine Learning in Crop Cultivation and Production: Systematic Review

Peter Appiahene[1], Ameyaw Adu Kofi[1], Kwadwo Gyasi Santo[2] and Richard Kwasi Bannor[3]

**Abstract**
Machine learning methods have a positive impact on crop predictions. This study is a systematic review of the application of machine learning in crop cultivation and production. The papers for this review have been extracted using electronic database algorithms, and features employed in crop selection analysis in precision agriculture are synthesised. Following the search criteria, 133 papers were found in Google Scholar, Elsevier, Springer Link, Wiley, Web of Science, and other electronic databases. Of these, 62 were chosen for further research and analysis using inclusion and exclusion criteria. The final papers were examined and the methodologies and features utilised were extensively analysed, suggesting additional research. Temperature, Rainfall, PH, Nitrogen, and Potassium are the most commonly used features, and KNN, SVM, ANN, and Linear Regression are the most commonly used ML algorithms. In the analysis, an analysis assessing how best ensemble methods can improve crop selection in this SLR is undertaken. In addition, a few papers that used Meta algorithms to develop their models performed better. The most commonly used ensembles in the SLR are XG boost and Bagged Tree, Ensemble of SVM & ANN, Ensemble of NB & SVM, and Ensemble of NB & SVM. Meta-algorithms, however, outscored single algorithms in the SLR.

**Keywords**
Machine Learning; Precision Agriculture; Electronic Databases; Meta-Algorithms; Ensemble

[1]Department of Computer Science and Informatics, University of Energy and Natural Resources, Sunyani, Ghana
[2]Department of Horticulture and Crop Production, University of Energy and Natural Resources, Sunyani, Ghana
[3]Department of Agribusiness Management and Consumer Studies, University of Energy and Natural Resources, Sunyani- Ghana
***Corresponding author**: peter.appiahene@uenr.edu.gh
**DOI:** 10.26796/jenrm.v8i2.209
**Received:** 5 October 2022 ; **Received in revised form:** 4 November 2022; **Accepted:** 8 November 2022; **Published:** 30 November, 2022

## Contents



## 1. Introduction

Agriculture is the backbone of most countries in the world by providing food and fabrics. The world population is increasing rapidly, and the increasing population must also correlate with food production. Therefore, there is a need to integrate precision agriculture and traditional farming (Rajak et al., 2017; Yesugade et al., 2018). Several models have been proposed and tested to improve prediction(Shahhosseini et al., 2021)(Palanivel et al., 2019). In precision agriculture, selecting crop(s) that a farmer can grow on a piece of land requires many factors to be considered(Naveen & Mohan, 2022; Liakos et al., 2018)(Klompenburg et al., 2020); for instance, soil minerals present, climatic conditions such as temperature, humidity, rainfall, etc. According to Thomas et al., any poor decision made by the farmer could result in financial strain and significant losses (Thomas et al., 2020). Machine learning introduces ways for defining rules and patterns in massive datasets and the ability to adjust the predictive model to make a successful crop selection decision(Prasad et al., 2019).

The rising deployment of distant and proximal sensing technologies has resulted from the necessity for timely and precise sensing of these inputs for large agricultural fields. Some of these studies looked at the Normalised Difference Vegetation Index (Aghighi et al., 2018; Chlingaryan et al., 2018; Nevavuori et al., 2019), Wi-Fi-Sensors (Majumdar et al., 2021; Sharifi, 2021; T.Banavlikar, A.Mahir, M.

Budukh, 2018; Yesugade et al., 2018), Satellite-remote sensing (Ferencz et al., 2010; Pham et al., 2022; Sonobe et al., 2017; Tatsumi et al., 2015; Vasavi et al., 2022; A. X. Wang et al., 2015; S. Wang et al., 2019; Zhou et al., 2016) and image processing (Classification, 2019; Danbaki et al., 2020; Hu et al., 2021; Ji et al., 2021; Nandhini & Shankar, 2020; Teng et al., 2016; Yasrab, 2021). Precision Agriculture relies heavily on machine learning and big data to increase farm output (Shruthi, 2020). Improving on-farm yield and quality while minimising expenses and environmental degradation is challenging.

In precision agriculture (PA), the decision to plant the correct crop for a piece of land and disease control determines the potential growth and yield (Patel & Sharaff, 2021). Precision agriculture has been one of the most significant agricultural breakthroughs in the previous two decades in terms of resource and fertiliser use, efficiency, and environmental impact(Babu, 2013)(Chougule et al., 2019). Precision agriculture aims to regulate both geographical and temporal variability of soil and crop factors by measurement, analysis, and appropriate intervention to maximise profitability, sustainability, and environmental protection (Danbaki et al., 2020; Beans & Writer, 2020). The use of features of Machine Learning (ML) in big data applications are employed in PA to increase productivity (Appiahene et al., 2020; Appiahene & Missah, 2019; Awoin et al., 2020; Junior et al., 2022; Kutiame et al., 2022; Peter et al., 2022) (Masood, 2021)(Cravero & Sep 2021) (Cai & Sharma, 2021). PA uses modern technologies to suggest a crop to grow on a plot of land. We have a large amount of data available in the form of ratings, reviews, graphs, photographs, and so on in this new era of digitalisation (K. Patel & Patel, 2020). PA is an example of technological innovation (Savla & Mandholia, 2015). People, on the other hand, have trouble extracting useful information or knowledge from these data.

Recommendation systems enter the scene to solve these challenges by presenting helpful material to users based on their history and similarity among users. Precision agriculture involves increasing farmer productivity and quality while lowering expenses, and environmental degradation is a big concern. The two primary building elements of recommendation systems are content-based and collaborative filtering. Recommendation systems have been used in various fields, including film, music, courses, literature, items, people, links, location, healthcare (Nica et al., 2021), and agriculture. In the agricultural arena, choosing proper crops to cultivate and picking appropriate pesticides depending on land quality and crop varieties are important variables for farmers to consider when making informed decisions about the crop to plant. This systematic literature review describes the use of machine learning in crop prediction. Discover potential research gaps to promote scholarship on the subject. The studied papers were retrieved via these electronic databases Google scholar, Wiley, Scopus and web of science. The study is structured as follows: Section one is the introductory part of the SLR, which covers the background and main objective. Section two discusses related work, which covers the explanation and description of the study's themes. Section three extensively details the approach to numerous scientific methodologies, strategies, and tools used to complete the study. In Section four, the outcomes of the study are analysed. Finally, Section five was on discussions while keeping in mind the specified objectives and offering the study's conclusion and recommendations for future research.

## 2. Review of related works

Crop prediction for a piece of land is a prudent decision to maximised farm out-put and minimise losses by farmers (R. K. Rajak, A. Pawar, M. Pendke, P. Shinde, S. Rathod, 2017) (Ujjainia et al., 2021). An accurate crop prediction model assists farmers with the type of crop to grow and the time to grow the crop. Although many models have been proposed, this systematic literature review identified an approach of using machine learning to predict the crops a farmer can grow on a piece of land in the literature. These discussed papers were closely related to the study. In a study, Balducci et al. (2018) published a paper on machine learning applications on smart farm enhancement. The paper addressed the challenges of smart farming to smaller farmers, that is, IoT in agriculture. Furthermore, Ali et al. (2020) used a machine learning application to design a recommendation system addressing local farmers' concerns to increase productivity. The model used temperature as a dataset.

Suresh et al. published a paper on a machine learning-based agricultural yield recommendation system for digital farming(Suresh, Senthil Kumar, & Lekashri, 2021). The paper addresses the challenge of wastage by farmers' inability to identify the right crop to grow on the piece of land. The paper proposed a framework to recommend the crop to cultivate depending on the information of the soil. In a study by Bakthavatchalam et al. (2022), the researchers devised a crop production measurement method in India that blends cutting-edge technology IoT in agriculture. Moreover Savla and Mandholia (2015) conducted a literature review on classification algorithms for yield forecast accuracy in precision agriculture

(Kumar et al., 2019)(Ahamed et al., 2015) released a review paper on data mining's applicability in precision agriculture that depended on using primary data acquired from individual farmers as well as soil samples to build a Simple Vector Machine and Artificial Neural Network model to select a crop for a specific location. Most importantly, according to the authors, the model was efficient and had higher accuracy in smaller farms than in bigger ones. In addition, to complement other studies on crop selection in Machine Learning(Chougule et al., 2019), developed an android-based application to address the challenge of precision agriculture. The application has a graphical user interface that receives input from the user. The authors concluded that the accuracy of the developed system is reasonably higher after the performance evaluation.

L. Naveen and H. S. Mohan researched the implications of using different weather parameters to select crops to be grown(Naveen & Mohan, 2022). The research was carried out entirely using the matrix laboratory (MATLAB) program to forecast weather factors needed for crop management.

A.R.M.S.Masood uses soft computing methods with a data mining approach for effective dimensionality reduction to be achieved(Masood, 2021). The paper uses PCA (principal component analysis) to achieve depth reduction and a decision tree to achieve prediction; the study centralises on agricultural yield analysis.

The application of Machine Learning on Satellite-Base Vegetation Health Indices to Improve Crop selection and Yield Prediction(Pham et al., 2022), the method of Independent Component Analysis and Principal Component Analyses, combined with the ML algorithm, are used to design a framework for improving the accuracy of estimating crop yield using Vegetation Condition Index and Thermal condition Index data.

A. Kamilaris and F. X. Prenafeta-Boldú Survey on Deep Learning in Agriculture(Kamilaris & Prenafeta-Boldú, 2018). The authors studied many research projects that use deep learning techniques to address various agricultural and food production concerns; the paper concluded that deep learning outperforms existing image processing approaches in terms of accuracy. Moreover, (Saiz-Rubio & Rovira-Más, 2020) review crop data management from smart farming to Agriculture. According to the report, current developments in data management are causing smart farming to grow dramatically, as data has become the basic aspect of modern agriculture to assist producers in making critical decisions.

Again, a study on precision agriculture for crop farming with machine vision systems(Mavridou et al., 2019) examined recent progress in the use of machine vision in agriculture with a focus on crop farming. The research covers a lot of plant and fruit detection techniques and harvesting support techniques, such as fruit grading, ripeness detection etc., to enhance farm output.

Also, a study on machine learning techniques in agricultural crop production combined computer science and agriculture aid in agricultural crop(s) forecasts (Mishra et al., 2016). Moreover, developing an appropriate model has advantages over standard forecasting methods. S. Kale et al. looked at predicting crop yield and success rate using Machine Learning(Kale, 2019). The paper suggests that by adding layers and tweak specific parameters, the ANN algorithm can be enhanced even more to provide more accurate predictions. The authors conclude that this article will assist farmers in decision-making.

T.Shakoor et al. look at Predicting Agricultural Output Using a Supervised Machine Learning technique(Shakoor et al., 2017). According to the authors, farmers typically design the cultivation procedure based on previous experience. They end up cultivating undesired crop(s) due to a lack of ideas regarding agriculture; the discovery gives a remedy to this problem which farmers in India desperately need. Synthetic aperture radar (SAR) research has progressed in Crop detection(Di & Zhongxin, 1980); according to the authors, accurate crop(s) selection allow for accurate estimates of crop area, planting structure, and geographic distribution, as well as critical input factors.

A. K. Saha et al. study looks at drones with the Internet of Things (IoT) for improving crop quality in agricultural fields(Saha et al., 2018). Authors argue that as the world's population grows, drone sensors or unmanned aerial vehicles (UAVs) will be invaluable in the field of agriculture, starting the crop cycle. Smart farming: opportunities, challenges and technology enablers(Bacco et al., 2014); the study provided an overview of the internet of things (IoT) technology in various smart agricultural settings and tried to improve upon it.

A. V. M. Ines et al. remotely sensed soil moisture and vegetation combined with a crop simulation model(Ines et al., 2013); the authors created and evaluated a data assimilation-crop model system for predicting crop yields at the aggregates soil texture and leaf area index.

A. T. M. S. Ahamed et al. data mining techniques were used to extract agricultural industry knowledge to estimate crop yields(Ahamed et al., 2015). A neural network-based crop recommender system is developed by (T.Banavlikar, A.Mahir, M. Budukh, 2018). The paper suggested a solution to the farmers' problem of today, whereby no farmland will go waste by selecting the wrong crop to cultivate on a specific land that can affect the crop yield. Furthermore, (Garanayak et al., 2021) published a paper on how to design a crop recommendation system using several machine-learning regression algorithms. According to the authors, farmers' inability to make good decisions led to the overripe of crops in India. The research helped the decision-makers collaborate with the alliance industries' ability to meet their occupation's needs.

D. Bhattacharya et al. systematic study of IoT for Agriculture Promotion in terms of weather monitoring, yield prediction, security standards, and cost-effective hardware. The authors used weather monitoring to implement modern IoT-based agriculture automation(Majumdar et al., 2021).

# 3. Methodology

## 3.1 Search strategy

In this study, a devised search strategy is defined in order to collect all the relevant materials published in relation to the study area. The search approach used in this project included searching resources, determining search terms and tactics, and the article selection procedure for the research. The electronic databases are used to retrieve the relevant studies. Scopus, Web of Science, Springer, Wiley, and Google Scholar were used in this research.

## 3.2 Identifying search terms

The steps below were utilised to find the relevant search terms for this study; 1. The study topic was used to find the most important terms associated with the topic; 2. Alternative spellings for all of the significant ones were discovered in the search terms; 3. The keywords were then checked in any relevant document; 4. To concatenate alternative spellings and major sentences, the Boolean OR and AND operators were utilised; 5. The search string was condensed to fulfill our requirements for literature.

## 3.3 Summarised limited search term with the search string

### 3.3.1 Resources to be searched

All relevant published material related to this topic was developed through a search strategy. We labored on a search strategy for the works that included determining search terms, resources to search, search techniques, and article selection. The search databases used for the study are:

1. Elsevier(https://www.elsevier.com),
2. Springer Link (https://link.springer.com)
3. Web of Science (https://webofscience.com)
4. Wiley (https://onlinelibrary.wiely.com)
5. GoogleScholar (https://scholar.google.com)

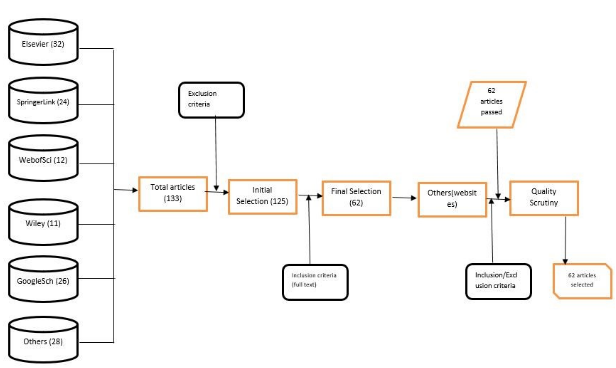


**Figure 1.** Article selection process graphical representation

Table 1 below shows the search results from the databases;

**Table 1.** Results and data source

| Database Name | Found Results in Total | Initial Selection | Final Selection | Percentage (%) |
|---|---|---|---|---|
| Elsvier | 32 | 28 | 20 | 32.26 |
| SpringerLink | 24 | 20 | 13 | 20.97 |
| Web of Science | 12 | 8 | 8 | 12.9 |
| Wiley | 11 | 10 | 6 | 9.68 |
| Google Scholar | 26 | 22 | 7 | 11.29 |
| Others | 28 | 26 | 8 | 12.9 |
| Websites | 8 | 0 | 0 | 0 |
| **Grand Total** | **141** | **195** | **62** | **100** |

### 3.3.2 Search process overview

The methods below were followed to discover all publicly available published publications relevant to the work

1. In section 3.2, the search string mentioned was used to collect the primary research available in the digital database library, as mentioned in 3.3. In the search process,133 were found regarding the query, as represented in Table 3.3.1;
2. Non-relevant papers were detected via the exclusion criteria. These were determined by reading the abstract, title, and keywords of the 133 papers found. All items that met this requirement were discarded. While the others were chosen for the following phase of the search process, resulting in the inclusion criteria. As a result of this procedure, we were able to narrow down the search string to 62 final publications (about 46.62 percent of the total papers found);
3. To identify new studies via other methods, websites were also searched for possible missing articles. About eight papers were found, and the website abstract titles, abstracts, and keywords were read. The exclusion/inclusion criteria were applied, and as a result, no paper was selected to be analysed further;
4. A quality evaluation process was conducted on all 62 articles utilised in this study to guarantee that relevant materials were added to the final list of sources. The decision was taken as to whether or not to include an article in the final selection based on the exclusion/inclusion criteria

In the end, 62 papers passed the steps above with the search string as well as the website search for the relevant literature. Figure 2 below shows the PRISMA diagram;

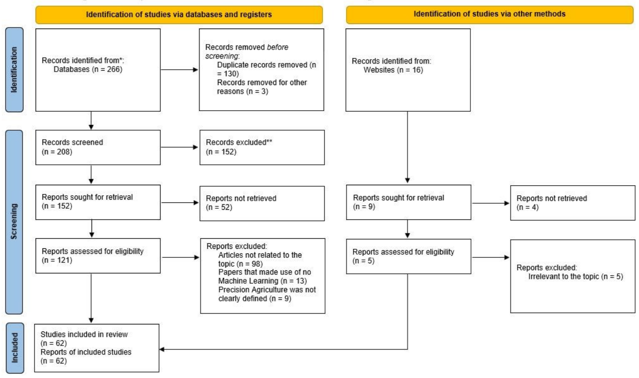


**Figure 2.** The Flow Diagram (PRISMA) diagram

#### 3.3.3 Inclusion and exclusion criteria

A paper passed the following criteria to better relate to our topic.

1. Exclusion criteria: Our analysis eliminated sources that met the limits listed below:
   - Articles that focused on methods other than the topic. As a result, 71 papers were rejected;
   - Articles written in a language other than English. A total of 7 papers (all written in Chinese) were disqualified;
   - Articles whose complete text is not available. 15 papers were further excluded
2. Inclusion criteria: Our analysis includes sources that met the limits listed below:
   - All articles are written in English and describe best for our topic;
   - Articles that introduce techniques or improve the efficacy of existing ones.

#### 3.3.4 Study quality scrutiny

The integrity and thoroughness of the chosen source were assessed after the final selection using the checklist below:

- Q1: Are the approaches using the explicitly specified technique proposed in the paper?
- Q2: Is the topic needed for this paper distinctly defined?
- Q3: Are the evaluation methods and metrics clearly stated?

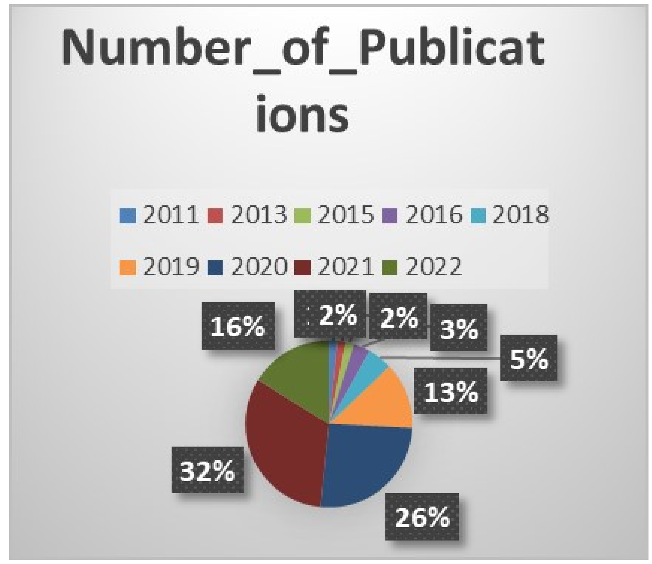


**Figure 3.** Number of publications in percentage

#### 3.3.5 Data extraction

We began extracting the data needed to meet the specifics of our topic after we had chosen the final papers to be used in the research. The data extraction was based on the programming language, machine learning employed, evaluation methodology, and categorisation type. It also highlighted any potential limitations of the research under consideration. This assisted in the detection of previous work's limits as well as instructions for future research.

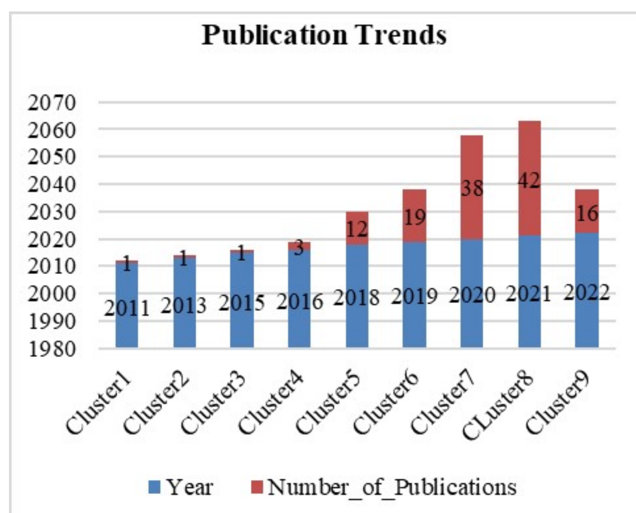


**Figure 4.** Publication trends

## 4. Results and analysis

This section contains the study's analysis. It also includes a study of the technique for predicting agricultural yields using Machine Learning. The quality of the publications that met the inclusion/exclusion criteria evaluation before providing the results of the systematic literature analysis concerning the selected research issues provided a quick overview of the demographics. In the selection process, 62 related articles were used based on crop prediction using Machine Learning.

### 4.1 Data extraction

In this study, papers that were published from 2011 to 2022 were considered. The results from publications between the years that we considered in this work are represented in figure 3 below;

In this study, 2021 represented the highest (32%) number of relevant publication trends, as represented in figure 3. below:

### 4.2 Bibliographic analysis

Bibliographic analysis helps the author/researchers to get an overview of the structure of the scientific literature in a certain domain, which identifies the main research area related to each other and gain insight. In this study, 9 different clusters were achieved, showing the co-occurrence in the considered papers between 2011 and 2022. Figure 6 below shows the bibliographic analysis.

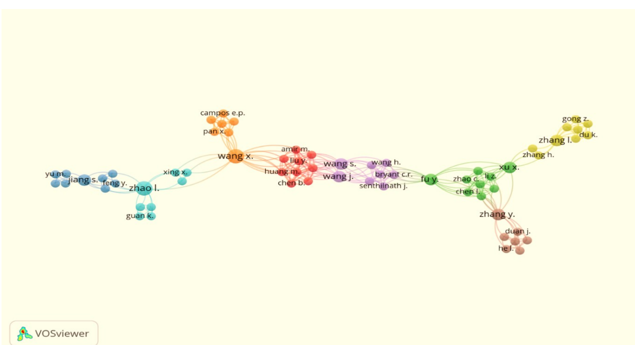


**Figure 5.** Bibliographic analysis of the relevant papers for the study

Table 2 below shows the samples of related literature used in the study The related literature was grouped in clusters. Figure 6 below shows the number of clusters.

**Table 2.** Samples of related literature

| Authors | Title | Year | Publication stage |
|---|---|---|---|
| Naveen L., Mohan H.S. | A novel weather parameters prediction scheme and their effects on crops | 2022 | Final |
| Dubois A., Teytaud F., Verel S. | Short-term soil moisture forecasts for potato crop farming: A machine learning approach | 2021 | final |
| Shi P., Wang Y., Xu J., Zhao Y., Yang B., Yuan Z., Sun Q. | Rice nitrogen nutrition estimation with RGB images and machine learning methods | 2021 | final |
| Mariano C., Mónica B.. | A random forest-based algorithm for data-intensive spatial interpolation in crop yield mapping | 2021 | final |
| Venkatesan S., Lim J., Ko H., Cho Y. | A Machine Learning Based Model for Energy Usage Peak Prediction in Smart Farms | 2022 | final |
| Huang M.-L., Chuang T.-C., Liao Y.-C. | Application of transfer learning and image augmentation technology for tomato pest identification | 2022 | final |
| Feng Z., Song L., Duan J., He L., Zhang Y., Wei Y., Feng W. | Monitoring wheat powdery mildew based on hyperspectral, thermal infrared, and RGB image data fusion | 2022 | final |
| Gong L., Yu M., Jiang S., Cutsuridis V., Pearson S. | Deep learning-based prediction on greenhouse crop yield combined TCN and RNN | 2021 | final |
| Nadeem R.M., Jaffar A., Saleem R.M. | IoT and machine learning-based stem borer pest prediction | 2022 | final |


**Figure 6.** Number of items in the Clusters

### 4.3 Independent variables considered in literature

The Features used in the papers are investigated and summarised in table 3.

**Table 3.** SAMPLES OF THE INDEPENDENT VARIABLE FEATURES USED

| **Features** | **Publication Times** |
|---|---|
| Temperature | 34 |
| rainfall | 15 |
| ph | 15 |
| climate | 6 |
| humidity | 12 |
| NDVI | 9 |
| precipitate | 5 |
| Soil type | 11 |
| Nitrogen | 11 |
| phosphorus | 6 |
| Soil water | 9 |
| magnesium | 5 |
| management | 1 |
| zinc | 3 |
| sulphur | 5 |

### 4.4 Machine learning algorithms in literature

The Machine Learning Algorithms were researched, summarised, and tallied in table 4.

**Table 4.** Algorithms used

| Algorithms | Times used | % |
|---|---|---|
| KNN | 11 | 26.83 |
| Support Vector Machine | 7 | 17.07 |
| ANN | 5 | 12.2 |
| Linear Regression | 4 | 9.77 |
| Ensemble | 4 | 9.77 |
| Naïve Bayes | 3 | 7.32 |
| Random Forest | 2 | 4.88 |
| Decision Tree | 1 | 2.44 |
| Gradient Boosting Tree | 1 | 2.44 |
| Neural network | 3 | 7.32 |
| Total | 41 | 100 |

### 4.5 Crop prediction evaluation parameters

The evaluation parameters and approaches utilised in the literature are summarised in table 5.

**Table 5.** Sample evaluation parameters useD

| Abbreviations | Evaluation parameters | Times used |
|---|---|---|
| ACC. | Accuracy | 22 |
| RRMSE | Relative Root Mean Square Error | 3 |
| RMSE | Root Mean Square Error | 26 |
| MAPE | Mean Average Percentage Error | 5 |
| RRSE | Relative Root Square Error | 1 |
| MBE | Mean Bias Error | 3 |
| MAE | Mean Average Error | 14 |
| MSE | Mean Square Error | 8 |
| REC. | Recall | 3 |

### 4.6 Ensemble algorithms used in the systematic literature reviews

To enhance crop prediction, ensemble algorithms which are Meta-algorithms, were used to enhance crop prediction. The ensembles used in the literature are summarised in table 6.

**Table 6.** Ensemble algorithms

| Algorithms used | Times used |
|---|---|
| XGboost | 3 |
| Bagged tree | 3 |
| Ensemble of SVM \& ANN | 1 |
| Ensemble of Naïve Bayes and SVM | 1 |
| Ensembles of Neural Network \& DT | 1 |
| AdaBoost | 1 |
| Gradient Boost of Decision Tree | 1 |
| Ensemble of DT \& SVM | 1 |

## 5. Discussion

Though vital documents may go missing, more closely related terms could have resulted in the new research. However, the search term yielded a large number of publications, showing that the search was thorough. Other obstacles that obstruct the SLR include the fact that not all publications mention what evaluation parameters were used. Some do not describe the characteristics employed, making it impossible to obtain the information needed to respond to the research questions. Many machine learning models were to see which one made the most accurate predictions. This SLR looked into the extent to which ensemble methods were employed to improve crop forecast.

In Research Question one: According to Table 4, the most commonly used algorithm for crop prediction is KNN, and the second most commonly used method is SVM (Marinkovi, 2011)(John et al., 2020). However, this does not imply that it is the most effective algorithm.

In this SLR, the ensemble algorithm, which is a combination of two or more methods, performed better(Heide et al., 2020). In research question two: Temperature, rainfall, and PH are the most commonly used features. Apart from these, others such as NDVI(Molero & Bort, 2014), zinc and nitrogen were also employed. In Research Question three: Accuracy (A), RMSE(Sharma et al., 2020)(Jeong et al., 2016), MAE, and R2 are the most commonly used evaluation parameters. Almost RMSE was used to measure the quality model, as shown in Table 5.

### 5.1 Ease of use

Given the field of study and data availability, this comprehensive literature review SLR found that many features were used in these articles. The kinds of characteristics employed in all crop prediction experiments with machine learning are different. The research also differs in diverse ways; for example, climatic conditions (Kamir et al., 2020)(Peng et al., 2018), geological position, and crops that survive there are all factors that affect the research. The availability of the dataset and the study's aims determine which Systematic Literature Review features are employed.

According to studies, some of the models could accurately predict crop yields. Though, many algorithms are used in different studies. No explicit conclusion can be taken about which model is the best among various machine learning algorithms. Ensemble algorithms were used to improve crop prediction. After scrutiny and synthesising the selected papers, the most used ensembles are XGboos and Bagged tree. However, there are also other kinds of ensemble algorithms used. AdaBoost(Balakrishnan & Muthukumarasamy, 2016), gradient boosting, an ensemble of neural network and decision tree (Moreno, 2020), ensemble of SVM and ANN(Khalil & Abdullaev, 2021). These meta-algorithms performed better (SMummaleti Keerthana, K J M Meghana, Siginamsetty Pravallika, 2021). This paper will induce researchers into solving problems related to crop cultivation and production prediction in the near future using the internet of Things (IoT).

**Funding** This research did not receive any specific grant from the public, commercial or non-governmental funding agencies.

**Competing interests** The authors declare that they have no competing interests.

**Data availability statement** Data will be made available upon request from the corresponding author.

**Author contribution** All authors participated in the manuscript's conceptualisation, design, analysis, writing and proofreading.

**Ethics approval statement** None

**Permission to reproduce material from other sources** Not needed

## References

[1] Aghighi, H., Azadbakht, M., Ashourloo, D., Shahrabi, H. S., & Radiom, S. (2018). Machine Learning Regression Techniques for the Silage Maize Yield Prediction Using Time-Series Images of Landsat 8 OLI. 1–15.

[2] Ahamed, A. T. M. S., Mahmood, N. T., Hossain, N., Kabir, M. T., Das, K., Rahman, F., & Rahman, R. M. (2015). Applying Data Mining Techniques to Predict Annual Yield of Major Crops and Recommend Planting Different Crops in Different Districts in Bangladesh.

[3] Aghighi, H., Azadbakht, M., Ashourloo, D., Shahrabi, H. S., & Radiom, S. (2018). Machine Learning Regression Techniques for the Silage Maize Yield Prediction Using Time-Series Images of Landsat 8 OLI. 1–15.

[4] Ahamed, A. T. M. S., Mahmood, N. T., Hossain, N., Kabir, M. T., Das, K., Rahman, F., & Rahman, R. M. (2015). Applying Data Mining Techniques to Predict Annual Yield of Major Crops and Recommend Planting Different Crops in Different Districts in Bangladesh.

[5] Ali, A. M., Darvishzadeh, R., Skidmore, A., Gara, T. W., Heurich, M., Mohammed, A., Darvishzadeh, R., Skidmore, A., & Skidmore, A. (2020). Machine learning methods ' performance in radiative transfer model inversion to retrieve plant traits from Sentinel-2 data of a mixed mountain forest. International Journal of Digital Earth, 0(0), 1–15. https://doi.org/10.1080/17538947.2020.1794064

[6] Appiahene, P., & Missah, Y. M. (2019). Predicting the Operational Efficiency of Banks in the presence of Information Technology Investment. International Journal of Advances in Electronics and Computer Science, 6(11), 1–6. https://doi.org/IJAECS-IRAJ-DOI-16504

[7] Appiahene, P., Missah, Y. M., & Najim, U. (2020). Predicting Bank Operational Efficiency Using Machine Learning Algorithm: Comparative Study of Decision Tree, Random Forest, and Neural Networks. Advances in Fuzzy Systems.

[8] Awoin, E., Appiahene, P., Gyasi, F., & Sabtiwu, A. (2020). Predicting the Performance of Rural Banks

in Ghana Using Machine Learning Approach. Advances in Fuzzy Systems, 2020.

[9] Babu, S. (2013). A Software Model for Precision Agriculture for Small and Marginal Farmers. 352–355.

[10] Bacco, M., Berton, A., Ferro, E., Gennaro, C., Gotta, A., & Matteoli, S. (2014). Smart Farming: Opportunities, Challenges and Technology Enablers.

[11] Bakthavatchalam, K., Karthik, B., Thiruvengadam, V., Muthal, S., Jose, D., Kotecha, K., & Varadarajan, V. (2022). IoT Framework for Measurement and Precision Agriculture: Predicting the Crop Using Machine Learning Algorithms. Technologies, 10(1), 13. https://doi.org/10.3390/technologies10010013

[12] Balakrishnan, N., & Muthukumarasamy, G. (2016). Crop Production - Ensemble Machine Learning Model for Prediction. International Journal of Computer Science and Software Engineering, 5(7), 148–153.

[13] Balducci, F., Impedovo, D., Informatica, D., & Moro, A. (2018). Machine Learning Applications on Agricultural Datasets for Smart Farm Enhancement. https://doi.org/10.3390/machines6030038

[14] Beans, C., & Writer, S. (2020). Crop researchers harness artificial intelligence to breed crops for the changing climate. 117(44), 27066–27069. https://doi.org/10.1073/pnas.2018732117

[15] Cai, Y., & Sharma, A. (2021). Swarm Intelligence Optimization: An Exploration and Application of Machine Learning Technology. 460–469.

[16] Chlingaryan, A., Sukkarieh, S., & Whelan, B. (2018). Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. Computers and Electronics in Agriculture, 151(November 2017), 61–69. https://doi.org/10.1016/j.compag.2018.05.012

[17] Chougule, A., Jha, V. K., & Mukhopadhyay, D. (2019). Crop Suitability and Fertilizers Recommendation Using Data Mining Techniques. Springer Singapore. https://doi.org/10.1007/978-981-13-0224-4

[18] Classification, D. (2019). CropDeep: The Crop Vision Dataset for Deep-Learning-Based Classification and Detection in. https://doi.org/10.3390/s19051058

[19] Cravero, A., & Sep, S. (2021). Use and Adaptations of Machine Learning in Big Data — Applications in Real Cases in Agriculture.

[20] Danbaki, C. A., Onyemachi, N. C., Gado, D. S. M., Mohammed, G. S., Agbenu, D., & Ikegwuiro, P. U. (2020). Precision Agriculture Technology: A Literature Review. Asian Journal of Advanced Research and Reports, 14(3), 30–34. https://doi.org/10.9734/ajarr/2020/v14i330335

[21] Di, W., & Zhongxin, C. (1980). Advances in research on crop identification using SAR.

[22] Ferencz, C., Bognár, P., Lichtenberger, J., Hamar, D., Tarcsai, G., Timár, G., Molnár, G., Pásztor, S. Z., Steinbach, P., & Székely, B. (2010). Crop yield estimation by satellite remote sensing. 1161. https://doi.org/10.1080/01431160410001698870

[23] G.Suresh, A.Senthil Kumar, S.Lekashri, R. M. (2021). Efficient Crop Yield Recommendation System Using Machine Learning For Digital Farming. 10(1), 906–914.

[24] Garanayak, M., Sahu, G., Mohanty, S. N., & Jagadev, A. K. (2021). Agricultural Recommendation System for Crops Using Different Machine Learning Regression Methods. 12(1), 1–20. https://doi.org/10.4018/IJAEIS.20210101.oa1

[25] Heide, E. M. M. Van Der, Kamphuis, C., Veerkamp, R. F., Athanasiadis, I. N., Azzopardi, G., Pelt, M. L. Van, & Ducro, B. J. (2020). Improving predictive performance on survival in dairy cattle using an ensemble learning approach. Computers and Electronics in Agriculture, 177(June), 105675. https://doi.org/10.1016/j.compag.2020.105675

[26] Hu, Y., Shen, J., & Qi, Y. (2021). applied sciences Estimation of Rice Biomass at Different Growth Stages by Using Fractal Dimension in Image Processing.

[27] Ines, A. V. M., Das, N. N., Hansen, J. W., & Njoku, E. G. (2013). Remote Sensing of Environment Assimilation of remotely sensed soil moisture and vegetation with a crop simulation model for maize yield prediction. Remote Sensing of Environment, 138, 149–164. https://doi.org/10.1016/j.rse.2013.07.018

[28] Jeong, J. H., Resop, J. P., Mueller, N. D., & Fleisher, D. H. (2016). Random Forests for Global and Regional Crop Yield Predictions. 1–15. https://doi.org/10.1371/journal.pone.0156571

[29] Ji, Y., Kumar, R., Singh, D., & Singh, M. (2021). Performance Analysis of Target Information Recognition System for Agricultural Robots. 12(2), 49–60. https://doi.org/10.4018/IJAEIS.20210401.oa4

[30] John, K., Isong, I. A., Kebonye, N. M., Ayito, E. O., Agyeman, P. C., & Afu, S. M. (2020). Using Machine Learning Algorithms to Estimate Soil Organic Carbon Variability with Environmental Variables and Soil Nutrient Indicators in an Alluvial Soil. 1–20.

[31] Junior, M. A., Appiahene, P., & Appiah, O. (2022). Forex market forecasting with two-layer stacked Long Short-Term Memory neural network ( LSTM ) and correlation analysis. Journal of Electrical Systems and Information Technology, 1, 1–24. https://doi.org/10.1186/s43067-022-00054-1

[32] Kale, S. S. (2019). A Machine Learning Approach to Predict Crop Yield and Success Rate. 1–5.

[33] Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. Computers and Electronics in Agriculture, 147, 70–90. https://doi.org/10.1016/j.compag.2018.02.016

[34] Kamir, E., Waldner, F., & Hochman, Z. (2020). ISPRS Journal of Photogrammetry and Remote Sensing Estimating wheat yields in Australia using climate records, satellite image time series and machine learning methods. ISPRS Journal of Photogrammetry and Remote Sensing, 160(November 2019), 124–135. https://doi.org/10.1016/j.isprsjprs.2019.11.008

[35] Khalil, Z. H., & Abdullaev, S. M. (2021). ScienceDirect ScienceDirect ScienceDirect Neural network for grain yield predicting based multispectral Neural network for grain yield predicting based multispectral satellite imagery: comparative study satellite imagery: a comparative study. Procedia Computer Science, 186, 269–278. https://doi.org/10.1016/j.procs.2021.04.146

[36] Klompenburg, T. Van, Kassahun, A., & Catal, C. (2020). Crop yield prediction using machine learning: A systematic literature review. Computers and Electronics in Agriculture, 177(January), 105709. https://doi.org/10.1016/j.compag.2020.105709

[37] Kumar, A., Sarkar, S., & Pradhan, C. (2019). Recommendation System for Crop Identification and Pest Control Technique in Agriculture. 2019 International Conference on Communication and Signal Processing (ICCSP), 185–189.

[38] Kutiame, S., Millham, R., Adekoya, A. F., Tettey, M., Weyori, B. A., & Appiahene, P. (2022). Application of Machine Learning Algorithms in Coronary Heart Disease: A Systematic Literature Review and Meta-Analysis. 13(6), 153–164.

[39] Liakos, K. G., Busato, P., Moshou, D., & Pearson, S. (2018). Machine Learning in Agriculture: A Review. Ml, 1–29. https://doi.org/10.3390/s18082674

[40] Majumdar, P., Mitra, S., & Bhattacharya, D. (2021). IoT for Promoting Agriculture 4.0: a Review from the Perspective of Weather Monitoring, Yield Prediction, Security of WSN Protocols, and Hardware Cost Analysis. Journal of Biosystems Engineering, 440–461. https://doi.org/10.1007/s42853-021-00118-6

[41] Marinkovi, B. (2011). Support Vector Machines with Features Contribution Analysis for Agricultural Yield Prediction. April. Masood, A. R. M. S. (2021). Effective dimensionality reduction by using soft computing methods in data mining techniques. Soft Computing, 25(6), 4643–4651. https://doi.org/10.1007/s00500-020-05474-7

[42] Mavridou, E., Vrochidou, E., Papakostas, G. A., Pachidis, T., & Kaburlasos, V. G. (2019). Machine vision systems in precision agriculture for crop farming. Journal of Imaging, 5(12). https://doi.org/10.3390/jimaging5120089

[43] Mishra, S., Mishra, D., & Santra, G. H. (2016). Applications of machine learning techniques in agricultural crop production: A review paper. Indian Journal of Science and Technology, 9(38). https://doi.org/10.17485/ijst/2016/v9i38/95032

[44] Molero, G., & Bort, J. (2014). NDVI as a Potential Tool for Predicting Biomass, Plant Nitrogen Content and Growth in Wheat Genotypes Subjected to Different Water and Nitrogen Conditions NDVI as a Potential Tool for Predicting Biomass, Plant Nitrogen Content and Growth in Wheat Genoty. March 2011. https://doi.org/10.1556/CRC.39.2011.1.15

[45] Moreno, T. R. (2020). Crop yield prediction with ensemble algorithms and Artificial Neural Crop yield prediction with ensemble algorithms and Artificial Neural Networks ( ANN ) Dissertation to apply for the degree: June 2019. https://doi.org/10.13140/RG.2.2.33145.90722

[46] Nandhini, N., & Shankar, J. G. (2020). PREDICTION OF CROP GROWTH USING MACHINE LEARNING BASED ON SEED FEATURES. 2232–2236. https://doi.org/10.21917/ijsc.2020.0319

[47] Naveen, L., & Mohan, H. S. (2022). A novel weather parameters prediction scheme and their effects on crops. 12(1), 639–648. https://doi.org/10.11591/ijece.v12i1.pp639-648

[48] Nevavuori, P., Narra, N., & Lipping, T. (2019). Crop yield prediction with deep convolutional neural networks. Computers and Electronics in Agriculture, 163(April), 104859. https://doi.org/10.1016/j.compag.2019.104859

[49] Nica, R. E., Sebastian, M., Lucian, Șerbănescu, Florescu, M., Camen, G. C., Streba, C. T., & Gheonea, I. A. (2021). Deep Learning: a Promising Method for Histological Class Prediction of Breast Tumors in Mammography. Journal of Digital Imaging, 1190–1198. https://doi.org/10.1007/s10278-021-00508-4

[50] Palanivel, K., Surianarayanan, C., & Author, C. (2019). AN APPROACH FOR PREDICTION

OF CROP YIELD USING MACHINE LEARNING AND BIG. 10(03), 110–118.

[51] Patel, B., & Sharaff, A. (2021). Rice Crop Disease Prediction Using Machine Learning Technique. 12(4), 1–15. https://doi.org/10.4018/IJAEIS.20211001.oa5

[52] Patel, K., & Patel, H. B. (2020). A state-of-the-art survey on recommendation systems and prospective extensions. Computers and Electronics in Agriculture, 178(June), 105779. https://doi.org/10.1016/j.compag.2020.105779

[53] Peng, B., Guan, K., Pan, M., & Li, Y. (2018). Benefits of seasonal climate prediction and satellite data for forecasting US maize yield. https://doi.org/10.1029/2018GL079291

[54] Peter, A., Yaw, M. M., Rose-Mary, O. M. G., Adu-Gyamfi, D., & Micheal, O. (2022). Predicting the operational efficiency of banks using their Information Technology: Decision tree algorithm approach. Information Technologist (The), 19(1), 37–57.

[55] Pham, H. T., Awange, J., Kuhn, M., & Nguyen, B. Van. (2022). Enhancing Crop Yield Prediction Utilizing Machine Learning on Satellite-Based Vegetation Health Indices. 1–19.

[56] Prasad, D., Goyal, S. K., Bindal, A., & Kushwah, V. S. (2019). System Model for Prediction Analytics Using K-Nearest Neighbors Algorithm System Model for Prediction Analytics Using K -Nearest Neighbors Algorithm. October. https://doi.org/10.1166/jctn.2019.8536

[57] R. K. Rajak, A. Pawar, M. Pendke, P. Shinde, S. Rathod, A. D. (2017). Crop Recommendation System to Maximize Crop Yield using Machine Learning Technique. 04(12 Dec, 2017).

[58] Saha, A. K., Saha, J., Ray, R., Sircar, S., Dutta, S., & Chattopadhyay, S. P. (2018). IOT-Based Drone for Improvement of Crop Quality in Agricultural Field. January. https://doi.org/10.1109/CCWC.2018.8301662

[59] Saiz-Rubio, V., & Rovira-Más, F. (2020). From smart farming towards agriculture 5.0: A review on crop data management. Agronomy, 10(2). https://doi.org/10.3390/agronomy10020207

[60] Savla, A., & Mandholia, A. (2015). Survey of classification algorithms for formulating yield prediction accuracy in precision agriculture.

[61] Shahhosseini, M., Hu, G., Huber, I., & Archontoulis, S. V. (2021). Coupling machine learning and crop modeling improves crop yield prediction in the US Corn Belt. Scientific Reports, 0123456789, 1–15. https://doi.org/10.1038/s41598-020-80820-1

[62] Shakoor, T., Rahman, K., Rayta, S. N., & Chakrabarty, A. (2017). Agricultural Production Output Prediction Using Supervised Machine Learning Techniques.

[63] Sharifi, A. (2021). Yield prediction with machine learning algorithms and satellite images. https://doi.org/10.1002/jsfa.10696

[64] Sharma, S., Rai, S., & Krishnan, N. C. (2020). Wheat Crop Yield Prediction Using Deep LSTM Model.

[65] Shruthi, G. (2020). Design And Implementation Of Crop Yield Prediction Model In Agriculture. 8(01).

[66] SMummaleti Keerthana, K J M Meghana, Siginamsetty Pravallika, D. M. K. (2021). An Ensemble Algorithm for Crop Yield Prediction. Icicv.

[67] Sonobe, R., Yamaya, Y., Tani, H., Wang, X., Kobayashi, N., Mochizuki, K., Sonobe, R., Yamaya, Y., Tani, H., Wang, X., & Kobayashi, N. (2017). Assessing the suitability of data from Sentinel-1A and 2A for crop classification. GIScience & Remote Sensing, 54(6), 918–938. https://doi.org/10.1080/15481603.2017.1351149

[68] T.Banavlikar, A.Mahir, M. Budukh, S. D. (2018). IRJET-Crop Recommendation System using Neural Networks. 05.

[69] Tatsumi, K., Yamashiki, Y., Angel, M., Torres, C., Leonidas, C., & Taipe, R. (2015). Crop classification of upland fields using Random forest of time-series Landsat 7 ETM + data. Computers and Electronics in Agriculture, 115, 171–179. https://doi.org/10.1016/j.compag.2015.05.001

[70] Teng, H., Viscarra, R. A., Shi, Z., & Behrens, T. (2016). Assimilating satellite imagery and visible – near-infrared spectroscopy to model and map soil loss by water erosion in Australia Environmental Modelling & Software Assimilating satellite imagery and visible e near-infrared spectroscopy to model and map so. Environmental Modelling and Software, 77(41461063), 156–167. https://doi.org/10.1016/j.envsoft.2015.11.024

[71] Thomas, K. T., Varsha, S., Saji, M. M., Varghese, L., & Thomas, E. J. (2020). Crop Prediction Using Machine Learning. 13(3), 1896–1901.

[72] Ujjainia, S., Gautam, P., & Veenadhari, S. (2021). A Crop Recommendation System to Improve Crop Productivity using Ensemble Technique. February. https://doi.org/10.35940/ijitee.D8507.0210421

[73] Vasavi, P., Punitha, A., & Rao, T. V. N. (2022). Crop leaf disease detection and classification using machine learning and deep learning algorithms by visual symptoms: a review. 12(2), 2079–2086. https://doi.org/10.11591/ijece.v12i2.pp2079-2086

[74] Wang, A. X., Lobell, D., & Ermon, S. (2015). Deep Transfer Learning for Crop Yield Prediction with Remote Sensing Data.

[75] Wang, S., Azzari, G., & Lobell, D. B. (2019). Crop type mapping without field-level labels: Random forest transfer and unsupervised clustering techniques. Remote Sensing of Environment, 222(May 2018), 303–317. https://doi.org/10.1016/j.rse.2018.12.026

[76] Yasrab, R. (2021). Predicting Plant Growth from Time-Series Data Using Deep Learning. 1–17.

[77] Yesugade, P. K. D., Kharde, A., Mirashi, K., Muley, K., & Chudasama, H. (2018). Machine Learning Approach for Crop Selection based on Agro-Climatic Conditions. 7(10), 103–106. https://doi.org/10.17148/IJARCCE.2018.71023

[78] Zhou, X., Zhu, X., Dong, Z., & Guo, W. (2016). ScienceDirect Estimation of biomass in wheat using random forest regression algorithm and remote sensing data. CJ, 1–8. https://doi.org/10.1016/j.cj.2016.01.008