Machine Learning Applications in Agriculture: A Review for Opportunities, Challenges, and Outcomes

1Neelam Bohra, 2Dr. Anil Gupta

1Research Scholar, 2Professor
Department of Computer Science & Engineering
M.B.M University, Jodhpur, Rajasthan

Abstract: Agriculture plays a very important role in our life, huge manpower of the country is engaged in the agriculture sector but they lack skills, they are not aware of the latest technologies and as a result, they are not able to improve the quality of crops. Finding solutions to various problems such as disease, soil nutrition, and water conservation in plants is very challenging. This work aims to conduct a systematic review of machine learning (ML) applications in agriculture and highlight the challenges farmers face to fully adopt machine learning in agriculture. Several research papers based on the applications of ML algorithms in the development of agriculture have been studied and a comparison has been made of various machine learning algorithms used in the agriculture sector. The main goal of this work is to provide detailed information about various machine-learning techniques that can be used in agriculture recently to improve agricultural production. Automation of agricultural activities is encouraged with the help of ML and IoT tools, for example, an automated disease detection and warning system is developed which sends alerts through buzzers when a disease is detected. The use of wireless sensor networks, IoT, robotics, drones, and AI is on the rise. A huge amount of data is collected from the fields with the help of various sensors and ML models are used on these data to extract useful information and improve the yield.

Keywords – Machine Learning, deep learning, digital transformation, Agriculture, IoT

I. Introduction

The history of agriculture in India dates to the Indus Valley Civilization and even earlier in some regions of southern India. The Indian economy's primary sector is agriculture, which is also one of the country's most significant sectors and a major source of employment. According to estimates, 60% of Indians work in this sector, which accounts for around 18% of the country's GDP. However, as other sectors of the economy grow, this sector's contribution gradually declines over time (Agriculture in India: Statistics and Facts | Statista). In addition, there is pressure to expand agricultural productivity by 70% over current levels by 2050, when the world's population is predicted to reach 9.7 billion. Due to urbanization, land availability for agriculture will decrease. It is expected that by 2050 India will become the most populated country and presently it is already facing a scarcity of food. The most common reasons for less food production are lack of planning, unpredictable weather conditions, and poor harvesting and irrigation techniques. Due to global warming, the earth’s average temperature has increased and because of this heavy rainfall and droughts are regular issues faced by poor farmers [1].

For crop monitoring, disease detection, and other activities farmers are still depending on manual methods, these methods are time-consuming, but with the help of new advanced techniques, farmers can get better results in a short period. It is the correct time when automation through artificial intelligence can enter the field of agriculture. The machine learning branch of Artificial Intelligence gives the skill to a machine to learn from experience. Its algorithm does not depend on a predefined model rather they learn from a dataset. Machine learning can take better decisions without involving humans to solve real-world problems. It is observed from the following statistics that artificial intelligence and machine learning are globally used to increase the GDP:

1. It is expected that there will be an increase in investment from 1 billion dollars to 4 billion dollars in 2020-2026 which is a 25.5 percent compound growth rate (CAGR) (www.marketsandmarkets.com/Market-Reports).
2. It is anticipated that by 2025, smart, connected agriculture’s fastest growing technology IoT-enabled Agriculture (IoTAg) monitoring reaches 4.5 billion dollars of market value (The State of AI and Machine Learning in Agriculture (twenty20solutions.com)).

Several technologies can be used in agriculture nowadays: To sense the weather situation IoT related sensors can be used, the soil moisture sensor is used to monitor the moisture percentage of soil, PIR (private infrared) sensor, temperature sensor, humidity sensor, pH sensor, and Zigbee, Xebec protocols are also using. Microcontrollers like Arduino, Raspberry Pi minicomputer, and camera modules will also use, Data Mining algorithms are used to find out the diseases in the plant, for crop classification, and Machine learning algorithms are used to extract the data. When images are used as data then image processing can be done. With the help of various sensors, large amounts of sensed data are generated and these sensed data are stored in the cloud.

Precision agriculture also known as digital agriculture is a required solution for food security, in it to decide agriculture various modern information technologies, software tools, and smart embedded farm management systems are used. The main goal of precision farming is to increase the profit of farmers and decrease the loss due to environmental effect variation. The major
contributions are given by AI, data analytics, cloud computing, and blockchain technology in precision agriculture. ANN and Data Mining algorithms are used to find out the annual yield of crop diseases in the plant. For obstacle identification, deep learning technology is also used for high-automation machines in farms. Computer vision approach and ML techniques are used for the recognition of medicinal plants and weed detection; feature extraction can be done using CNN.

This review article draws on previous years' papers to explore the role of machine learning in agriculture and other challenges, which explain how machine learning can be used in agriculture to address the problems and challenges in this sector. The rest of the paper is organized as follows. II. Research Goal III. Research Field in Agriculture IV. Machine Learning Approaches V. Challenges of Machine Learning in Agriculture VI. Related work VII. Recent Work Done in Agriculture Using Machine Learning Techniques VIII. Summary Table IX. Conclusion.

II. Research Goal

This work aims to review the available research that is being done to determine how machine learning affects agriculture. To help farmers, become more productive and economically viable, multiple machine learning approaches are utilized to boost production as well as to identify solutions to a range of challenges they confront. The following research issue will be the primary subject of this review paper:

RQ1. What methods of machine learning are employed in agriculture?
RQ2. How much recent work has been done?
RQ3: What aspects of agricultural management use machine learning?
RQ4: How are datasets gathered for machine learning processing?

III. Research Field in Agriculture

In agriculture constant monitoring of the field is required for taking decision hence all machine learning application depends on real-time data therefore all other technology and tools like IOT, sensors, big data, and drones must cooperate with machine learning and give maximum benefit to farmers. There are so many fields in agriculture where machine learning works as a catalyst and work 24 hours for the farmer but still, some problems remain unsolved, or smart machine-learning solution is yet to be designed. In the following field machine learning can be employed and improve production:

1. **Crop Management**: It is responsible for future yield production; the main aim of crop management task is to produce higher yield by controlling the pest and weeds. Machine learning can be used for complete crop management, it helps to identify weeds and the occurrence of diseases in a crop.

2. **Insect Detection**: Insects are the main reason for demolishing crop production, so to prevent it farmers use pesticides to kill the insects, overuse of pesticides also affects the crop, so machine learning with drones is used to identify insects so that farmers can use pesticides in control manner.

3. **Livestock Management**: Machine learning techniques help farmers to update the well-being of their livestock by monitoring daily activity. With the help of a chip with a sensor that is connected to the user's mobile phone or computer farmers can find out their food intake and diseases can be detected 2 days before they manifest.

4. **Price Forecasting for Crops**: Machine learning and applied economics help farmers to understand cost fluctuation and other risk management factors. Price prediction helps the farmer to decide the type of crop for harvesting, for this task machine learning, depends on the agriculture dataset, and a large number of factors are input in the algorithm.

5. **Automatic Weeding**: Weed control is very important in agriculture, before the use of machine learning farmers manually detect weeds and remove them but it is very time-consuming so now machine learning is used to detect weeds and control them.

6. **Soil Classification**: Soil is a very important factor in agriculture, there are so many parameters in soil that decides which soil is best for which type of crop. Machine learning algorithm helps to find out the soil type best suited for desired crop.

8. **Water/Irrigation Management**: In India, rainfall is seasonal so Irrigation is very important here. For some crops water provided by rainfall is not sufficient, hence farmers must depend on irrigation. Machine learning algorithms help the farmer in irrigation management, and it is also used to find the pH level of water.

9. **Seed Quality**: To get the best crop quality, the quality of the seed is also very important, before using machine learning quality of the seed can be found out manually which is very time-consuming and tedious work, now machine learning is used in germination process so better quality of seed can be germinated and due to automation in this task farmer get good quality of seed in a short time.

Above are a few applications of machine learning in agriculture, but in the current scenario machine learning is the fastest growing area in agriculture whose range is from simple analytical systems to complex robotic hardware.

IV. Machine Learning Approaches
Machine learning is a branch of artificial intelligence. It is used for data analysis. This branch of artificial intelligence systems can learn, identifying the patterns between the data and making decisions with less involvement of the human.

Machine Learning takes out the knowledge from data and makes a framework for taking an intelligent decision, so we can say that the ML process is divided into three main parts i.e., data input, model building, and generalization. We can define generalization as a process that is used to find out the output for the input for which the algorithm has not been trained before [2].

Machine learning is broadly classified into mainly three categories; the following diagram describes the various subtypes of machine learning.

---

**Figure 4.1 Machine learning process**

**Figure 4.2 Categorization of Machine learning algorithm**
Classification, regression, and clustering are three types of problems that are generally handled by machine learning algorithms. The following are the commonly used machine learning algorithms and techniques:

1. **Deep Learning Techniques**
   It is a process to train multi-ANN using little data. Machine learning algorithm helps to recognize part of a face like eyes and nose but deep learning will use to get more or deep knowledge like the distance between eyes and nose for face detection and the length of the nose. It can be defined to be ANN with input and output and at least one hidden layer of units between input and output layers, the level of abstraction increases by adding this extra layer. Convolutional Neural Networks or ConvNets are the most popular type of Deep learning. [3].

2. **Gradient Descent Algorithm**: It is an iterative method whose aim is to reduce a cost function; it can be achieved by calculating the partial derivative of the function (i.e., slope or gradient). In this method, at each iteration, the coefficients are calculated by taking the negative of the derivative and to get local minima in a few iterations coefficients are reduced at each step by a learning rate multiplied by the derivative. When it reaches the minimum value of the cost function, the iterations are stopped and then there is no more reduction in the cost function. “Stochastic Gradient Descent” (SGD), “Batch Gradient Descent” (BGD), and “Mini Batch Gradient Descent” (MBGD) are three different types of methods used in Gradient Descent Algorithm. [4].

3. **Linear Regression Algorithm**: It is the simplest type of regression; in it when the relationship between the variables of the dataset is linear an effort is made to fit a straight line to the dataset. In this approach, SGD can also be used to update linear models with new data. It is not useful for non-linear relationships.

4. **Multivariate Regression Analysis**: In linear regression, there is a one-to-one mapping from the input variable to the output variable while in multiple linear regressions, there is a many-to-one mapping from many independent variables to one dependent variable. It is very complex and this technique also required sufficient knowledge of statistical techniques.

5. **Logistic Regression**: In it, the binomial outcome is obtained when an event will occur or not (0 or 1) Based on input values, for example, consider the case for predicting whether a tumor is malignant or benign is considered as a binomial outcome of Logistic Regression.

6. **Decision Tree**: It is a supervised Machine Learning algorithm that is used to solve classification and regression problems. Based on some parameters in it there is a continuous splitting of data and in this data is split into the nodes and decisions are shown in leaves. The decision variable is in the form of Yes/No in ClassificationTree and the decision variable is continuous in the Regression tree. A Decision Tree is useful for solving regression as well as a classification problem.

7. **Support Vector Machines**: It is used to solve both classifications as well as regression problems. A hyperplane is defined which is also called a decision boundary, this boundary is used to separate a set of objects belonging to different classes. The main goal of SVM is to classify the objects according to the example in the training dataset.

8. **Bayesian Learning**: In this, a prior probability distribution is selected and updated to obtain posterior distribution, and when a new observation is obtained the posterior distribution becomes a prior Bayesian Network that can handle an incomplete dataset. Over-fitting of data can also prevent by this method. In this, it is very difficult to select the prior, and posterior distribution completely depends on the prior, if a selection of the prior is not correct then there will be a wrong prediction.

9. **Naive Bayes**: Conditional probability is the base of this algorithm and in this algorithm, a probability table is a model and it is updated by training data. To predict a new observation, one must look up the class probabilities, the "probability" table is based on its feature values. The basic hypothesis is conditional independence so it is called "naive". In a real-world scenario, this assumption that all input features are independent of one another can be difficult to hold. This algorithm can be used for both continuous and discrete data.[4].

10. **K Nearest Neighbor Algorithm**: It is a classification algorithm and this algorithm uses a database in which there are data points grouped in too many classes. It classifies the sample data point given it as a classification problem. This algorithm is easy to implement and it is very economical to build a model.[4].

11. **K Means Clustering Algorithm**: This is an unsupervised type of learning. It is often used for solving clustering problems as it is more efficient computationally than hierarchical clustering when variables are large. It produces tighter clusters than a hierarchical cluster with a globular cluster and small k, but the prediction of K and its performance degrade when clusters are globular and there is a difference in size and density in the clusters [4].

V. **Challenges of Machine Learning in Agriculture**
Some challenges are faced by farmers and this limitation prevents the complete adoption of Machine Learning in the Agriculture sector; they are as follows:

5.1 Rate of literacy is very poor for Indian farmers so it is very difficult to reduce the gap between farmers and technology.
5.2 Farmers are not ready to leave their comfort zone and are ready to learn new technology but it is very difficult to raise the standard of farming.
5.3 Farming is done in a rural area where implementation of IOT and WSN, needs cloud service. Storage is a big issue as internet connectivity is not available in rural areas.
5.4 Due to varying geographic conditions accurate prediction through the machine is difficult.
5.5 There is a large capital required for hardware and software setup for digital farming.
5.6 There is huge power required for the deployment of smart sensors and electronic gadgets.

VI. **Related Work**
This section presents the research that has already been done by several researchers in the field of applying machine learning to agriculture. This review points out the important role that agriculture plays in the world economy. To figure out the role of machine learning in the agriculture sector, previous years’ papers are reviewed. It has been observed that a region's predominant weather
conditions have a major impact on crop production in general. This section presents the research that has already been done by several researchers in the field of applying machine learning to agriculture: This analysis illustrates the significant impact of agriculture also on the world economy.

In 2011 Seshadri Baral1, Asis Kumar Tripathy, and Pritirajan Bijayasingh used Artificial Neural Network (ANN) technology with PSO as an optimization technique for the approximation and prediction of paddy yield at 3 different districts in different climatic zones based on 10 years of historical data sets of yields of paddy, daily temperature (mean and maximum) and precipitation (rainfall). The overall error rate for the model is 8% which is in the allowable error range. The model can be given extra input parameters like soil types, nutrient contents, and variety of crops to get more accurate results with reduced error rates [5].

In 2012 Rozman Črtomir, Cvelbar Urška, and Tojanko Stanislav developed a hybrid model based on image analysis and neural network which showed that early information on yield has special importance in intensive apple production. After performing several ANN training procedures using the number of fruits at four or five different dates obtained by image analysis it can be concluded that the ANN was able to improve yield prediction accuracy for both analyzed varieties: ‘Golden Delicious’ and ‘Braeburn’ results and conclusion of this study provide a new possible approach for apple yield predictions before harvesting [6].

In 2013 M. Gunasundari Ananthara and his co-authors suggested a crop yield prediction model (CRY) which works on an adaptive cluster approach over a dynamically updated historical crop data set to predict the crop yield and improve the decision-making in precision agriculture. They proposed a relational cluster Bee Hive algorithm for extracting yield patterns across multiple data sets. From the extracted patterns, graphs are plotted to illustrate yield variation at any region of discussion or type of crop in different regions. The outcome helped in the identification of an investigation of areas of unusually high or low yield. This study also shows that many approaches are used to predict crop yield and this improves decision-making in precision agriculture [7].

In 2015 authors applied clustering techniques to divide regions and then suitable classification techniques were applied to obtain crop yield predictions. The dataset used in this research work has been collected from BARI (Bangladesh Agricultural Research Institute). The authors applied the K-means clustering technique through Rapid Miner Studio software. Linear Regression, k-NN (k-nearest neighbor algorithm), and Neural Net (an artificial neural network (ANN) is a mathematical model) were used to obtain the crop yield prediction results. Clustering was applied by authors to find out if any strong correlation exists between crop yield and different attributes (i.e., weather attributes, soil PH and soil salinity attributes, and area cultivated attributes). There were some similarities between the clusters obtained using different attributes with the clusters obtained according to crop yields, but any exact or strong correlation between crop yield with weather/ soil/ cultivate area attribute was not observed. ANN provides a better prediction for some of the crops, which have more missing values than others. In this work 5 environmental variables, 3 biotic and 2 areas-related variables were used to determine crop yield in different districts, for future work they proposed geospatial analysis can be added for data processing which helps to improve accuracy, and a recommendation system can be enhanced, and incorporate the time between seeding and harvesting for different crop species across different districts [8].

In 2016 Kim Arild Stee, Peter Christansen, and co-author proposed an algorithm for obstacle detection in the agricultural field. The presented algorithm in this paper is based on an existing deep convolutional neural network, which is fine-tuned for the detection of a specific obstacle. It was observed that fine-tuned deep convolutional network can detect obstacles with a precision of 99.9% in row crops and 90.8% in grass mowing, while simultaneously not detecting people and other very distinct obstacles in the image frame. Deep convolutional neural networks have demonstrated outstanding performance in various vision tasks such as image classification, object classification, and object detection. The authors concluded that the standardized obstacle presented in the standard is not fully sufficient to ensure safe operations with highly autonomous machines in agriculture and further work should be conducted to describe an adequate procedure for testing the obstacle detection performance of highly autonomous machines in agriculture [9].

In 2017 Lots of work has been done to generate an automated system to recognize species of plants by analyzing the digital images of their leaf. Amala Sabu, Sreekumar K, and Rahul R Nair had done work to focus on the importance of medicinal plants and how to preserve the knowledge of medicinal plants in the form of digital form by using the concept of machine learning. Their work proposed a computer vision approach for the identification of ayurvedic medicinal plant species found in the Western Ghats of India by applying machine learning which uses features extracted from the leaf images with a classification algorithm. They created their dataset for experiments. The images of leaves were captured using a digital camera of resolution 14.1 MP. The data set contains a total of 200 images spread across 20 classes. Each class contains images of 10 distinct leaves collected from different plants. Their proposed system makes use of computer vision and machine learning approaches to identify a pre-trained medicinal plant from its leaf. The main highlight of this work is the non-use of typical shape and color features of leaves which are computationally expensive to extract as they are spatial features. Their proposed system provided nearly 100% of accuracy when experimented with KNN classifier. Scope for future work includes expanding the leaf dataset, changing the classifier from KNN to SVM or ANN, and experimenting with a blend of more features added with HOG and SURF [10].

In 2017 Adams Begue, Venitha Kovlessur, and co-authors presented a paper “Automatic Recognition of Medicinal Plants using Machine Learning Techniques” The main highlight of this work was the creation of a database of medicinal plants which was found on the island of Mauritius. For creating this database leaves from 24 different medicinal plant species were collected and photographed using a smartphone. A Java programming environment with the open-source Weka machine learning workbench was then used to assess the performance of the system. Five different machine learning classifiers i.e., random forest, k-nearest neighbor, naive Bayes, support vector machine, and neural network were used to assess the recognition rate. Many features were extracted from each leaf such as its length, width, perimeter, area, number of vertices, color, perimeter, and area of the hull. Several derived features were then computed from these attributes. The best results were obtained from a random forest classifier using a 10-fold cross-validation technique with an accuracy of 90.1%. Authors suggested that for future research, to achieve higher accuracies, probabilistic neural networks and deep learning neural networks would be used and a web-based or
mobile computer system for the automatic recognition of medicinal plants could be designed which help the local population to improve their knowledge on medicinal plants, help taxonomists to develop more efficient species identification techniques [11].

In 2018, Anna X. Wang, Caelin Tran, and co-author showed results in predicting soybean crop yield in Argentina using deep learning techniques. To perform this crop yield prediction task with remotely sensed data, they used Moderate Resolution Imaging Spectroradiometer (MODIS) satellite imagery, which provides free and easy-to-access coverage. They also achieved satisfactory results with a transfer learning approach to predict Brazil’s soybean harvests with a smaller amount of data. The reason for encouraging transfer learning was that the success of deep learning models was largely dependent on abundant ground truth training data. Successful crop yield prediction with deep learning in regions with little training data relies on the ability to fine-tune pre-trained models [12].

In 2018 Fabrizio Balducci, Donato Impedovo, and Giuseppe Pirlo showed how to manage heterogeneous information and data coming from real datasets that collect physical, biological, and sensory values. The objective of this study was to design and deploy practical tasks, ranging from crop harvest forecasting to missing or wrong sensors data reconstruction, exploiting and comparing various machine learning techniques to suggest toward which direction to employ efforts and investments. In this work, three different datasets were used that differ from each other by origin; structure; organization; and availability of their values. The following dataset was used in this work:

1. The well-structured and publicly available Istat dataset.
2. The second one related to industrial IoT sensors, the reconstruction, and forecasting of IoT missing or wrong data, as well as the detection of faulty hardware sensors from monitoring stations, is performed by exploiting several machine learning methods.
3. The third one was the publicly available scientific National Research Council (CNR) dataset.

In this work, intelligent systems were developed with machine learning algorithms (supervised and non) to manage fault tolerance and hardware malfunction prediction [13].

In 2018 Olakunle Elijah, Tharek Abdul Rahman, and co-author presented a paper that highlighted the many benefits and challenges of IoT. This paper presented the IoT ecosystem and how the combination of IoT and data analytics was enabling smart agriculture. This work also discussed future trends and opportunities which are categorized into technological innovations, application scenarios, and business. In this work, the IoT ecosystem for agriculture was presented. It has consisted of four major components: 1. IoT devices 2. Communication technology 3. Internet and 4. data storage and processing. Many benefits can be obtained from the use of IoT in agriculture. Following are the few benefits: 1. Community Farming 2. Safety Control and Fraud Prevention 3. Competitive Advantages 4. Wealth Creation and Distributions 5. Cost Reduction and Waste 6. Operational Efficiency 7. Awareness 8. Asset Management In this work so many challenges related to the deployment and application of IoT had also discussed like Business Issue 1) Cost 2) Business Models 3) Lack of Adequate Knowledge 2. Technical Issues: 1) Interference, 2) Security and Privacy. 3) Choice of Technology 4) Reliability 5) Scalability 6) Localization 7) Optimization of Resources. 3. Sectorial Issues: 1) Regulatory Challenges 2) Interoperability. This paper also presented the future trends based on the following areas: 1) technological innovations; 2) application scenarios; and 3) business and marketability [14].

In 2018 Manjunath Jogan and co-author had done work on feature extraction of an image using a convolutional neural network by using the concept of deep learning and classification algorithms were also implemented for various applications. Image classification is a process that involves the following tasks: pre-processing the image (normalization), image segmentation, extraction of key features, and identification of the class. The authors proved with the idea of deep learning, the models can be trained better and able to identify different levels of image representation. CNNs is an often-used architecture for deep learning and has been widely used in computer vision and audio recognition. The Convolutional neural networks from AlexNet to GoogLeNet have tremendously developed and have accuracy greater than humans. In this work, various classification algorithms are implemented by authors using the concept of deep learning. Results showed that convolutional Neural Networks will give 85.97% accuracy for image classification. Further, it can be used for various video surveillance and security-related applications for feature extraction and classification [15].

In 2019 Om Tiwari, Vidit Goyal, and co-author presented a paper at the IEEE 4th International Conference on Internet of Things: Smart Innovation and Usages (IoT-SIU) which was held in Ghaziabad having the title “An experimental set up for utilizing a convolutional neural network in automated weed detection”. From this work, it was observed that one of the major factors that lead to a decrease in crop production is the presence of weeds. The weeds absorb the nutrients and water which results in weight reduction in plants and also decreases the grains per ear and grain yield. The traditional methods used for weed destruction is based on manual spraying of herbicides, due to which overall cost of the crop increased this was the reason that there was an urgent need for developing a method which would automatically discover these weeds in the field and then herbicide was sprayed on them to destroy them by using new drone technology and deep learning in the field of convolutional neural networks. In this work, the dataset from the fields of the Indian Agriculture Research Institute (IARI) was used. This work also showed that the Inception model (V2) used in transfer learning was very effective for detecting weeds in crops. This technique helped us to remove the difficult part and time-consuming feature extraction from images to train models. Transfer learning uses pre-trained models which do not need a lot of data to train and enhance the performance of neural networks. The future provision can be added as the model which is trained for detecting weeds can also be trained for detecting Diseases, Nutrients, and water requirement deficiency in plants by deep learning (transfer learning) neural networks. A robotic arm can be added to the Drone which can be used to destroy weeds and diseases by different chemical use to destroy or kill other organisms [16].
In 2019 Saptarshi Sengupta and co-author presented a review on Deep learning, this review presented several Deep Learning architectures starting from the beginning to recent development. The author also discussed the blend of swarm intelligence in Deep Learning approaches and how they can be used in real-world problems. The use of deep learning architectures is increasing, so the question comes to our mind about the reliability of architecture, to answer this question this review gives an overview of testing neural network architecture and different methods for adversarial test generation as well as countermeasures to be adopted had also discussed. This review also focused on specific applications of deep learning including Medical Imaging, Prognostics and Health Management, Applications in Financial Services, Financial Time Series Forecasting, and lastly the applications in Power Systems. It also highlighted a few open areas of research and addressed challenges that lie within. The following challenges are discussed: 1. Challenges with the scarcity of data 2. Adopting unsupervised approaches, 3. Influence of cognitive neuroscience 4. Neural networks and reinforcement learning. The main goal of this work is to help the beginner as well as the practitioner in the field of deep learning architectures as well as an exploratory dissection of some applicable application areas. In this paper, a system was proposed which can identify the plant species based on the input leaf sample. This work marked the problem of identifying the medicinal plant species by the analysis of leaf images obtained directly from their habitat and irrespective of lighting conditions. The result showed that the algorithm can adequately segment the leaf region. This method worked well in images with reflection. Classification of the plant species based on the leaf requires some preprocessing. The classification of medicinal plant species is done by using Weka and the accuracy of 93.3% is measured [17].

In 2019 Sivaranjani. C and co-author proposed a system that can identify the plant species based on the input leaf sample. The plant species were classified by the color and texture features on each extracted leaf using a Logistic Regression classifier with an accuracy of 93.3%. Classification of the plant species based on the leaf requires some preprocessing. In this work, the author highlights the problem of identifying the medicinal plant species by the analysis of leaf images obtained directly from their habitat and irrespective of lighting conditions. This method worked well in images with reflection. The feature extraction based on the color and texture features had done. The classification of medicinal plant species is done by using Weka and an accuracy of 93.3% is measured. In the future author suggested planning a design and developing a system that automatically identifies plant species through the analysis of not only the leaf images but also the other parts of the plant acquired directly in their habitat irrespective of complex backgrounds and various lighting conditions.[18].

In 2019 V S Magomadov presented a paper in the Journal of Physics on deep learning the objective of this paper was to explore deep learning. This paper covered some major DL-based technologies and their applications in the field of agriculture. In this work it was observed that DL has been applied in various areas of agriculture, for example, plant classification, plant disease detection, fruit counting, etc. [19].

Yemeserach Mekonnen and coauthors presented a Review paper on —Machine Learning Techniques in Wireless Sensor Network Based Precision Agriculture in 2019, in this review paper a case study had done on an IoT-based data-driven smart farm prototype as an integrated food, energy, and water (FEW) systems. This paper reviewed the Machine learning methods mostly used by researchers in the past two years in conjunction with wireless sensor networks. This paper had shown different Machine learning models applied in multiple applications within the precision agriculture ecosystem, including yield prediction, weed, and disease detection. By applying Machine learning Techniques to sensor data, farm management systems were evolving into real Artificial Intelligence systems, providing optimal insights for decisions and actions to be made. The implementation of smartphone applications and the back-end data analysis framework for the prediction of weather, crop yield, and crop quality were also presented [20].

Abirami Devaraj and coauthor in the year 2019 developed a software system that finds and classify disease. The leaf pictures were used for detecting plant diseases. Therefore, the use of the image process technique to find and classify diseases in agricultural applications is useful. This proposed automatic illness detection used image processing techniques in MATLAB. The development of an automatic disease detection system using advanced technology like image processing facilitates support the farmers in the identification of diseases at an early or initial stage and supplies helpful data for its management. Authors suggested in the future to extend their work additional on a lot of disease detection [21].

In 2019 Vanessa Rezende, Michel Costa, and coauthors focus on performing image processing on a plant disease database for enhancing the classification of 20 different diseases from 10 distinct plant species. In this work processing encompassed two pre-processing activities (image selection and resizing) and the application of two modified VGG architectures, VGG16 and VGG19, along with pre-trained weights from the ImageNet database. A comparison study was carried out based on classification metrics, such as accuracy, precision, recall, and F1-score. The obtained results demonstrated that for this database, a pre-trained CNN with depth equal to or smaller than VGG16 can compute disease-sensitive features which aggregate more refinement to recognize plant pathologies, which justified the better performance of VGG16 against VGG19 on training, validation, and test data [22].

Mukesh Kumar Tripathi and Dr. Dhananjay D. Maktekar in the year 2020 studied the different types of diseases present in various fruit and vegetable. The authors also focused on the comparison of different machine-learning approaches for different performance metrics on the same dataset. Thus, it found that among all existing machine learning techniques, SVM gives better classification accuracy [23].

In 2020 Pradeep Kumar Mugithe and coauthors had done work which showed that in recent years, crop yield was not as
expected because the plants are drastically affected by diseases therefore the farmer had to apply the pesticides for the entire farm periodically but it did not only result in wastage of manpower, money, and time, but also in the increase of toxic levels in vegetables. Consuming such vegetables affect both animal and human life. Hence it was necessary to obtain a balance between high-yield and less toxic vegetables. It was possible only when pesticides were applied to disease-affected plants only. Due to the large size of the farm, manual inspection was not possible therefore it was necessary to develop an automated disease-detecting and alerting system. Authors proposed such a system that after detecting the disease, sends alerts through the buzzer [24].

K Dokic, I Blaskovic, and D Mandusic in 2020 had done a review on the topic: From machine learning to deep learning in agriculture this work, it had found that the development of artificial intelligence and machine learning in the last two decades had led to a significant increase in the number of projects in the field of agriculture. Artificial intelligence is a general term that includes machine learning, neural networks, and deep learning. It is observed if we look at the production of scientific papers by country, it is evident that China dominates in production. A particularly significant difference was observed in the number of published scientific papers about deep learning in the Web of Science citation database. Papers about deep learning published in the first half of 2020 and the trends that are looming are particularly attractive. It can be concluded that TensorFlow is a dominant framework, but most researchers use Keras to simplify the process of constructing neural networks. A camera had become the dominant input device, and images are the most common input medium used by neural networks. Wheat, corn, and grape are most often the focus of scientists, and crop analysis is the most common area of deep learning application. Finally, the use of convolutional neural networks has become commonplace in solving problems in agriculture. Authors conclude that “the future in agriculture has already come,” and we just need to use the available resources and knowledge [25].

In 2020 Abhinav Sharma and his coauthor presented an article that demonstrated how knowledge-based agriculture can improve the sustainable productivity and quality of the product. In this article, the authors presented a systematic review of ML applications in the field of agriculture. The areas that were focused on are the prediction of soil parameters such as organic carbon and moisture content, crop yield prediction, disease, and weed detection in crops, and species detection. ML with computer vision was reviewed for the classification of a different set of crop images to monitor crop quality and yield assessment. This approach can be integrated for enhanced livestock production by predicting fertility patterns, diagnosing eating disorders, and cattle behavior based on ML models using data collected by collar sensors, etc. Intelligent irrigation which includes drip irrigation and intelligent harvesting techniques was also reviewed that can reduce human labor to a great extent [26].

In 2020 Nikita Genze and coauthor presented a paper that emphasized that for getting a better-quality crop, the quality of seed is also very important and for this seed assessment task has been done which is done manually so obviously there are higher chances of getting errors. In this paper, the author proposed a machine-learning approach using modern artificial neural networks for three different crops, Zea mays (maize), Secale cereale (rye), and Pennisetum glaucum (pearl millet), with a total of more than 23,000 images with region proposals for precise seed germination detection. In this work, Faster R-CNN is used to find out whether a seed is germinated or not. R-CNN uses two neural networks Region Proposal Network (RPN) and convolutional neural network (CNN). RPN is used to find out the location of the seed and CNN is used to find out whether a seed is germinated or not. Their proposed models attained a high mean average precision (mAP) on a Zea mays, Secale cereale, and Pennisetum glaucum approximately 97.9%, 94.2%, and 94.3% respectively [27].

In 2021 at the University of Bahrain A. AlKameli and M. Hammad presented a review of existing applications of machine learning in agriculture with a focus on the applications of Deep Reinforcement Learning techniques in agriculture. They gave conventional solutions for agricultural machine learning decision-making problems using supervised approaches. In supervised learning, the machine needs to be trained on samples of inputs and outputs to support decision-making. While, in reinforcement learning, sequential decision-making happens and the next input depends on the decision of the machine. In this paper, they performed a survey of 49 publications of which 10 were secondary research reports that discussed a variety of machine learning approaches applications in agriculture and 39 research reports that used machine learning approaches to support the automation of agricultural activities [28].

VII. Recent Work Done in Agriculture Using Machine Learning Techniques

After reviewing so many research papers, it is observed that in previous years research on agriculture using machine learning techniques was limited to predict crop yield, identification of plant species, and automatic object detection, but in the last two decades, machine learning shows its usefulness in other areas of agriculture like processing, production, marketing, and distribution of crops and livestock production and cost estimation. Now it is clear that agricultural activities serve as the basic source of income, expand GDP, providing raw materials for production in industries so overall continuous contribution to GDP. Machine learning with big data and high-performance computing are used for data-intensive processing in agriculture. In the past machine learning algorithm like artificial neural networks. K-means, support vector machines, decision trees, etc. are commonly used algorithms. In this section, some recent work done in 2022 is highlighted. In this section work done in various areas of agriculture is also reviewed.

In Disease Detection:
Javeria Amin, Muhammad Almas Anjum, and coauthor work on cotton crop in their paper “Explainable Neural Network for Classification of Cotton Leaf Diseases.” In this, they concluded that various diseases that affect cotton crop yield occur on the leaf for example powdery mildew, leaf curl, leaf spot, target spot, bacterial blight, and nutrient deficiencies. It is observed that if diseases were detected at early-stage crops can be protected from harm. Various computerized methods play a very important role in cotton leaf disease detection at an early stage. The method given by authors in this work consists of two core steps i.e., feature extraction and classification. The author proposed a method, in which data augmentation is applied to balance the input data then features are extracted from a pre-trained VGG-16 model and passed to 11 fully convolutional layers, which freeze the majority and randomly initialize convolutional features to subsequently generate a score of the anomaly map, which defines the probability of the lesion region. This work proposed a model which was trained on the selected hyperparameters that produce great classification results. The proposed model performance is evaluated on two publicly available Kaggle datasets, Cotton Leaf and Disease. The proposed method provides 99.99% accuracy [29].

Hamna Waheed and Noureen Zafar and others also work for the early identification of disease, to find pest patterns, and if there is any deficiency in nutrition present in the ginger plant. For this purpose, the authors used deep artificial neural networks and deep learning-based methods. A real-field dataset was used in this work which consists of both healthy and affected leaves of the ginger plant. They concluded from this work that for disease detection convolutional neural network (CNN) was best-having accuracy of 99%. Pest pattern leaves can be identified by VGG-16 with an accuracy of 96% and ANN was considered best for the identification of nutritional deficiency-affected leaves but results obtained in this work was completed for the dataset used in this work and if a dataset is changed may be different results will be obtained [30].

Md. Reduanul Haque and Ferdous Sohel proposed a system that automatically detects disease in eggplant crop leaves. In this work, authors used two-stream deep fusion architecture of CNN-SVM and CNN-Softmax pipelines, with an inference model to conclude the nine diseases of an eggplant crop. The authors used a dataset that consist of 2284 images of nine disease classes of which 1842 images were taken from various sites in Bangladesh and the source of 442 images was the internet. In this work all images collected were infectious and they are categorized into the following classes: aphids (class 0), bacterial wilt (class 1), Cercospora melongenae (class 2), collar rot (class 3), eggplant–Colorado potato beetle (class 4), little leaf disease (class 5), spider mites (class 6), Phomopsis blight (class 7), and tobacco mosaic virus (class 8). In this work, it was concluded that the proposed methods had achieved better accuracy and fewer false-positive results compared to VGG16, Inception V3, VGG 19, MobileNet, NasNetMobile, and ResNet50 [31].

Changguang Feng, Minlan Jiang, and co-authors worked on rice crops to identify rice blast disease which is the most destructive disease. The size of the area occupied by the disease spot on leaves showed the severity of the disease and traditionally it was checked manually which was time-consuming so correct and fast segmentation of rice blast is of much beneficial for improving the rice yield. It was found that most of the research had been done for the identification of rice disease types only a few works had been done on disease segmentation. In this work, authors proposed a real-time semantic segmentation method called DFFANet network which consists of modules like feature extraction module, feature fusion module, and attention mechanism module. In this work, authors used the following evaluation criteria for evaluating the segmentation model: the average intersection rate (abbreviated as MioU), inference speed (abbreviated as FPS), and several parameters (abbreviated as Params), and IoU is the ratio of the intersection and union of the prediction result. They concluded that the proposed method gains 96.15% MioU, a speed of 188 FPS, and the number of parameters is only 1.4 M, which results in a better balance between accuracy and speed hence decreasing the requirement for hardware and giving support for real-time fast detection of rice diseases [32].

Damar Novtahaning, Hasnain Ali Shah, and Jae-Mo Kang had done a research study intending to identify the four main types of disease phoma, miner, rust, and Cercospora present on coffee leaves using deep learning. In their proposed method, they were able to find out the deep characteristics of coffee plant leaves by using transfer learning and various pre-trained CNN networks. To determine the results authors, choose the best three models and concatenated them for building an ensemble architecture along it with increasing the quality and data sample quantity data pre-processing and augmentation methods were also applied. The outcome of this study was that when compared with other DL models this proposed ensemble architecture achieved 97.31% validation [33].

Insect Detection:

Tiago Domingues, Tomás Brandão, and others worked for insect detection on a tomato crop field, in this work authors showed the use of the YOLOv5 object detection model on images collected from a tomato crop field for insect detection in yellow sticky traps. For finding the insect population, assigning management, and evaluating the effectiveness of treatment in projects Insect traps were used. The sliding window approach was used in the insect detection process to minimize the duplicate detection on yellow sticky card IT images. The YOLOv5 is doing well in differing element exposure conditions. The work done in this paper focused on insect detection and monitoring as well as it also helped to identify diseases and pests. The authors proposed a method that shows good results compared to related works, 94.4% for mAP_0.5, with a precision and recall of 88% and 91%, respectively, using YOLOv5x [34].

Wei Li, Tengfei Zhu, Xiaoyu Li, Jianzhang Dong, and Jun Liu also worked on insect detection. For this work, they used...
In Determination of Growing Stages:

Rodney Tai-Chu Sheng, Yu-Hsiang Huang, and coauthors had done work for rice growth stages. In this paper they highlighted the issue related to the complete growth of rice crop, which must pass through various stages like a vegetative stage, reproductive stage, and ripening stage, it is very difficult for a new cultivator to observe the conversion from one stage to other and along it aging and labor cost also increase the overall cost of agriculture, to solve this problem authors proposed a random forest (RF)-based machine learning (ML) classification model for rice growth stages. To conduct the experiment a setup was installed in the experiment fields which consisted of an HD smart camera (Speed-dome) to collect the image and video data, along with Internet of Things (IoT) devices such as 7-in-1 soil sensors, a weather monitoring station, flow meter, and millimeter connected with LoRa base station for numerical data. After that various image processing techniques like object detection, object classification, instance segmentation, excess green index (EGI), and modified excess green index (EGI) were used to calculate the paddy height and canopy cover (CC) or green coverage (GC). These values are used as input in the proposed ML model and, to develop a model various growth-related factors such as height, CC, accumulative temperature, and DAT were considered the result obtained from this optimal model has an accuracy of 0.98772, and a macro F1-score of 0.98653, so authors concluded that this developed model can be used in real-world scenarios [36].

In Yield Prediction:

Mario Lillo-Saavedra, Alberto Espinoza-Salgado, and another coauthor presented a work in which a methodology was proposed for the early prediction of tomato yield by decision tree ensembles, vegetation spectral indices, and shape factors from images since early yield prediction helps a farmer to take a decision and do required planning. For this task images were captured by multispectral sensors board on an unmanned aerial vehicle (UAV) during different phenological stages of crop development. In this work, the predictive model was developed and the tomato yield was estimated for a 0.4 ha plot, having an error rate of 9.28%. The estimation was done based on training characteristics for 6 weeks before harvesting. The authors used a DTE algorithm for forecasting tomato yields. The approach behind generating this model using DTE was obtaining a set of individual decision trees (DT), each of which is trained with a sample slightly different from the training data. The prediction of a new observation is obtained by adding the predictions of all individual trees that form the model and creating ensembles bagging and boosting techniques were used Bagging and boosting are common ways to create ensembles of DTs. Bagging is a technique used to reduce the variation of predictions via a combination of results from various classifiers, each modeled with different subsets taken from the same population and boosting consists of changing the results from various weak classifiers to obtain a robust classifier [37].

Classifying the Factors for Regional Raw Milk Prices and Weedy rice seed:

Svetlana Kresova and Sebastian Hess had done a study regarding the regional raw milk price variation using the machine learning algorithm, the objective of this work was to find out the factors for the variation of milk prices in Russia between 2015-2019. For this study, authors took official data from Russia from 2015 to 2019 and 12 predictor variables were used to analyze these data. The authors used a machine learning algorithm in which model training and hyper-optimization were done with a spatiotemporal cross-validation technique. After this study, results show that the following factors were responsible for variation in raw milk prices in Russia: milk production, income, livestock numbers, and population density. This study also concluded that an increase in drinking milk production and livestock numbers in a region was a reason for a decrease in raw milk price, but an increase in income results in an increase in milk price. It was also observed that the relationship between milk prices and population density was non-linear. This study helps stakeholders decide on the dairy supply chain in Russia, milk producers, and dairy companies [38].

Rashidah Ruslan, Siti Khairunniza-Bejo, and others worked on the classification of weedy rice seeds and cultivated rice seeds. It was difficult to classify weedy seeds because these seeds have the same taxonomic and physiological features as cultivated rice seeds. It was important to identify the weedy seeds because they reduced the rice yield by up to 60%. The certified authority checks the seed quality manually by counting and visually inspecting so to address this problem authors presented this work and they used image processing and machine learning techniques to classify weedy rice seeds. In this work, authors set up
a machine vision unit for image acquisition using an area scan camera for the Red, Green, and Blue (RGB) and monochrome images of five cultivated rice varieties and a weedy rice seed. For this work dataset which consists of five varieties of cultivated seed and weedy seed were taken from the Department of Agriculture, Teluk Cengai, Kedah, a total of 7350 cultivated seeds were used, with 1470 seed kernels for each variety and about 895 weedy seeds were obtained during the search for weedy from the cultivated seed samples. In this proposed work 70% of the data was used for training and 30% for testing datasets. The proposed method used only three parameters morphology, color, and texture, and sixty-seven features were extracted and used as the input for machine learning. In this work, seven machine learning classifiers were used, and the performance metrics: sensitivity, specificity, accuracy, and the average correct classification were used for analyzing classifiers. They concluded that the best optimum model was developed by the RGB image using the logistic regression (LR) model that achieved 85.3% sensitivity, 99.5% specificity, 97.9% accuracy, and 92.4% average correct classification utilizing all the 67 features [39].

Determination of Optimal Pollination Time

Yali Zhang, Luchao Bai, and others addressed the problem of detection of rice spikelet flowering and find out that it is difficult to determine the optimal pollination timing for hybrid rice seed production. This task is accomplished by farmers by manual observation obviously which has low efficiency and high error rates. The authors proposed a model which completes this task automatically with faster speed and for this, they obtained spikelet flowering information by using a hyperspectral technique and machine learning. In this work, Hyperspectral data of rice male parents with flowering and non-flowering were collected by two experimental sites with an ASD FieldSpec® HandHeld™2 spectrometer and three classifiers, Random Forest (RF), Support Vector Machine (SVM) and Back Propagation (BP) neural network, and Convolutional Neural Network (CNN), was used to build classification models for rice spikelets flowering detection. Three data processing methods, PCA feature extraction, GA feature selection, and the PCA and GA combination algorithm, were used for data dimensionality reduction. They compare the precision and recall rate of different algorithms and data processing methods. They concluded that the optimal model for rice spikelets flowering detection is the BP model with PCA feature extraction. The accuracy of the model was up to 96–100%. Hyperspectral technology and machine learning algorithm are capable of effective detection of rice spikelet flowering. Their work helped farmers to determine the optimal operation time for supplementary pollination of hybrid rice[40].

In Quality Checking

Ting-Wei Huang, Showkat Ahmad Bhat, and others worked to help the farmers find out the quality of pineapple which is classified based on the percentage of water, there are two types of pineapple drum sound pineapple and meat sound pineapple, meat sound pineapple contains more percentage of water compared to drum sound pineapple so there are more chances of meat sound pineapple to get rotten. A farmer must do extra care for these pineapples and to sell them overseas it is required to classify them so the authors proposed a method that classifies the pineapple automatically. The authors proposed an automatic method that was based on embedded onboard computing processors, a servo, and an ultrasonic sensor, they built a hitting machine and combined it with a conveyor to automatically separate pineapples. To classify pineapples, the proposed method was associated with acoustic spectrogram spectroscopy, which used acoustic data to generate spectrograms. In the acoustic data collection step, they used the hitting machine mentioned before and collected many groups of data with different factors; some groups also included the noise in the farm. With these differences, authors tested deep learning-based convolutional neural network (CNN) performances. They observed that Group V data had the best accuracy of the developed CNN model and it was 0.97. They concluded that their proposed hitting machine and the CNN model can assist in the classification of pineapple fruits with high accuracy and time efficiency [41].

VIII. Summary Table

In the following table, we try to summarize the work done by various researchers from previous years and we conclude that new advanced technology with machine learning continuously contributes to the field of agriculture and try to increase production. It is observed that since 2018 concepts of deep learning are also implemented in various applications.

<table>
<thead>
<tr>
<th>S. No</th>
<th>Title</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Yield Prediction Using Artificial Neural Networks[5] in 2011.</td>
<td>In this work, the data obtained for the model is processed for missing values using a neighboring mean method in which the missing value is the mean of its adjacent values.</td>
</tr>
<tr>
<td>2</td>
<td>Application of Neural Networks and Image Visualization for Early Forecast of Apple Yield[6] in 2012.</td>
<td>In this paper, an attempt was made to improve image analysis-based apple yield forecasts with an application of artificial neural networks (ANN).</td>
</tr>
<tr>
<td>3</td>
<td>CRY – An improved Crop Yield Prediction model using Bee Hive Clustering Approach for agricultural data sets [7] in 2013.</td>
<td>In this paper, a relational cluster Bee Hive algorithm is proposed for extracting yield patterns across multiple data sets.</td>
</tr>
<tr>
<td>4</td>
<td>Applying Data Mining Techniques to Predict the Annual Yield of Major Crops and Recommend Planting Different Crops in Different Districts in Bangladesh [8] in 2015.</td>
<td>In this, the effects of environmental (weather), biotic (pH, soil salinity), and area of production are considered as factors towards crop production into clustering techniques to divide regions; and then apply suitable classification techniques to obtain crop yield predictions.</td>
</tr>
<tr>
<td>5</td>
<td>Using Deep Learning to Challenge Safety Standards for Highly Autonomous Machines in Agriculture [9] in 2016.</td>
<td>In this, an algorithm is proposed for obstacle detection in an agricultural field. The algorithm is based on an existing deep CNN and it can detect obstacles with a precision of 99.9% in row crops and 90.8% in grass mowing.</td>
</tr>
<tr>
<td>6</td>
<td>Recognition of Ayurvedic Medicinal Plants from Leaves: A Computer Vision Approach [10] in 2017.</td>
<td>The proposed system makes use of computer vision and machine learning approaches to identify a pre-trained medicinal plant from its leaf. It uses a blend of SURF and HOG features which provided nearly 100% of accuracy when experimented with a k-NN classifier.</td>
</tr>
<tr>
<td>8</td>
<td>Deep Transfer Learning for Crop Yield Prediction with Remote Sensing Data [12] in 2018.</td>
<td>It encourages transfer learning as the success of deep learning models is largely dependent on abundant ground truth training data. Successful crop yield prediction with deep learning in regions with little training data relies on the ability to fine-tune pre-trained models.</td>
</tr>
<tr>
<td>9</td>
<td>Machine Learning Applications on Agricultural Datasets for Smart Farm Enhancement [13] in 2018.</td>
<td>It design and deployment of practical tasks, ranging from crop harvest forecasting to missing or wrong sensors data reconstruction, exploiting and comparing various machine learning techniques to suggest toward which direction to employ efforts and investments.</td>
</tr>
<tr>
<td>10</td>
<td>An Overview of Internet of Things (IoT) and Data Analytics in Agriculture: Benefits and Challenges [14] in 2018.</td>
<td>An IoT ecosystem for agriculture is presented. It consists of four major components: 1) IoT Devices 2) communication technology 3) Internet and 4) data storage and processing.</td>
</tr>
<tr>
<td>11</td>
<td>Feature Extraction using Convolution Neural Networks (CNN) and Deep Learning” [15] in 2018.</td>
<td>The concept of deep learning and classification algorithms are also implemented for various applications.</td>
</tr>
<tr>
<td>12</td>
<td>An experimental setup for utilizing a convolutional neural network in automated weed detection [16] in 2019.</td>
<td>Weed detection using Transfer learning in Convolution Neural Network (CNN) It helps us to remove the difficult part and time-consuming of feature extraction from images to train Model.</td>
</tr>
<tr>
<td>Table 8.1 Summary Table</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Time Identification of Medicinal Plants using Machine Learning Techniques [18] in 2019.</td>
<td>A system is proposed which can identify the plant species based on the input leaf sample. The classification of medicinal plant species is done by using the Weka tool and an accuracy of 93.3% is a measure.</td>
</tr>
<tr>
<td>15</td>
<td>Deep learning and its role in smart agriculture [19] in 2019.</td>
<td>DL-based technologies and their applications in the field of agriculture. In this work, it can be observed that DL has been applied in various areas of agriculture.</td>
</tr>
<tr>
<td>16</td>
<td>Review—Machine Learning Techniques in Wireless Sensor Network Based Precision Agriculture [20] in 2019.</td>
<td>ML methods have mostly been used by researchers in the past two years in conjunction with Wireless sensor networks. In the future, there may be an increased use of more advanced techniques like distributed (or edge) deep learning. AI must be a strength to increase the automation of tasks in agriculture.</td>
</tr>
<tr>
<td>17</td>
<td>Identification of Plant Disease using Image Processing Technique [21] in 2019.</td>
<td>The leaf pictures are used for detecting plant diseases. Therefore, the use of the image process technique to find and classify diseases in agricultural applications is useful.</td>
</tr>
<tr>
<td>18</td>
<td>Image Processing with Convolutional Neural Networks for Classification of Plant Disease [22] in 2019.</td>
<td>This work focuses on the image processing of a plant disease database for enhancing the classification. This processing uses two pre-processes in activities (image selection and resizing) and the application of two modified architectures, VGG16 and VGG19, along with pre-trained weights from the ImageNet database.</td>
</tr>
<tr>
<td>19</td>
<td>A role of computer vision in fruits and vegetables among various horticulture products of agriculture fields: A survey [23] in 2020.</td>
<td>In it, different machine learning approaches for different performance metrics on the same dataset has considered and found that SVM gives better classification accuracy.</td>
</tr>
<tr>
<td>20</td>
<td>Image Processing Technique for Automatic Detection of Plant Disease and Alerting System in Agricultural Farms [24] in 2020.</td>
<td>Due to the large size of the farm, manual inspection is not possible therefore it is necessary to develop an automated disease-detecting and Alerting System.</td>
</tr>
<tr>
<td>22</td>
<td>Machine Learning Applications for Precision Agriculture: A Comprehensive Review [26] in 2020.</td>
<td>This article shows how knowledge-based agriculture can improve the productivity and quality of the product.</td>
</tr>
<tr>
<td>23</td>
<td>Accurate machine learning-based germination detection, prediction, and quality assessment of three-grain crops [27] in 2020.</td>
<td>This paper proposed a machine learning approach using ANN for three different crops, Zea mays (maize), Secale cereale (rye), and Pennisetum glaucum (pearl millet), with a total of more than 23,000 images with region proposals for seed germination detection. In this work, Faster RCNN is used to find out whether a seed is germinated or not.</td>
</tr>
<tr>
<td>24</td>
<td>Automatic Learning in Agriculture: A Survey [28] in 2021.</td>
<td>This paper presents a review of existing applications of machine learning in agriculture with a focus on the applications of Deep Reinforcement Learning techniques used in agriculture, the conventional solutions for agricultural machine learning decision-making problems are using supervised approaches.</td>
</tr>
</tbody>
</table>
IX. Conclusion

This review shows that agriculture plays a very important role in a global economy, in this work to find out the role of machine learning in the agriculture sector so many papers are reviewed. It has been observed that Crop production in general mostly affected by the prevailing weather condition of a region [5]. This work also states that early information on yield has special importance in intensive apple production. The focus is to improve image analysis-based apple yield forecasts with an application of artificial neural networks (ANN) [6].

This study also shows that so many approaches are used to predict crop yield and this improves decision-making in precision agriculture [7]. There are so many data mining techniques used to extract knowledge from agricultural data to estimate crop yield for major cereal crops [8]. This work also focuses on the importance of medicinal plants and how to preserve the knowledge of medicinal plants in the form of digital form by using the concept of machine learning. Lots of work has been done to generate an automated system to recognize species of plants by analyzing the digital images of their leaf for this recognition, so many methodologies have been proposed to analyze plant leaves in an automated fashion. This work also emphasizes that the success of the deep learning approach mainly depends on the training dataset, so data can be collected either use of remote satellite, Machine-generated (MG): data coming from sensors and intelligent machines (drones, Unmanned Aerial Vehicles (UAVs), Global Positioning System (GPS)), Process-mediated (PM): traditional commercial data coming from business processes referencing corporate events such as purchases and orders Human- sourced (HS): attestation of human experiences recorded in books, photos, audio, and video; they are now almost digitized in digital devices [12],[13]. There are so many benefits as well as challenges for the deployment of IoT in the agriculture sector all issues are described in detail [14]. A study report which focused on various reasons for the downfall of agriculture has also been addressed [15]. It also shows that CNN can be used to determine the weed which is harmful to plants [16]. A review of Deep Learning with a focus on Architecture, Applications, and Recent Trends has also been done [17]. In the coming years, we may see an increased use of more advanced techniques like distributed (or edge) deep learning. AI must be a strength to increase the automation of tasks in agriculture and improve yield while optimizing the use of natural resources [20]. It also observed that the use of image process techniques to find and classify diseases in agricultural applications is also useful [21], [22]. This review also studies the work done in the detection of different types of disease present in various fruit and vegetable and focused on the comparison of different machine-learning approaches for different performance metrics on the same dataset [23]. The research was also done for finding a complete solution to the problem of detecting disease and a solution proposed a system that after detecting the disease, sends alerts through the buzzer [24].

This work also highlighted the recent contribution of machine learning in agriculture and concluded that along with disease and insect detection, machine learning plays a vital role in the determination of various growth stages of the crop so that new cultivators can easily determine the transmission of the crop from one stage to another [36] and machine learning can also be used for classification of weedy seed which improves the quality of crops [39].

Therefore, after conducting this review, it has been determined that agriculture is crucial for the development of any country. In recent years, many new, advanced techniques have been adopted to boost productivity, but because this field requires a lot of labor and time, it is a crucial consideration, we must use advanced techniques to boost crop growth and quality. This is especially true if we use new technology to track plant development from the seedling stage to harvest. After examining so many publications, it was discovered that each paper generally addressed a problem for a particular crop. As a result, a generalized system is required that offers. Solution for all problems.

REFERENCES

35. Wei Li, Tengfei Zhu, Xiaoyu Li, Jianzhang Dong, and Jun Liu: Recommending Advanced Deep Learning Models for


