This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3048415, IEEE Access

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Machine Learning Applications for Precision Agriculture: A Comprehensive Review

Abhinav Sharma[1], Arpit Jain[2], Prateek Gupta[3], Vinay Chowdary[4]
[1,2,3,4] University of Petroleum and Energy Studies (UPES), Dehradun-248007, Uttarakhand, India
[1,2,4] Department of Electrical and Electronics Engineering, School of Engineering,
[3]Department of Systemics, School of Computer Science

Corresponding author: Arpit Jain ajain@ddn.upes.ac.in ; arpit.eic@gmail.com

Abstract Agriculture plays a vital role in the economic growth of any country. With the increase of population, frequent changes in climatic conditions and limited resources, it becomes a challenging task to fulfil the food requirement of the present population. Precision agriculture also known as smart farming have emerged as an innovative tool to address current challenges in agricultural sustainability. The mechanism that drives this cutting edge technology is machine learning (ML). It gives the machine ability to learn without being explicitly programmed. ML together with IoT (Internet of Things) enabled farm machinery are key components of the next agriculture revolution. In this article, authors present a systematic review of ML applications in the field of agriculture. The areas that are focused are 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 are reviewed for the classification of a different set of crop images in order to monitor the crop quality and yield assessment. This approach can be integrated for enhanced livestock production by predicting fertility patterns, diagnosing eating disorders, cattle behaviour based on ML models using data collected by collar sensors, etc. Intelligent irrigation which includes drip irrigation and intelligent harvesting techniques are also reviewed that reduces human labour to a great extent. This article demonstrates how knowledge-based agriculture can improve the sustainable productivity and quality of the product.

Index Terms: Agricultural Engineering; Machine Learning; Intelligent Irrigation; IoT; Prediction

I. INTRODUCTION

The population of the world will increase to 9.1 billion approximately thirty-four percent as of today by the end of 2050. Food requirement will increase by 70 percent and due to rapid urbanization, land availability for agriculture will decrease drastically in the coming years. India will be the most populated country by 2050 and presently it is already lagging the domestic food production. The main reason for reduced food production is the lack of planning, unpredictable weather conditions, improper harvesting and irrigation techniques and livestock mismanagement. In the last few years, nature has experienced a drastic change in weather conditions due to global warming. The average temperature of the earth has been increased due to which there is uncertainty in climatic conditions. Frequent droughts, heavy rainfall are the biggest challenge for poor farmers. According to the government of India annual economic survey, adverse climatic conditions, reduce the farmer’s income by 20-25%.

NOMENCLATURE

AI	Artificial Intelligence
ML	Machine Learning
DL	Deep Learning
IoT	Internet of Things
GPS	Global Positioning System
UAV	Unmanned Aerial Vehicle
ASC	Agriculture Supply Chain
NLP	Natural Language Processing
SI	Swarm Intelligence
ANN	Artificial Neural Network
NN	Neural Network
kNN	K-Nearest Neighbour
SVM	Support Vector Machines
RNN	Recurrent Neural Network
ELM	Extreme Learning Machines
RELM	Regularized Extreme Learning Machine
XGBoost	Extreme Gradient Boosting
MLP	Multi-Layer Perceptron
CNN	Convolutional Neural Network
PCA	Principal Component Analysis
RBFN	Radial Basis Function Network
RF	Random Forest
GBM	Gradient Boosting Model
SVR	Support Vector Regression
BPNN	Back Propagation Neural Network

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3048415, IEEE Access

LS-SVM	Least square support vector machine
GRNN	Generalized Regression Neural Networks
RELM	Residual Maximum Likelihood
DBN	Deep Belief Network
RT	Regression Tree
MLR	Multiple Linear Regression
LASSO	Least Absolute Shrinkage and Selection Operator Regression
RIDGE	Ridge Regression
SNN	Shallow Neural Network
GCN	Graph Convolutional Network
GEP	Gene Expressions Programming
RCNN	Regions-CNN
GA	Genetic Algorithm
PSO	Particle Swarm Optimization
PLSR	Partial Least Square Regression
ANFIS	Adaptive Neuro Fuzzy Inference System
TCN	Temporal Convolution Network
SCC	Somatic Cell Count
OPF	Optimum-Path Forest
BVDV	Bovine Viral Diarrhea Virus
MC	Moisture Content
SOC	Soil Organic Carbon
TN	Total Nitrogen
SOM	Soil Organic Matter
NDVI	Normalized Difference Vegetation Index
CEC	Cation Exchange Capacity
ETc	Estimation of evapotranspiration
SOM	Soil Organic Matter

LAI	Leaf-Area Index
RGB	Red Green Blue
DW	Accumulated Dry Weight
VRI	Variable Rate Irrigation
ET	Evapo-Transpiration
EC	Electrical Conductivity
SCM	Sub-Clinical Mastitis
SI	Scatter Index
AWM	Attribute Weighting Model
AUC	Area Under the Curve
R	Correlation Coefficient
R^2	Coefficient of Determination
MSPE	Mean Squared Prediction Error
MAPE	Mean Absolute Percentage Error
MAE	Mean Absolute Error
RMSE	Root Mean Square Error
RRMSE	Relative Root Mean Square Error
RPD	Residual Prediction Deviation
ROC	Receiver Operating Characteristic
RMSD	Root Mean Square Difference
NS	Nash–Sutcliffe coefficient
WSN	Wireless Sensor Network
GWO	Grey Wolf Optimization
SPI	Serial Peripheral Interface
I2C	Inter-Integrated Circuit
UART	Universal Asynchronous Receiver Transmitter
USB	Universal Serial Bus
BLE	Bluetooth Low Energy

Precision agriculture [1-2] is one of the solutions to ensure food security for the entire world [3]. Precision agriculture also abbreviated as digital agriculture is a technology-enabled data-driven sustainable farm management system. It is basically the adoption of modern information technologies, software tools, and smart embedded devices for decision support in agriculture [4] as shown in figure 1. Mechanized agriculture and the green revolution are the two key components of the first and second agriculture revolution. Precision farming is an important part of the third agriculture revolution [5].

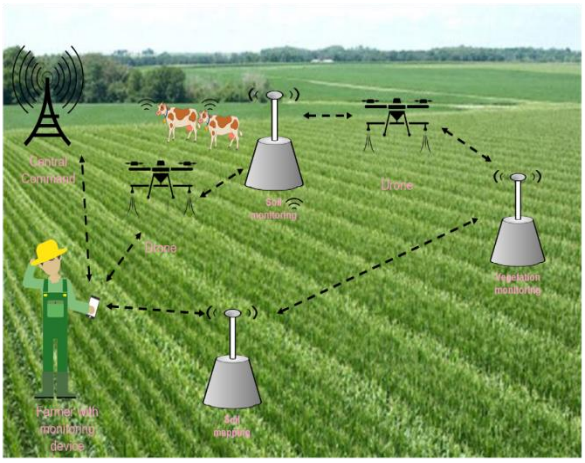


Figure 1. **Precision Agriculture**

John Deere introduced this technology in 1990 for the sowing of seeds and spraying of fertilizers using global positioning system (GPS) controlled tractors. The main focus of precision farming is to reduce the production cost and environmental effects to increase the farm's profitability. Digital technologies such as IoT [6], AI, data analytics, cloud computing, and block-chain technology play a key role in precision agriculture. In precision farming, IoT based smart sensors are deployed in the agriculture land for collecting data related to soil nutrients, fertilizers, and water requirements as well as for analysing the crop growth. Autonomous and semi-autonomous devices such as an unmanned aerial vehicle (UAV) [7] and robots are used for identifying weed and disease in the plants using computer vision techniques. Satellite images are also used in precision agriculture for monitoring the field and identifying the diseases in the plants. The data obtained from the deployed sensors [8] are processed and analyzed using ML algorithms to make farming practice more controlled and optimized. ML algorithms are also used for weather and rainfall prediction based on the data obtained from sensors, climatic records, and satellite images. This could save the lives of thousands of farmers who commit suicide

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3048415, IEEE Access

because of crop loss due to uncertainty in weather conditions. Smart livestock management is an important component of precision agriculture. It helps in monitoring the health, welfare, productivity, and reproduction of animals throughout their life cycle. Sensors and cameras monitor animal's health and computer vision techniques help in making intelligent decisions such as stopping the communal spread of diseases. Autonomous tractors and automated irrigation systems provide modern farming solutions to farmers. The widespread utilization of precision farming across the world is due to the presence of innovative machine and deep learning (DL) algorithms, high-speed internet access, and efficient computational devices. In [10] authors have discussed applications of ML for sustainable agriculture supply chain (ASC) performance. Authors have presented a unique ML-ASC framework that can guide researchers and agriculture practitioners to understand the role and importance of digital technologies in the agriculture industry. In [11] authors reviewed different ML applications in agriculture and discussed how digital technologies will benefit the agriculture industry. In this paper, the authors have presented a comprehensive review of the ML application for precision agriculture. This review article will provide an insight into the research community about the adoption of digital practices in the agriculture management system. It is anticipated that government agencies will frame policies to promote precision farming across the world. The main contribution of the article is outlined as follows:

- Applications of artificial intelligence and IoT in precision agriculture are discussed along with their practical implications.
- Foundation of ML and DL algorithms which find their application in precision agriculture has been discussed.
- Performance comparison for various ML, DL algorithms in precision farming has been carried out based on the state-of-art literature.
- Assessment of artificial intelligence techniques in precision agriculture is outlined along with its statistical and performance analysis.
- Comparison of performance parameters of sensors used in IoT applications in precision agriculture is presented.
- Integration of wireless sensor network (WSN) with IoT and artificial intelligence in precision agriculture is discussed.
- Challenges and future trends of artificial intelligence in precision farming are briefly outlined.

Table 1 highlights the major differences of this review article with other articles published in this field. The paper is organized as follows. Section 2 presents the impact of artificial intelligence (AI) and IoT in the field of agriculture. Section 3 briefly elucidate ML algorithms. In section 4 different ML applications in precision farming are briefly reviewed. Section 5 presents the IoT application in precision agriculture. Section 6 evaluates and access the knowledge-based agriculture system. Section 7 outlines the challenges and limitation of AI in precision agriculture. Section 8 presents the future trends of AI in precision agriculture. Section 9 provides conclusive remarks to summarize the paper.

II. IMPACT OF ARTIFICIAL INTELLIGENCE AND IoT IN AGRICULTURE

The term AI was first coined in the Dartmouth conference in the year 1956 by John McCarthy and he defined it as a science and engineering of making intelligent machines or more specifically intelligent computer programs. AI technology provides computational intelligence to machines so that they can learn, understand and react according to the situation. ML, DL, natural language processing (NLP), swarm intelligence (SI), expert systems, fuzzy logic, and computer vision are the subfields of AI as shown in figure 2. This field finds endless applications across different sectors of human life. Intelligent AI programs are widely explored in health-care, agriculture, finance, robotics, e-commerce and the automation industry. Samsung, Apple, and other electronics giant companies announced that they will be utilizing this technology in every device they will manufacture in the near future. IoT is another emerging technology in which smart sensors, devices are interconnected through the internet. These smart sensors can be utilized to gather data across different disciplines such as solar plants, agriculture fields, disaster-prone areas, manufacturing industry for efficient resource utilization. With the increase in population over the year’s demand for agriculture products is increasing day by day. However, with limited land availability for farming and reduce interest among the young generation to adopt farming as their profession, it has become a challenging task for the agriculture industry to satisfy the food requirement of millions of people. Now, the agriculture industry is widely adopting smart technologies like IoT and AI to efficiently cultivate

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3048415, IEEE Access

organic products in limited land areas as well as to overcome the traditional challenges of farmers.

TABLE 1.
Key differences of the article with published articles

S. No.	Paper	Key Differences
1	Jawad et. al. [9]	Ref. [9] focuses entirely on the WSN and IoT applications in precision agriculture and does not include ML applications. This article focuses on the applications of ML in precision agriculture, hence there is a difference in scope of both the research articles.
2	Sharma et. al. [10]	Ref. [10] focuses on ML applications in agriculture supply chain and only a brief overview of precision agriculture is discussed. This article focuses on application of ML and IoT in each cycle of precision agriculture in needed detail, along with the guidelines for future researchers, advantages, challenges and future trends.
3	Liakos et. al. [11]	Ref. [11] focuses on the application of AI and ML in precision agriculture, however does not include: (a) The application of IoT/WSN in precision agriculture (b) Guidelines for future practitioners, advantages, challenges and future trends. This article provide a detail overview and analysis of application of ML in precision agriculture. The article focuses on application of IoT/WSN in precision agriculture along with challenges and future trends of precision agriculture.
4	Chlingaryan et. al. [31]	Ref. [31] focuses on subfields of precision agriculture: Crop yield and nitrogen estimation. The paper does not discuss application of ML in soil parameters (other than nitrogen), disease and weed identification, drip irrigation, livestock management, intelligent harvesting. The application of WSN/IoT has not been discussed in the paper. This article focuses on all aspects of precision agriculture which includes soil properties and weather prediction, crop yield prediction, disease and weed identification, drip irrigation, livestock management and intelligent harvesting.
5	Jha et. al. [151]	Ref. [151] has narrow scope as the literature covered in the paper is limited. This paper does not discuss application of ML in drip irrigation, livestock management, intelligent harvesting. The application of WSN/IoT has not been discussed in the paper. This article focuses on application of ML and IoT in each cycle of precision agriculture in needed detail, along with the guidelines for future researchers, advantages, challenges and future trends.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3048415, IEEE Access

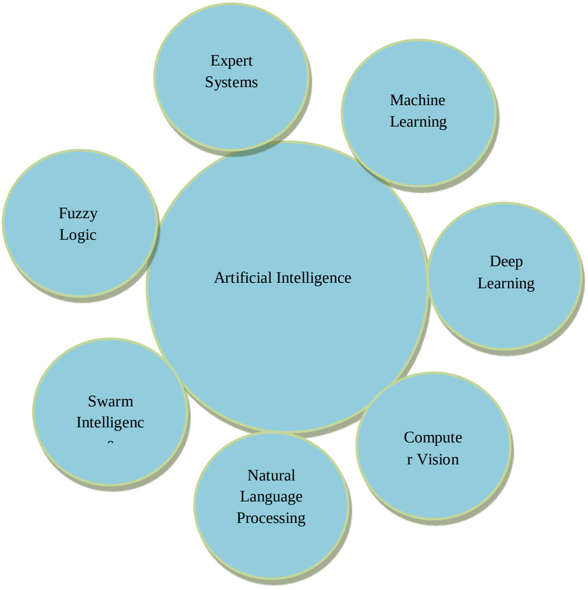


Figure 2. Artificial Intelligence Techniques

IoT based smart farming system is built for monitoring soil nutrients and soil moisture using sensors. ML algorithms are explored for determining the optimum amount of fertilizers required for soils before the sowing of crops.

Drones are revolutionizing the agriculture industry. These drones are cameras enabled and are used for different applications such as field and crop monitoring, spraying of pesticides, and drip irrigation. The images captured by the drones over the entire lifecycle of crops can be examined using DL and computer vision algorithms for disease and weed identification. Thereafter, these drones are used for spraying pesticides over the weeds and infected crops. Over the years uncertainty in weather conditions is the main concern of farmers. Drip irrigation using drones is an efficient AI-empowered irrigation system which is basically trained on weather pattern and can effectively reduce the water problems of farmers.

AI-enabled robots can be used for harvesting the crops at a much faster pace and in large volumes. Robots can reduce human labour to a large extent and can be used along with drones for monitoring the field. Livestock management is another major concern for farmers. IoT based sensors can be deployed in the field for health monitoring of cattle. This information can be utilized for protecting the bunch of cattle from diseased cattle.

NLP based virtual assistant applications like chatbots can update the farmers with the latest advancement in technologies for agriculture. Farmers can finds solutions for their problems and incorporate the latest technology in their farming for improving their field productivity. Thus, AI and IoT are the two major technologies that will play a vital role in the agriculture industry.

III. MACHINE LEARNING ALGORITHMS

ML is the subfield of computer science that gives computers the ability to learn without being explicitly programmed. (Arthur Samuel, 1959) [12]. Alan Turing in the year 1950 proposed the concept of learning machines and wrote a research article "The Turing Test for Machine Intelligence" [13]. He performed a test and examined the machine's ability to demonstrate intelligent behaviour similar to humans. A machine or intelligent computer program learns and extract knowledge from the data, builds a framework for making predictions or intelligent decisions. Thus, the ML process is divided into three key parts, i.e. data input, model building, and generalization as shown in figure 3. Generalization is the process for predicting the output for the inputs with which the algorithm has not been trained before. ML algorithms are mainly used to solve complex problems where human expertise fails such as weather prediction, spam filtering, disease identification in plants, pattern recognition.

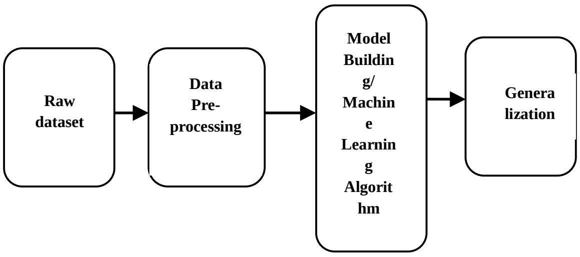


Figure 3. A Machine learning process

Today, due to the availability of innovative algorithms and large data sets through internet resources industries and research communities are widely using ML algorithms for solving a diverse set of problems. DL is the subfield of the family of ML algorithms which is trained from large sets and uses an artificial neural network (ANN) to make intelligent decisions.

ML algorithms are categorized as supervised learning, unsupervised learning, and reinforcement learning as shown in figure 4. Supervised learning as the name suggests is learning with the supervisor or teacher. This set of algorithms works with labeled data-set which means corresponding to each input there are outputs. The algorithm builds an input-output relationship with this labeled data set and thereafter generalize or predicts outputs for unseen inputs. Supervised learning algorithms used for predicting the categorical value are known as classification algorithms and the algorithms that are used for predicting the numerical value are known as regression algorithms. Unsupervised learning algorithms works with unlabelled data and discovers unknown objects by grouping similar objects. The goal of an unsupervised learning algorithm is to extract hidden knowledge from the training data set thus this approach is difficult to implement than supervised learning algorithms. Reinforcement learning is another approach that learns from the environment through reward and punishment. AlphaGo, a chess-playing game developed by DeepMind utilized reinforcement learning for defeating the world's best chess-playing computer program.

In this paper the performance of different ML algorithms are analysed and discussed in the field of agriculture. Table 2 presents different types of supervised, unsupervised and reinforcement learning algorithms utilized for soil and weather prediction, disease and weed identification, intelligent irrigation and harvesting techniques as well as livestock management.

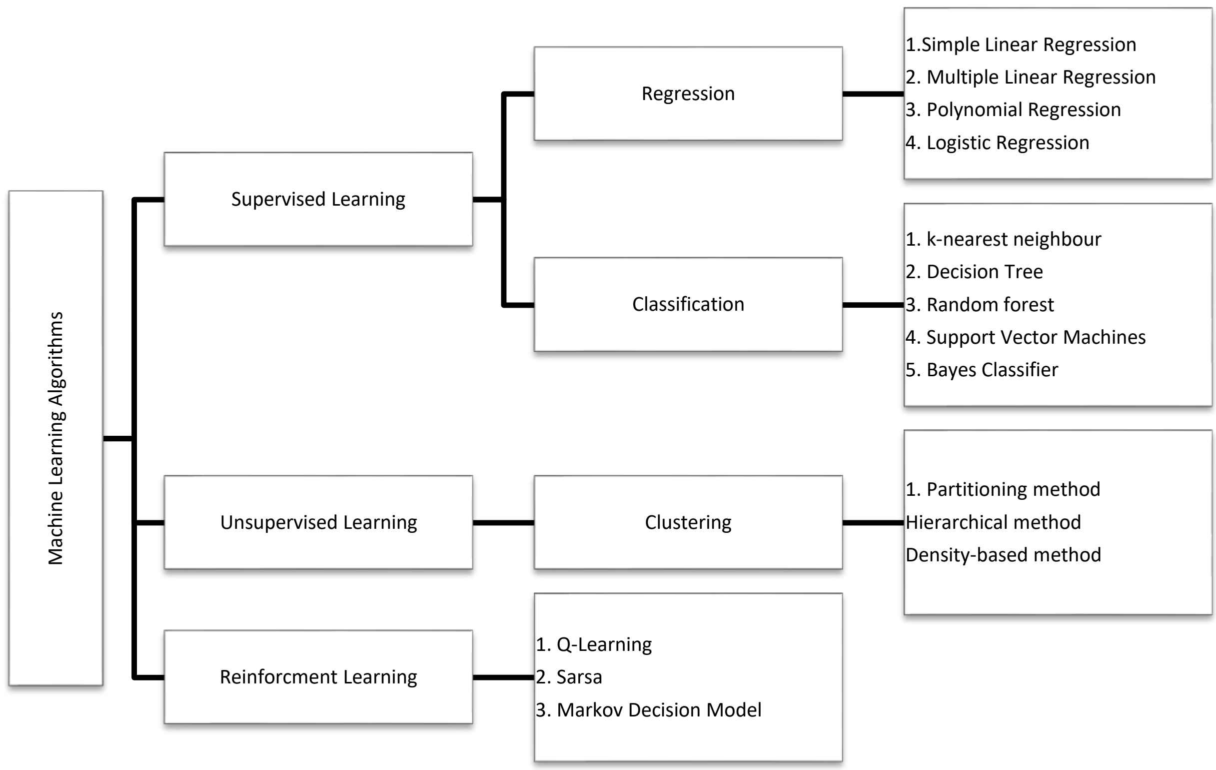


Figure 4. Categorization of Machine learning algorithms

TABLE 2.
Machine Learning Algorithms

Machine Learning Algorithm	Algorithm Description
Regression Algorithm	Regression algorithms are the supervised learning algorithms in which the relationship between input and output is based on the training data and it predicts the output numerical value for the unseen input. Simple and multiple linear regression, polynomial regression, and logistic regression are some of the common regression algorithms.
kNN	kNN is a simple supervised classification algorithm. In this algorithm first, the labelled dataset is divided into different classes based on their outputs. Thereafter, a new sample object is assigned a particular class based on its k-nearest neighbours.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3048415, IEEE Access

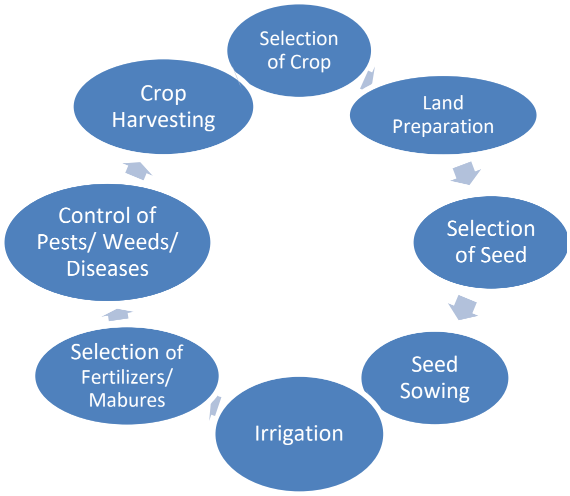
Random Forest	Random forest is the ensemble classification model which combines a number of decision tree classifier. The final class of a new object is found out based on the majority class predicted from different decision trees classifiers.
SVM	SVM is a classification and regression algorithm that builds multi-dimensional boundaries between data points in the feature space. The output of the SVM is predicted based on the classes divided using the training data.
RNN	RNN is a feedforward artificial neural network with feedback from the output layers of neurons to the input layer. The network also consists of self-loops.
ELM	ELM are feedforward NN with single or multiple layers of neurons. It is a non-iterative approach and tuning of parameters is completed in a single run thus finds useful applications in real-time regression and classification problems.
MLP NN	MLP NN is a feedforward biologically inspired artificial neural network which has multiple layers of neurons. The synaptic weights of the network are optimized with the training dataset and later the network is used for generalization.
CNN	CNN is the most widely used deep neural network. This network consists of a number of layers of neurons in which network use mathematical operation convolution instead of matrix multiplication in at least one of the network layers.

IV. MACHINE LEARNING APPLICATIONS IN PRECISION AGRICULTURE

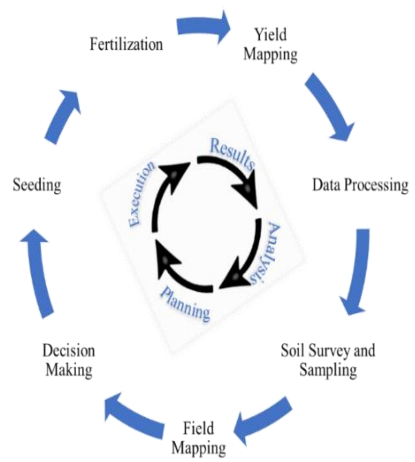
In many countries, the farmers rely on the traditional ways of farming which is based on the reliability of the suggestions from the elderly and their experience. This method leaves farmers at the mercy of random climatic conditions which are already getting random due to global warming and uneven rainfall patterns. The manual spraying method for pesticides led to improper usage of resources and harms the environment. AI and IoT enabled precision agriculture removes the randomness and assist new age farmer to optimize every step of the farming process. Figure 5 (a) and (b) presents a pictorial view of traditional agriculture and technology enabled farm management system.

Gaitán [14] provided a systematic study of the impact of extreme weather events, such as hail events, cold waves, heat waves, and their impact on agricultural practices. The author reported floods, droughts, frost, hail, heatwaves, and pest outbreaks are impacted by climatic conditions.

The AI systems are applicable in each farming operation as depicted in figure 4 and some of them even extend beyond the conventionally recognized steps. In this section we will discuss the state of art techniques proposed/implemented by various researchers and practitioners worldwide.



(a)



(b)

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3048415, IEEE Access

Figure 5. **(a) Traditional agriculture cycle (b) Precision agriculture cycle**

A. Soil Properties and Weather Prediction

Prediction of soil properties is the first and the most crucial step which influences the selection of crop, land preparation, selection of seed, crop yield, and selection of fertilizers/manure. The soil properties are directly related to the geographic and climatic conditions of the land in use and hence is an important factor to take into consideration. The soil properties prediction mostly consists of predicting nutrients in the soil, soil surface humidity, weather conditions during the lifecycle of the crop. Human activities have highly affected the properties of soil and hence our ability to cultivate the crops [15]. In general, there are 17 essential elements as listed in table 3 which play an important role in plant growth [16]. The growth of crops depends on the nutrients available in a particular soil. The soil nutrients are mostly monitored by electric and electromagnetic sensors [17]. Depending on the nutrients farmers make informed decisions as to which crop is optimal for the land. However, the nutrients can be added through fertilizers, manure, etc. but with an additional cost. Some of them may also damage the environment and have an adverse effect on the soil cycle.

TABLE 3.
Essential Plant Nutrients [2]

Essential plant element		Symbol	Primary form
Non-Mineral Elements	Carbon	C	CO_2 (g)
	Hydrogen	H	H_2O (l), H^+
	Oxygen	O	H_2O (l), O_2(g)
Mineral Elements			
Primary Macronutrients	Nitrogen	N	NH_4^+, NO_3^-
	Phosphorus	P	HPO_4^{2-}, $H_2PO_4^-$
	Potassium	K	K^+
Secondary Macronutrients	Calcium	Ca	Ca^{2+}
	Magnesium	Mg	Mg^{2+}
	Sulfur	S	SO_4^{2-}
Micronutrients	Iron	Fe	Fe^{3+}, Fe^{2+}
	Manganese	Mn	Mn^{2+}
	Zinc	Zn	Zn^{2+}
	Copper	Cu	Cu^{2+}
	Boron	B	$B(OH)_3$
	Molybdenum	Mo	MoO_4^{2-}
	Chlorine	Cl	Cl^-
	Nickel	Ni	Ni^{2+}

A scientific analysis of soil nutrients, soil moisture, pH is important for determining the soil properties. Acar et al. [18] employed an extreme learning machine (ELM) based regression model for prediction of soil surface humidity. The author selected two terrains having area 4 KM2 and 16 KM2 located in Dicle university campus for experimental analysis. The real-time field data was extracted using polarimetric Radarsat-2 data, which was pre-processed using the SNAP toolbox [19] and features were added with the help of local measurements by separating the field into square grids. Once the pre-processing and feature extraction is done the data is passed to ELM based regression model to predict the soil surface humidity. The algorithm was tested with 5 different kernel functions and the prediction was validated using leave-one-out cross-validation technique. The experimental results confirmed the lowest root mean square error (RMSE) of 2.19% when using ‘sine’ kernel function.

Wang et al. [20] deployed soft sensors based on ELM for the measurement of nutrient solution composition in the soilless cultivation method. The soilless cultivation method is an emerging planting method. It is imperative to monitor the pH value, temperature and concentration changes in nutrient solution composition as the performance of soilless cultivation is highly dependent on these parameters. The significant variables in a nutrient solution cannot be measured directly hence these are determined with the help of auxiliary variables. The authors utilized conductivity, pH value, flow rate, and temperature measurements for auxiliary measurements. These auxiliary measurements are fed to a deep belief network-based ELM which predicts the values of significant variables. For experimental analysis, the authors deployed the model to measure the concentration of SO_4^{2-}, and $H_2PO_4^-$ in a nutrient solution. The authors reported an average RMSE of 1.2414 for predictions in SO_4^{2-} and RMSE of 0.8892 for prediction of $H_2PO_4^-$. Park et al. [21] utilized ML algorithms to predict the soil moisture using data from MODIS. The authors downscaled the AMSR2 soil moisture to 1KM using random forest (RF) and Cubist algorithms. An ensemble of these algorithms was used to obtain soil moisture data. The results obtained through the ML methods were compared with the statistical ordinary least squares technique. The ML model exhibited a R^2 (coefficient of determination) of 0.96 and an RMSE of 0.06, whereas a $R^2$of 0.47 and a RMSE of 0.16 was associated with the statistical ordinary least squares.
Reda et al. [22] explored ML algorithms to estimate soil organic carbon (SOC) and total nitrogen (TN) in soil samples collected from four agricultural lands of Moroccan. Data set of near-infrared spectroscopy is utilized in comparison to traditional chemical methods as this technique reduces the computation time and

resource utilization. The ensemble learning modeling algorithm presents the best performance among other regression models and back-propagation neural networks (BPNN) algorithm. The proposed algorithm presents R^2 of 0.96, RMSE of 1.92, performance to deviation (RPD) of 4.87 for SOC and R^2 of 0.94 and RMSE of 0.57, RPD of 4.91 for TN prediction. Morellos et al. [23] also utilized visible and infrared spectroscopy to determine TN, SOC, and moisture content (MC) in the arable field in Premslin Germany. Spectroscopy dataset is used for building the predictive ML model for estimating all three soil properties. Least square support vector machine (LS-SVM) and cubist ML algorithms outperform principal component regression and partial least square regression multivariate methods in terms of RMSE and residual prediction deviation (RPD). LS-SVM best predict SOC and MC with RMSE of 0.062 and 0.457, RPD of 2.24, and 2.20. Cubist best predicts for TN with RMSE of 0.071 and RPD of 1.96. Andrade et al. [24] build ordinary least square regression, RF, cubist regression, XGboost prediction model for determining soil properties from portable X-ray fluorescence (pXRF) spectrometry dataset in Brazilian coastal plains. Three soil properties total nitrogen, soil organic matter (SOM), cation exchange capacity (CEC) were analyzed using RF, ordinary least squares regression (OLS), cubist regression (CR), XGBoost (XGB). RF algorithm gives the best performance with R^2 of 0.50 for TN, R^2 0.75 for CEC and 0.56 for SOM. Deiss et al. [25] estimated the soil properties (clay, sand, pH, SOC) in northern Tanzania and USA Midwest from the spectroscopy dataset using ML algorithms. THE tuned SVM model outperforms the partial least square (PLS) regression model in terms of predicting all the soil parameters.

Mahmoudzadeh et al. [26] explored the ML algorithm to predict SOC in the Kurdistan province of Iran. The simulation results suggest that RF accurately predicts SOC with R^2 of 0.60 and RMSE of 0.35% in comparison to SVM, kNN, Cubist, and Extreme Gradient Boosting (XGBoost) ML algorithm. The study also suggests that air temperature, annual rainfall, valley depth, texture of terrain surface are some of the important factors that influence SOC spread over the Kurdistan region. Veres et al. [27] explored DL architecture such as CNN for predicting the soil properties from the infra-red spectroscopy dataset. Benke et al. [28] predict soil electrical conductivity (EC) and SOC in different regional locations of Victoria, Australia using pedotransfer function (PTF) based on ML algorithm. PTF basically converts soil measurement into soil properties and provides inputs for ML simulation algorithms. In the proposed approach PTF use Generalised Linear Mixed Effects Model (GLMM) model and Residual Maximum Likelihood (REML) to predict` the soil properties. Traditional approaches to soil properties and crop yield prediction require time-consuming field surveys and the deployment of expensive sensors. Khanal et al. [29] proposed an alternative approach in which the dataset for the prediction of soil properties and crop yield is generated using remotely sensed aerial images of agricultural land. Five soil properties, viz. pH value, magnesium (Mg), potassium (K), SOM, CEC, and crop yield were predicted using RF, SVM, Cubist, NN, Gradient Boosting Model (GBM) ML algorithms. NN presents the highest prediction accuracy for SOM having R^2 of 0.64, RMSE of 0.44 and CEC having R^2 of 0.67, RMSE of 2.35; SVM best predicts K having R^2 of 0.21, RMSE of 0.49 and Mg having R^2 of 0.22, RMSE of 4.57; and GBM best predicts pH having R^2 of 0.15, RMSE of 0.62. RF outperforms other algorithms in terms of crop yield prediction and presents higher accuracy having R^2 of 0.53 and RMSE of 0.97. Labrador et al. [30] estimate calcium and Mg content in soil using generalized regression NN and genetic algorithm (GA). The digital elevation model and satellite images were used as input to the prediction model for estimating the soil properties. Chlingaryan et al. [31] discussed different ML approaches used in precision agriculture for accurate crop yield prediction and soil nitrogen estimation over the last fifteen years.

Ju-Young et al. [32] investigated a seasonal climatic forecasting model using regularized ELM to predict day-wise mean air temperature at field level for a period of 90 days. The authors selected data from Korea Metrological Administration using the Met GloSea5GC2 model [33]. The authors fed 240 days of forecast data and hindcast data from the ensemble based model to the RELM algorithm. The algorithm performance was evaluated by measuring: RMSE, mean absolute error (MAE) the model prediction vs the actual values. The authors achieved an RMSE in the range of 1.02 to 3.35 which outperformed the meteorological data which has an RMSE range of 1.61 to 3.37.

Soil moisture content is an important parameter to acknowledge in the agriculture industry as it addresses precise irrigation scheduling. Stamenkovic et al. [34] build a support vector regression (SVR) prediction model to predict soil moisture content from remotely sensed hyperspectral images. Song et al. [35] proposed a macroscopic cellular automata (MCA) model and combined its deep belief network (DBN) to predict soil moisture content over a cornfield in northwest china. The simulation results of DBN-MCA outperforms the multi-layer perceptron (MLP)-MCA in terms of RMSE. Acheing [36] explored the SVR model (i.e. RBF), ANN, DNN for simulating soil water retention curve (SWRC) curve of loamy sand. Dataset of loamy sand subjected to wetting and drying condition is collected using a reflectometer and tensiometer. RBF based SVR model best predicts SWRC under both wet and dry conditions. Feng et al. [37] estimate soil temperature at various soil depths of Loess Plateau of China. Four different ML algorithms ELM, generalized regression neural networks (GRNN), backpropagation neural networks (BPNN), and RF were investigated for predicting the soil temperature. ML algorithms were trained with air temperature, wind speed, relative humidity, and vapour pressure and solar radiation as input parameters and the simulations show that ELM outperforms other ML algorithms in terms of RMSE, MAE, Nash–Sutcliffe coefficient (NS) and concordance correlation coefficient. Mohammadi et al. [38] explored ELM for predicting daily dew point temperature in different parts of Iran. This part of the world experience different climatic conditions throughout the year. The proposed model accurately predicts dew point temperature than SVM and ANN algorithms. Zhu et al. [39] accurately predict daily evapotranspiration in Northwest China using hybrid particle swarm optimization (PSO)-ELM model to optimize crop water requirement in agriculture. Alizamir et al. [40] accurately predicts soil temperature at different depths of 5, 10, 50 and 100 cm using ELM, ANN, classification and regression trees, group method of data handling using dataset obtained from Mersin station operated by Turkish Meteorological Service. The simulation results suggest that soil temperature can be estimated easily using air temperature upto depth of 50 cm while for depth of 100 cm additional information of solar radiation and wind speed is required.

Rainfall prediction plays a critical role in the water resource management system, flood risk assessment, and the agriculture industry. Acknowledging the chaotic nature of rainfall, it is very difficult for statistical approaches to accurately predict the rainfall. Cramer et al. [41] evaluated the performance of seven ML algorithms for rainfall prediction. The statistical results show that the radial-basis function neural network (RBFNN) shows the best performance among other state of the art algorithms. Sierra and Jesus [42] predicted the rainfall in Tenerife, an island in Spain based on atmospheric synoptic patterns using different ML algorithms and found NN gave the best performance among other ML algorithms. Kamatchi and Parvathi [43] employed NN for weather prediction and proposed a hybrid recommender system for enhancing the success ratio of the system. Lazri et al. [44] build a multi-classifier model for estimating precipitation using MSG images (Meteosat second generation) and dataset obtained using radar. The proposed approach shows that the proposed multi-classifier improves the standard of classification. Shardoor and Rao [45] surveyed three different approaches i.e. ML techniques, data mining techniques, and satellite forecasting techniques for rainfall prediction. Table 4 presents a comparative study of different ML algorithms for prediction of soil properties and weather prediction.

TABLE 4.
Different ML algorithms for Prediction of Soil properties and weather conditions

Reference	Attributes	Experimental Place	ML Algorithm	Accuracy Measure
Acar et al. [18]	Soil Surface Humidity	Polarimetric measurements using Radarsat-2 for a field located in Dicle university	Extreme learning Machine based Regression (ELM-R) model	RMSE of 2.19%
Wang et al. [20]	SO_4^{2-},$H_2PO_4^-$	Auxiliary measurement of: conductivity, temperature, pH and flow rate	ELM	RMSE of 1.2414 for SO_4^{2-} prediction and 0.8892 for $H_2PO_4^-$ prediction
Park et al. [21]	Soil moisture	AMSR2 soil moisture data from MODIS (satellite)	RF and Cubist algorithm	RMSE of 0.06 and R^2 of 0.96
Reda et al. [22]	SOC and TN	Four agricultural lands of Moroccan	Ensemble Learning algorithm	For SOC R^2 of 0.96, RMSE of 1.92 and RPD of 4.87 For TN R^2 of 0.94, RMSE of 0.57, RPD of 4.91
Morellos et al. [23]	TN, SOC MC	Arable field of Premslin Germany	LS-SVM and Cubist method	LS-SVM best predicts

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3048415, IEEE Access

				MC RMSE of 0.457% and RPD of 2.24 and SOC RMSE 0.062% and RPD of 2.20, Cubist method best predicts TN RMSEP 0.071 and RPD of 1.96
Andrade et al. [24]	TN, CEC, and SOM	Brazilian Coastal Plains	OLS, CR, XGB and RF	RF best predicts with R^2 of 0.50, CEC of 0.75, and SOM of 0.56
Mahmoudzadeh et al. [26]	SOC	Kurdistan province of Iran	Cubist method, kNN , XGBoost and SVM	RMSE of 0.35% and R^2 of 0.60
Benke et al. [28]	Soil EC and SOC	Regional places in the state of Victoria, Australia	GLMM and parameters are estimated by REML	Prediction error for EC, MSPE is 0.686 and MAPE is 0.635 Prediction error for EC, MSPE is 0.413 and MAPE is 0.474
Khanal et al. [29]	SOM, Mg, CEC, K, pH	North western part of Madison County, Ohio, USA	RF, SVM, Cubist, NN, GBM	NN best predicts SOM with R^2 of 0.64, RMSE of 0.44 and CEC with R^2 of 0.67, RMSE of 2.35; SVM for K with R^2 of 0.21, RMSE of 0.49 and Mg with R^2 of 0.22, RMSE of 4.57); and GBM for pH with R^2 of 0.15, RMSE of 0.62
Ju-Young et al. [32]	Mean air temperature	Korea climatic model from GloSea5GC2	Regularized Extreme Learning Machine (RELM)	RMSE in the range of 1.02 to 3.35 was reported
Feng et al. [37]	Soil temperature	Maize field, located in Shouyang County, Shanxi Province, northern P. R. China	ELM, GRNN, BPNN, and RF	ELM best predict with RMSE of 2.26-2.95oC, MAE of 1.76–2.26oC, NS of 0.856–0.930 and concordance correlation coefficient of 0.925–0.965
Mohammadi et al. [38]	Daily dew point temperature	Bandar Abass and Tabass stations situated in Iran	ELM, SVM, ANN	ELM best predict with Tabass station having MABE of 0.3240ºC, RMSE of 0.5662ºC, and R of 0.9933 Bandar Abass having MABE of 0.5203ºC, RMSE of 0.6709ºC, and R of 0.9877
Zhu et al. [39]	Reference ETc	Arid region of Northwest China	Hybrid PSO-ELM model	R^2, MAE, RRMSE and NS ranging from 0.886–0.969, 0.268–0.536 mm d−1, 14.2–23.8 % and 0.886–0.969 at the four stations
Alizamir et al. [40]	Soil temperature Prediction at different depths	Turkey	ELM, ANN, Classification and Regression trees, Group method of data handling	With air temperature as input ELM best predicts with RMSE, of 1.914, 1.429, 1.456 and NSE of 0.965, 0.977, 0.966 and R^2 of 0.977, 0.980, 0.984 for depth of 5, 10, 50 cm
Sierra and Jesus [42]	Rainfall Prediction	Tenerife Spain	SVM, kNN, RF, k-means clustering, NN	NN best predict the rainfall occurrence and intensity with average F score close to 0.4 and R 0.1-0.8 depending on the month and gauge analyzed
Lazri et al. [44]	Estimation of Precipitation	Northern region of Algeria	Level 1 classification combines RF, ANN, SVM, NB, weighted kNN, K-means algorithm Level 2 classification based on RF	R^2 of 0.93, Bias of 2.2 mm and RMSD 12 mm

B. Crop Yield Prediction

A significant piece of information for any farmer is the prediction of crop yield and how the yield can be increased. pH value, soil type, and quality, weather pattern: temperature, rainfall, humidity, sunshine hours, fertilizers, and harvesting schedules are some of the parameters which play an important role in predicting the crop yield [46]. Scientifically manual farming can be considered as a feedback control system in which the corrective action is taken once a setback in a crop is observed. The crop yield will highly depend on the efficiency of the optimal utilization of the above-mentioned resources. If some

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3048415, IEEE Access

kind of anomaly goes undetected in the initial stage may harm the crop yield in an unprecedented way. Singh et al. [47] assessed hailstorms on India's wheat production and observes that in February and March 2015 alone the hailstorm events caused a decline of 8.4% in national wheat production. For financially weak farmers in a country such as India, where intermittent storage of harvested crops is a rare resource, accurate weather predictions may turn to be miraculous for farmers. ML models when systematically applied to a system act as feedforward control. With the help of accurate ML models, we can anticipate the factors which are going to affect the crop yield. Hence the corrective action can be taken before even an anomaly hits the crop production.

Kamir et al. [48] used ML models to identify the yield gap hotspots in wheat production. Authors generated very high-resolution yield maps using data from various sources between 2009 and 2015. The data was collected from various sources:(a) NDVI time-series data across Australia using the MOD13Q1 data set [49], (b) rainfall and temperature data were collected from historic climate data at Australia bureau of metrology, (c) maps for observed grain yield were collected at source using intelligent harvesting machines. The dataset generated were tested with 9 ML algorithms: RF, XGBoost, Cubist, MLP, SVR, Gaussian Process, k-NN, and Multivariate Adaptive Regression Splines. The authors combined predictions from each of the algorithms into ensembles for prediction optimization [50]. Out of these algorithms, SVR with RBFNN outperformed other algorithms and investigators were able to achieve the yield estimate with an R^2 of $0.77$and an RMSE value of $0.55\ \mathrm{t\ ha^{-1}}$. The results were validated using 10-fold cross-validation techniques applied to the full data set.

Aghighi et al. [51] used various advanced regression algorithms to predict the yield of silage maize crops. The authors selected maize fields located at Moghan Agro-Industrial and Animal Husbandry Company (MAIAHC), which is about 28,000 hectares' area and located in Iran. The crop yield dataset was collected for around 40 silage maize fields were collected for a period from 2013-2015. In addition to it, the historic crop yield data the authors also gathered time-series NDVI data from Landsat 8 OLI satellite. The data was fed to advanced regression algorithms: (a) Gaussian Process Regression, (b) SVR, (c) Boosted Regression tree (d) RF Regression models and the prediction form each of the regression models were compared and evaluated. Authors found out the boosted regression tree reported best evaluation parameters with an average R-value of higher than 0.87, and RMSE in a range of 8.5 to 11.10, with a mean value of 9.5 during the period 2013-14. Kuwata and Shibasaki [52] employed DL models to estimate the crop yield. Authors deployed SVR for predicting the yield of corn in Illinois. For input following dataset was employed by the authors: (a) 5 year moving average of corn crop yield, (b) The enhanced vegetation index is obtained using the MOD09A1 dataset MODIS satellite, and (c) Historic climatic data. The dataset was fed to support the vector regression model and the authors reported an RMSE of 8.204 and a correlation coefficient of 0.644 for the model. For result, validation authors conducted 10-fold cross-validation on the full data set. Kulkarni et al. [53] utilized DL models to predict rice crop yield. The authors utilized soil properties and nutrients measurements recorded over 31 years and historic rainfall data. The input data was fed to recurrent neural network models for crop yield prediction. For effective prediction the authors explored different activation functions viz. sigmoid, reLu, and linear in the neural network.

Chu and Yu [54] builds an end to end summer and winter rice prediction model in 81 counties in the Guangxi Zhuang Autonomous Region, China. The proposed BBI model works in three stages, in the first stage the original area data and time series metrological data is pre-processed and its output works as input for the second stage where BPNN and RNN (recurrent neural network) learns deep spatial and temporal features from the input data. In the third stage, BPNN learns the relationship between deep features and rice yield to predict the summer and winter rice yields. The performance of the model is evaluated in terms of error and rate of convergence, the model presents lowest error values with MAE and RMSE of 0.0044 and 0.0057 for summer rice prediction and 0.0074 and 0.0192 for winter rice prediction while the algorithm converges within 100 iterations. Feng et al. [55] proposed a hybrid approach for wheat yield prediction in new-south Wales in southeastern Australia. Multiple growth specific indicators, viz. agricultural production system simulators (APSIM), NDVI, and SPEI (Standardized Precipitation and Evapotranspiration Index) are used before the prediction of wheat yield using regression models (multiple linear regression (MLR) and RF). APSIM+ RF hybrid model presents the best performance among other predictors in terms of prediction accuracy. Cai et al. [56] integrated two data sources, i.e. climate data and satellite data over fourteen years to predict the wheat yield in Australia using ML algorithms (SVM, RF, and NN). Simulation results show that climate data provides distinctive information in comparison to satellite data for yield prediction with R^2 of around 0.75. Planting the crops on accurate date plays an important role in improving productivity and reducing financial loss. Gumuscu et al. [57] explored three supervised ML algorithms; kNN, SVM, and decision trees for predicting planting dates; early, normal, and late for wheat crops in Turkey. The authors utilized climate data of the last 300 days to train ML algorithms and explored GA for feature selection. kNN classification ML algorithm shows robust performance and best predicts the planting dates of wheat crops. Several African, American, and Asian countries are the major producer of coffee in the world. Nevavuori et al. [58] explored a deep learning approach, i.e. CNN for wheat and barley yield prediction in the agriculture field of Pori, Finland. NDVI and RGB dataset obtained from cameras installed in UAV is used to train the six-layer CNN. RGB dataset best predicts the crop yield in CNN with MAE of 484.3 $kgha^{-1}$ and mean absolute percentage error (MAPE) of 8.8%. Koirala et al. [59] reviewed deep learning approaches for fruit detection and yield estimation. CNN in the context of computer vision is widely used for feature extraction from images that provide useful insight to object detection and yield estimation.

Kouadio, et al. [60] predicted the Robusta coffee yield using ML techniques from soil fertility dataset of Vietnam. ELM model outperforms multiple linear regression and RF algorithm with RMSE of 496.35 kg ha^{-1} and MAE of 326.40 kg ha^{-1}. Gamboa et al. [61] predict the cocoa yield in Santander, Columbia using a generalized linear model (GLM) and SVM. In recent decades researchers have explored statistical and probabilistic models for crop yield prediction. Gyamerah et al. [62] proposed a novel robust probabilistic forecasting model based on quantile random forest and Epanechnikov kernel function (QRF-E) for crop yield prediction in Ghana. The proposed approach didn't only predict discrete yield values but completely showcase probability descriptions for prediction interval for the two crops groundnut and millet. The simulation result shows the superior performance of the proposed algorithm in terms of prediction intervals coverage probability and prediction interval normalized average width under uncertain weather conditions.

Peng et al. [63] explored remote sensed satellite-based Solar-Induced Chlorophyll Fluorescence (SIF) dataset for training ML algorithms to predict maize and soybean yield in the mid-west region of the United States. Simulation results show that non-linear algorithms such as SVM, ANN, RF best predict the crop yield in comparison to least absolute shrinkage and selection operator regression (LASSO) and ridge regression (RIDGE) algorithm. Khaki and Wang [64] predicted the hybrid maize yield with a dataset of 2,267 locations of the United States and Canada between the years 2008 to 2016 using deep neural networks (DNN). Genotype, weather, and soil properties were the three components used to train DNN. The proposed model accurately predicts the maize yield with RMSE of 12% of the average yield for predicted weather dataset and 11% of the average yield for perfect weather dataset and outperforms LASSO, shallow neural network (SNN) and regression tree (RT). Simulation results show that environmental factors have a large impact on the prediction accuracy of crop yield. In most areas of Africa, agriculture field data is scarcely available thus remotely sensed dataset is widely used for monitoring the field. Leroux et al. [65] explored the ML algorithm for predicting the maize yield in Burkina Faso with a remotely sensed dataset. A process-based crop model SARRO which is basically designed to simulate attainable agricultural yields under tropical conditions is used in this study. RF outperforms MLR in maize yield prediction with R^2 of 0.59 at the end of the season and 0.49 before two months of harvest. Li et al. [66] build a statistical model for predicting the rain-fed crop yield using climate, satellite, and country-specific datasets in the mid-west region of the USA. Maimaitijiang et al. [67] explored the potential of UAV with DNN for soybean yield prediction from the fields of Columbia, Missouri, USA. Multi-modal information such as canopy spectral, structural, and thermal features extracted from images obtained from the sensors installed on UAV is used as the input dataset for training DNN. The simulation result shows that DNN accurately predict the crop yield and outperforms partial least square regression (PLSR), RF, SVR algorithms with R^2 of 0.720 and RMSE of 15.9%. Zhang et al. [68] explored ANN for prediction of annual crop planting utilizing a historical cropland data layer (CDL) dataset of corn-belt of mid-west, USA.

Kocian et al. [69] utilized both approaches to predict crop growth in greenhouses. IoT smart sensors are installed in the greenhouses to monitor different environmental parameters, soil properties and plant growth parameters such as leaf area index (LAI), accumulated dry weight (DW) and evapo-transpiration (ET). These parameters are in real-time send to the cloud through IoT devices and permit the implementation of an agriculture decision-support system. The probabilistic Bayesian network is explored in the proposed system to predict crop development parameters. Shahinfar and Kahn [70] explored ML algorithms for early prediction of adult wool growth in Merino sheep of Australia. Model Tree algorithms best predict the wool growth in comparison with NN with a correlation coefficient of 0.93, 0.90, 0.94, 0.81 and 0.59, MAE of 0.48 kg, 0.41 kg, 0.92 µm, 6.91mm and 6.82 N/ktex, for predicting Greasy Fleece Weight, adult Clean Fleece Weight, adult Fibre Diameter and adult Staple Length. Table 5 presents a comparative study of different ML algorithms for crop yield prediction.

TABLE 5.
Different ML Algorithms for Crop Yield Prediction

Reference	Attributes	Experimental Place	ML Algorithm	Accuracy Measure
Kamir et al. [48]	Determine yield for wheat crop	High-resolution yield maps using: weather data, NDVI time-series data, yield data from harvesting equipment, crop type, and geolocation data	Ensemble-based learners using 9 ML algorithms	Support vector regression with radial basis function displayed lowest RMSE of 0.59 and R^2 score of 0.73
Aghighi et al. [51]	Determine yield for Silage Maize crop	Historic crop yield data for from 40+ silage maize fields and time-series data for surface reflectance and climate data record using Landsat 8 OLI satellite	Boosted Regression Tree, Random Forest Regression, Support Vector Regression, and Gaussian Process Regression	Boosted Regression Tree model predicted satisfactorily values for 2013-14 with an R-value of 0.87
Kuwata and Shibasaki [52]	Determine crop yield for corn crops	Enhanced vegetation index from MODIS, climatic data, and 5 years moving average for corn yield in the U.S.	DL, SVR	DL model presents the best results with R of 0.810 and RMSE is 6.298
Kulkarni et al. [53]	Determine crop yield for rice crops	Historic soil properties and rainfall data (31 years period)	Recurrent Neural Networks	RMSE of 41.497 and MAE of 41.6
Chu and Yu [54]	Summer and winter rice prediction	Meteorology data, and area data from 81 counties in the Guangxi Zhuang Autonomous Region, China	BBI model fusion of two BPNN with independently RNN (IndRNN)	Summer Rice Prediction with MAE of 0.0044 and RMSE of 0.0057 Winter Rice Prediction with MAE of 0.0074 and RMSE of 0.0192
Feng et al. [55]	Wheat Yield Prediction	New South Wales (NSW) wheat belt located in southeastern Australia	A hybrid APSIM, NDVI, SPEI with MLR and RF	APSIM+ RF hybrid model best predicts one month before harvest with MAPE of 17.6%, RMSE of 0.70 t ha^{-1}, and ROC score of 0.90, and two months before MAPE = 27.1%, RMSE = 1.01 t ha^{-1}, and ROC score = 0.88)
Cai et al. [56]	Wheat Yield Prediction	Climate and Satellite data of Australia	LASSO, SVM, RF, NN	SVM best predicts with R^2 around 0.75
Nevavuori et al. [58]	Wheat and barley yield prediction	NDVI and RGB data acquired from UAVs in Pori Finland	CNN	With data acquired in early June 2017, MAE is 484.3 kg/ha and MAPE is 8.8% With data acquired around July and August 2017, MAE is 624.3 kg/ha and MAPE is 12.6%
Kouadio, et al. [60]	Robusta coffee yield prediction	Lam Dong Province, Vietnam	ELM, MLR, and RF	ELM best predicts with RMSE of 496.35 kg ha^{-1} or ±13.6%, and MAE of 326.40 kg ha^{-1} or ±7.9%
Gyamerah et al. [62]	Groundnut and millet yield prediction	Statistics, Research and Information Directorate (SRID) of the Ministry of Food and Agriculture, Ghana	Quantile RF and Epanechnikov kernel function (QRF-E) model	QRF (mean) best predicts for Groundnut with RMSE of 0.3787 t/ha, MAPE of 24.0026%, R2 of 0.9830 and Bias of 0.3394t/ha. QRF (mean) best predicts for Millet with RMSE of 0.0173 t/ha, MAPE of 0.9090%, R2 of 0.9805 and Bias of 0.0100 t/ha
Khaki and Wang [64]	Hybrid maize yield prediction	U.S. and Canada	Deep NN, Shallow NN, Regression Tree	Deep NN best predicts with RMSE of 12% of average yield and 50% of the standard deviation for predicted weather dataset. RMSE reduced to 11% of average yield and 46% of the standard deviation for perfect weather dataset

Leroux et al. [65]	Maize and cotton yield prediction	Burkina Faso, West Africa	RF, MLR	RF best predicts with (R^2 of 0.59) at the end of season and R^2 of 0.49 approximately two months before harvest
Maimaitijiang et al. [67]	Soybean yield prediction	RGB, multispectral, and thermal images were collected through UAV in Columbia, Missouri, USA	Partial Least Squares Regression (PLSR), RF, SVR, input-level feature fusion based DNN (DNN-F1) and intermediate-level feature fusion based DNN (DNN-F2)	DNN-F2 best predicts with an R^2 of 0.720 and an RMSE of 15.9%
Shahinfar and Kahn [70]	Wool growth prediction	Australian Merino sheep	ANN, Model Tree, Bagging, Linear Regression	Model Tree best predict with R of 0.93, 0.90, 0.94, 0.81 and 0.59 and MAE of 0.48 kg, 0.41 kg, 0.92 μm, 6.91mm and 6.82 N/ktex for prediction of adult Greasy Fleece Weight, adult Clean Fleece Weight, adult Fibre Diameter, adult Staple Length, and adult Staple Strength.

C. Disease and Weed Detection

Disease fungi, microorganisms, and bacteria take their energy from the plants they live on, which in turn affects the crop yield. If not detected at the right time may account for a huge economic loss to farmers. A lot of financial burden goes to a farmer in the form of pesticides, to get rid of diseases and restore the functioning of crops. Excessive use of pesticides also leads to environmental damage and the effects of the water and soil cycle of the agricultural land.

Using an optimally designed AI system during crop growth period not only reduces the risk of crop disease and minimizes the economic impact, but it also results in minimizing the adverse impact of unsystematic farming on the environment. Sambasivan and Opiyo [71] used a CNN based DL model to detect disease in cassava crops for imbalanced datasets. The authors took a database of 10,000 labeled images that were pre-processed to improve the image contrast using contrast limited adaptive histogram equalization algorithm. The model evaluation was done using the performance metrics: confusion matrix, accuracy measure, precision measure, sensitivity, and F1 score. The authors reported a best-case accuracy of 99.30% and the lowest accuracy was reported as 76.9%. Ramcharan et al. [72] used DL algorithms to detect diseases in cassava crops. Authors deployed deep CNN to identify three different diseases and two types of pests from a set of 11,670 images dataset. Author's utilized GoogLeNet algorithm based Inception v3 in Tensor Flow. The authors achieved efficiency in a range of 80% to 93.0%, and the validation of the results was done with the help of the confusion matrix.

Mohanty et al. [73] employed DL methods to detect crop disease from the image dataset of plant leaves. The authors used a public database consisting of smartphone generated 54,306 images of diseased and healthy plants leaves. These images were resized to 256×256 pixels and were assigned 38 different class labels of crop-disease pair, and transformed into 3 datasets color, grayscale and segmented. The dataset was then fed to two of the most common deep CNNs: AlexNet [74] and GoogLeNet [75]. The authors achieved an accuracy of 99.34% for GoogLeNet, and an accuracy of 85.53% for AlexNet network. The results were validated using F_1 score, authors achieved a mean F_1 score of 0.9886 for GoogLeNet, and a mean F_1 score of 0.9848 for AlexNet.

Amara et al. [76] used LeNet based CNN architecture for disease detection in banana leaves. Authors utilized data from open source local and digital libraries which were pre-processed and resized to 60×60 pixels, and the model was implemented for RGB as well as grayscale images. Hughes and Salathe [77] utilized this developed model for the identification of diseases in the images dataset. The authors achieved the best F_1 score of 0.9971 for detection in RGB images and a score of 0.976 for grayscale images.

Ferreira et al. [78] deployed CNN for the identification of weeds in soybean crops. The image dataset for soy plantation located at São José farm, Campo Grande Brazil was acquired using phantom DJI3 drone. The images are segmented using the SLIC algorithm into square grids. For training, the segmented images were manually annotated to their class. The segmented images dataset was fed to AlexNet (a convolution neural network) for classification. The performance of the AlexNet was compared with SVM, AdaBoost, and RF. To evaluate the performance of the AlexNet the model was fed with a balanced dataset and the authors reported an overall accuracy of above 90% and 96.3% images were correctly classified. Waheed et al. [79] proposed a cost-effective optimized dense CNN (DenseNet) for disease detection in corn leaves with an accuracy of 98.0%. Simulation results show that the proposed model outperforms other CNN models such as EfficientNet, VGG19Net, NASNet, and XceptionNet in terms of fewer parameters, accuracy, computation complexity, and computation time. Pereira et al. [80] proposed an expert system for identifying three species of aquatic weeds from aquatic weed leaves dataset based on their shape and supervised pattern recognition techniques. The author explored five shape descriptors with different shape-based skills viz. Beam Angle Statistics (BAS), Fourier Descriptors (FD), Moment Invariants (MI), Multi-scale Fractal dimension (MS), and Tensor Scale Descriptor (TSD) along with five ML algorithms viz. Optimum-Path Forest (OPF), SVM, Naive Bayes, ANN, MLP. Simulation results show that OPF using the BAS-100 descriptor presents the best results with a recognition rate of 96.41% in comparison to other approaches. Jiang et al. [81] proposed a semi-supervised CNN feature-based Graph Convolutional Network (GCN) for identifying weeds utilizing 6000 images of corn, lettuce, radish, and mixed weed dataset. The proposed approach works in two parts, i.e. in the first part CNN model is used for feature extraction thereafter in the second part GCN graph is explored utilizing CNN feature dataset for extracting feature of an unlabelled dataset using labelled dataset. The proposed approach shows the best results in comparison with AlexNet, VGG16, and ResNet-101 approaches with recognition accuracies of 97.80%, 99.37%, 98.93%, and 96.51% on four different weed datasets.

Oppenheim and Shani [82] explored CNN for identifying four different types of diseases in potatoes. The simulation result shows that the model trained on 90% of the images and tested on 10% of images give the best results with 96% accuracy. Sugar beet contributes around 30% of world sugar production. Leaf spot diseases in sugar beet can create a loss of around 10 % to 50 % of yearly sugar yield. Rumpf et al. [83] proposed SVM with a radial basis function as a kernel based model for early detection and classification of three diseases Cercospora, leaf spot, leaf rust, and powdery mildew in sugar beet leaves. Diseased and non-diseased leaves were classified with an accuracy of 97% and three diseases were identified with accuracy higher than 86%. Ozguven and Adem [84] proposed updated faster R-CNN for leaf spot disease identification and classification in sugar beet. Leaf spot disease initially generates as small circular spots and later spread over the entire leaf surface. The proposed R-CNN architecture changes its parameters according to the images and the disease infected regions, which improves the overall classification rate to 95.48%. Bah et al. [85] explored CNN for weed detection in images obtained using UAV from bean and spinach fields. The proposed model first identifies the crop rows and then identifies the inter-crop row weeds which are used as a training dataset for CNN for crop and weed identification and classification. Kerkech et al. [86] identified the vine diseases from visible and infrared UAV images obtained in the Center Val de Loire region in France. A CNN model is trained with this dataset of images to classify each pixel according to different instances, namely, shadow, ground, healthy, and symptom. The model identifies with an accuracy of 92% at grapevine-level and 87% at leaf level. Oslen et al. [87] explored robust deep learning models Inception-v3 and ResNet-50 for weed species identification and classification from a dataset of images collected in Australian rangeland. Simulation results show that the average classification performance of both the models is 95.1% and 95.7%. These results found fruitful for automatic real-time robotic weed control in the agricultural field.

Sudars et al. [88] establish an experimental set up with RGB digital cameras in Latvia to collect images of the field having 6 food crops and 8 weed species grown in normal field conditions and controlled environment. This dataset can be utilized by deep learning algorithms for weed identification and classification. Sethy et al. [89] identified the rice leaf disease based on a hybrid CNN and SVM. In this model, CNN is explored for deep feature extraction from 5932 diseased rice leaf images and this data is used as input for SVM classifier. The resnet50 with SVM classification model best classify with respect to other models with F1 score of 0.9838. Garcia et al. [90] proposed an ML and DL learning hybrid approach for weed and crop identification in the agriculture fields of Greece. Image dataset of two crops tomato and cotton and two weeds black nightshade and velvetleaf was generated for training and testing of the model. Initially CNN (Xception, Inception-Resnet, Vignette's, Mobilenet, and Densenet)) is used for feature extraction and this feature set is later used to train ML classifier (SVM, XGBoost and Logistic Regression) for classification. The simulation result shows that Densenet and Support Vector Machine outperforms other approaches with F1 score of 99.29%. Shah and Jain [91] identified the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3048415, IEEE Access

disease in cotton leaf through ANN with some image pre-processing techniques. Yu et al. [92] explored deep learning algorithms with a dataset of images for identifying dandelion, ground ivy, and spotted spurge in perennial ryegrass. Parraga-Alava et al. [93] generated a robusta coffee leaf image dataset (RoCoLe) for disease identification using ML algorithms.

Glezakos et al. [94] proposed an innovative method to identify two viruses Tobacco Rattle Virus (TRV) and the Cucumber Green Mottle Mosaic Virus (CGMMV) in plants. In the proposed research Bio-Electric Recognition Assay (BERA) technique is utilized to obtain time-series information of the two viruses by measuring the waves through biosensors for 331s. This time-series data is preprocessed using GA to eliminate noise and for dimensionality reduction of a large dataset. Thereafter this meta-data is used to train MLP neural network classifier. The proposed model is tested against other ML classifiers via cross-validation. Ramesh and Vydeki [95] explored optimized deep NN with the Jaya algorithm for the identification of paddy leaf diseases. A dataset of rice plant leaves was taken from the agricultural field to identify and classify normal, bacterial blight, brown spot, sheath rot, and blast diseases. Simulation results show that the proposed model accurately classifies the diseased and normal images with an accuracy of 98.9%, 95.78%, 92%, 94%, and 90.57% for blast affected, bacterial blight, sheath rot, brown spot, and normal rice leaf images. Chechlin´ ski et al. [96] explored CNN for weed identification in four plant species at different growth level and under varying light conditions. CNN architecture combines U-Net, MobileNets, DenseNet, and ResNet models for classification of weeds in crops. In [97-100] author has reviewed machine and deep learning techniques for weed, pests and disease identification, and classification in crops at different growth stages. Table 6 presents a comparative study of different ML algorithms for disease and weed identification.

TABLE 6.
Different ML Algorithms for Disease and Weed Identification

Reference	Attributes	Experimental Place	ML Algorithm	Accuracy Measure
Sambasivan and Opiyo [71]	Disease detection in cassava crops	10,000 labeled high-resolution image database collected from Uganda	CNN based DL model	Best case accuracy is 99.30%
Ramcharan et al. [72]	Disease and pest detection in cassava crops	Digital image dataset collected from Tanzania	CNN Inception v3 based on GoogLeNet	Best case accuracy is 93.0%
Mohanty et al. [73]	Twenty-six disease detection from 54,306 leaf image	Smartphone gathered image dataset of diseased and healthy plant leaves images	CNN based GoogleNet and AlexNet	99.34% accuracy for GoogLeNet, 85.53% of accuracy for AlexNet
Amara et al. [76]	Classification of banana leaf diseases	Open course local and global repositories for training the model.	CNN based LeNet	Performance measures for RGB images: Accuracy of 99.72%, F1-score of 0.9971, Precision of 99.70%, Recall of 0.9972 is achieved for dataset split into 50% for train and 50% for test. Performance measure for Gray Scale images: Accuracy of 97.57%, F1-score of 0.9764, Precision of 97.6%, Recall of 0.975 is achieved for dataset split into 40% for train and 60% for test.
Ferreira et al. [78]	Weed detection in Soybean crops	A dataset of high-resolution field images collected through phantom DJI3 drone in Brazil	CNN based CaffeNet	Accuracy higher than 98% for the classification of all classes.
Waheed et al. [79]	Corn Leaf disease identification and classification	A dataset of 12,332 images of 250 X250 dimension is collected from different sources	Dense CNN (DenseNet)	Accuracy of 98.06% is achieved in recognizing three corn disease
Jiang et al. [81]	Weed and Crop Recognition	A dataset of Corn, Lettuce, Radish and Mixed Weed	CNN based GCN ResNet-101	Recognition accuracy of 97.80%, 99.37%, 98.93% and 96.51% is achieved for four dataset
Oppenheim and Shani [82]	Potato disease classification	A dataset of contaminated images of potatoes with different shape and size and tone is acquired under normal conditions	CNN	The model trained with 90% of images and tested with 10% of images gives the best results with 96% of accuracy
Rumpf et al. [83]	Early detection and identification of Sugar beet disease	Hyperspectral data is obtained for healthy and leaves with the disease for 21 days	SVM with radial basis function as a kernel-based model	Classification accuracy of 97% is obtained to classify healthy and diseased leaves and accuracy between 65 % to 90% is obtained

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3048415, IEEE Access

				for Presymptomatic disease identification
Ozguven and Adem [84]	Detection and classification of leaf spot disease in sugar beet	A dataset of healthy and diseased sugar beet leaf images	R-CNN	The proposed model presents an accuracy of 95.48% in disease detection and classification
Bah et al. [85]	Weed detection	A dataset of UAV images obtained from bean and spinach fields	CNN based ResNet	In bean field AUCs is 88.73% for unsupervised data and 94.84% for supervised data In spinach field, the AUCs is 94.34% for unsupervised data and 95.70% for supervised data
Kerkech et al. [86]	Vine disease detection	A dataset of visible and infrared UAV images obtained in Center Val de Loire region in France	CNN	Accuracy of 92% at grapevine-level and 87% at leaf level is achieved
Oslen et al. [87]	Weed identification	A dataset of 17,509 labelled images is obtained of eight weed species across 8 locations in northern Australia	Two deep learning models Inception-v3 and ResNet-50	Average classification accuracy of 95.1% and 95.7% is obtained for Inception-v3 and ResNet-50
Sethy et al. [89]	Rice leaf disease identification	A dataset of 5932 diseased rice leaf images are obtained from fields of western Odisha and database http://bcch.ahnw.gov.cn/Right.aspx	Deep features ResNet50 plus SVM model	F1 score of the proposed model is 0.9838
Garcia et al. [90]	Crop and weed identification	A dataset of visible images of tomatoes and cotton crops and two weed species is obtained from fields across Greece	A hybrid approach of CNN (Xception, Inception-Resnet, VGNets, Mobilenet, and Densenet) with classifier (SVM, XGBoost and Logistic Regression)	Densenet and SVM best classify with a micro F1 score of 99.29%
Glezakos et al. [94]	Identify two viruses Tobacco Rattle Virus (TRV) and the Cucumber Green Mottle Mosaic Virus (CGMMV) in plants	Bio-Electric Recognition Assay (BERA) technique is utilized to obtain time-series information of the two viruses by measuring the waves through biosensors	Genetic algorithm and ANN	NN achieves a success rate of 93% in 234th generation to identify two viruses
Ramesh and Vydeki [95]	Recognition and classification of disease in rice leaf	A dataset of leaf images with different disease spread is taken from the fields of ayikudi and panpoli, Tirunelveli District, Tamilnadu	Jaya Algorithm Optimized Neural network	Accuracy of 98.9% for the blast affected, 95.78% for the bacterial blight, 92% for the sheath rot, 94% for the brown spot and 90.57% for the normal leaf image is achieved
Chechlin´ ski et al. [96]	Weed and Crop Identification	A dataset of beet, cauliflower, cabbage, strawberry images is obtained at different growth stages	CNN architecture U-Net, MobileNets, DenseNet and ResNet is used	Classify 47–67% of weed area and misclassifies as weed around 0.1–0.9% of crop area

D. Drip Irrigation

In the modern era, irrigation for crops has been improvised using the concept of drip irrigation [101], where the system consists of thin plastic tubes placed in or above the soil along the vertical rows of the plants for nurturing the water supply to the crops. Employing the proper operational management of drip irrigation, minimizes the utilization of water supply for crop production, and provides a better yield of crops. Socio-economic and environmental demands have widely appreciated in use of drip irrigation on farmlands for agriculture, especially for the high cost valued crops i.e., vegetables and fruits. Furthermore, drip irrigation is based on the low-pressure watering system in comparison to sprinkler systems; this makes the system more efficient in terms of energy consumption [102]. Various advantages have been observed using drip irrigation in agriculture over other irrigation systems which include sub-irrigation systems or sprinkler irrigation systems. These advantages are entitled to minimal usage of water supply, usage of soluble fertilizers through a drip irrigation system, automated system, minimization of soil erosion, uninterrupted activities, minimized weed problems, facilitation of double-cropping. Precision irrigation is another innovative approach in intelligent farming where it uses the water intelligently that further helps the farmers to achieve better yield in crops with minimal water usage. It can also be featured as providing the right amount of water, at the right time and the right place in the field. It focuses its implementation based on variable rate irrigation (VRI) methods employing drips or sprinklers. [103-105]. Advancements in the field of on-farm sensor technologies, weather forecasting, IoT based sensor detection system of vegetation and precision-based smart irrigation produces a huge size of data that ultimately benefits the farmers in optimizing the usage of water resources, improve the yield of crops and maximizes the profit of farmers [106].

ML and DL and reinforcement learning are employed on the historical data and it provides various opportunities for real-time prediction and decision making purposes for smart irrigation which are solely based on the data collected by the sensors and IoT enabled systems [107-110]. Roberts et al. [111] have discussed that a sensor-based control system might create some bottleneck in terms of reducing the reliability of decision support tools on process-based crop models, which further may require costly calibration and affect in generating an uncertain representation of soil-plant-atmosphere processes. Further, ML techniques have been employed extremely well for protection analysis of hydrological processes i.e., soil moisture and groundwater levels [112-113]. Li et al. [114] utilized ANN for estimating nitrate distribution in different types of soils under a drip irrigation system.

Kavianand et al. [115] proposed a fully automated drip irrigation system based on the ARM9 processor along with different kinds of sensors equipped for monitoring the PH content and nitrogen content of the soil and controlling the irrigation of the field. Emmanuel et al. [116] establishes an experimental set-up in a greenhouse in Malaysia to monitor the growth of mustard leaf vegetable plants through IoT devices and alongside developed a data-driven model of drip irrigation system. Soil moisture, irrigation volume, evapotranspiration were measured through sensors and were given to the Raspberry Pi 3 controller for storing it in the cloud. This data was utilized by different predictive models ARX, BJ, and state-space models to predict soil moisture content for an optimized drip irrigation system. ARX model outperforms other predictive models in terms of MSE and response time. Seyedzadeh et al. [117] explored ML algorithms to optimize the uniform emitter discharge rate of drip irrigation system under varying pressure and temperature conditions. In this model operating pressure, water temperature, discharge coefficient, pressure exponent, and nominal discharge were taken as input parameters while ration of emitter discharge to nominal discharge is taken as output temperature. Authors explored four different ML algorithms for optimizing emitter discharge rate and simulation results show that LS-SVM presents best results with the least error of mean absolute error. Peng et al. [118] utilize soil moisture, soil electrical conductivity, air temperature, and light intensity parameters to build an optimized irrigation prediction model using backpropagation NN in China. The proposed prediction model presents good results with MSE of 0.00857724. The authors also identified an optimized layout and network arrangement for pipe in a drip irrigation system using computational fluid dynamics (CFD) software. The simulation results present that the H-shaped network layout is more suitable for field crop irrigation than the comb-shaped and fish bone-shaped layout. Drip irrigation system gives the best performance when the wetting front dimension, i.e. diameter, depth, and upward movement are optimized. Shiri et al. [119] explored soft-computing approaches viz., gene expressions programming (GEP), and RF techniques in modeling wetting front dimensions over different soil types for surface and sub-surface irrigation system. Proposed model, best predicts ETc with an improved correlation coefficient and decreases MSE and MAE. Elnesr and Alazba [120] explored ANN for predicting the wetting front dimensions from the dataset of a well-tested HYDRUS 2D/3D model. The simulation results show that the proposed model has a good correlation of 0.93-0.99.

Chang et al. [121] developed a smart irrigation model based on ML with the LoRa P2P network to learn the irrigation experiences from the expert farmers working on greenhouse organic crops. Singh et al. [122] have discussed an ML and IoT based model for soil moisture prediction during irrigation. Torres-Sanchez et al. [123] proposed a decision support system for irrigation management of citric crops in southeast Spain. In the proposed model smart sensors are deployed in the field to monitor water supplied previous week, weather data, soil water status, and based on this data three regression models SVM, RF, and Linear regression was trained to build the irrigation decision support system. RF best predicts with comparatively less prediction error.

Hellín et al. [124] explored the Partial Least Square Regression (PLSR) and Adaptive Neuro-Fuzzy Inference System (ANFIS) model for building a smart irrigation decision support system crops in southeast Spain. Goumopoulos et al. [125] proposed a real-time adaptable intelligent and autonomous closed-loop irrigation management system. The authors built an experimental set-up in a greenhouse and deployed wireless sensors for monitoring the plant growth and environmental conditions along with plant growth control actuators. Estimation of evapotranspiration (ETc) plays a vital role in water resource management system. Chen et al. [126] estimate crop actual ETc from temporal convolution network (TCN) from the dataset of lysimeters for maize under drip irrigation with film mulch. Simulation results show that the proposed model best predicts ETc with an improved correlation coefficient and decreases MSE and MAE. Table 7 presents a comparative study of different ML algorithms for drip irrigation.

TABLE 7.
Different ML Algorithms for Drip Irrigation

Reference	Attributes	Experimental Place	ML Algorithm	Accuracy Measure
Li et al. [114]	Estimation of nitrate distribution in the soil under fertigation through drip-irrigation systems	A dataset is created through an experiment on loam soil	ANN	R^2 of 0.83 is obtained
Seyedzadeh et al. [117]	Estimation of Emitter outflow discharge under varying temperature and pressure	A dataset is obtained from a physical model of drip irrigation built at Water and Soil Research Laboratory in the University of Kurdistan	ANN, neuro-fuzzy sub-clustering (NF-SC), neuro-fuzzy c-Means clustering (NF-FCM) and LS-SVM	LS-SVM1 model best predicts with minimum SI and MAE values of 0.073 and 0.054 and the maximum R^2 of 0.942
Peng et al. [118]	Water demand prediction and drip irrigation network optimization	Agriculture field in China	BPNN	R and MSE value of 0.98963 and 0.00857724 was obtained for the water demand prediction and drip irrigation pipe network in the H-shaped arrangement is most suitable for drip irrigation
Shiri et al. [119]	Wetting front dimension estimation in surface and sub-surface irrigation system	A dataset is obtained from the laboratory experiments conducted in the central laboratory of the Agricultural College of University of Kurdistan	Gene expressions programming (GEP) and RF	Surface irrigation system ΔSI and ΔNS is 0.098 and 0.094 for GEP and 0.068 and 0.032 for RF for estimation of D_h and ΔSI and ΔNS values were 0.114 and 0.068 for GEP and 0.115 and 0.099 for RF for D_V estimation Subsurface irrigation system ΔSI and ΔNS is 0.051 and 0.025 for GEP and 0.037 and 0.052 for RF for estimation of D_h
Chang et al. [121]	Determination of the soil prospect for smart irrigation	Light intensity and soil humidity	Multiple linear regression algorithm	MSE of 91 and R^2 of 0.74
Singh et al. [122]	Determination of soil moisture	Dataset of air temperature, air humidity, soil moisture, soil temperature, radiation, and the weather forecast is created	Multiple Linear Regression (MLR), Elastic Net Regression (ENR), Gradient Boosting Regression Trees (GBRT) and Random Forest Regressor (RFR)	GBRT best predicts with MSE of 4.04 and R^2 of 0.94 considering soil temperature
Hellín et al. [124]	A decision support smart irrigation system based on climate and soil variables	Data obtained from soil sensors & weather information obtained from citrus trees located in the South-East of Spain	PLSR and ANFIS	Accurately measures the irrigation requirement when soil sensor information is added to weather information

E. Livestock Production and Management

Livestock production is basically related to the production and management of cattle i.e., sheep, pigs, etc. for human consumption in terms of meat. Livestock production and their management are based on the farming parameters of these cattle i.e., health, food, nutrition, and behaviour to optimize their production in such a way that the economic efficiency of this livestock can be maximized. In the present

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3048415, IEEE Access

scenario, Artificial intelligence, IoT and Blockchain technologies [127] are widely explored to improves livestock sustainability and for analysis of their chewing habits, eating patterns, their movement patterns i.e., standing, moving, drinking and feeding habits, indicate the amount of stress the animal is going through which in turn helps in predicting the vulnerability to disease, weight gain, and production of the livestock. Furthermore, an ML-based weight predicting system can help in the estimation of their body weight 90-180 days before the slaughtering day. According to these analyses and estimations, farmers can change their diet plans and living conditions for their better growth in terms of health, behaviour, and weight gain which in turn will improve the economic efficiency of these livestock [128, 129]. Villeneuve et al. [130] build a decision support system that encounters not only real-time data but also expert knowledge for precision sheep farming.

Livestock production and management can be further classified into two sub-categories, i.e., animal welfare and livestock production. Animal welfare generally deals with the animal's health and their well-being; for this ML techniques are applied to their health monitoring feature for prospective of early disease detection. Whereas, livestock production employs the ML on the estimation of the balanced production of livestock for the producers to achieve economic benefits. Dutta et al. [131] described a procedure for the classification of cattle behaviour employing the ML techniques for data collection using collar-based sensors i.e., magnetometers and three-axis accelerometers. In this study, events such as oestrus and dietary changes on cattle have been analyzed for their well-nutrition. Pegorini et al. [132] presented an automatic identification and classification of chewing habits of claves employing ML-based techniques for analysing their health and behavioural patterns. Ebrahimie et al. [133] proposed ML predictive model for estimating Sub-Clinical Mastitis (SCM) from milking parameters in dairy herds. Mastitis is an inflammatory disease that is widely affecting the dairy industry. Author's explored four classification models decision trees, stump decision trees, parallel decision trees, and random forest to discover SCC independent of Somatic Cell Count (SCC) which is widely used to measure SCM worldwide. RF with Gini Index criteria best predicts SCM with an accuracy of 90%. Ebrahimie et al. [134] explored the attribute weighting model (AWM) for identifying lactose concentration and electrical conductivity in milk, which are two of the major indicators of SCM in dairy cattle. Hyde et al. [135] also explored RF to predict the route of transmission of germs and classify them into contagious (CONT) or environmental (ENV) with ENV further sub-classified into non-lactating "dry" period (EDP) or lactating period (EL). The simulation results show that an accuracy of 98% was achieved for discovering CONT vs ENV and 78% for discovering EDP and EL. Esener et al. [136] utilized spectral profiles dataset to discriminates CONT and ENV strains using GA, NN, and quick classifier. Ebrahimi et al. [137] predicted sub-clinical bovine mastitis using a large milking dataset collected through an automated in-line monitoring system in commercial New Zealand dairy farm. The simulation results show that GBM outperforms other ML model and best predict sub-clinical bovine mastitis with an accuracy of 84.9%. Sharifi et al. [138] explored meta-analysis and decision trees data mining tools to discover genes that can help to find mastitis in dairy cattle. Machado et al. [139] explored the RF model to identify factors influencing the occurrence of Bovine viral diarrhea virus (BVDV) viral disease in cattle in southern Brazil. The proposed approach identifies that insemination, the number of cattle in neighbouring farms, and routine rectal palpation are among the main factors of the occurrence of this disease.

Matthews et al. [140] developed an ML-based automated monitoring system for tracking animal behaviour and movement i.e., standing, moving, feeding, and drinking by employing the depth video cameras and sensors. Qiao et al. [141] explored the DL technique Mask R-CNN for examining cattle health and welfare information in precision livestock management. The proposed model extract key features from image frames, enhance the image to remove non-uniform illumination shadow influences, segment image of cattle from the background image using the Mask R-CNN DL tool, and lastly extract cattle contour lines from the segmented image. The proposed approach outperforms SharpMask and DeepMask image segmentation models with mean pixel accuracy of 0.92 and an average distance error of 33.56 pixels.

Liakos et al. [142] explored ML model for predicting healthy cattle and cattle suffering from lameness utilizing basic features of cattle which includes per day habits of cattle like steps taken, overall walking, lying, and eating habits. Morales et al. [143] employed a method based on SVM for early detection, warnings, and production issues of eggs in the poultry farms. The simulation results show that the proposed technique alerts a day before with an estimation accuracy of 0.9854. The identification of livestock is an important aspect of monitoring growth and animal welfare. Hansen et al. [144] explored deep learning techniques CNN for identifying pigs faces from the dataset of digital images of pigs obtained from commercial farm environment where the parameters such as dirt and lighting are highly unpredictable. The proposed approach accurately predicts the faces with an accuracy of 96.7%.

Fenlon et al. [145] build a decision support system using predictive ML algorithms to provide calving assistance in the dairy industry. Four ML techniques multimodal regression, decision trees, RF, and NN were explored to predict three calving difficulties unassisted, slight assistance, and veterinary assistance. The simulation result shows that NN and multimodal regression models accurately classify 75% of calving difficulties with an average prediction error of 3.7 % and 4.5%. Fenlon et al. [146] analyzed calving difficulties in dairy herds in Ireland using ML algorithms. A dataset of parity, log days in milk, inter-service interval, difficulties faced in the last calving, herd body conditions were built to predict conceptions using artificial insemination in the Iris dairy industry. Logistic regression outperforms RF, decision trees, and Naive Bayes in predicting conception using artificial insemination. Borchers et al. [147] explored RF, linear discriminant, and NN for calving prediction in dairy cattle by examining their behaviour which includes number the of steps, lying time, standing time, transition from one state to other and total motion 14 days before the predicted calving date. Although, the innovative algorithms play a crucial role in livestock management but combining livestock data with public data will improve precision livestock farming standards [148]. Table 8 presents a comparative study of different ML algorithms for livestock production and management.

TABLE 8.
Different ML Algorithms for Livestock Production and Management

Reference	Attributes	Experimental Place	ML Algorithm	Accuracy Measure
Dutta et al. [131]	Cattle behavior classification	Dataset is collected from the Tasmanian Institute of Agriculture Dairy Research Facility at Elliott	Binary Tree, Linear Discriminant Analysis, Naive Bayes, k-NN, ANFIS classifier	Bagging ensemble classification with Tree learner best classify with an accuracy of 96%, sensitivity of 97%, specificity of 89%, F1 score of 89% and false discovery rate of 9%
Pegorini et al. [132]	Automatic identification and classification of chewing habits of livestock	Dataset of stress is created which is measured by FBG sensor attached to the animal's jaw by surgical screws	Decision Trees Algorithm	Chewing process is classified with an overall accuracy of 94%
Ebrahimie et al. [133]	Prediction of SCM	Dataset of 346,248 recorded samples with eight variables recorded at different milking time in Australia	Decision Tree, Stump Decision Tree, Parallel Decision Tree, and RF Decision Tree	RF, Decision Tree with Gini Index criterion best predicts with an accuracy of 90%
Hyde et al. [135]	Prediction of mastitis infection patterns in dairy herds	Dataset of herd mastitis is obtained from 1000 UK dairy herd between 2009 and 2014	RF	An accuracy of 98%, PPV of 86% and NPV of 99% is obtained for the diagnosis of CONT vs ENV (with CONT as a positive diagnosis), and an accuracy of 78%, PPV of 76% and NPV of 81% for the diagnosis of EDP vs EL (with EDP as a positive diagnosis)
Ebrahimi et al. [137]	Prediction of sub-clinical bovine mastitis	Data is collected from a commercial New Zealand dairy farm in Ongaonga, Hawkes Bay, from July 2011 to June 2013 and then it is transformed by Z-Standardization	DL, Naïve Bayes, Generalized Linear Model, Logistic Regression, Decision Tree, GBT, and RF	GBT best predicts with an accuracy of 84.9%
Machado et al. [139]	Identification of factors associated with the occurrence of Bovine viral diarrhea virus (BVDV)	Dataset of dairy herds is obtained from the Rio Grande do Sul, a southernmost state of Brazil	RF	RF presents an average error rate of 32.03% for the negative class of BVDV and 36.78% for the positive class of BVDV
Qiao et al. [141]	Cattle segmentation and contour extraction	Dataset of cattle images is collected from Brisbane, Australia	Mask R-CNN	A mean pixel accuracy of 0.92 is obtained for cattle segmentation and contour extraction average distance

				error of 33.56 pixels is obtained
Hansen et al. [144]	Identification of livestock such as pigs and cows	Dataset of digital images of pig face is created	Conventional neural networks	Pigs faces are recognized with an overall accuracy of 96.7%
Fenlon et al. [145]	Predicting calving difficulty in dairy heifers and cows	Dataset is obtained from 2,076 calving records in 10 Irish dairy herds	Multimodal regression, Decision Trees, RF, NN	Multimodal regression and NN best classify 75% of calving cases with an average error in the predicted probability of 3.7% and 4.5%
Fenlon et al. [146]	Prediction of insemination outcome in Irish dairy cows	Dataset of 2,723 artificial insemination records from Irish research farms and 4,205 breeding events from commercial dairy farms	Logistic regression, Naive Bayes, Decision Tree, and RF	Logistic regression best predicts with a precision of 57.89%, Recall of 44.82%, F-score of 50.52% and Matthews correlation coefficient of 0.16
Borchers et al. [147]	Prediction of Calving in dairy cattle	Dataset is obtained from 20 primiparous and 33 multiparous Holstein dairy cattle from University of Kentucky Coldstream Dairy between September 2011 to May 2013	RF, linear discriminant analysis and NN	NN best predicts in 2-h periods in the 8 h before calving with 82.8% sensitivity and 80.4% specificity

F. Intelligent Harvesting Techniques

Smart harvesting systems helps the farmers to harvest agriculture goods by reducing human efforts. In this approach, technologies such as smart sensors, robotics, UAVs, and IoT devices [149], AI, and ML-based computer vision techniques are employed to intelligently harvest the crops. The research community has provided a comprehensive

review of different intelligent techniques used to automate the agriculture industry [150-152] and have analyzed the potential and challenges of this decision support system [153]. In the last few years, different robots have been developed for harvesting fruits and vegetables [154]. Smart harvesting offers better insight into the crops and helps farmers to achieve the potential harvest of crops which leads to increased productivity. Smart harvesting system has numerous advantages in comparison to traditional harvesting approaches like it requires less labour, optimized crop yield, maximum probability, better insight into crops, reduced cost of harvesting, and cost-efficient production.

A significant problem in the Japanese agriculture industry is a labour shortage. Sakai et al. [155] utilized machine vision for asparagus robot harvesting in Nagasaki prefecture. The speed of asparagus robot harvesting is three times faster than the human being. Since asparagus harvesting is modeled on their size and doesn't require color properties thus laser sensor is used to collect 3D distance information in the proposed work. Monta et al. [156] also explored laser sensors along with color cameras for tomato harvesting through robots. Preter et al. [157] developed an autonomous system consisting of e-vehicle, cameras, robotic arm, localization system, gripper, quality monitoring, and logistic handling system, which can efficiently detect, plucks, and puts the strawberries in a box. The proposed robot prototype is fast enough to pluck the fruit in just 4 seconds. Hayashi et al. [158] practically evaluated the performance of strawberry harvesting robots in a greenhouse test field. The proposed autonomous system efficiently access the fruit position and its maturity level and pick the fruit with and without suction in a duration of two to three weeks without damaging the fruit. Horng et al. [159] proposed a smart harvesting system that employs IoT and smart image recognition systems for the detection of mature crops using object detection feature trained on MLP neural network. The mature crop can be harvested using a robotic arm whose movement is predicted using ML algorithms.

Zhang et al. [160] explored Regions-CNN (RCNN) for multi-class canopy object detection in shake and catch the apple harvesting system. A dataset of RGB images was created in the commercial orchard using a Kinect v2 sensor and pre-trained RCNN is utilized for real-time detection of apple, branches, and trunks. The authors also developed an estimation algorithm to predict shaking location based on the results of RCNN. Spectral and thermal images have also been explored for the detection of fruits and vegetables [161, 162]. Zhang et al. [163] investigated eleven canopy parameters using principal component analysis (PCA)

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3048415, IEEE Access

and classified the removal status of apples into mechanically harvested and mechanically unharvested. Zhang et al. [164] reviewed technology progress in the mechanical harvesting of apples which includes shake and catch, robots, and harvest assist platforms. Pise and Upadhye [165] explored Naive Bayes and SVM ML techniques for grading of harvested mangoes based on their color, size, features, quality, and maturity. Grading of fruits increases the profit of the agriculture and food industries. A mango image dataset comprising of three different colors red, green, and yellow is created and is used for training and testing the ML algorithm. The proposed approach presents limited scope as it can detect defects in a particular surface area which can be overcome by creating a dataset of rotational view images. Wu et al. [166] explored NN for recognition, classification of fruits and vegetables, and obstacle avoidance in a harvesting robot. Table 9 presents a comparative study of different ML algorithms for intelligent harvesting.

TABLE 9.
Different ML Algorithms for Intelligent Harvesting

Reference	Attributes	Experimental Place	ML Algorithm	Accuracy Measure
Horng et al. [159]	Intelligent harvesting system using IoT and smart image recognition	Dataset of digital images from the agriculture land	MLP, CNN	Object detection model has a mean average precision of 84% and arm movement prediction model has a mean picking accuracy of 89%
Zhang et al. [160]	Multi-class canopy object detection and estimation of shaking location in apple harvesting system	Dataset of color images and corresponding 3D point cloud data of apple trees were acquired in commercial orchards	Regions-CNN based AlexNet, VGG16, and VGG19	VGG19 best predicts with mean Average Precision of 82.4%, average computation time is 0.45s per image
Zhang et al. [163]	Determination of key canopy parameters for mass mechanical apple harvesting	Dataset is collected from the two-year field trials in two commercial apple orchards	k-NN, PCA	In training prediction accuracies is 76-92% and 62-74% for “Scifresh” and “Envy” in testing dataset, the overall test accuracies were 81-91% on “Scifresh” but only 36-79% on “Envy”
Pise and Upadhye [165]	Grading of harvested mangoes based on their quality and maturity	Dataset of mango images	Naive Byes and SVM	Results are more accurate

V. IoT APPLICATIONS IN PRECISION AGRICULTURE

Precision agriculture refers to a system with minimizing direct involvement of the caretaker/farmer except when there is an urgent need or an emergency i.e. when there is a failure in the system. IoT helps in maintaining the defined standards of parameters needed for day to day work in agriculture. The parameters can be measured using the required sensors and can be uploaded to an IoT cloud for remote monitoring so that the direct involvement of farmers is minimized. The IoT cloud can be used for control purposes also, say for example in detecting and avoiding animal intrusion in the agriculture field. Sensors are an integral part of IoT for precision agriculture without which the monitoring and controlling becomes next to impossible task. Figure 6 shows the trend search of keywords “IoT in agriculture” and “sensor in agriculture” on google in the last 10 years. Apart from monitoring and controlling, IoT in agriculture is also used as data-storage technology. Parameters like properties of soil, crop yield, seasonal behaviour data, temperature changes, etc can be stored on the IoT cloud which will be helpful in analyses, prediction, and deciding on estimated crop production.

A. Sensors for IoT in Precision Agriculture

IoT is defined as the interconnection of things, where one example of a thing is a sensor. A group of sensors can communicate with every other sensor and thereby with the control center. A WSN in IoT has the benefits of increasing the efficiency of production, enhancing the yield quality, detecting and avoiding plant-eating pets, detecting the fires in the farms [167]. IoT has helped in increasing the scope of farming, animal, and pet rearing along with smart irrigation [168]. Sensors form an integral part of IoT architecture in agriculture. A sensor is defined as a transducer that converts the sensed parameter (soil moisture, for example) into the equivalent electrical signal.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3048415, IEEE Access

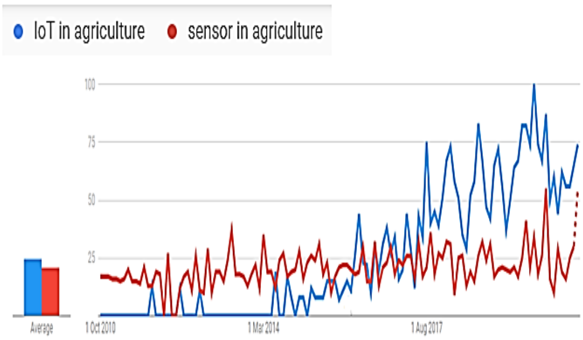


Figure 6. Google trend response for keywords IoT in agriculture and sensor in agriculture for the last 10 years.

Depending on the nature of the output signal they generate, sensors are classified as analog or digital sensors. An analog sensor's output needs to be converted to digital before it is being fed, processed by any IoT system. On the other hand, sensors that generate signals in digital form can be directly connected to any IoT system. Table 10 compares the list of some important sensors applicable in precision agriculture. Addressing the complete list of sensors available for precision agriculture is beyond the scope of this article, although, table 10 provides the list of sensors and their parameters that are very widely used and covers almost every aspect of IoT in agriculture. A pair of sensors and actuators can be used to collect information about some of the vital parameters of precision agriculture and react to perform predefined action whenever required. IoT plays an important role in assuring that the action performed happens instantaneously with minimum delay. The factors that can affect the real-time decision making and causes a delay is the tolerance of the measuring parameter and the communication protocol used. The operating temperature where the sensors are placed have a proportional effect on tolerance. An increase temperature on either side will increase the tolerance of the measuring parameter and sensor reading will deviate the value of the measurand from the actual value. The communication protocol is used to send the readings of the sensor to the microcontroller from where the value will be uploaded to the IoT cloud. The data rate of communication protocol decides the time required for this data transfer.

TABLE 10.
Sensor Parameters used in Precision Agriculture

Sensor	Measuring Parameter	Output (A/D)	Power Consumption (uW)	Operating Temperature Range (°C)	Tolerance	Communication protocol
Accelerometer (ADXL345), [169]	Static and Dynamic acceleration	Digital	62.5-325 (at V= 2.5V)	-40 to 85	± 2g	SPI, I2C
Air Humidity (HTU21D), [170]	Relative Humidity & Temperature	Digital	0.06 to 1350 (at V = 3.0V)	-40 to125	± 2%	I2C
Air Temperature (SHT11), [171]	Relative Humidity & Temperature	Digital	50 (at V=5V & 25°C)	-40 to 123.8	±3.5(%RH) & ±0.5(°C)	I2C
Camera (PixyCam, Raspberry Pi), [172]	Image	Both	0.7	-30 to 70	-	SPI, I2C, UART, Analog, Digital
Capacitive Touch (MPR121), [173]	Capacitance	Digital	104.4	-40 to 85	3.3V	I2C
Depth (Microsoft Kinect 1 & 2), [174]	Depth	Digital	3.3W	5 to 30	-	USB
Force (Flexiforce A20-25) [175]	Force	Analog	0.03	-40 to 60	±3%	USB
Indoor Localization (iBeacon) [176]	Proximity	Digital	0.3W	-	-	BLE
Light (GL5528) [177]	Light intensity	Both	0.1	-30 to 70	±1%	USB, Bluetooth
Proximity (HC-SR501) [178]	Light intensity	Digital	0.065	-15 to 70	±3%	I2C, USB
RGB Colour (TCS34725) [179]	Photodiode current	Digital	705	-40 to 85	±5%	I2C

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3048415, IEEE Access

Soil Moisture (SEN0114) [180]	Resistance	Analog	0.175	-10 to 55	±1%	USB, I2C
Sound (SEN-12642) [181]	voltage	Both	0.1	-30 to 65	±2.5dB	USB
Distance (HC-SR04) [182]	Time	Both	0.075	-15 to 70	±3mm	UART
Pulse Oximeter (SP02) [183]	Current	Digital	<0.007	-10 to 40	±2%	SPI
Temperature (MAX30205) [184]	Temperature	Digital	1800	0 to 50	±0.1°C	I2C
Piezoelectric (PZT) [185]	Voltage difference	Analog	0.001	-70 to 120	±20%	BLE

B. Wireless Sensor Networks in Precision Agriculture

WSN is the collection of spatially displaced sensor deployed to monitor the physical parameters of the environment and coordinating the collected data at central location. IoT transfers the recorded data to cloud which is further processed and analyzed through intelligent algorithms. In precision agriculture integration of artificial intelligence with WSN allows real time monitoring and intelligent decision making in agriculture fields. IoT sensor network which includes soil moisture senor, electrochemical sensor, optical sensors, etc. continuously monitor the field data and works as a training data for ML and DL algorithms. Edge computing enabled AI systems assist in reducing the amount of data to be uploaded to IoT cloud by identification of meaningful data to be communicated and discarding the redundant data.

Intelligent processing of data generated from nodes result in better management of sensor network In [186] author utilized AI driven sensor network to classify land as suitable, more suitable, moderately suitable and unsuitable after every cultivation. In [187] author developed a power efficient WSN using Arduino microcontroller and ZigBee module to monitor and control essential parameters that effect crop growth such as soil and weather conditions in Florida, USA. In [188] author integrating sensor nodes with AI systems to reduce the power consumption of nodes by optimizing the performance and data transmission of respective nodes. RNN based Long-Short term (LSTM) network was built which increases the runtime of a single sensor and guarantees 180 days autonomous operation using Li-ion battery. The proposed system continuously monitors the growth dynamics of plant leaves. In [189] author presents an autonomous system built with low power sensor nodes and IoT based cloud platform to estimate level of phosphorous in soil through ANN. Author incorporates dynamic power management system to maintain balance between energy consumption and estimation accuracy. In [190] author presents GA optimized WSN for precision agriculture applications. Thus, we conclude that integrating artificial intelligence with WSN, IoT plays a key role in assuring the best yield of crops.

VI. ASSESSMENT AND EVALUATION OF KNOWLEDGE-BASED AGRICULTURE SYSTEM

In this section ML algorithms used by different researchers in the precision agriculture system are analyzed. The agriculture industry is facing many challenges across the world, and a knowledge-based agriculture system allows sustainable use of resources by the farmers aiming to get maximum output from the agriculture land. There are two basic stages in precision agriculture, i.e. pre-processing stage and processing stage. In the pre-processing phase market trends are studied and based on geographical conditions and soil properties of the land seeds are selected and the land is prepared for precision agriculture system. In the post-processing stage machine vision techniques are explored for disease and weed identification while intelligent techniques are used for irrigation and harvesting. In this article, author reviewed and discussed 70 articles where multiple ML algorithms are presented for performance optimization of the agricultural cycle. Figure 7 shows the classification of articles based on different applications of precision agriculture.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3048415, IEEE Access

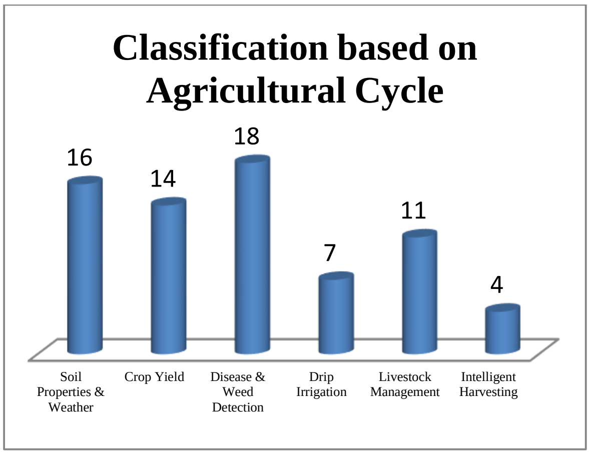


Figure 7. Classification based on agriculture cycle

Figure 8 depicts the cumulative distribution of the ML and DL models used by researchers in precision agriculture. The graph depicts the broad categorisation of the techniques with their applications to agricultural cycle. It has been observed that in majority of the literature the researchers have applied multiple algorithms for classification and parameter prediction. Regression models and ANN together make up around 65% of the AI techniques employed by researchers. Hence, it is important to investigate the techniques used and compared by the authors. The individual best performing algorithms have already been covered in appropriate sections, however the figure 8 depicts the distribution of the various regression algorithms and DL models throughout the literature. ELM algorithm is widely explored in prediction of soil properties such as soil moisture, soil temperature, surface humidity, ETc. ANN accurately predicts the rainfall and crop yield across different regions of globe. DL based CNN model finds wide applications for accurate disease and weed classification in agriculture crops. ANN model best predicts the nitrate content and water requirement in drip irrigation system. SVM regression model estimates the emitter outflow discharge under varying temperature and pressure conditions. Decision Tree algorithm accurately identify the chewing habits and predicts SCM in dairy herds. CNN have widely explored for livestock identification. Metaheuristic optimized ML algorithms are also explored by researchers in precision agriculture.

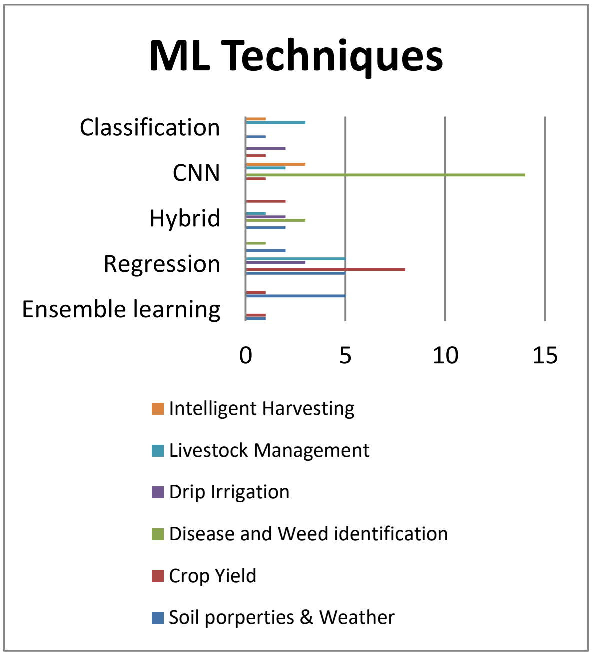


Figure 8. ML techniques used in Precision agriculture applications

In the reviewed articles, authors have used around 22 different regression algorithms for prediction, however 5 most commonly used algorithms are identified and depicted in the figure 9. Remaining 17 algorithms which are used either only for comparison or employed as a support algorithms have been classified into others.

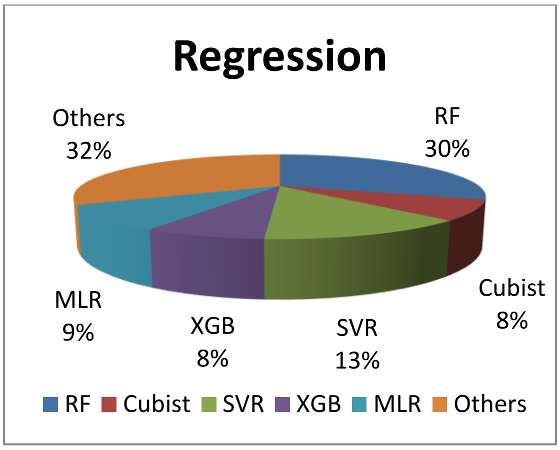


Figure 9. Regression algorithm in Precision agriculture

DL models have contributed significantly and outperforms ML classification algorithms in classification of crop disease and weed as well as for livestock diseases identification. Figure 10 shows CNN, ANN and RNN algorithms explored in precision agriculture. In the reviewed articles, authors have used around 10 different DL/NN algorithms for prediction/classification, however 8 most commonly

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3048415, IEEE Access

used algorithms are identified and depicted in the figure 10. Remaining 2 (LeNet, and Caffee) algorithms which are used either only for comparison or employed as a support algorithms have been classified into others.

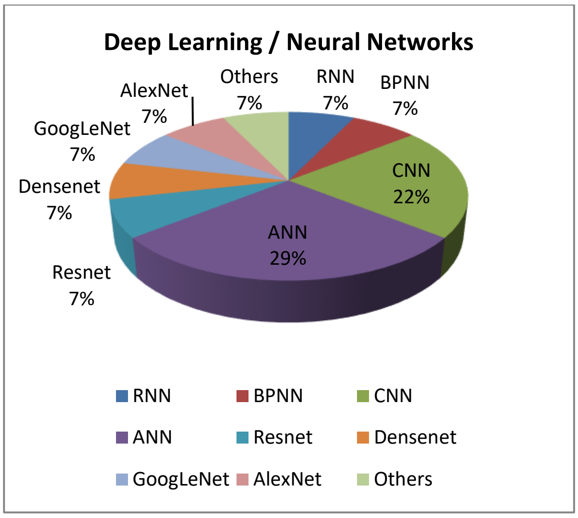


Figure 10. Classification algorithm in Precision agriculture

Performance Comparison of ML Algorithms in Precision Agriculture

The application of ML and DL algorithms highly depends on the agriculture cycle and the dataset involved. This section discusses the advantages and limitations of various ML and DL algorithms such as regression and classification algorithms based on the agriculture cycle involved.

A. Soil Properties and Weather Prediction

The application of AI techniques in prediction of soil parameters and weather is dependent on various factors. The researchers generally employ around 3 to 4 algorithms in for prediction and select the algorithms which has most accurate prediction and is robust to factors such as: noise, non-linearity, outliers etc. the most commonly employed algorithms are ELM, RF, SVR, and cubist algorithm.

Advantages of using ML in prediction of soil properties and weather pattern:

1. Non-linear dataset – these predictions often attributes a non-linear dataset which can be utilized for accurate prediction by regression algorithms such as: ELM, RF, SVR
2. Large dataset – the dataset for is often obtained from satellite which can be well handled by the regression algorithms with less convergence time and accurate predictions.
3. Insensitivity to outliers – Weather patterns often encounter outlier events which may affect the prediction accuracy, however algorithms such as ELM, NN are robust to outliers and provide accurate predictions.
4. Accurate prediction – prediction of parameters using ML exhibit low error indices such as RMSE, and R^2 which are standard measures of accuracy for statistical analysis.

Challenges and limitations in prediction of soil properties and weather pattern:

1. Varying geographical conditions poses a challenge for universal design of the prediction algorithms.
2. Soil parameters prediction is highly dependent on the sample selection philosophy.
3. Dataset selection and filtering is a challenge for researchers with non-computing background.

B. Crop Yield Prediction

The application of AI techniques in prediction of crop yield is a mammoth task and lack of availability of a universal model makes designing of the algorithm challenging. The most promising algorithms for crop yield prediction are regression algorithms, and neural networks.

Advantages of using ML in crop yield prediction:

1. Complex dataset – crop yield prediction involves enormous dataset composing of satellite data and/or historic data. Faster and accurate predictions can be made by utilizing the AI techniques such as regression algorithms (SVR, RF) Neural networks (CNN).
2. Parameter variation – the crop yield depends on a lot of parameters, like climatic factors, soil quality, NDVI, altitude, air parameters. The AI based prediction systems handle the parameters dependency efficiently.
3. Accurate prediction – prediction of parameters using ML exhibit low error indices such as RMSE, and R^2 which are standard measures of accuracy for statistical analysis.

Challenges and limitations in prediction of crop yield:

1. Varying parameters and complex datasets pose a challenge for universal design of the prediction algorithms.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3048415, IEEE Access

2. Dataset selection is critical due to the complexity; as an improper selection of data may result in underfit/overfit prediction pattern.

C. Disease and Weed Detection

The applications of AI techniques in disease and weed detection primarily depends on the advances in image processing. CNN's are the most prominent choice for building a disease identification system. Training dataset will govern the performance of the algorithm, although these are available in open-source format, users have to be cautious while using the dataset.

Advantages of using ML in detection of weed and disease in a crop field:

1. Prediction accuracy – AI offer accurate detection of disease and weeds with an accuracy of 99% which is better compared to manual/classical techniques.
2. Robust prediction – the algorithms can predict the disease/weed even with smartphone images, which is commonly available with farmers.
3. Easy configuration – with CNN being the most common and reliable technique, designing a disease/weed detection system is not a complex job unlike other systems discussed in text.

Challenges and limitations in detection of weed and disease in a crop field:

1. The accuracy of prediction depends on the quality of training dataset some of which is available as an open-source dataset, but is applicable to only a limited number of crops.
2. Improperly labelled data may result in a disastrous prediction system, as the training of the system plays a major role in the performance of the system.
3. Overtraining the model may result in a sensitive prediction system.

D. Drip Irrigation

Smart irrigation systems are not only crop friendly but are environmental friendly too. The combination of IoT with the AI not only reduces the manual intervention but also utilizes the available in an optimum way to ensure no adverse effect to environment. Regression and

Advantages of using ML in drip irrigation for an agricultural field:

1. Optimum resource utilization – accurate estimation of irrigation requirements results in a system which optimizes the resource (water, electricity) utilization (NN algorithms).
2. Crop protection – optimized irrigation practices minimizes water related damage to the crops and hence increases the crop yield.
3. Robust to weather variations – an accurately designed AI based (Regression algorithms) irrigation system handles the random weather events in a better way when compared with the non-AI based irrigation methods.

Challenges and limitations in drip irrigation for an agricultural field:

1. Accurate prediction sometimes depends on the number of sensors and hence increases the initial investment of the farmers.
2. An incorrect sensor placement in the filed affects the accuracy of the system, hence sensor optimization becomes imperative in designing a smart irrigation system.
3. The architecture of prediction system highly depends on the dataset; hence no universal guidelines can be laid out for system design.

E. Livestock Production and Management

The livestock management primarily focuses on the well-being of the farm animals and uses advanced image recognition (CNN) algorithms, and regression techniques to detect and predict the disease/ disease spread.

Advantages of using ML in livestock production and management:

1. Decreased risk of diseases – AI systems assists in identifying the livestock diseases and also helps in combating the disease, by predicting the root of diseases and transmission (Regression algorithms).
2. Minimization of disease spread – timely diagnosis and treatment reduces the risk of spreading the disease.
3. Psychological analysis – advanced image recognition and behavioural analysis (CNN techniques) help is detecting the stress in animals ensuring heath of the livestock.

Challenges and limitations in drip irrigation for an agricultural field:

1. With varying geographic and climatic conditions the attributes of the cattle and diseases changes hence, no universal system can be designed to cater to the diversities.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3048415, IEEE Access

2. Some viruses are difficult to predict even using the state-of-art prediction algorithms.

F. Intelligent Harvesting

The applications of AI techniques in harvesting is primarily an assistive technology for automatic harvesting systems. Harvesting prediction system largely relies on the advances in image processing and CNN's are the most prominent choice for building these systems.

Advantages of using ML in intelligent harvesting:

1. Assistive technology – AI in conjunction with existing harvesting robots exhibit high accuracy in harvesting.
2. Image processing – the identification of harvesting relies on the state-of-art image processing algorithm (CNNs) and hence the developments in the image processing algorithms result in direct accuracy enhancement of intelligent harvesting techniques.
3. Universal algorithms – the AI harvesting techniques largely depend on image recognition methods, hence CNNs can easily be deployed for implementing intelligent harvesting techniques.

Challenges and limitations in intelligent harvesting:

1. The accuracy of the prediction systems largely depends on the training dataset, hence accurately labelled dataset is a primary requirement of implementing an intelligent harvesting system.
2. Inaccurate harvesting recognition system result in economic loss for farmers, as a delay in harvesting might lead to an overripe crop or early harvesting might lead to rejection of the product.

VII. CHALLENGES AND LIMITATIONS OF ARTIFICAL INTELLIGENCE IN PRECISION AGRICULTURE

Artificial intelligence has the potential of playing an important role in meeting the food requirement of entire world. However, there are certain challenges which are hampering its adoption in agriculture industries which are outlined as follows:

- A recent government survey in India estimated that literacy rate of Indian farmers is very low therefore bridging the gap between farmers and technology is a challenging task.
- Farmers are less motivated to come out from their comfort zone and learn digital skills to improve their farming standards.
- Agriculture lands are mostly situated in rural areas. Implementation of IoT architecture and WSN which requires cloud services for data storage and analysis is a big issue in rural areas where reliable internet connectivity is not available.
- Accurate prediction and classification through cognitive ability of machines is difficult in varying geographical conditions.
- Initial set up of digital farming which includes hardware and software requires huge investment.
- Deployment of smart sensors and other electronic gadgets requires heavy energy consumption.

VIII. FUTURE TRENDS OF ARTIFICAL INTELLIGENCE AND IoT IN PRECISION AGRICULTURE

Agriculture industry is globally US$5 trillion industry and now it has been revolutionized with artificial intelligence and IoT technologies. These innovative tools are assisting famers to improve crop yield, monitor soil parameters, livestock health and temperature conditions, control pests and improve other agriculture related tasks. Conventional ML and DL models such as SVM, RF, ANN finds difficult to accurately estimate soil parameters and weather conditions in varying ecosystem. Therefore, swarm intelligence optimized robust and adaptive ML and DL algorithms such as SVM-PSO, ANN-GWO algorithms can be explored to effectively forecast different parameters in precision agriculture. In large agriculture fields swarm intelligence inspired autonomous system can be built for crop health and growth monitoring. UAV swarm can be utilized for near real time field and livestock monitoring through computer vision and DL algorithms and accordingly swarm of UAV can be used for spraying of pesticides and fertilizers in the infected crops. Greenness of crops can be identified through UAVs installed cameras and an automated irrigation system can be built in large agriculture fields. Swarm of mobile robots can be used in the agriculture fields to efficiently automate task such as harvesting, weed identification and elimination, etc. Metaheuristic algorithms can be explored for nodes localization in agriculture fields in order to optimize the sensor deployment in the field and keep the minimize cost to farmers. Offline service chatbots can be built to assist farmers in developing countries where farmers don't have good internet connectivity. These chatbots can assist farmers by providing timely advice based on expert recommendations and will help to resolve their specific farming problems. Artificial intelligence assisted renewable energy plants can be installed in agriculture lands to maximize the power output of clean energy in unpredictable weather conditions. This will allow for sustainable agricultural practices. Artificial intelligence can also be explored in vertical and soilless agriculture. In near future artificial intelligence systems, robotics and smart sensor technology will automate the whole farming process starting from seed sowing to intelligent fruits and vegetables harvesting and packaging.

IX. CONCLUSION

Precision agriculture is empowering the farmers with technology intending to get optimum outputs with precise inputs. IoT enabled smart sensors, actuators, satellite images, robots, drones are some of the key technological revolutions that boosted the agriculture industry. These components play a vital role in collecting real-time data and accordingly making decisions without human support. Artificial intelligence which is the automation of intelligent behaviour is continuously benefiting our planet and helping humans in various aspects of life. In this paper, authors have reviewed ML applications for precision agriculture. The impact of AI and IoT in smart farm management is discussed with a brief introduction to ML algorithms which are most commonly used in precision agriculture. Regression algorithms are the backbone for soil properties, weather, and crop yield prediction. DL algorithms such as CNN and ML classification algorithms such as SVM, Decision trees, and RF were explored for the identification of disease and weeds in the plants. Smart irrigation systems and harvesting techniques play an important component in precision agriculture as these techniques quickly complete the work and reduces human labour. Drones and robots enabled with a digital camera are employed for this work. Livestock management is an important concern for farmers across the world. Knowledge-based agriculture system which includes smart IoT devices and AI tools efficiently handle livestock management.

As a scope of future work, NLP based chatbots can be built for famers and more ML, DL and hybrid algorithms can be explored in the agriculture industry for sustainable use of available resources.

REFERENCES

1. Zhang, Naiqian, Maohua Wang, and Ning Wang. "Precision agriculture—a worldwide overview." *Computers and electronics in agriculture* 36.2-3 (2002): 113-132.
2. V. Hakkim, E. Joseph, A. Gokul and K. Mufeedha, "Precision farming: the future of Indian agriculture," *Journal of Applied Biology and Biotechnology*, pp. 68-072, 2016.
3. Gebbers, R., & Adamchuk, V. I. (2010). Precision agriculture and food security. *Science*, *327*(5967), 828-831.
4. Pierce, F. J., & Nowak, P. (1999). Aspects of precision agriculture. In *Advances in agronomy* (Vol. 67, pp. 1-85). Academic Press.
5. Stafford, J. V. (2000). Implementing precision agriculture in the 21st century. *Journal of Agricultural Engineering Research*, *76*(3), 267-275.
6. Gubbi, J., Buyya, R., Marusic, S., & Palaniswami, M. (2013). Internet of Things (IoT): A vision, architectural elements, and future directions. *Future generation computer systems*, *29*(7), 1645-1660.
7. Nex, F., & Remondino, F. (2014). UAV for 3D mapping applications: a review. *Applied geomatics*, *6*(1), 1-15.
8. Adamchuk, V. I., Hummel, J. W., Morgan, M. T., & Upadhyaya, S. K. (2004). On-the-go soil sensors for precision agriculture. *Computers and electronics in agriculture*, *44*(1), 71-91.
9. Jawad, H. M., Nordin, R., Gharghan, S. K., Jawad, A. M., & Ismail, M. (2017). Energy-efficient wireless sensor networks for precision agriculture: A review. Sensors, 17(8), 1781.
10. Sharma, R., Kamble, S. S., Gunasekaran, A., Kumar, V., & Kumar, A. (2020). A systematic literature review on machine learning applications for sustainable agriculture supply chain performance. *Computers & Operations Research*, 104926.
11. Liakos, K. G., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine learning in agriculture: A review. *Sensors*, *18*(8), 2674.
12. Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. *IBM Journal of research and development*, *3*(3), 210-229.
13. Turing, A. M. (2009). Computing machinery and intelligence. In *Parsing the Turing Test* (pp. 23-65). Springer, Dordrecht.
14. C. F. Gaitán, "Machine learning applications for agricultural impacts under extreme events," in *Climate Extremes and Their Implications for Impact and Risk Assessment*, Elsevier, 2020, pp. 119-138.
15. S. J. Parikh and B. R. James, "Soil: The Foundation of Agriculture," *Nature Education Knowledge*, vol. 3, no. 10, 2012.
16. F. R. Troeh and L. M. Thompson, Soils and Soil Fertility, Oxford University Press, 1993.

17. Adamchuk, V. I., Hummel, J. W., Morgan, M. T., & Upadhyaya, S. K. (2004). On-the-go soil sensors for precision agriculture. *Computers and electronics in agriculture*, *44*(1), 71-91.
18. E. Acar, M. S. OZERDEM and B. B. USTUNDAG, "Machine Learning based Regression Model for Prediction of Soil Surface Humidity over Moderately Vegetated Fields," in *8th International Conference on Agro-Geoinformatics (Agro-Geoinformatics)*, Istanbul, Turkey, 2019.
19. E. Tarihi, "Polarimetric Decompositions Tutorial," 2016. [Online]. Available: https:// earth.esa.int/ documents/653194/656796/ Polarimetric_Decompositions.pdf. [Accessed 13 02 2017].
20. X. Wang, W. Hu, K. Li, L. Song and L. Song, "Modelling of Soft Sensor Based on DBN-ELM and Its Application in Measurement of Nutrient Solution Composition for Soilless Culture," in *IEEE International Conference of Safety Produce Informatization (IICSPI)*, Chongqing, China, 2018.
21. S. Park, J. Im, S. Park and J. Rhee, "AMSR2 soil moisture downscaling using multisensor products through machine learning approach," in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Milan, Italy, 2015.
22. Reda, R., Saffaj, T., Ilham, B., Saidi, O., Issam, K., & Brahim, L. (2019). A comparative study between a new method and other machine learning algorithms for soil organic carbon and total nitrogen prediction using near infrared spectroscopy. *Chemometrics and Intelligent Laboratory Systems*, *195*, 103873.
23. Morellos, A., Pantazi, X. E., Moshou, D., Alexandridis, T., Whetton, R., Tziotzios, G., ... & Mouazen, A. M. (2016). Machine learning based prediction of soil total nitrogen, organic carbon and moisture content by using VIS-NIR spectroscopy. *Biosystems Engineering*, *152*, 104-116.
24. Andrade, R., Silva, S. H. G., Weindorf, D. C., Chakraborty, S., Faria, W. M., Mesquita, L. F., ... & Curi, N. (2020). Assessing models for prediction of some soil chemical properties from portable X-ray fluorescence (pXRF) spectrometry data in Brazilian Coastal Plains. *Geoderma*, *357*, 113957.
25. Deiss, L., Margenot, A. J., Culman, S. W., & Demyan, M. S. (2020). Tuning support vector machines regression models improves prediction accuracy of soil properties in MIR spectroscopy. *Geoderma*, *365*, 114227.
26. Mahmoudzadeh, H., Matinfar, H. R., Taghizadeh-Mehrjardi, R., & Kerry, R. (2020). Spatial prediction of soil organic carbon using machine learning techniques in western Iran. *Geoderma Regional*, *21*, e00260.
27. Veres, M., Lacey, G., & Taylor, G. W. (2015, June). Deep learning architectures for soil property prediction. In *2015 12th Conference on Computer and Robot Vision* (pp. 8-15). IEEE.
28. Benke, K. K., Norng, S., Robinson, N. J., Chia, K., Rees, D. B., & Hopley, J. (2020). Development of pedotransfer functions by machine learning for prediction of soil electrical conductivity and organic carbon content. *Geoderma*, *366*, 114210.
29. Khanal, S., Fulton, J., Klopfenstein, A., Douridas, N., & Shearer, S. (2018). Integration of high resolution remotely sensed data and machine learning techniques for spatial prediction of soil properties and corn yield. *Computers and electronics in agriculture*, *153*, 213-225.
30. Labrador, Y., Chang, C., & Viloria, J. (2012, December). Prediction of the calcium and magnesium content in soils through a generalized regression neural networks and genetic algorithms. In *2012 11th International Conference on Machine Learning and Applications* (Vol. 1, pp. 661-666). IEEE.
31. 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*, 61-69.
32. S. Ju-Young, K. K. Rang and H. Jong-Chul, "Seasonal forecasting of daily mean air temperatures using a coupled global climate model and machine learning algorithm for field-scale agricultural management," *Agricultural and Forest Meteorology*, vol. 281, 2020.
33. C. MacLachlan, A. Arribas, K. Peterson, A. Maidens, D. Fereday, A. Scaife, M. Gordon, M. Vellinga, A. Williams, R. E. Comer, J. Camp, P. Xavier and G. Madec, "Global Seasonal forecast system version 5 (GloSea5): a high-resolution seasonal forecast system," *Quarterly Journal of the Royal Meteorological Society*, vol. 141, 2015.
34. Stamenkovic, J., Tuia, D., De Morsier, F., Borgeaud, M., & Thiran, J. P. (2013, June). Estimation of soil moisture from airborne hyperspectral imagery with support vector regression. In *2013 5th workshop on hyperspectral image and signal processing: evolution in remote sensing (WHISPERS)* (pp. 1-4). IEEE.
35. Song, X., Zhang, G., Liu, F., Li, D., Zhao, Y., & Yang, J. (2016). Modeling spatio-temporal distribution of soil moisture by deep learning-based cellular automata model. *Journal of Arid Land*, *8*(5), 734-748.
36. Achieng, K. O. (2019). Modelling of soil moisture retention curve using machine learning techniques: Artificial and deep neural networks vs support vector regression models. *Computers & Geosciences*, *133*, 104320.
37. Feng, Y., Cui, N., Hao, W., Gao, L., & Gong, D. (2019). Estimation of soil temperature from meteorological data using different machine learning models. *Geoderma*, *338*, 67-77.
38. Mohammadi, K., Shamshirband, S., Motamedi, S., Petković, D., Hashim, R., & Gocic, M. (2015). Extreme learning machine based prediction of daily dew point temperature. *Computers and Electronics in Agriculture*, *117*, 214-225.
39. Zhu, B., Feng, Y., Gong, D., Jiang, S., Zhao, L., & Cui, N. (2020). Hybrid particle swarm optimization with extreme learning machine for daily reference evapotranspiration prediction from limited climatic data. *Computers and Electronics in Agriculture*, *173*, 105430.
40. Alizamir, M., Kisi, O., Ahmed, A. N., Mert, C., Fai, C. M., Kim, S., Kim, N. & El-Shafie, A. (2020). Advanced machine learning model for better prediction accuracy of soil temperature at different depths. Plos one, 15(4), e0231055.
41. Cramer, S., Kampouridis, M., Freitas, A. A., & Alexandridis, A. K. (2017). An extensive evaluation of seven machine learning methods for rainfall prediction in weather derivatives. *Expert Systems with Applications*, *85*, 169-181.
42. Diez-Sierra, J., & del Jesus, M. (2020). Long-term rainfall prediction using atmospheric synoptic patterns in semi-arid climates with statistical and machine learning methods. *Journal of Hydrology*, 124789.
43. Kamatchi, S. B., & Parvathi, R. (2019). Improvement of Crop Production Using Recommender System by Weather Forecasts. *Procedia Computer Science*, *165*, 724-732.
44. Lazri, M., Labadi, K., Brucker, J. M., & Ameur, S. (2020). Improving satellite rainfall estimation from MSG data in Northern Algeria by using a multi-classifier model based on machine learning. *Journal of Hydrology*, *584*, 124705.

45. N.K. Shardoor and M. V. Rao, Analysis of rainfall prediction using machine learning data mining and satellite techniques, International Journal of Engineering and Technology, Vol. 7, Issue 4, 4362-4367, 2018.
46. S. Bhanumati, M. Vineeth and N. Rohit, "Crop Yield Prediction and Efficient use of Fertilizers," in *International Conference on Communication and Signal Processing*, Chennai, 2019.
47. S. Singh, R. Saxena, A. Porwal, N. and S. Ray, "Assessment of hailstorm damage in wheat crop using remote sensing," *Current Science,* vol. 112, no. 10, pp. 2095-2099, 2015.
48. E. Kamir, F. Waldner and Z. Hochman, "Estimating wheat yields in Australia using climate records, satellite image time series and machine learning methods," *ISPRS Journal of Photogrammetry and Remote Sensing,* vol. 160, pp. 124-135, 2020.
49. K. Didan, "MOD13Q1 MODIS/Terra Vegetation Indices 16-Day L3 Global 250m SIN Grid V006," NASA EOSDIS Land Processes DAAC, 2015.
50. K. Tumer and J. Ghish, "Error Correlation and Error Reduction in Ensemble Classifiers," *Connection Science,* vol. 8, no. 3-4, pp. 385-404, 1996.
51. H. Aghighi, M. Azadbakht, D. Ashourloo, H. S. Shahrabi and S. Radiom, "Machine Learning Regression Techniques for the Silage Maize Yield Prediction Using Time-Series Images of Landsat 8 OLI," *IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING,* vol. 11, no. 12, pp. 4563 - 4577, 2018.
52. K. Kuwata and R. Shibasaki, "Estimating crop yields with deep learning and remotely sensed data," in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Milan, Italy, 2015.
53. S. Kulkarni, S. N. Mandal, G. S. Sharma, M. R. Mundada and Meeradevi, "Predictive Analysis to Improve Crop Yield using a Neural Network Model," in *International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, Bangalore, 2018.
54. Chu, Z., & Yu, J. (2020). An end-to-end model for rice yield prediction using deep learning fusion. *Computers and Electronics in Agriculture*, *174*, 105471.
55. Feng, P., Wang, B., Li Liu, D., Waters, C., Xiao, D., Shi, L., & Yu, Q. (2020). Dynamic wheat yield forecasts are improved by a hybrid approach using a biophysical model and machine learning technique. *Agricultural and Forest Meteorology*, *285*, 107922.
56. Cai, Y., Guan, K., Lobell, D., Potgieter, A. B., Wang, S., Peng, J., ... & Peng, B. (2019). Integrating satellite and climate data to predict wheat yield in Australia using machine learning approaches. *Agricultural and forest meteorology*, *274*, 144-159.
57. Gümüşçü, A., Tenekeci, M. E., & Bilgili, A. V. (2019). Estimation of wheat planting date using machine learning algorithms based on available climate data. *Sustainable Computing: Informatics and Systems*.
58. Nevavuori, P., Narra, N., & Lipping, T. (2019). Crop yield prediction with deep convolutional neural networks. *Computers and electronics in agriculture*, *163*, 104859.
59. Koirala, A., Walsh, K. B., Wang, Z., & McCarthy, C. (2019). Deep learning–Method overview and review of use for fruit detection and yield estimation. *Computers and Electronics in Agriculture*, *162*, 219-234.
60. Kouadio, L., Deo, R. C., Byrareddy, V., Adamowski, J. F., & Mushtaq, S. (2018). Artificial intelligence approach for the prediction of Robusta coffee yield using soil fertility properties. *Computers and Electronics in Agriculture*, *155*, 324-338.
61. Gamboa, A. A., Cáceres, P. A., Lamos, H., Zárate, D. A., & Puentes, D. E. (2019, April). Predictive model for cocoa yield in Santander using Supervised Machine Learning. In *2019 XXII Symposium on Image, Signal Processing and Artificial Vision (STSIVA)* (pp. 1-5). IEEE.
62. Gyamerah, S. A., Ngare, P., & Ikpe, D. (2020). Probabilistic forecasting of crop yields via quantile random forest and Epanechnikov Kernel function. *Agricultural and Forest Meteorology*, *280*, 107808.
63. Peng, B., Guan, K., Zhou, W., Jiang, C., Frankenberg, C., Sun, Y., ... & Köhler, P. (2020). Assessing the benefit of satellite-based Solar-Induced Chlorophyll Fluorescence in crop yield prediction. *International Journal of Applied Earth Observation and Geoinformation*, *90*, 102126.
64. Khaki, S., & Wang, L. (2019). Crop yield prediction using deep neural networks. *Frontiers in plant science*, *10*, 621.
65. Leroux, L., Castets, M., Baron, C., Escorihuela, M. J., Bégué, A., & Seen, D. L. (2019). Maize yield estimation in West Africa from crop process-induced combinations of multi-domain remote sensing indices. *European Journal of Agronomy*, *108*, 11-26.
66. Li, Y., Guan, K., Yu, A., Peng, B., Zhao, L., Li, B., & Peng, J. (2019). Toward building a transparent statistical model for improving crop yield prediction: Modeling rainfed corn in the US. *Field Crops Research*, *234*, 55-65.
67. Maimaitijiang, M., Sagan, V., Sidike, P., Hartling, S., Esposito, F., & Fritschi, F. B. (2020). Soybean yield prediction from UAV using multimodal data fusion and deep learning. *Remote Sensing of Environment*, *237*, 111599.
68. Zhang, C., Di, L., Lin, L., & Guo, L. (2019). Machine-learned prediction of annual crop planting in the US Corn Belt based on historical crop planting maps. *Computers and Electronics in Agriculture*, *166*, 104989.
69. Kocian, A., Massa, D., Cannazzaro, S., Incrocci, L., Di Lonardo, S., Milazzo, P., & Chessa, S. (2020). Dynamic Bayesian network for crop growth prediction in greenhouses. *Computers and Electronics in Agriculture*, *169*, 105167.
70. Shahinfar, S., & Kahn, L. (2018). Machine learning approaches for early prediction of adult wool growth and quality in Australian Merino sheep. *Computers and Electronics in Agriculture*, *148*, 72-81.
71. G. Sambasivam and G. D. Opiyo, "A predictive machine learning application in agriculture: Cassava disease detection and classification with imbalanced dataset using convolutional neural networks," *Egyptian Informatics Journal*, 2020.
72. A. Ramcharan, K. Baranowski, P. McCloskey, B. Ahmed, J. Legg and D. P. Hughes, "Deep Learning for Image-Based Cassava Disease Detection," *Frontiers in Plant Science,* vol. 8, 2017.
73. S. P. Mohanty, D. P. Hughes and M. Salathé, "Using Deep Learning for Image-Based Plant Disease Detection," *Frontiers in Plant Science,* vol. 7, pp. 1-10, 2016.
74. A. Krizhevsky, I. Sutskever and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in *Advances in Neural Information Processing Systems 25 (NIPS 2012)*, 2012.
75. C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke and A. Rabinovich, "Going deeper with convolutions," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA, 2015.
76. J. Amara, B. Bouaziz and A. Algergawy, "A Deep Learning-based Approach for Banana Leaf Diseases Classification," in *Lecture Notes in Informatics*, 2017.

77. D. P. Hughes and M. Salathe, "An open access repository of images on plant health to enable the development of mobile disease diagnostic," Computers and Society: Cornell university, 2016.
78. A. d. S. Ferreiraa, D. M. Freitasa, G. G. d. Silvaa, H. Pistori and M. T. Folhes, "Weed detection in soybean crops using ConvNets," *Computers and Electronics in Agricultre*, vol. 143, pp. 314-324, 2017.
79. Waheed, A., Goyal, M., Gupta, D., Khanna, A., Hassanien, A. E., & Pandey, H. M. (2020). An optimized dense convolutional neural network model for disease recognition and classification in corn leaf. *Computers and Electronics in Agriculture*, *175*, 105456.
80. Pereira, L. A., Nakamura, R. Y., De Souza, G. F., Martins, D., & Papa, J. P. (2012). Aquatic weed automatic classification using machine learning techniques. *Computers and electronics in agriculture*, *87*, 56-63.
81. Jiang, H., Zhang, C., Qiao, Y., Zhang, Z., Zhang, W., & Song, C. (2020). CNN feature based graph convolutional network for weed and crop recognition in smart farming. *Computers and Electronics in Agriculture*, *174*, 105450.
82. Oppenheim, D., & Shani, G. (2017). Potato disease classification using convolution neural networks. *Advances in Animal Biosciences*, *8*(2), 244.
83. Rumpf, T., Mahlein, A. K., Steiner, U., Oerke, E. C., Dehne, H. W., & Plümer, L. (2010). Early detection and classification of plant diseases with support vector machines based on hyperspectral reflectance. *Computers and electronics in agriculture*, *74*(1), 91-99.
84. Ozguven, M. M., & Adem, K. (2019). Automatic detection and classification of leaf spot disease in sugar beet using deep learning algorithms. *Physica A: Statistical Mechanics and its Applications*, *535*, 122537.
85. Bah, M. D., Hafiane, A., & Canals, R. (2018). Deep learning with unsupervised data labeling for weed detection in line crops in UAV images. *Remote sensing*, *10*(11), 1690.
86. Kerkech, M., Hafiane, A., & Canals, R. (2020). Vine disease detection in UAV multispectral images using optimized image registration and deep learning segmentation approach. *Computers and Electronics in Agriculture*, *174*, 105446.
87. Olsen, A., Konovalov, D. A., Philippa, B., Ridd, P., Wood, J. C., Johns, J., ... & Calvert, B. (2019). DeepWeeds: A multiclass weed species image dataset for deep learning. *Scientific reports*, *9*(1), 1-12.
88. Sudars, K., Jasko, J., Namatevs, I., Ozola, L., & Badaukis, N. (2020). Dataset of annotated food crops and weed images for robotic computer vision control. *Data in Brief*, 105833.
89. Sethy, P. K., Barpanda, N. K., Rath, A. K., & Behera, S. K. (2020). Deep feature based rice leaf disease identification using support vector machine. *Computers and Electronics in Agriculture*, *175*, 105527.
90. Espejo-Garcia, B., Mylonas, N., Athanasakos, L., Fountas, S., & Vasilakoglou, I. (2020). Towards weeds identification assistance through transfer learning. *Computers and Electronics in Agriculture*, *171*, 105306.
91. Shah, N., & Jain, S. (2019, February). Detection of Disease in Cotton Leaf using Artificial Neural Network. In *2019 Amity International Conference on Artificial Intelligence (AICAI)* (pp. 473-476). IEEE.
92. Yu, J., Schumann, A. W., Cao, Z., Sharpe, S. M., & Boyd, N. S. (2019). Weed Detection in Perennial Ryegrass With Deep Learning Convolutional Neural Network. *Frontiers in plant science*, *10*.
93. Parraga-Alava, J., Cusme, K., Loor, A., & Santander, E. (2019). RoCoLe: A robusta coffee leaf images dataset for evaluation of machine learning based methods in plant diseases recognition. *Data in brief*, *25*, 104414.
94. Glezakos, T. J., Moschopoulou, G., Tsiligiridis, T. A., Kintzios, S., & Yialouris, C. P. (2010). Plant virus identification based on neural networks with evolutionary preprocessing. *Computers and electronics in agriculture*, *70*(2), 263-275.
95. Ramesh, S., & Vydeki, D. (2019). Recognition and classification of paddy leaf diseases using Optimized Deep Neural network with Jaya algorithm. *Information Processing in Agriculture*.
96. Chechliński, Ł., Siemiątkowska, B., & Majewski, M. (2019). A System for Weeds and Crops Identification—Reaching over 10 FPS on Raspberry Pi with the Usage of MobileNets, DenseNet and Custom Modifications. *Sensors*, *19*(17), 3787.
97. Wang, A., Zhang, W., & Wei, X. (2019). A review on weed detection using ground-based machine vision and image processing techniques. *Computers and electronics in agriculture*, *158*, 226-240.
98. Kim, Y. H., Yoo, S. J., Gu, Y., Lim, J., Han, D., & Baik, S. (2014). Crop pests prediction method using regression and machine learning technology: Survey. *IERI Procedia*, *6*, 52-56.
99. Saleem, M. H., Potgieter, J., & Arif, K. M. (2019). Plant disease detection and classification by deep learning. *Plants*, *8*(11), 468.
100. Liu, B., & Bruch, R. (2020). Weed Detection for Selective Spraying: a Review. *Current Robotics Reports*, *1*(1), 19-26.
101. Goldberg, D., Gornat, B., & Rimon, D. (1976). Drip irrigation: principles, design and agricultural practices. *Drip irrigation: principles, design and agricultural practices.*
102. G. Hochmuth, "Drip irrigation", in A Guide to the Manufacture Performance, and Potential of Plastics in Agriculture, M. D. Orzolek, Elsevier, 2017, pp-1-197.
103. Olutobi Adeyemi, Ivan Grove and Sven Peets, "Dynamic Neural Network Modelling of Soil Moisture Content for Predictive Irrigation Scheduling", in Sensors, 2018.
104. D. Tseng, D. Wang and C. Chen, "Towards Automating Precision Irrigation: Deep Learning to Infer Local Soil Moisture Conditions from Synthetic Aerial Agricultural Images", in 14th International Conference on Automation Science and Engineering (CASE), IEEE, 2018.
105. R. Berenstein, R. Fox, S. McKinley, S. Carpin, and K. Goldberg, "Robustly adjusting indoor drip irrigation emitters with the toyotahsr robot", in ICRA, IEEE, 2018.
106. Evans, R. G., & Sadler, E. J. "Methods and technologies to improve efficiency of water use". Water Resources Research, 44(7): W00E04, 2008.
107. Janani M and Jebakumar R, " A Study on Smart Irrigation Using Machine Learning ", Cell & Cellular Life Sciences Journal, Vol. 4, Issue 2, 2019.
108. Jha, K., Doshi, A., & Patel, P. (2018). Intelligent irrigation system using artificial intelligence and machine learning: a comprehensive review ISSN: 2320-5407. *Int. J. Adv. Res.(Indore)*, *6*(10), 1493-1502.
109. Vij, A., Vijendra, S., Jain, A., Bajaj, S., Bassi, A., & Sharma, A. (2020). IoT and Machine Learning Approaches for Automation of Farm Irrigation System. *Procedia Computer Science*, *167*, 1250-1257.
110. Kondaveti, R., Reddy, A., & Palabtla, S. (2019, March). Smart Irrigation System Using Machine Learning and IOT. In *2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN)* (pp. 1-11). IEEE.
111. Roberts, M. J., et al. (2017). Comparing and combining process-based crop models and statistical models with some implications for climate change. Environmental Research Letters, 12(9): 095010.
112. Karandish, F., &Šimůnek, J. "A comparison of numerical and machine-learning modeling of soil water content with limited input data". Journal of Hydrology, 543: 892-909, 2016.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3048415, IEEE Access

113. Sahoo, S., et al. (2017). Machine learning algorithms for modeling groundwater level changes in agricultural regions of the US. Water Resources Research, 53(5): 3878-3895.
114. Li, J., Yoder, R. E., Odhiambo, L. O., & Zhang, J. (2004). Simulation of nitrate distribution under drip irrigation using artificial neural networks. *Irrigation Science*, *23*(1), 29-37.
115. G. Kavianand, V. M. Nivas, R. Kiruthika and S. Lalitha, "Smart drip irrigation system for sustainable agriculture," *2016 IEEE Technological Innovations in ICT for Agriculture and Rural Development (TIAR)*, Chennai, 2016, pp. 19-22, doi: 10.1109/TIAR.2016.7801206.
116. Emmanuel, A. A., Abidin, M. S. Z., Mahmud, M. S. A., Buyamin, S., AbdRahman, M. K. I., Otuoze, A. O., ... & Ijike, O. D. (2020). IoT Monitoring and Data-Driven Modelling of Drip Irrigation System for Mustard Leaf Cultivation Experiment. *Information Processing in Agriculture*.
117. Seyedzadeh, A., Maroufpoor, S., Maroufpoor, E., Shiri, J., Bozorg-Haddad, O., & Gavazi, F. (2020). Artificial intelligence approach to estimate discharge of drip tape irrigation based on temperature and pressure. *Agricultural Water Management*, *228*, 105905.
118. Peng, Y., Xiao, Y., Fu, Z., Dong, Y., Zheng, Y., Yan, H., & Li, X. (2019). Precision irrigation perspectives on the sustainable water-saving of field crop production in China: Water demand prediction and irrigation scheme optimization. *Journal of Cleaner Production*, *230*, 365-377.
119. Shiri, J., Karimi, B., Karimi, N., Kazemi, M. H., & Karimi, S. (2020). Simulating wetting front dimensions of drip irrigation systems: Multi criteria assessment of soft computing models. *Journal of Hydrology*, 124792.
120. Elnesr, M. N., & Alazba, A. A. (2017). Simulation of water distribution under surface dripper using artificial neural networks. *Computers and Electronics in Agriculture*, *143*, 90-99.
121. Y. Chang, T. Huang and N. Huang, "A Machine Learning Based Smart Irrigation System with LoRa P2P Networks," *2019 20th Asia-Pacific Network Operations and Management Symposium (APNOMS)*, Matsue, Japan, 2019, pp. 1-4, doi: 10.23919/APNOMS.2019.8893034.
122. G. Singh, D. Sharma, A. Goap, S. Sehgal, A. K. Shukla and S. Kumar, "Machine Learning based soil moisture prediction for Internet of Things based Smart Irrigation System," *2019 5th International Conference on Signal Processing, Computing and Control (ISPCC)*, Solan, India, 2019, pp. 175-180, doi: 10.1109/ISPCC48220.2019.8988313.
123. Torres-Sanchez, R., Navarro-Hellin, H., Guillamon-Frutos, A., San-Segundo, R., Ruiz-Abellón, M. C., & Domingo-Miguel, R. (2020). A Decision Support System for Irrigation Management: Analysis and Implementation of Different Learning Techniques. *Water*, *12*(2), 548.
124. Navarro-Hellín, H., Martínez-del-Rincon, J., Domingo-Miguel, R., Soto-Valles, F., & Torres-Sánchez, R. (2016). A decision support system for managing irrigation in agriculture. *Computers and Electronics in Agriculture*, *124*, 121-131.
125. Goumopoulos, C., O'Flynn, B., & Kameas, A. (2014). Automated zone-specific irrigation with wireless sensor/actuator network and adaptable decision support. *Computers and electronics in agriculture*, *105*, 20-33.
126. Chen, Z., Sun, S., Wang, Y., Wang, Q., & Zhang, X. (2020). Temporal convolution-network-based models for modeling maize evapotranspiration under mulched drip irrigation. *Computers and Electronics in Agriculture*, *169*, 105206.
127. Alonso, R. S., Sittón-Candanedo, I., García, Ó., Prieto, J., & Rodríguez-González, S. (2020). An intelligent Edge-IoT platform for monitoring livestock and crops in a dairy farming scenario. *Ad Hoc Networks*, *98*, 102047.
128. Berckmans, D., 2017. General introduction to precision livestock farming. Animal Front. 7(1), 6–11. https://academic.oup.com/af/article/7/1/6/4638786.
129. Halachmi, I., Guarino, M., 2016. Editorial: Precision livestock farming: a 'per animal' approach using advanced monitoring technologies. Animal 10, 1482–1483. https:// doi.org/10.1017/S1751731116001142.
130. Villeneuve, E., Abi Akle, A., Merlo, C., Masson, D., Terrasson, G., & Llaria, A. (2019). Decision support in precision sheep farming. *IFAC-PapersOnLine*, *51*(34), 236-241.
131. Dutta, R.; Smith, D.; Rawnsley, R.; Bishop-Hurley, G.; Hills, J.; Timms, G.; Henry, D. Dynamic cattle behavioural classification using supervised ensemble classifiers. Comput. Electron. Agric. 2015, 111, 18–28.
132. Pegorini, V.; Karam, L.Z.; Pitta, C.S.R.; Cardoso, R.; da Silva, J.C.C.; Kalinowski, H.J.; Ribeiro, R.; Bertotti, F.L.; Assmann, T.S. In vivo pattern classification of ingestive behavior in ruminants using FBG sensors and machine learning. Sensors 2015, 15, 28456–28471.
133. Ebrahimie, E., Ebrahimi, F., Ebrahimi, M., Tomlinson, S., & Petrovski, K. R. (2018). Hierarchical pattern recognition in milking parameters predicts mastitis prevalence. *Computers and electronics in agriculture*, *147*, 6-11.
134. Ebrahimie, E., Ebrahimi, F., Ebrahimi, M., Tomlinson, S., & Petrovski, K. R. (2018). A large-scale study of indicators of sub-clinical mastitis in dairy cattle by attribute weighting analysis of milk composition features: highlighting the predictive power of lactose and electrical conductivity.
135. Hyde, R. M., Down, P. M., Bradley, A. J., Breen, J. E., Hudson, C., Leach, K. A., & Green, M. J. (2020). Automated prediction of mastitis infection patterns in dairy herds using machine learning. *Scientific reports*, *10*(1), 1-8.
136. Esener, N., Green, M. J., Emes, R. D., Jowett, B., Davies, P. L., Bradley, A. J., & Dottorini, T. (2018). Discrimination of contagious and environmental strains of Streptococcus uberis in dairy herds by means of mass spectrometry and machine-learning. *Scientific Reports*, *8*(1), 1-12.
137. Ebrahimi, M., Mohammadi-Dehcheshmeh, M., Ebrahimie, E., & Petrovski, K. R. (2019). Comprehensive analysis of machine learning models for prediction of sub-clinical mastitis: Deep Learning and Gradient-Boosted Trees outperform other models. *Computers in biology and medicine*, *114*, 103456.
138. Sharifi, S., Pakdel, A., Ebrahimi, M., Reecy, J. M., Fazeli Farsani, S., & Ebrahimie, E. (2018). Integration of machine learning and meta-analysis identifies the transcriptomic bio-signature of mastitis disease in cattle. *PLoS One*, *13*(2), e0191227.
139. Machado, G., Mendoza, M. R., & Corbellini, L. G. (2015). What variables are important in predicting bovine viral diarrhea virus? A random forest approach. *Veterinary research*, *46*(1), 85.
140. Matthews, S.G.; Miller, A.L.; PlÖtz, T.; Kyriazakis, I. Automated tracking to measure behavioural changes inpigs for health and welfare monitoring. Sci. Rep. 2017, 7, 17582.
141. Qiao, Y., Truman, M., & Sukkarieh, S. (2019). Cattle segmentation and contour extraction based on Mask R-CNN for precision livestock farming. *Computers and Electronics in Agriculture*, *165*, 104958.
142. Liakos, K., Moustakidis, S. P., Tsiotra, G., Bartzanas, T., Bochtis, D., & Parisses, C. (2017). Machine Learning Based Computational Analysis Method for Cattle Lameness Prediction. In *HAICTA* (pp. 128-139).
143. Morales, I.R.; Cebrián, D.R.; Fernandez-Blanco, E.; Sierra, A.P. Early warning in egg production curves from commercial hens: A SVM approach. Comput. Electron. Agric. 2016, 121, 169–179.

144. Hansen, M.F.; Smith, M.L.; Smith, L.N.; Salter, M.G.; Baxter, E.M.; Farish, M.; Grieve, B. Towards on-farm pigface recognition using convolutional neural networks. Comput. Ind. 2018, 98, 145–152.

145. Fenlon, C., O'Grady, L., Mee, J. F., Butler, S. T., Doherty, M. L., & Dunnion, J. (2017). A comparison of 4 predictive models of calving assistance and difficulty in dairy heifers and cows. *Journal of dairy science*, *100*(12), 9746-9758.

146. Fenlon, C., O'Grady, L., Dunnion, J., Shalloo, L., Butler, S. T., & Doherty, M. L. (2016). A comparison of machine learning techniques for predicting insemination outcome in Irish dairy cows.

147. Borchers, M. R., Chang, Y. M., Proudfoot, K. L., Wadsworth, B. A., Stone, A. E., & Bewley, J. M. (2017). Machine-learning-based calving prediction from activity, lying, and ruminating behaviors in dairy cattle. *Journal of dairy science*, *100*(7), 5664-5674.

148. Bahlo, C., Dahlhaus, P., Thompson, H., & Trotter, M. (2019). The role of interoperable data standards in precision livestock farming in extensive livestock systems: A review. *Computers and Electronics in Agriculture*, *156*, 459-466.

149. Boursianis, A. D., Papadopoulou, M. S., Diamantoulakis, P., Liopa-Tsakalidi, A., Barouchas, P., Salahas, G., ... & Goudos, S. K. (2020). Internet of Things (IoT) and Agricultural Unmanned Aerial Vehicles (UAVs) in Smart Farming: A Comprehensive Review. Internet of Things, 100187

150. Barhate, D., & Nemade, V. (2019, June). Comprehensive Study on Automated Image Detection by using Robotics for Agriculture Applications. In 2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA) (pp. 637-641). IEEE.

151. Jha, K., Doshi, A., Patel, P., & Shah, M. (2019). A comprehensive review on automation in agriculture using artificial intelligence. Artificial Intelligence in Agriculture, 2, 1-12.

152. Tian, H., Wang, T., Liu, Y., Qiao, X., & Li, Y. (2020). Computer vision technology in agricultural automation—A review. *Information Processing in Agriculture*, *7*(1), 1-19.

153. Zhai, Z., Martínez, J. F., Beltran, V., & Martínez, N. L. (2020). Decision support systems for agriculture 4.0: Survey and challenges. *Computers and Electronics in Agriculture*, *170*, 105256.

154. Kondo, Naoshi, Mitsuji Monta, and Noboru Noguchi. "Agri-robot (I)-fundamentals and theory." Chinese translation from Japanese language) Corona Publishing Co., Ltd (2004): 56-58.

155. Sakai, H., Shiigi, T., Kondo, N., Ogawa, Y., & Taguchi, N. (2013). Accurate position detecting during asparagus spear harvesting using a laser sensor. Engineering in agriculture, environment and food, 6(3), 105-110.

156. Monta, M., & Namba, K. (2003, July). Three-dimensional sensing system for agricultural robots. In *Proceedings 2003 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM 2003)* (Vol. 2, pp. 1216-1221). IEEE.

157. De Preter, A., Anthonis, J., & De Baerdemaeker, J. (2018). Development of a robot for harvesting strawberries. IFAC-PapersOnLine, 51(17), 14-19.

158. Hayashi, S., Shigematsu, K., Yamamoto, S., Kobayashi, K., Kohno, Y., Kamata, J., & Kurita, M. (2010). Evaluation of a strawberry-harvesting robot in a field test. *Biosystems engineering*, *105*(2), 160-171.

159. Horng, G. J., Liu, M. X., & Chen, C. C. (2019). The Smart Image Recognition Mechanism for Crop Harvesting System in Intelligent Agriculture. *IEEE Sensors Journal*, *20*(5), 2766-2781.

160. Zhang, J., Karkee, M., Zhang, Q., Zhang, X., Yaqoob, M., Fu, L., & Wang, S. (2020). Multi-class object detection using faster R-CNN and estimation of shaking locations for automated shake-and-catch apple harvesting. *Computers and Electronics in Agriculture*, 105384.

161. Kane, K. E., & Lee, W. S. (2006). Spectral sensing of different citrus varieties for precision agriculture. In *2006 ASAE Annual Meeting* (p. 1). American Society of Agricultural and Biological Engineers.

162. Silwal, A., Karkee, M., & Zhang, Q. (2016). A hierarchical approach to apple identification for robotic harvesting. *Transactions of the ASABE*, *59*(5), 1079-1086.

163. Zhang, X., He, L., Zhang, J., Whiting, M. D., Karkee, M., & Zhang, Q. (2020). Determination of key canopy parameters for mass mechanical apple harvesting using supervised machine learning and principal component analysis (PCA). *Biosystems Engineering*, *193*, 247-263.

164. Zhang, Z., Igathinathane, C., Li, J., Cen, H., Lu, Y., & Flores, P. (2020). Technology progress in mechanical harvest of fresh market apples. *Computers and Electronics in Agriculture*, *175*, 105606.

165. D. Pise and G. D. Upadhye, "Grading of Harvested Mangoes Quality and Maturity Based on Machine Learning Techniques," *2018 International Conference on Smart City and Emerging Technology (ICSCET)*, Mumbai, 2018, pp. 1-6, doi: 10.1109/ICSCET.2018.8537342.

166. Wu, W., Ding, S., & Lv, J. (2012, July). Application research of neural networks in fruit and vegetable harvesting robot. In *Proceedings of the 10th World Congress on Intelligent Control and Automation* (pp. 1790-1795). IEEE.

167. Dlodlo, N.; Kalezhi, J. The internet of things in agriculture for sustainable rural development. In Proceedingsof the International Conference on Emerging Trends in Networks and Computer Communications (ETNCC), Windhoek, Namibia, 17–20 May 2015; pp. 13–18.

168. Ojha, T.; Misra, S.; Raghuwanshi, N.S. Wireless sensor networks for agriculture: The state-of-the-art in practice and future challenges. Comput. Electron. Agric. 2015, 118, 66–84. [CrossRef].

169. Analog Devices Inc., “Digital Accelerometer,” ADXL345 datasheet, 2009.

170. Measurement Specialties, “HTU21D(F) sensor digital relative

171. Humidity sensor with temperature output,” HTU21D datasheet, 2013.

172. Sensirion, “Humidity and Temperature Sensor IC,” SHT10, SHT11, SHT15 datasheet, 2011.

173. Camera Module, “Raspberry Pi,” https://www.raspberrypi.org/products/camera-module.

174. MPR 121, “Proximity capacitive touch sensor controller,” MPR 121: Proximity Capacitive Touch Sensor Controller, Freescale Semiconductor Inc, http://www.nxp.com/pages/proximity-capacitive-touch-sensor-controller:MPR121.

175. Microsoft, “Meet kinect for windows,” https://developer.microsoft.com/en-us/windows/kinect. Tekscan, “FlexiForce Standard Model A201,” A201 datasheet.

176. Apple, “iBeacon for Developers,” https://developer.apple.com/ibeacon.

177. Sparkfun, “CdS photoconductive cells,” GL5528 datasheet.

178. Unique-Bargains, “HC-SR501 PIR motion detector,” HC-SR501 datasheet, 2016.

179. TAOS, “Color light-to-digital converter with IR filter,” TCS3472 datasheet, 2012.

180. DFRobot, “Moisture Sensor (SKU:SEN0114),” https://www.dfrobot.com/wiki/index.php/Moisture_Sensor_(SKU:SEN0114).

181. Sparkfun, "SparkFun Sound Detector," https://www.sparkfun.com/products/12642.
182. Elec Freaks, "Ultrasonic ranging module HC - SR04," HC-SR04 datasheet, 2016.
183. Pulse and Oxygen in Blood Sensor (SPO2) for e-Health Platform, "Cooking hacks," https://www.cooking-hacks.com/pulse-and-oxygen-in-blood-sensor-spo2-ehealth-medical.
184. Body Temperature Sensor for e-Health Platform, "Cooking Hacks," https://www.cooking-hacks.com/body-temperature-sensor-ehealth-medical.
185. Piezo Systems Inc., "Standard quick-mount bending generator," Piezoelectric sensor datasheet, 2011
186. Durai Raj Vincent, N Deepa, Dhivya Elavarasan, Kathiravan Srinivasan, Sajjad Hussain Chauhdary, Celestine Iwendi, Sensors Driven AI-Based Agriculture Recommendation Model for Assessing Land Suitability. Sensors Vol 19, 2019.
187. Yemeserach Mekonnen, Srikanth Namuduri, Lamar Burton, Arif Sarwat, Shekhar Bhansali, Review—Machine Learning Techniques in Wireless Sensor Network Based Precision Agriculture. Journal of the Electrochemical Society, Vol. 167, 2020
188. Dmitiri Shardin, Alexander Menschikov, Andrey Somov, Gerhild Bornemann, Jens Hauslage, Maxin Fedorov, Enabling Precision Agriculture Through Embedded Sensing With Artificial Intelligence. IEEE Transactions on Instrumentation and Measurement Vol 69, No.7, 2020
189. Estrada-López, J. J., Castillo-Atoche, A. A., Vázquez-Castillo, J., & Sánchez-Sinencio, E. (2018). Smart soil parameters estimation system using an autonomous wireless sensor network with dynamic power management strategy. IEEE Sensors Journal, 18(21), 8913-8923.
190. Ferentinos, K. P., & Tsiligiridis, T. A. (2007). Adaptive design optimization of wireless sensor networks using genetic algorithms. Computer Networks, 51(4), 1031-1051.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3048415, IEEE Access

Author Biography

Abhinav Sharma received his B. Tech from H. N. B. Garhwal University, Srinagar, India, in 2009 and the M. Tech from Govind Ballabh Pant University of Agriculture and Technology, Pantnagar, India, in 2011. He did his Ph. D. from Govind Ballabh Pant University of Agriculture and Technology, Pantnagar, India, in 2016. He is working as an Assistant Professor (Senior Scale) in the Department of Electrical and Electronics Engineering in University of Petroleum and Energy Studies, Dehradun.
He has published number of research articles in Journals and Conferences. His field of research includes Adaptive array signal processing, Smart Antennas, Artificial Intelligence and Machine Learning.

Arpit Jain received his B.Eng. from SVITS, Indore, India in 2007, received his M.Eng. from Thapar university, Patiala, India in 2009. He did his Ph.D. from UPES, Dehradun, India in 2018. He is currently working as an Assistant Professor (Selection Grade) in the department of Electrical and Electronics Engineering in University of Petroleum & Energy Studies (UPES).
He has over 10 years of teaching and research experience and a rich experience in curriculum design. The areas of his research interests include Real-Time Control System, Fuzzy Logic, Machine Learning, and Neural Networks.

Prateek Gupta received his B. Tech in Computer Science & Engineering from Uttar Pradesh Technical University, Lucknow, India in 2008 and M. Tech in Computer Science & Engineering from National Institute of Technology, Hamirpur, India in 2011. He has received the full-time PhD from the Department of Computer Science & Engineering, Dr. B R Ambedkar National Institute of Technology, Jalandhar, India in 2019. He is currently working as an Assistant Professor in School of Computer Science, UPES, Dehradun. His professional research activities lies in the field of optimization algorithms, wireless communications and sensor networks. He is a student member of IEEE.

Vinay Chowdary received his B. Tech. from MIST, JNTUH Hyderabad, India in 2005, received his M. Tech. from MIST, JNTUH Hyderabad, India in 2012. He did his Ph.D. from UPES, Dehradun, India in 2020. He is currently working as an Assistant Professor (Selection Grade) in the department of Electrical and Electronics Engineering in University of Petroleum & Energy Studies (UPES). He has an over 13 years of teaching and research experience and a rich experience in curriculum design. The areas of his research interests include Embedded Systems, Internet of Things, Wireless Sensor Networks, and Advanced Microcontrollers.