Sustainability insights on learning-based approaches in precision agriculture in internet-of-things

Kiran Muniswamy Panduranga¹, Roopashree Hejjaji Ranganathasharma²
¹Department of Computer Science and Engineering, Government Engineering College, Hassan, Affiliated to Visvesvaraya Technological University, Belagavi, India
²Department of Artificial Intelligence and Data Science, GSSS Institute of Engineering and Technology for Women, Mysuru, Affiliated to Visvesvaraya Technological University, Belagavi, India

ABSTRACT

Precision agriculture (PA) is meant to automate the complete agricultural processes with the sole target of enhanced crop yield with reduced cost of operation. However, deployment of PA in internet-of-things (IoT) based architecture demands solutions towards addressing various challenges where most are related to proper and precise predictive management of agricultural data. In this perspective, it is noted that learning-based approaches have made some contributory success towards addressing different variants of issues in PA; however, such methods suffer from certain loopholes, primarily related to the non-inclusion of practical constraints of IoT infrastructure in PA and lack of emphasis towards bridging the trade-off between higher accuracy and computational burden that is eventually associated with this. This paper contributes towards highlighting the strengths and weaknesses of recent learning approaches and contributes towards novel findings.

Keywords: Accuracy, Agriculture, Internet-of-things, Precision agriculture, Predictive management, Architectures

Corresponding Author:
Kiran Muniswamy Panduranga
Department of Computer Science and Engineering, Government Engineering College, Hassan, Affiliated to Visvesvaraya Technological University, Belagavi-590018, Karnataka, India
E-mail: mtech.kiran@gmail.com

1. INTRODUCTION

The evolution of the precision agriculture (PA) concept leads to a novel and innovative, intelligent farming notion that targets enhancing crop yields [1]. There has been an increasing level of attention towards PA for multiple reasons. The first reason is associated with a maximized efficiency of the PA concept-based operation that leads to the optimal usage of seeds, pesticides, fertilizers, and water [2]. Owing to maximized efficiency, the crop's productivity is eventually enhanced by adopting various data-driven techniques and precise technologies [3]. The second essential reason is associated with cost reduction and conservation of resources. Owing to the involvement of data analytical operation, PA can assist the farmers by providing calculative measures of input, precisely avoiding the wastage of resources. This directly affects minimizing the operational cost and downtime [4]. Adopting sophisticated analytical operations leads to effective decision-making towards pest management, fertilization, irrigation, and planting [5].

Further, PA assists in adequately managing and monitoring fields in real-time, improving field management overall [6]. PA is also closely linked with internet-of-things (IoT), where varied ranges of IoT devices were utilized for data collection and offering improved connectivity among the field sensors. The adoption of IoT-based technologies also assists PA to gain more control over transparency, traceability, automation, and optimization. From the viewpoint of the state-of-the-art in PA, global navigation satellite

Journal homepage: http://ijece.iaescore.com
systems (GNSS) are reported to be extensively used for variable rate technology, machinery guidance, and field mapping [7]. Apart from this, briefing of various state-of-art PA technologies has been reported towards monitoring technologies (e.g., weather monitoring, canopy sensing, soil sensing), guidance technologies (e.g., driver assistance, machine guidance, controller traffic farming), and acting technologies (variable rate assessment for defoliation, fertilization, weeding, treatments, seeding, and irrigation) [8]. However, there is another side of IoT applications when used with PA-based on-field sensors. Deploying IoT in PA will eventually call for the collection of massive data. This data is subjected to various sophisticated operations led by the big data approach and various variants of artificial intelligence [9]. In this perspective, it is noted that machine learning (ML) and deep learning (DL) are the two most widely adopted schemes in existing research work [10]. Adopting such approaches leads to correctly identifying anomalies or detecting some critical behavior directly linked with yield production [11]. Machine and DL-based approaches are further evolving in the current era with the evolution of various problem-solving capabilities. However, these learning-based methodologies are also associated with potential issues [12]. From these above studies, it is now noted that the main contributors towards PA development are based on big data and different variants of artificial intelligence. These contributing technologies found a computational mechanism towards predicting the prominent indicators as the outcome to enhance PA-based operation performance.

The prime outline of the research problem in PA is that there has been an evolution of various research work towards PA adopting these learning-based approaches with claimed benefits towards their predictive operation; however, the reality is that the commercial application of PA is yet to be known as actual practice in the existing farming era. Another problem is the presence of various technological constraints with such learning-based techniques that existing researchers consistently address. However, the questions associated with research challenges are manifold and yet to be answered. Hence, the statement of the problem undertaken in this study is “exploring the beneficial and limiting perspective of state-of-art literature towards improving the operational task of PA-based framework to find a potential gap is quite challenging and not reported in existing studies”.

Therefore, this paper contributes towards highlighting the strengths and weaknesses of recently published learning-based approaches for addressing varied PA issues. The paper also highlights the frequently used dataset and contributes towards showcasing the research trend and reaching a significant research gap. The new contribution of this paper is towards pinpointed discussion of the strengths and weakness of ML and DL approaches in PA-based IoT environments. The outline of this manuscript is as follows: section 2 offers a brief discussion about the strategy used towards carrying out the proposed review work to ensure presenting a concise and enriched learning outcome in the form of gap and trend, section 3 offers a brief background of insights on PA in IoT, followed by a discussion of recent work of ML-based approaches in section 4. A discussion of DL-based approaches is carried out in section 5, followed by a discussion of datasets frequently adopted in section 6. Section 7 highlights some critical research trends of manuscript publication patterns and research gaps. The conclusion of this paper is done in section 8.

2. METHOD

Before briefing the adopted research methodology, it is essential to understand that the proposed study is basically a review, especially emphasizing the adoption of learning-based approaches in PA. This motive is essential to understand which will potentially influence the data collection task. Figure 1 highlights the adopted method to carry out the research work.

![Methodology adopted towards presented review work](image-url)
According to the adopted methodology presented in Figure 1, the first process is to explore the database of scientific journals related to the usage of learning models in PA, where Institute of Electrical and Electronics Engineers (IEEE) and Multidisciplinary Digital Publishing Institute (MDPI) publications are found rich enough to give the information. This process leads to a massive number of 8,626 articles where learning-based methods have been reported to be used. **Primary shortlisting** follows the next step, where the title and abstract are reviewed to confirm that the collected dataset adheres to the inclusion and exclusion criteria. The inclusion criteria consist of i) only artificial intelligence-based PA schemes, ii) usage of a maximum number of recent publications, and iii) computational framework-based modelling. The exclusion criteria consist of i) papers published before 2016 are considered, ii) theory or conceptual manuscript, and iii) hardware-based model or non-artificial intelligence-based schemes. The next step is to eliminate replicated studies, where the dataset of scientific papers is deemed replicated if i) both bears exactly similar methodology in their implementation and ii) if the same authors have written multiple papers published on different portals. This led to the generation of 520 articles to find that most existing learning-based methods are nearly identical in different forms. Finally, all the filtered manuscripts are thoroughly studied concerning implementation sections, methodology, and accomplished results in a final shortlisting stage in adherence to inclusive and exclusive criteria to arrive at 119 research papers. This process leads to the extraction of the research gap and identification of trend that finally contributes towards extracting learning outcomes of this review work. This method is slightly manual and time-consuming, but a researcher gains a potential insight into the information provided in current studies, which assists in better decision-making and arrives at conclusive remarks. Further, the knowledge gap raised in the prior introduction section is addressed in the form of learning outcomes while adopting the proposed research method.

3. **INSIGHTS OF PA IN IoT**

PA can be stated to offer an innovative mechanism of smart farming by utilizing various types of advanced technologies, e.g., IoT, sensor technology, remote sensing, geographic information systems, artificial intelligence (AI), data analytics, automation and robotics, cloud computing, and web/mobile applications [13]. The prime reason for using all the technologies mentioned above in PA is to ensure that maximum and diversified information can be gathered and subjected to processing with a common agenda to increase yield and retain a healthy agricultural environment. To develop a research model of PA towards an IoT, it is essential to consider some of PA's essential aspects from the IoT viewpoint as shown in Figure 2.

- **Sensor network**: the involvement of sensors in IoT is quite common in agricultural fields. Various information associated with pests, nutrient content, potential of hydrogen (pH) level, humidity, temperature, and soil moisture can be collected with the presence of sensors in intelligent farming. Linking all these sensors in a network will eventually assist in aggregating information and transmission in IoT.
Analysis and aggregation of data: the agricultural data collected by farming-based sensors are meant to be subjected to further analytical proceedings. Various AI, ML, and DL based approaches are further applied to gain insight into the essential knowledge that helps make specific decisions towards smart farming in IoT towards crop management, pest control, fertilizer usage, and irrigation.

Remote control and monitoring: with the adoption of various technologies and communication, it is now possible for farmers to acquire information about their crops and farming areas remotely. By acquiring real-time feed of agricultural land, farmers can control the irrigation system and remotely control the usage of pesticides/fertilizers.

Predictive operation: various sophisticated analytical operations can be carried out on intelligent agricultural data. Big data can further facilitate predictive operation using environmental conditions and historical data. The predictive outcomes can be used for assessing varied challenges, variations in yield, and diseases in crops that further assist in planning necessary steps to resist potential threats to crops.

Automated system: the inclusion of IoT assists in the complete automation of specific processes of intelligent farming in PA. Deployment of various robots and drones can be used not only for monitoring but also for undertaking specific actions, e.g., targeted spraying of fertilizer/pesticides, and assessing the health of certain crops. Hence, more precise information can be acquired while tasks can be carried out with better accuracy and less demands of labor that conventional farming cannot offer in current times.

Optimization of resources: one of the prime contributions of PA compared to conventional farming is the proper optimization of varied resources, e.g., energy, fertilizer, and water. The better decision towards reducing the water waste and estimation of water in the irrigation process can be carried out in PA based on accurate real-time data of farming areas. This controls the surplus utilization of resources and reduces the environmental impact.

Direct access to decision support: the involvement of advanced technologies in PA based on an IoT environment provides a simplified interface via varied devices using a unique interface and dashboard. This facilitates farmers to access information from varied sources of IoT devices (in the form of sensors). The same interface can act as a controller too to managing various agricultural-related operations.

Apart from this, various devices are being used exclusively for PA. Apart from this, some applications mainly seeking their usage are equipment for variable rate technology (VRT) [14]. Such a form of equipment is capable of finetuning the input rate of various resources, e.g., seeds, and pesticides/fertilizers, based on variable demands in farming areas. Geospatial information system (GIS) also plays a significant role in PA, where data is analyzed and processed with a target to construct a comprehensive field map of their farming area. This visualization offers a simplified output of crop health, soil condition, and disease, to make more informed decisions depending on the spatial data. Irrespective of such immense beneficial factors of PA, it is also characterized by various challenging conditions, viz. data management, technology integration, investment and cost, training and technical expertise, infrastructure and connectivity, data privacy and security, acceptance of farmers, mitigating variable and uncertain condition in the farming area. These problems are massive and vast, filled with various specific sets of problems. It is noted that learning-based approaches are one of the most frequently adopted solutions for mitigating such issues. The following section discusses the effectiveness of such learning-based approaches in PA-based IoT deployed environments.

4. ML APPROACHES IN PA-IoT

Most applications towards PA in IoT are operated based on the environmental data it captures, while this data is subjected to further analytical operation. Apart from this, various functional operations carried out by devices in PA call for decision-making, where ML can be deployed. ML has been used in the existing design of PA from different perspectives. The work carried out by Bashir et al. [15] has used ML for the evaluation of the level of salinity in soil. The study aims to meet the demands of leaching water using an IoT-based environment in agriculture considering two attributes, i.e., temperature attribute and level of salinity. Using the standard reference score, the system uses the naïve Bayes classifier to perform predictive analysis of leaching requirements assessed on cotton crops. Further, the work carried out by Pal et al. [16] has developed a scheme towards assessing the density and elevation of grass vegetation in IoT-based infrastructure for crop monitoring. The study uses a path loss model to track network dysconnectivity. Considering the sugarcane and paddy medium grass fields, the system uses a formulated path loss model to identify the signal strength. A further multiple regression model is utilized to construct a generalized path loss model with an enhanced data transmission environment.

The study towards predictive operation in ML is further discussed by Diedrichs et al. [17] associated with a frost event for sensing devices deployed in IoT-based agriculture. For this purpose, the authors have used multiple ML algorithms concerning classification and regression forms towards the thermodynamic
condition to predict front events. The experiments on multiple locations show that random forest is the best-
performing ML algorithm from the classification outcome. At the same time, the performance of regression
and classifications is also improved. ML approach is also used in livestock management in smart farming of
IoT. According to this study model presented by Chatterjee et al. [18], the system continuously monitors
behavioral patterns and cows’ health using a predictive unsupervised ML algorithm. This is a multiclass
classifier; hence, it helps identify specific diseases in livestock, too. The work carried out by Chakraborty
et al. [19] has addressed an issue associated with the processing of image data from the agricultural farm. The
authors have introduced a unique image compression mechanism that uses ML to retain maximal image
quality. Liu et al. [20] have developed a predictive model to diagnose plant disease using ML in the
environment of an IoT. The authors use multiple linear regression for this predictive objective to understand
the connectivity between environment and disease threats.

ML has been adopted to improve the productivity score, as noted in Reyana et al. [21]. According to
the study model, agriculture data is collected from multiple sensors, fused, and subjected to ML. This model
has used random forest (RF), Hoeffding tree (HT), and J48 decision tree (JDT) as ML approaches for
classifying multiple crops, viz. wheat, sugarcane, paddy, maize, groundnut, and cotton. The
outcome shows RF to offer better performance. Another practically viable model is presented by Khan et al.
[22], where multiple sets of machine-learning approaches have been used for recommending the amounts of
fertilizers required in IoT-based smart farming. The study model uses k-nearest neighbor (KNN), Gaussian
naïve Bayes (GBN), support vector machine (SVM), and logistic regression (LR) by mapping soil fertility.
The study outcome exhibits GNB to offer better accuracy.

A similar form of study is also carried out by Garg et al. [23], where prediction of damage inflicted
on crops is carried out by assessing fertilizer proportion using KNN, decision tree (DT), XGBoost (XGB),
light gradient boosting (LGB), and RF. Apart from ML, this model has also deployed densely connected
convolution neural network (DenseNet121), residual network (ResNet50), and visual geometry group-16
(VGG-16) (pretraining model) to cross-check the outcomes of ML model. The recommended essential soil
nutrient was also carried out by Senapaty et al. [24] using an experimental approach. The model uses directed
acyclic graph (DAG) and SVM, further integrated with a unique bioinspired fruit fly algorithm (FFA)
approach. The study outcomes reveal SVM performs better towards the recommendation of soil nutrients.
The work carried out by Thilakarathne et al. [25] has developed another recommendation model towards PA
using ML. The work carried out by Elashmawy and Uysal [26] has presented a sensor-based model towards
studying the condition of the soil. The authors have considered using a regression framework using the
Gaussian process and neural network to conduct predictive soil analysis.

The work carried out by Bakthavatchalam et al. [27] used multiple sets of machine-learning
approaches towards crop recommendation. According to this study, multiple attributes have been selected to
recommend multiple crops in IoT-based smart farming. The study has used supervised learning using the
Waikato environment for knowledge analysis (WEKA) tool and a decision table with JRip (a multilayer
perceptron-based classification tool) for a predictive recommendation system. Another unique
implementation is carried out by Talaat [28], where big data and ML have been used to mitigate critical
decision-making issues for crop yield prediction.

The ensemble approach using the voting mechanism for ML technique was introduced by Peppes
et al. [29], where the notion of study is to identify the presence of threat in intelligent farming. The
framework has evaluated multiple ML models, including stochastic gradient descent (SGD) and soft/hard
ensembled voting techniques. Cadenas et al. [30] have presented a temporal scheme where data is subjected
to preprocessing and a soft computing-based ML approach—the idea to perform early prediction of frost in
farming. A unique work by Rokade et al. [31] has used a supervised learning scheme to improve crop yield.
The study has used artificial neural networks (ANN) and support vector machine (SVM), where decision-
making is carried out on the fog layer. The study has analyzed classification and regression models to assess
an effective predictive model. Hence, various ML-based approaches are being implemented to address
multiple issues based on the IoT environment in the PA framework. Table 1 (in appendix) offers a
summary of these approaches.

5. DL APPROACHES IN PA-IoT

Existing schemes have also reported extensive usage of DL towards addressing multiple predictive
problems associated with PA in IoT environments. The probable reason is that DL-based approaches were
claimed of higher accuracy and reduced dependency towards feature extraction compared to ML. The work
carried out by Pal et al. [32] has addressed the issue of minimal energy and cost towards tracking the
moisture content of soil. The study of water movement in soil is captured using conditioning circuits. The
information obtained from correlated data is subjected to a neural network framework on multiple depth
measures. The scheme uses bidirectional long short-term memory (BiLSTM) and ANN for predictive
analysis of soil movement. The work carried out by Lin et al. [33] presents the detection of anomalies in the trend of electric energy in IoT-based agricultural farms. The study model used long short-term memory (LSTM) integrated with a revised convolution neural network (CNN) towards the data captured from smart meters. Cruz et al. [34] have designed an experimental prototype using Raspberry Pi to detect heterogeneous forms of data in strawberry farming. Different types of sensors have been used to capture farm data. Analysis has been carried out overleaf images of standard datasets, while CNN is used for classification.

Further, leaf-based analysis of images concerning the classification of leaf disease is carried out in the work of Barbureanu et al. [35]. This classification is based on texture features using multiple pre-trained models. The species and leaf diseases are identified using DL where the textures are subjected to CNN, and later ML approach is applied. Study towards leaf diseases using the DL approach is also witnessed in the work of Ramana et al. [36]. An experimental prototype was designed using the Raspberry Pi module, where multiple sensory information was gathered, followed by the classification of leaf diseases using a deep neural network.

A study towards greenhouse monitoring is reported in the work of Castillo et al. [37] using DL. The authors have used CNN to predict the need for essential supplies and resources to improve the yield. The study outcome shows 90% of classification accuracy with reduced packet loss and minimal energy consumption. Zheng et al. [38] have presented a detection and classification scheme towards PA using DL. The authors have developed a new dataset of different crops with annotated instances of many images of different classes. The study uses a CNN-based model to carry out classification.

Jin et al. [39] have developed a predictive scheme for ensuring sustainable PA performance using DL. Weather data is acquired from IoT devices planted in farming areas, followed by applying a gated recurrent unit (GRU) for training. Varied predictive scores were acquired from the GRU-based training. The adoption of a GRU-based approach was also witnessed in the following work of Jin et al. [40] using a hybrid deep learning scheme. In this scheme, the decomposition of the empirical mode is used to break the climatic data and apply GRU. Sigalingging et al. [41] have used a unique deep-learning scheme to improve predictive performance in PA. This scheme selects an adaptive network using contextual information to improve the features. The DL-based approach is also witnessed towards classifying weeds in intelligent farming, as noted in the work of Lammie et al. [42]. The authors have addressed the higher resource and processor dependency issues in DL and presented a field programmable gate array (FPGA) using the OpenCL framework for performing binary classification of weeds images. The study model benefits agricultural robots with lesser power consumption than conventional DL methods. Horng et al. [43] have developed a scheme to recognize images of farming areas where the object detection method is used to identify crop maturity. A neural network is further used for training the identified images. The study uses a multi-layered perceptron (MLP) to forecast the possible position and mobility of robotic arms.

Further, MobileNetv2-based CNN is used for extracting image features that is further integrated with detection using a single shot. Shafi et al. [44] have used a multimodal approach towards remote sensing for monitoring crop health. According to this study model, temporal data is acquired from IoT devices, while crop health maps are generated from transformed multispectral images. Further DL is applied for the final steps of predictive classification. Kashyap et al. [45] have developed a scheme towards facilitating intelligent irrigation systems in PA. The scheme uses LSTM to carry out predictive analysis of soil moisture contents before the irrigation period, followed by water distribution towards arable land.

The work carried out by Kim et al. [46] has developed a DI-based filter to suppress the noise present in various agricultural data. LSTM has been used for designing this filter, which is reported to improve production with reduced cost. Zhang et al. [47] have developed an autonomous method of scouting a farming area using DL. The technique uses CNN and reinforcement learning (RL) for this purpose. The work carried out by Sarayana et al. [48] has presented an analysis of DL methods on PA. According to this model, transfer learning (TL) has been adopted along with the VGG16 model towards detecting pests in PA.

Further, Manikandan et al. [49] have used the DL approach towards smart farming, where a controller design uses fuzzy logic to facilitate an intelligent irrigation system. Studies in PA are not only restricted to monitoring or improving yield but also towards improving security. Kumar et al. [50] have developed a security scheme where DL improves privacy preservation. According to this model, data authentication uses blockchain and DL, while data transformation is carried out using sparse auto encoder (SAE). Finally, stacked LSTM is used for detecting anomalies. The adoption of auto encoder (AE) is witnessed in the work of Adkisson et al. [51] for similar security objectives.

Hence, multiple variants of DL approaches have been witnessed in recent studies towards PA based on the IoT environment. These approaches report various beneficial features addressing different challenges, as well as having limitations. Table 2 summarizes DL-based approaches.
TABLE 2. SUMMARY OF DL-TECHNIQUES IN PA-IoT

<table>
<thead>
<tr>
<th>Authors</th>
<th>Problems</th>
<th>Mechanism</th>
<th>Advantage</th>
<th>Limitation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pal et al. [32]</td>
<td>Tracking soil movement</td>
<td>BiLSTM, ANN</td>
<td>Highly reduced error score</td>
<td>The model assessed a smaller set of data</td>
</tr>
<tr>
<td>Lin et al. [33]</td>
<td>Prediction of electric energy</td>
<td>LSTM, revised CNN</td>
<td>99% accuracy in anomaly detection</td>
<td>Heterogeneity in devices and the impact of the environment on energy is not involved.</td>
</tr>
<tr>
<td>Cruz et al. [34]</td>
<td>Disease detection</td>
<td>Edge computing</td>
<td>Accomplishes 92% detection accuracy</td>
<td>Need extensive analysis further to prove its generalized applicability</td>
</tr>
<tr>
<td>Barburiceanu et al. [35]</td>
<td>Texture classification</td>
<td>Pre-trained CNN model, ML classifier</td>
<td>Discriminative power, reduced processing time</td>
<td>No benchmarking</td>
</tr>
<tr>
<td>Ramana et al. [36]</td>
<td>Leaf disease detection</td>
<td>Raspberry Pi, deep neural network, CNN</td>
<td>Accomplishes 96% accuracy</td>
<td>Model lacks interpretability</td>
</tr>
<tr>
<td>Castillo et al. [37]</td>
<td>Early disease detection, controlling supplies and resources in the greenhouse</td>
<td>CNN</td>
<td>Satisfactory detection performance</td>
<td>Low accuracy of 90%</td>
</tr>
<tr>
<td>Zheng et al. [38]</td>
<td>Greenhouse monitoring</td>
<td>Constructed new dataset, multiple CNN based models, only look once (YOLOv3)</td>
<td>99% classification accuracy</td>
<td>Computationally extensive process</td>
</tr>
<tr>
<td>Jin et al. [39], [40]</td>
<td>Sustainable production in variable weather conditions</td>
<td>Decomposition of climate data by empirical mode, GRU deep learning using contextual information, adaptive network formation</td>
<td>Effective predictive performance for humidity and temperature improves the classification of an image</td>
<td>The scope of the study is limited to a specific geographic area.</td>
</tr>
<tr>
<td>Sigalingging et al. [41]</td>
<td>Sustainable production in variable weather conditions</td>
<td>Decomposition of climate data by empirical mode, GRU deep learning using contextual information, adaptive network formation</td>
<td>Effective predictive performance for humidity and temperature improves the classification of an image</td>
<td>Low accuracy (88%) Applicable for specific crop datasets only.</td>
</tr>
<tr>
<td>Lammie et al. [42]</td>
<td>Weed classification</td>
<td>FPGA, OpenCL framework</td>
<td>Better energy efficiency</td>
<td>Scalability analysis towards FPGA modules not carried out</td>
</tr>
<tr>
<td>Horng et al. [43]</td>
<td>Robotic harvesting</td>
<td>Object detection, MLP, MobileNetv2, CNN</td>
<td>Effective crop detection</td>
<td>Assessed on narrowed test-cases</td>
</tr>
<tr>
<td>Kashyap et al. [44]</td>
<td>Crop health mapping in IoT</td>
<td>ML, DL approach</td>
<td>Accomplishes 98% of accuracy</td>
<td>The computational burden is high</td>
</tr>
<tr>
<td>Kashyap et al. [45]</td>
<td>Smart irrigation</td>
<td>LSTM</td>
<td>Simplified learning</td>
<td>The model lacks dynamic constraint modelling</td>
</tr>
<tr>
<td>Kim et al. [46]</td>
<td>Noise suppression in agricultural data</td>
<td>LSTM</td>
<td>Increase yield</td>
<td>Not applicable to online learning</td>
</tr>
<tr>
<td>Zhang et al. [47]</td>
<td>Autonomous scouting in PA</td>
<td>CNN, RL</td>
<td>Reduced labor cost</td>
<td>Reduced accuracy (89%), modelling uses less sample of data.</td>
</tr>
<tr>
<td>Sarayana et al. [48]</td>
<td>Pest detection in PA</td>
<td>TL, VGG16</td>
<td>Accomplishes 96% accuracy</td>
<td>Overfitting issues not addressed</td>
</tr>
<tr>
<td>Manikandan et al. [49]</td>
<td>Improving field information</td>
<td>Deep learning, fuzzy logic</td>
<td>Sustainable for multiple environmental conditions</td>
<td>Needs extensive ruleset for optimal accuracy</td>
</tr>
<tr>
<td>Kumar et al. [50]</td>
<td>Privacy preservation</td>
<td>Blockchain, SAE, stacked LSTM</td>
<td>Offers substantial security</td>
<td>The model does not support distributed computing</td>
</tr>
<tr>
<td>Adkisson et al. [51]</td>
<td>Cyber attacks</td>
<td>Unsupervised AE</td>
<td>Accomplishes 98% accuracy</td>
<td>Higher likelihood of improper training</td>
</tr>
</tbody>
</table>

6. DATASET OF PA

Various applications are associated with PA-based IoT services; hence, such architecture is designed and tested on specific standard datasets. Before understanding the available dataset in PA, it is necessary to realize that there are various acquisition techniques of the dataset in PA, e.g., i) using artificial intelligence and data analytics on collected data from various sources [52], ii) farm management software [53], iii) analysis and sampling of soil that consists of information of organic matter, pH level, and nutrient level [54], iv) yield monitoring system where harvesting machines are mounted with yield monitors [55], v) global positioning system (GPS) for tracking and precise positioning, vi) IoT and sensors for acquiring various physical attributes, e.g. nutrient level, light intensity, humidity, temperature, and soil moisture [56], and vii) remote sensing using aerial imagery, drones, and satellites. to capture information of soil moisture, vegetation indices, and crop health [57]. At present, all these repositories are arranged in the publicly available dataset for facilitating researcher towards their investigation, e.g.:

- Yield dataset: a reputed organization, e.g., the United States Department of Agriculture, offers various data towards crop yield from varied regions. These datasets can be used for predictive PA yield analysis [58].
- Crop disease dataset: the dataset of image-based crop disease [59] and PlantVillage [60] associated with various standard crops offers information about plant diseases. Sensory information or image-based data
are used for extracting disease information of soybeans, maize, and wheat. The dataset is labelled with a
degree of disease severity.

- Soil data: soil data mart provides exclusive information associated with soil properties, viz., organic
  matter, pH level, and nutrient content. Such dataset is organized from laboratory analysis and soil survey
  [61].

- Climate data: PA also requires information about climatic data, e.g., solar radiation, humidity, rainfall,
  and temperature. The National Centers for Environmental Information (NCEI) offers the data necessary
to connect influencing factors of climatic attributes with crop production [62].

- Indices of vegetation MODIS: the existing research often considers the enhanced vegetation index (EVI)
  and normalized difference vegetation index (NDVI) to offer essential information associated with vigor,
growth, and crop health. Such information is offered by moderate resolution imaging spectroradiometer
  (MODIS) [63].

- National agricultural imagery program (NAIP): this dataset consists of high-resolution multispectral
  images that target identifying anomalies, tracking crop health status, and classifying crops [64].

Table 3 (see in appendix) highlights some of the datasets used in the research work of PA in IoT
apart from those mentioned above. It can be seen that most of the data are in image form and captured from
ground-based devices in red, green, and blue (RGB) mode. A closer look into this dataset shows that varied
numbers of data are not enough for analysis using AI-based schemes. Hence, data augmentation is applied to
scale up the dataset to overcome the data shortage for predictive analysis. Various new form of data
augmentation techniques using data warping and generative adversarial network (GAN) is currently used in
PA, as seen in Barth et al. research [65]. Existing studies on datasets have also been extended towards
performing image annotation by providing a label to describe the specific region of an image with ground
truth. The work of Bhagat and Choudhary [66] has presented a mechanism of annotation towards a learning
model subjected to training and allocating autonomously semantic labels. For this purpose, various forms of
annotators have been used, viz. Openlabelling [67], LabelMe [68], labelling [69], imglab [70], ImageTagger
[71], computer vision annotation tool (CVAT) [72], common objects in context (COCO) annotator [73],
visual object tagging tool (VoTT) [74], VGG image annotator (VIA) [75], and and Yolo_mark [76]. It should
be noted that usage of such annotators is suitable for smaller datasets and not for bigger-sized datasets. The
complete process of annotation takes usually 15-30 minutes. The research work carried out by Kovashka
et al. [77], Wiesner-Hands et al. [78], and Zhou et al. [79] presented a crowdsourcing-based annotation
 technique for larger-sized datasets.

After the annotation is carried out, these datasets are required to be reused for further adoption in
research work, and the existing system finds its sources from various publicly available platforms, e.g.,
Zenodo, Mendley Data, Open Science Framework, Harvard Dataverse, GitHub, Figshare, Dryad, and
CyVerse. At present, there are a total of 34 image dataset that has been classified into 10 datasets for research
on fruit detection, 15 dataset that are used for research on weed control, nine datasets for other different
research areas of PA open image V4 dataset is another new dataset which is witnessed to be adopted by
various research work as witnessed in study of Kuznetsova et al. [80], while PlantVillage (Mohanty et al.
[59]) and Leafsnap (Redmon and Farhadi [81]) are another dataset which is seen to be frequently adopted in
existing research work. Table 3 highlights the dataset characteristics increasingly adopted for research work
in PA to address various issues.

7. CRITICAL DISCUSSION OF FINDINGS

The prior sections have an elaborate discussion associated with the findings of the machine and deep
learning-based methods deployed in PA environments. From the distinctive highlights of beneficial and
limiting factors exhibited in Tables 1, 2, and 3, an insight towards comparisons can be seen from existing
studies. This section discusses the interpretation of critical discussion of accomplished findings concerning
the existing research trends towards adopting various methodologies in improving PA-based services. The
outcome presented in this section can be considered a ramification of the review findings. For this purpose,
the research journals published between 2018 to 2023 have been considered reputed publications. The
following are the observations towards existing research trends.

7.1. Research trends

Figure 3 highlights 24,019 research publications towards PA, with 14781 publications towards
generic PA-based approaches. In contrast, ML-based approaches (n=3,266) are slightly higher than DL-based
approaches (n=3111), whereas 2,249 publications have been seen towards TL based approach. Figure 4 states
that there are 28,272 research publications where ML approaches have been used towards PA and intelligent
farming. A close observation shows that regression-based methods (n=5,332), RF based model (n=4,996),
and LR (n=4,427) are in higher adoption. This score is followed up by the next set of frequently used methods, e.g., SVM (n=3,011), k-means clustering (KMC) (n=3,027), DT (n=2,523) and KNN (n=2,217). Other methods, e.g., Q-learning, R-learning, and TD learning, have not received much attention, while 391 publications are witnessed for NB. The overall score of the complete ML-based techniques eventually shows that its adoptions are much higher in existing research trends addressing various sets of research problems in IoT.

Figure 3. The trend towards research methods in PA

Figure 4. The trend towards ML methods in PA
Figure 5 highlights that there is a total of 6,036 research journals using different forms of DL-based approach, where ANN (n=1,926) has extensively used while is further followed up by another set of frequently used DL approaches, e.g., self-organizing map (SOM) (n=1,698) and CNN (n=1,274). It can also be seen that the adoption of RNN (n=672), LSTM (n=309), and Autoencoder (AE) (n=157) has received significantly less attention towards PA-based research work.

Figure 6 showcases that there are 4,913 research articles witnessed to discuss transfer learning (TL)-based approaches in PA where feature extraction (n=1,489) and multi-task learning (n=1,207) has been more adopted while finetuning (n=734), domain adaptation (n=921), and Zero-shot learning (n=562) have not witnessed much research publications in last five years. Following are the learning outcomes witnessed after insight towards current research trends: i) studies towards non-IoT-based infrastructure and technologies in PA is relatively higher compared to IoT-based deployment scenario in PA; ii) the adoption of ML techniques is slightly higher and eventually increases its pace towards research publication compared to DL-based approaches in PA. Further, TL-based techniques are much less research areas in this regard; iii) from the perspective of ML-based approaches, higher adoption of RF and regression-based techniques are witnessed compared to other techniques, e.g., SVM, KNN, K-means clustering, and DT. Much less research work is carried out considering naïve Bayes algorithms; iv) from the perspective of DL-based approaches, more models have used ANN, while a closer look into core implementation in DL shows less adoption of other approaches, e.g., CNN, RNN, and LSTM; and v) the TL-based approach is considered a better version of DL-based approach in many aspects, has witnessed a much smaller number of adoptions towards implementation models.
7.2. Research gaps

After reviewing the existing variants of learning approaches in PA, it is noticed that ML approaches have been extensively chosen while DL and TL-based approaches are yet to receive the equivalent acceptance in PA-based IoT. It is also observed from the prior section that PA is studied in existing research work, sometimes considering the essential background of IoT infrastructure. Sometimes, IoT has never been considered. Almost all the existing studies have potential contributions and certain limitations are associated with existing learning-based approaches. Further, it is noticed that various survey work is carried out towards addressing various PA-based issues via. Privacy (Demestichas et al. [101], Gupta et al. [102]), control strategies in PA (Hassan et al. [103]), emerging technologies in PA (Mesias-Ruiz et al. [104], Singh et al. [105]), ML-based approaches (Benos et al. [106], Sharma et al. [107]), ML approaches for pest and disease identification (Domingues et al. [108]), IoT based approaches (Dhanaraju et al. [109]), AI-based approaches for vertical farming (Holzinger et al. [110], Siregar et al. [111]), DL-based approaches used in PA (Altalak et al. [112]). However, it has been noted that some review works have emphasized IoT-based deployment approaches towards PA [113]-[119]. Such form of research trends showcases those studies towards adoption of IoT-based deployment scenarios are increasing in pace compared to other forms of methodologies or techniques in PA. From the insights of all these research work evidences, there are yet some of the open-ended gaps that are essentially required to be considered as follows:

- Reliability and quality of data: The majority of the existing approaches have directly fed the data to the learning models and proved the effectiveness of existing models. However, these models [9]–[13] don not consider significant challenges associated with considered data that is not only massive in size but also highly complex in its structure. Without the inclusion of customized and dedicated processing towards data quality, there is a higher likelihood of incorrect decision-making.

- Ambiguity towards real-time processing: existing approaches have also been reported to claim their applicability towards supporting real-time processing in PA. However, no substantial evidence supports this claim, as data captured from IoT devices are highly fluctuating and error-prone, often leading to scalability issues. Due to a lack of benchmarking, existing ML or DL approaches reported by [15]-[51] don not prove this phenomenon.

- Less emphasis on interoperability of data and services: a sophisticated form of PA will be adopting multiple IoT devices to capture the event and data. Hence, existing studies do not discuss the process of integrating them to make them suitable for learning-based PA methods. This causes interoperability problems, rendering the model inapplicable in real-time operations.

- Lack of dynamic characteristic involvement: a PA-based application will encounter various uncertain changes in its environment, weather or soil conditions. Existing studies [15]-[51] using learning-based approaches do not consider this uncertainty or dynamic characteristic of the environment while performing training operations. It will eventually infer that existing learning-based models in PA are not adaptable to dynamic alterations.

- Issues in ML-based approaches: despite frequent adoption of ML-based approaches in PA, the models don not encounter substantial scalability performance from the training perspective [15]-[31]. Higher accuracy can be only claimed in the presence of higher epoch values of ML models while stringent evaluation of response time of such models is not evaluated.

- Issues in DL-based approaches: there is no doubt that CNN offers better predictive performance in accuracy with no dependency on feature extraction [33]–[37]. However, none of the CNN or other DL-based approaches have proven its interpretability issues owing to its theoretical adoption of the Blackbox-based notion while performing analysis. The outcome obtained by these models is at the cost of higher resources, which cannot be stated as a practical solution. This problem exists for TL-based approaches, but they have not been deeply investigated.

- Low emphasis towards computational burden: a closer look into existing models showcases that accuracy is the pivotal aim to be accomplished using learning-based methods. However, gaining a higher accuracy with a low computational burden is significantly lacking to be reported in existing studies. Hence, a better and more innovative learning model is demanded, with low resource utilization, scalability, better interpretability, and low computational burden in PA based on the IoT environment.

8. CONCLUSION AND FUTURE SCOPE

PA is an innovative and futuristic smart farming technology that will eventually substitute the conventional agriculture mechanism. However, such a futuristic model will demand the inclusion of sophisticated technologies with a higher emphasis towards accuracy and deployment of cost-effective models. From a practical viewpoint, the cost and installation of PA, considering the IoT deployment environment, calls for a heavy expenditure; hence, its return on investment should be at par. At present, the
incorporation of sophisticated technologies of AI has already started to be investigated with a certain degree of claimed success reports. Still, seeing the full-fledged theoretical benefits in the real world has to go a long way. The following is the novelty of the proposed review work: i) unlike existing review work, the proposed review discusses only recent methodologies crisply for ML, DL, and TL approaches with pinpointed highlights of their advantages and limitations; ii) the proposed review offers some essential learning outcomes to state the degree of adoption of various learning-based approaches. The study finds that adopting ML techniques is considerably more than DL-based approaches, despite knowing that DL-based approaches are known for their higher accuracy; iii) By reviewing the methodologies of existing learning-based methods, an exciting finding states that ML techniques are known for their cost-effective deployment in large scenarios of PA implementation in an IoT environment, and there is a higher scope in this regard. Although there are some severe issues in existing ML approaches, sorting them out will eventually be an optimal solution for fulfilling the actual objectives of PA; and iv) Another essential novel finding of this study was the usage trend of the PA-based dataset. There is various dataset available that are eventually annotated using various available annotators. However, the findings showcase the ineffective applicability of such annotated datasets when such models are subjected to many dynamic instances in PA. Although the TL-based approach can overcome these issues, the cost of such an optimal solution is more to bear computationally.

Therefore, future work will be in the direction of mitigating the identified research gap. The first line of future work can be deployed towards a novel PA-based dataset management to be subjected to a training algorithm. The notion will be to make a novel training algorithm that can offer predictive consistency over dynamic ranges of test data in PA. The second line of future work will address the current issues in ML-based approaches and develop a novel framework that can consider a specific use case of PA to offer a better solution. The prime notion will balance higher accuracy with an optimally cost-effective computational process. Developing comprehensive and extensive test cases is necessary to prove the futuristic model’s applicability towards supporting real-time PA-based applications without much re-engineering.

**APPENDIX**

Table 1. Summary of ML-techniques in PA-IoT

<table>
<thead>
<tr>
<th>Authors</th>
<th>Problems</th>
<th>Mechanism</th>
<th>Advantage</th>
<th>Limitation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bashir et al. [15]</td>
<td>Prediction of leaching water requirements</td>
<td>Naïve Bayes classifier</td>
<td>Satisfactory accuracy</td>
<td>Study specific and applicable only to cotton plants</td>
</tr>
<tr>
<td>Pai et al. [16]</td>
<td>Connectivity issues in grass field</td>
<td>Path loss model, multiple regression</td>
<td>Enhanced coverage</td>
<td>No benchmarking</td>
</tr>
<tr>
<td>Diedrichs et al. [17]</td>
<td>Prediction of frost events</td>
<td>Regression and classification, RF</td>
<td>Effective decision making</td>
<td>It does not address data quality issues</td>
</tr>
<tr>
<td>Chatterjee et al. [18]</td>
<td>Prediction for the health of livestock</td>
<td>Unsupervised multiclass classifier</td>
<td>Simplified architecture</td>
<td>Poor model validation</td>
</tr>
<tr>
<td>Chakraborty et al. [19]</td>
<td>Image compression in farm</td>
<td>Machine learning</td>
<td>Higher retention of signal quality</td>
<td>Constraints associated with hyperspectral images or visual sensor-based images are not considered.</td>
</tr>
<tr>
<td>Liu et al. [20]</td>
<td>Early disease prediction</td>
<td>Multiple linear regression</td>
<td>Accomplishes 91% of predictive accuracy</td>
<td>Model lacks interpretability</td>
</tr>
<tr>
<td>Reyana et al. [21]</td>
<td>Multi-crop classification</td>
<td>Multiple machine learning model (RF, HT, JDT)</td>
<td>RF to offer better performance, highly practically viable model</td>
<td>Overfitting issues not analyzed</td>
</tr>
<tr>
<td>Khan et al. [22]</td>
<td>Fertilizer recommendation</td>
<td>KNN, GNB, SVM, LR</td>
<td>GNB to perform better, 90% accuracy</td>
<td>Demands higher iteration</td>
</tr>
<tr>
<td>Garg et al. [23]</td>
<td>Predicting crop damage</td>
<td>KNN, DT, XGB, LGB, RF</td>
<td>Simplified architectural modelling</td>
<td>Higher accuracy demands higher training efforts</td>
</tr>
<tr>
<td>Senapaty et al. [24]</td>
<td>Soil nutrient recommendation</td>
<td>Experimental approach, machine learning (DAG, SVM), Bio-inspired algorithm (FFA)</td>
<td>SVM to excel better predictive performance</td>
<td>Low convergence performance for dynamic constraints</td>
</tr>
<tr>
<td>Thilakaratne et al. [25]</td>
<td>Crop recommendation</td>
<td>SVM, XGB, RF, DT, KNN</td>
<td>Sustainable framework</td>
<td>Demands extensive training</td>
</tr>
<tr>
<td>Elashmawy and Uysal [26]</td>
<td>Controlling and monitoring soil condition</td>
<td>Regression (Gaussian), neural network</td>
<td>Effective predictive capability</td>
<td>Demands higher iteration for better predictive accuracy</td>
</tr>
<tr>
<td>Bakhtavachalam et al. [27]</td>
<td>Crop recommendation</td>
<td>Multilayer perceptron, decision table</td>
<td>Accomplishes 98% accuracy</td>
<td>Consistency reduces with an increase in data</td>
</tr>
<tr>
<td>Talaat [28]</td>
<td>Predicting crop yield</td>
<td>Big data, regression model (RF, DT)</td>
<td>RF offers 99% accuracy, minimizes dependency on labelled trained data</td>
<td>Consumes higher processing time</td>
</tr>
<tr>
<td>Peppes et al. [29]</td>
<td>Threat detection</td>
<td>SGD, RF, FT, KNN, ensembled voting</td>
<td>Simplified model implementation</td>
<td>Data constraints and quality are not considered</td>
</tr>
<tr>
<td>Cadenas et al. [30]</td>
<td>Early prediction of frost</td>
<td>Preprocessing, soft computing</td>
<td>Effective decision support</td>
<td>Narrowed evaluation</td>
</tr>
<tr>
<td>Rokade et al. [31]</td>
<td>Improving prediction</td>
<td>ANN, SVM (classification and regression)</td>
<td>SVM found better performance than ANN</td>
<td>No Benchmarking</td>
</tr>
</tbody>
</table>
Table 3. Summary of currently available dataset

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Application</th>
<th>Annotation</th>
<th># data</th>
<th>Platform</th>
<th>Modality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mortensen et al. [82] for growth of oil radish</td>
<td>Estimation of yield</td>
<td>Pixel level</td>
<td>129</td>
<td>Ground vehicle</td>
<td>RGB</td>
</tr>
<tr>
<td>Skovsen et al. [83] GrassClove</td>
<td>Prediction of biomass, canopy species</td>
<td>Pixel level</td>
<td>More than 10,000</td>
<td>Ground-based</td>
<td>RGB</td>
</tr>
<tr>
<td>Jiang et al. [84] DeepSeeding</td>
<td>Counting seeding</td>
<td>Bounding box</td>
<td>5743</td>
<td>Ground-based</td>
<td>RGB</td>
</tr>
<tr>
<td>Wiesner-Hanks et al. [85] Maize disease</td>
<td>Counting seeding</td>
<td>Line level</td>
<td>More than 10,000</td>
<td>Multiple platforms</td>
<td>RGB</td>
</tr>
<tr>
<td>Hāni et al. [86] Sugarcane</td>
<td>Detection of disease</td>
<td>Line level</td>
<td>156</td>
<td>Ground-based</td>
<td>RGB</td>
</tr>
<tr>
<td>Dias et al. [87] Fruit flower dataset</td>
<td>Flower detection</td>
<td>Pixel level</td>
<td>190</td>
<td>Hand-held, ground vehicle</td>
<td>RGB</td>
</tr>
<tr>
<td>Dutta and Zisserman [88]</td>
<td>-</td>
<td>Pixel level</td>
<td>More than 10,000</td>
<td>No imaging platform</td>
<td>RGB</td>
</tr>
<tr>
<td>Capricorn Annunun</td>
<td>-</td>
<td>Pixel level</td>
<td>More than 10,000</td>
<td>No imaging platform</td>
<td>RGB</td>
</tr>
<tr>
<td>Akbar et al. [89] Apple Trees</td>
<td>Tree pruning</td>
<td>No annotation</td>
<td>NA</td>
<td>Hand-held</td>
<td>RGBD</td>
</tr>
<tr>
<td>Kusumam et al. [90] 3D Broccoli</td>
<td>Flower detection</td>
<td>No annotation</td>
<td>NA</td>
<td>Ground vehicle</td>
<td>RGBD</td>
</tr>
<tr>
<td>Hani et al. [91] Minne Apple</td>
<td>Fruit health</td>
<td>Pixel level</td>
<td>More than 10,000</td>
<td>Hand-held</td>
<td>RGB</td>
</tr>
<tr>
<td>Gené-mola et al. [92] LFuji-air dataset</td>
<td>Fruit health</td>
<td>Bounding-box</td>
<td>NA</td>
<td>Ground-based</td>
<td>LiDAR</td>
</tr>
<tr>
<td>Gené-Mola et al. [93], Fuji-StM</td>
<td>Fruit health</td>
<td>Bounding</td>
<td>288</td>
<td>Ground-based</td>
<td>RGB</td>
</tr>
<tr>
<td>Bhushal et al. [94], W.S.U. Apple dataset</td>
<td>Fruit health</td>
<td>Bounding box</td>
<td>2298</td>
<td>Ground vehicle</td>
<td>RGB</td>
</tr>
<tr>
<td>Koirala et al. [95], Mango YOLO</td>
<td>Fruit health</td>
<td>Bounding box</td>
<td>1730</td>
<td>Ground vehicle</td>
<td>RGB</td>
</tr>
<tr>
<td>Kestur et al. [96], MangoNet</td>
<td>Fruit health</td>
<td>Pixel level</td>
<td>49</td>
<td>Hand-held</td>
<td>RGB</td>
</tr>
<tr>
<td>Gene-Mola et al. [97], KFuji RGB-DS</td>
<td>Fruit health</td>
<td>Bounding box</td>
<td>967</td>
<td>Ground vehicle</td>
<td>RGB-D</td>
</tr>
<tr>
<td>Altaheri et al. [98], Date fruit</td>
<td>Fruit health</td>
<td>Image level</td>
<td>More than 10,000</td>
<td>Unknown</td>
<td>RGB</td>
</tr>
<tr>
<td>Bargoti and Underwood [99], Orchard Fruit</td>
<td>Fruit health</td>
<td>Circle, bounding box</td>
<td>3704</td>
<td>Ground vehicle</td>
<td>RGB</td>
</tr>
<tr>
<td>Sa et al. [100] DeepFruits</td>
<td>Fruit health</td>
<td>Bounding box</td>
<td>587</td>
<td>Ground-based</td>
<td>RGB</td>
</tr>
</tbody>
</table>

REFERENCES


BIOGRAPHIES OF AUTHORS

Kiran Muniswamy Panduranga has completed BE (information science and engineering) in the year 2005 from Sri Siddhartha Institute of Technology Tumkur, affiliated to Visvesvaraya Technological University, Belagavi, Karnataka, India and MTech (computer science and engineering) in the year 2007 from AMC Engineering College Bangalore, affiliated to Visvesvaraya Technological University, Belagavi, Karnataka, India. He has around 12 years of teaching experience. He is presently working as assistant professor in the Department of Computer Science and Engineering at Government Engineering College Hassan, Karnataka, India. He can be contacted at email: mtech.kiran@gmail.com.

Roopashree Hejjaji Ranganathasharma has completed BE (electronics and communication engineering) and MTech (computer science and engineering) from VTU, Belagavi, Karnataka, India and PhD from CHRIST (Deemed to be University) Bengaluru, Karnataka, India. She has around 13 years of Industrial experience and 4 years of teaching experience. She is presently working as professor and head in the Department of Artificial Intelligence and Data Science at GSSS Institute of Engineering and Technology for Women, Mysuru, Karnataka, India and supervising 6 PhD research scholars in Visvesvaraya Technological University. She can be contacted at email: roopashreehr@gsss.edu.in.