

Received March 17, 2020, accepted April 15, 2020, date of publication May 6, 2020, date of current version May 20, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2992480

Crop Yield Prediction Using Deep Reinforcement Learning Model for Sustainable Agrarian Applications

DHIVYA ELAVARASAN AND P. M. DURAIRAJ VINCENT^{ID}, (Member, IEEE)

School of Information Technology and Engineering, Vellore Institute of Technology (VIT), Vellore 632014, India

Corresponding author: P. M. Durairaj Vincent (pmvincent@vit.ac.in)

ABSTRACT Predicting crop yield based on the environmental, soil, water and crop parameters has been a potential research topic. Deep-learning-based models are broadly used to extract significant crop features for prediction. Though these methods could resolve the yield prediction problem there exist the following inadequacies: Unable to create a direct non-linear or linear mapping between the raw data and crop yield values; and the performance of those models highly relies on the quality of the extracted features. Deep reinforcement learning provides direction and motivation for the aforementioned shortcomings. Combining the intelligence of reinforcement learning and deep learning, deep reinforcement learning builds a complete crop yield prediction framework that can map the raw data to the crop prediction values. The proposed work constructs a Deep Recurrent Q-Network model which is a Recurrent Neural Network deep learning algorithm over the Q-Learning reinforcement learning algorithm to forecast the crop yield. The sequentially stacked layers of Recurrent Neural network is fed by the data parameters. The Q-learning network constructs a crop yield prediction environment based on the input parameters. A linear layer maps the Recurrent Neural Network output values to the Q-values. The reinforcement learning agent incorporates a combination of parametric features with the threshold that assist in predicting crop yield. Finally, the agent receives an aggregate score for the actions performed by minimizing the error and maximizing the forecast accuracy. The proposed model efficiently predicts the crop yield outperforming existing models by preserving the original data distribution with an accuracy of 93.7%.

INDEX TERMS Crop yield prediction, deep recurrent Q-network, deep reinforcement learning, intelligent agrarian application.

I. INTRODUCTION

Agriculture is the one amongst the substantial area of interest to society since a large portion of food is produced by them. Currently, many countries still experience hunger because of the shortfall or absence of food with a growing population. Expanding food production is a compelling process to annihilate famine. Developing food security and declining hunger by 2030 are beneficial critical objectives for the United Nations. Hence crop protection; land assessment and crop yield prediction are of more considerable significance to global food production [1]. A country's policymaker depends on precise forecast, to make appropriate export and import assessments to reinforce national food security. Cultivators

The associate editor coordinating the review of this manuscript and approving it for publication was Dongxiao Yu.

and farmers further benefit from yield forecast to make financial and management decisions. Agricultural supervision, especially the observation of crop yield, is indispensable to determine food security in a region [2]. On the other hand, crop yield forecasting is exceedingly challenging because of various complex aspects. Crop yield mainly depends upon climatic conditions, soil quality, landscapes, pest infestations, water quality and availability, genotype, planning of harvest activity and so on [3]–[5].

The crop yield processes and strategies vary with time and they are profoundly non-linear in nature [6], and intricate due to the integration of a wide extent of correlated factors [7], [8] characterized and impacted by non-arbitrate runs and external aspects. Usually, a considerable part of the agricultural framework cannot be delineated in a fundamental stepwise calculation, especially with complex, incomplete, ambiguous

and strident datasets. Currently, many studies demonstrate that machine learning algorithms have comparatively more improved potential than conventional statistics [9]–[12]. Machine learning belongs to the field of artificial intelligence by dint of which computers can be instructed without definite programming. These processes resolve non-linear or linear based agricultural frameworks with remarkable forecasting ability [13]. In Machine learning agricultural frameworks, the techniques are obtained from the learning process. These processes demand over train to perform a specific task. After the completion of the training process, the model makes presumptions to test the information.

Further, machine learning resembles an umbrella that holds various significant strategies and methodologies. On observing the most prominent models in agriculture, we can see the utilization of artificial and deep neural networks [14]. Deep learning is a subgroup of machine learning that can determine outcomes from varying arrangements of raw data. Deep learning algorithms, for example, can develop a probability model by taking a decade of field data and providing insights about crop performance under various climatic conditions [15]. Data scientists utilize various machine learning algorithms to derive actionable insights from the available information. Another intriguing area of artificial intelligence is reinforcement learning [16]. These can be examined as an essential class of algorithms that can be utilized for streamlining logic for dynamic programming. Reinforcement learning is the preparation of machine learning models to make decision sequences [17]. The agent learns to accomplish an objective in an ambiguous, potentially complex environment. Based on the agent's action, the environment rewards it. This scenario depicts the machine as the agent and its surroundings as the environment.

In recent times advanced and progressive artificial intelligence technique named, deep reinforcement learning (DRL) is profound for intelligent decision making in various domains like energy management [18], robotics [19], health care [20], smart grid, game theory [21], [22], finance, computer vision [23], Natural Language Processing [24], Sentiment analysis [25] and so on with an extensive combination of reinforcement learning methods with deep learning models [26], [27]. This model has been efficient to resolve a wide extent of complicated decision-making tasks that were formerly beyond the bounds for the machine. As a result, it is a convincing model endorsed for developing intelligent agricultural frameworks. The characteristic models of deep reinforcement learning include deep successor network, multi-agent deep reinforcement learning and deep Q-network.

In this paper, we propose a supervised smart agriculture framework based on the deep reinforcement learning algorithm. A deep Q-Learning based DRL algorithm is used to strengthen the crop yield forecasting efficiency with the best rewarding iterations. There exist several other deep learning algorithms that may not be bounded by the biases or require huge manual effort in label creation deriving the insights

directly from the data like, Autoencoders [28], deep belief networks [29], Gaussian Bernoulli RBM's [30], Bayesian Neural Nets [31], Deep Generative models [32]. These models can sometimes fail to account for uncertainty while interpreting ambiguous inputs. Most of these approaches follow greedy procedures that are sub-optimal, learning a single layer of features at a time without updating its lower-level parameters resulting in slow and inefficient computations. The proposed work overcomes the above-mentioned shortcomings promoting the advancement of smart agriculture and thereby leading to increased food production. The rest of the paper is organized as follows. Section 2 presents the literature review of the existing works. Section 3 briefs about the Deep Q-learning algorithm and the proposed Deep Recurrent Q-Network (DRQN) model for forecasting the crop yield.. Section 4 explains the agriculture dataset and study area description. Section 5 presents the experimental results and frameworks, and the performance of the DRL model over the other machine learning algorithms. Section 6 wraps up with the conclusion and future works.

II. RELATED WORK

The potential growths in Artificial Intelligence have undoubtedly endless potential results [33], [34]. For creating new opportunities, deep learning has surged together with enormous data advancement [35]. This result in need of improved measures to envision, determine and assess data exhaustive strategies in agricultural frameworks [36], [37]. Crop yield prediction can be considered as a pattern recognition problem where AI has shown notable efficiency for agricultural applications [38]. Abrougui *et al.* have proposed yield prediction of potato crop the soil properties and tillage system by the ANN. The ANN model showed great potential to estimate yield [39]. Haghverdi *et al.* have defined the prediction of cotton lint from the phenology of crop indices using ANN. The ANN approach is used to generate 61200 models relating to individual crop indices to field estimates of cotton yield to be predicted [40]. Byakatonda *et al.* explained an ANN-based yield forecast for the maize crop based on the climatic indices and the precipitation length. In order to facilitate agricultural planning, yield predictions are made using ANN models [41]. In the approaches as mentioned earlier, the ANN's were used for the processing, which relied on feature extraction by time-domain and frequency-domain processing methods. This results in the drawback of manual feature extraction mainly depending on the prior knowledge of the data for predicting yield, and the ANN's shallow architecture in learning the complex non-linear relationships in the yield prediction system. With the advent of deep learning, such problems are handled to a certain extent.

Yang *et al.* have proposed a deep convolution neural network model to predict the crop yield estimation of the rice crop at the ripening stage. The CNN network learns the significant spatial features concerning the crop yield from the high spatial resolution RGB image [42]. Deep learning enabled the crop mapping strategy to identify the crop yield in

a respective region. Winter wheat mapping using the ground data statistical references employing the artificial neural networks and the deep CNN are modeled by Zhong *et al.* This enables automatic identification of wheat seasonality without using samples [43]. Ramesh *et al.*, proposed an optimized deep neural network algorithm recognize and classify crop yield based on the diseased leave images obtained by the image processing method [44]. Babak *et al.* computed a numerical deep learning model of crop growth by incorporating the DSSAT model's rainfall and irrigation inputs to predict maize yield [45]. An efficient automatic rice crop yield heading date estimation method through deep learning CNN network using time series RGB images of the crop [46] has been proposed by Desai *et al.* Koirala *et al.*, proposed a two-staged deep learning method using CNN for mango fruit yield estimation [47]. From the literature, the ANN-based process can be efficiently identified as a primary predictor, whereas deep learning approaches can recognize adaptive crop feature extraction by the hierarchical representation of DNN architecture. DNN architecture, however, needs a great deal of experience and prior knowledge which limits its generalization capability. Therefore it is essential to organize deep reinforcement learning (DRL) based smart architecture to examine crop yield prediction. In the DRL framework, deep learning provides the agent with the ability to sense the environment and reinforcement learning provides the ability to learn the best strategy for real-time problems [48]. DRL enables creating an agent that can generalize to an environment that is examined as meta-learning [49]. As a generic way of solving optimization problems through trial and error, DRL finds its application in several fields like agriculture [50], health care [51], energy management [52], robotic system [53] and game theory [54]. The following section provides a brief introduction to the Deep Q-Network DRL algorithm and the proposed methodology.

III. DEEP Q-NETWORK ALGORITHM BACKGROUND

Deep reinforcement learning has advanced together with enormous data growth and improved measure persistence to make new opportunities to determine, evaluate and acknowledge extensive data procedures for agricultural frameworks. Some of the essential factors that need to be analyzed in structuring the deep reinforcement learning models are:

- Understanding the patterns and basic structures from the restricted sample space.
- We are reviewing the objective functions with constant representations of events.
- The performance of the framework must be adequately viable to embrace consistently dynamic actions.

This section explains in detail the reinforcement learning, Q-learning and the deep Q-Network algorithm.

A. REINFORCEMENT LEARNING

Reinforcement learning (RL) is a framework in artificial intelligence with a dynamic programming concept that develops and trains algorithms utilizing a strategy of reward and

penalty. RL differs from other machine learning algorithms by the way that, it is not explicitly advised in performing a task, but it solves through the problem on its own [55]. For the RL study process, a Markov choice Procedure (MDP) is characterized that endorses the formalism where the reinforcement learning problems are embraced. The RL algorithm, which is an agent learns by collaborating and interacting with the environment. The agent will get rewards on the correct actions performed and penalties for the wrong actions. The agent learns by itself without human intervention by increasing its rewards and limiting its penalties. The process of reinforcement learning is presented in Fig.1.

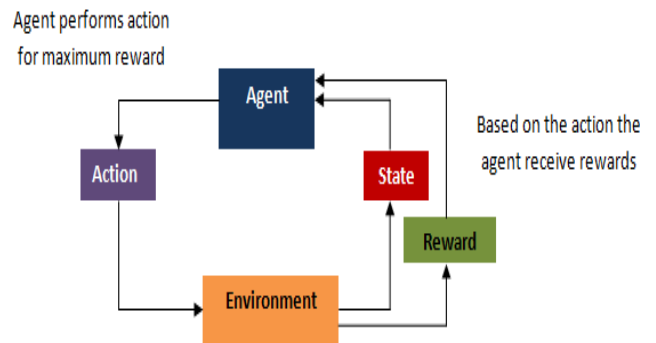


FIGURE 1. Reinforcement learning process.

An agent that is present in a state 's' performs an action 'a'. On Performing an action the agent attains a reward $R(s,a)$ and moves into a new state s' . The policy is a function that maps the states and the actions. In each state, a policy π is determined to specify the action to be carried out by an agent. In an agent's lifetime, its key objective is to identify an optimal policy π^* which magnifies the total discounted reward. The optimal policy π^* is defined in equation (1).

$$\pi^*(s) = \underset{a \in A}{\operatorname{argmax}} \gamma \sum_{s' \in S} P_{sa}(s', a) V^*(s', a) \quad (1)$$

A value function $V^\pi(s,a)$ [56] is defined for each state-action pair is an estimate of the expected reward following a policy π . The most optimal value function is attained from the best optimal policy, which is identified by the highest reward obtained by an agent from all the other states. This optimal value function is represented in equation (2).

$$V^*(s, a) = R(s, a) + \underset{a \in A}{\operatorname{max}} \gamma \sum_{s' \in S} P_{sa}(s', a) V^*(s', a) \quad (2)$$

Thus reinforcement learning agent learns from the environment through interactions. They maximize their rewards by determining the best bellman optimal policy and value function using dynamic programming functions.

B. Q-LEARNING

Q-Learning is a method that assesses which action to take by an agent, depending on an action-value function. It decides the value of being in a specific state and making a specific

action at that state. It is one of the most significant progress in reinforcement learning by the development of an off-policy temporal difference control algorithm. Q-Learning evaluates a state-action value function for a target policy that ascertains in choosing the action of maximum value. The function Q takes the input as the current state 's' and an action 'a' and returns an expected reward of that action in that state. In the initial steps before analyzing the environment, Q functions give the arbitrary fixed values. Later with better analysis, Q function provides a better approximation of the value function for the action 'a' in the state 's'. The Q function goes on updating providing the optimal value. The agent will perform a series of actions that will ultimately generate the total maximum reward.

C. DEEP Q-NETWORK

Deep Q-networks is an advanced reinforcement learning agent that uses a Deep Neural Networks (DNN) to map the connections among the states and the actions analogous to a Q-Table in Q-Learning. DNNs like Convolution Neural Networks (CNN), Recurrent Neural Networks (RNN) and sparse auto-encoder can directly learn the abstract representations of the raw data from the sensors. A DQN agent communicates with the environment through a series of observations, actions and rewards which is identical to the task of Q-Learning agent. Fig. 2 depicts the generic structure of deep Q-Network.

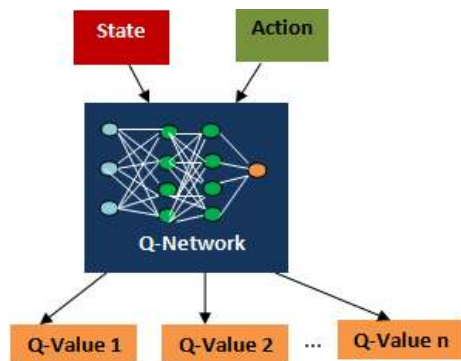


FIGURE 2. Structure of deep Q-Network.

The network takes a state as an input and for each action in the action space, the Q-Values are generated. The objective of the neural network is to learn and train the parameters. During the prediction process, this trained network is used to predict the next best action to occur in the environment. Basically, Q-Learning determines the state-action value function for a specific target policy that ultimately chooses an action of best value. It works fine for a restricted state and action space. However, for a huge set of action space may require millions of records to be stored in program memory. This results in the inflation of memory volume leading to curse of dimensionality or an unstable representation of a Q-Function.

The instability in Q-Learning arises due to the correlations existing in the series of observations. The relative small updates in the Q-value can result in the drastic change in

the policy of the agent and also the correlation between the target and Q-Value. These inadequacies are overcome in Deep Q-Network using two strategies, namely, experience replay and iterative updates. Iterative updates minimize the correlation between the target and the Q-values through consistently revising the Q-values towards the target values. While experience replay tends to solve the correlation problem by smoothing over the data distribution changes through data randomization. In the proposed work during the enhancement of the DQN agent, the experience replay randomly selects the experience from the memory and the Deep Q-network utilized is the RNN, which acts as a function approximation with weights θ . Hence the Q-Network can be prepared by revising the parameters θ_i in the 'i'th iteration by diminishing the mean squared error in the Bellman equation. The loss function, which is the squared difference between the Target Q and the Predicted Q is defined in equation (3) as follows:

$$Loss = (r + \gamma \max_a Q(s', a'; \theta') - Q(s, a; \theta))^2 \quad (3)$$

Gradient descent for the actual parameters can be performed in order to reduce this loss.

D. PROPOSED DEEP REINFORCEMENT LEARNING MODEL FOR CROP YIELD PREDICTION

Reinforcement learning is broadly designed in areas such as operations research, game theory, multi-agent systems, and control theory. In the proposed work, forecasting crop yield is studied as a regression problem that is resolved by supervised learning. This supervised learning-based crop yield prediction process needs to consider the crop yield data and its corresponding parameters as the inputs to determine the crop yield in the concerned region. In the RL based methods, the learning efficiency of the yield predicting agents is determined by the overall rewards. It results in unsteady feedback for the agents to adapt their performance along with the supervised learning methods. In other words, the agents will not be able to recognize from the inputs which samples are not efficiently learned during the learning process. Such a component enforces the agent to be more efficient by uncovering the deep characteristic contrasts among the crop yield.

In order to understand the yield forecast method based on DRL, a yield forecasting environment is designed based on the input parameters that converts the supervised learning process to the reinforcement learning process. The environment can be determined as a 'yield prediction game'. Every game incorporates certain parametric feature combinations and thresholds that aids in crop yield and each combination has a set of samples and its corresponding labels. When the agent starts playing, it determines the crop yield parameter values by performing the actions to attain the rewards. For every nearby predicted value of the target, the agent gets a positive reward, otherwise a negative reward. After completing the entire process, the agent will receive an aggregate score for its actions performed. This flow of yield prediction is presented in Fig. 3.

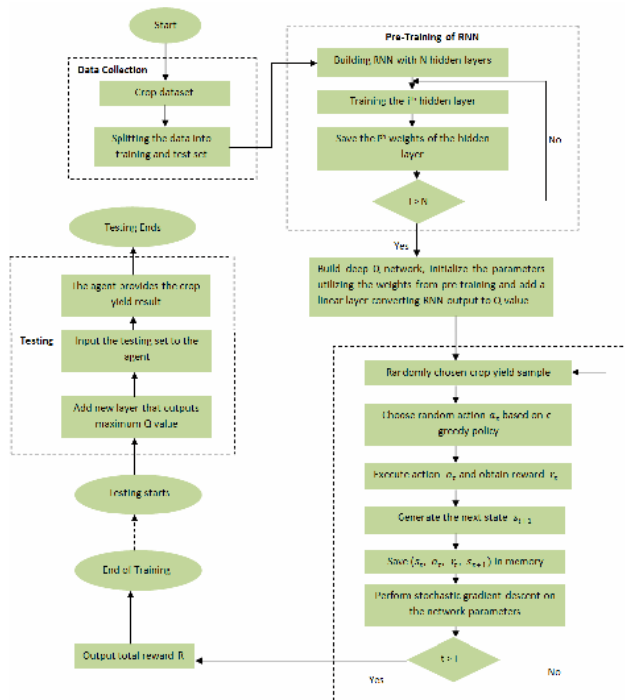


FIGURE 3. The flow diagram of the proposed deep recurrent Q-Network model for crop yield prediction.

For the actual reinforcement learning methods like the Q-learning, it is challenging to discriminate and analyze the crop yield prediction due to the restricted ability of those methods to describe the states. Inspired by the DQN concept of processing huge information, a Recurrent Neural network Based DNN is used in the proposed method to predict the crop yield using the various environmental, soil and groundwater parameters. It is termed as the Deep Recurrent Q-Learning model which is basically an RNN on top of the DQN. RNN can assist in mining temporal and semantic data and has advanced in time series analysis, language modeling and speech recognition. RNN is a variant of the ANN, where the current state input is connected to the output of the previous state. The definite explanation is that the network will recollect the previous information and apply it to the present network calculation.

In our proposed method the DQN agent is framed by stacking the RNN layers sequentially, initializing the parameters utilizing the weights saved in the RNN Pre-training process and adding a linear layer mapping the RNN output to Q-Values. Fig. 4 shows the structure of the RNN used in the DQN. x^t denotes the training data input at time t , H^t defines the hidden state at time t . Current input x^t and the previous hidden layer state H^{t-1} determines H^t . Here, O^t represents the output of the current layer at time t . The training data original output Y^t and the current output O^t determines the error L at time t . The weights shared across the RNN's are represented as u , v , and w .

F depicts the activation function of the hidden layers. The thresholds shared across the RNN's are defined as b_1 and b_2 .

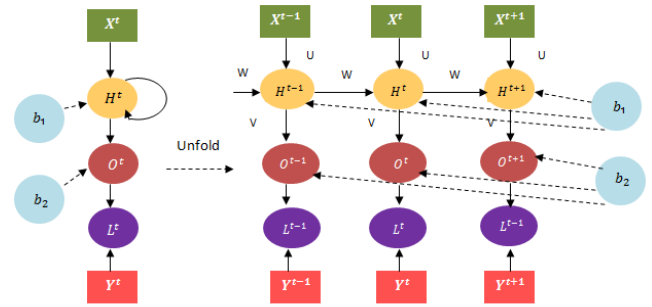


FIGURE 4. Propagation process of recurrent neural network.

The value of the hidden state at time 't' is given in equation (4) as follows:

$$H^t = f(u \times x^t + w \times H^{t-1} + b_1) \quad (4)$$

The predicted output O^t of the RNN at time 't' is given as follows in equation (5):

$$O^t = f(v \times H^t + b_2) \quad (5)$$

The error L of the RNN at time 't' is given as follows in equation (6):

$$L = O^t - y^t \quad (6)$$

The crucial aspects of the RNN, which can efficiently determine the crop yield, are the representation of the actual features self-learning layer after layer and the sparse constraint that limits the parameter space preventing over-fitting. The RNN in the proposed work consists of three hidden layers between the input layer and the output Q-value layer. For each RNN layer, a ReLU [57] activation function and L1 regularization [58], [59] is used. It results in penalizing the absolute values of the data parameters in the neural network when they are huge. Before the training process of the DRL, a pre-training process is applied to all the training data samples. Then the agent's yield prediction perception is built by stacking the input layers and the fully connected layer to output the final Q-values.

During the training process of the DRL framework, a huge set of state and action space is processed which can result in instability due to data correlations. Hence in the training process of the DQN, two alterations of the Q-Learning are made to ensure non-divergence of DRL's training process. The first is terms of experience replay, where the agent's experience is saved in the replay memory (D) by means of state, action and reward of the present time stamp and the state of the next timestamp. Say initially at each time step t , the experience replay saves the agent's experience resulting in a collection of specific sets of experiences. An individual experience e_t at a time t is described as $e_t = (s_t, a_t, r_t, s_{t+1})$ and the memory at time t is defined as D_t , where $D_t = \{e_1 \dots e_t\}$. Experience replay is an effective technique in eliminating the divergence in the parameters enabling the agents to recognize its experience in the learning process. The second alteration

of Q-Learning is to utilize an independent network for generating the targets during the Q-Learning update process. These alterations can substantially improve the DRL stability. Also, it is observed that usually RL algorithms iteratively update the action-value function using a Bellman equation. As this approach is tedious in practice the action-value function is estimated using an RNN function approximator with weight θ . Hence the Q-Network can be prepared by revising the parameters θ_i in the 'i'th iteration by diminishing the mean squared error in the Bellman equation.

The training process comprises two steps. The first step involves the pre-training of the RNN and the second step is the training of the DQN agent. The agent selects and executes an action based on a ϵ -greedy policy. Here the action is selected randomly with a probability ϵ , while the probability $1-\epsilon$ chooses the action representing the maximum q value. The optimization algorithm utilized in the proposed study is the stochastic gradient descent algorithm. The optimization algorithm updates the network weights iteratively based on the training data. The algorithm for training the RNN based Deep Recurrent Q-Network is defined in follows:

Algorithm Training of RNN Based DQN

Step 1: Pre-training of the RNN.

- (a) Initialize the replay memory capacity as N;
- (b) Initialize the RNN network with random weights θ_i .

For $i = 1, I$ **do**

- (c) Train the i th hidden layer.
- (d) Save the parameters of the i th hidden layer.

End For

(d) Initialize action-value network Q with the parameters of the hidden layer, other than the input and the output layer.

(e) Initialize the target action-value function Q' with the same parameters as Q.

Step 2: Training of the DQN agent

For event = I, M **do**

(a) Initialize the observation sequence s_1 by outputting the predicted yield randomly.

For $t = 1, T$ **do**

- (b) Select a random action a_t , with probability ϵ .
- (c) Perform the action a_t and obtain the reward r_t .
- (d) Randomly generate the next state $s_{(t+1)}$.
- (e) Save the memory D as $(s_t, a_t, r_t, s_{(t+1)})$.
- (f) With respect to the network parameters θ , perform gradient descent on $(r_t - Q(s_t, a_t; \theta))^2$
- (g) Reset $Q' = Q$.

End For

End For

The following section explains a brief definition of the various agrarian factors that influence the crop yield, and evaluation of various crop parameters to be considered for the construction of the learning models.

IV. DATASET AND STUDY AREA DESCRIPTION

Deep learning models demand huge data volume for efficient processing. Information with adaptable characteristics streamlines the effort of finding regularities by removing the irrelevant features for the learning objective. Fabricating a deep reinforcement learning model for the agricultural framework is highly tedious since they are extremely unsteady and possess a dynamic non-linear behavior. This section explains in detail the dataset used for the study for predicting the crop yield.

The proposed study investigates the yield prediction of paddy crop for the Vellore district in the southern part of India. Here, the block of district considered for the study include Ponnai, Arcot, Sholinghur, Ammur, Thimiri, and Kalavai. Paddy is one of the prevailing monetary crops cultivated in this region and hence this area is considered for investigation. In varying to the typical climatic and soil parameters, the dataset incorporates specific climate, soil and groundwater properties along with the volume of fertilizers consumed by the crops of the study area. Some of the parameters analyzed in the current study include evapotranspiration, ground frost frequency, groundwater nutrients, wet day frequency, aquifer characteristics which are not recognized together in the existing literature. Table 1 represents concise information about the various crop parameters utilized in the study. The data is taken for 35 years. The paddy crop yield is estimated in terms of area cultivated (in hectares), paddy production (in tons) and yield acquired (in kg/hectare). The knowledge pertinent to regular climatic factors like temperature, precipitation, reference crop evapotranspiration, potential evapotranspiration, humidity and distinctive climatic parameters like ground frost frequency, diurnal temperature range, and wind speed has been utilized. The climatic data are provided by the Indian Meteorological department from its portal metadata tool. The soil parameters comprise topsoil density, soil PH and the amount of the soil macronutrients (Nitrogen, Phosphorus and Potassium) present. Distinctive hydro-chemical properties of groundwater like transmissivity, aquifer type, permeability, electrical conductivity, pre-monsoon and post-monsoon micro-nutrients (calcium, potassium, sodium, magnesium, and chloride) content in groundwater are considered for the study.

The following section presents the experimental results obtained for predicting the crop yield using the DRQN model and comparison of the results with the existing models.

V. RESULTS AND DISCUSSION

The efficiency of a learning model is determined by evaluating the model various execution measures or by monitoring the performance by various evaluation metrics. For the proposed work the model is validated in terms of:

- Performance estimation
- Comparison of various other algorithms in terms of:
 - Evaluation metrics
 - Data distribution properties
 - Model accuracy measures

TABLE 1. List of dataset parameters for the proposed DRL framework.

S.No	Parameters Name	Description	Normal Range	Units
1	QK2O	Quantity of Potassium fertilizers used	15-20	Integer (Kilogram / Hectare)
2	Quantity Rice	Total production of Rice in the study area	2.37 - 2.5	Integer (Ton / Hectare)
3	QNitro	Quantity of nitrogen fertilizers used	15-20	Integer (Kilogram / Hectare)
4	QP2O5	Quantity of Phosphorus fertilizers used	2-3	Integer (Kilogram / Hectare)
5	Vapour Pressure	The pressure exerted by water vapor in thermodynamic equilibrium with its condensed phase (i.e) solid or liquid form at a given temperature	23.8 - 41.2	Integer (Hectopascal)
6	Gross Cropped Area	Total area planted to crops during all growing seasons of the year	195 - 220	Integer (Hectare)
7	Net Irrigated Area	Total geographic area which has received irrigation throughout the year	80 - 110	Integer (Hectare)
8	Ground Frost Frequency	Number of days that refers to the condition when the temperature of the upper layer of the soil falls below the freezing point of water	5-7	Integer (Number of Days)
9	Diurnal Temperature range	Difference between the daily maximum and minimum temperature	90 - 130	Integer (°C)
10	Net Cropped Area	Total geographic area on which the crop has been planted at least once during a year	175 - 200	Integer (Hectare)
11	Precipitation	Rain or water vapor condensation from the atmosphere	1400 - 1800	Integer (Millimetre / year)
12	Gross Irrigated Area	Total area under crops that have received irrigation during all the growing seasons of the year.	90 - 116	Integer (Hectare)
13	Average temperature	The average air temperature recorded in a particular location	21 - 25	Integer (°C)
14	Wet Day Frequency	The number of days in which an amount of 0.2 mm or more of rain is recorded in an observed period of 24 hrs.	45 - 55	Integer (Number of Days)
15	Area Rice	Total area planted for Rice Crop	35 - 40	Integer (Hectare)
16	Potential Evapotranspiration	Amount of evaporation that would occur in an area if a sufficient water source were available	27 - 35	Integer (Millimetre / Day)
17	Reference Crop Evapotranspiration	The evapotranspiration rate from a crop reference surface that is not short of water	25 - 30	Integer (Millimetre / Day)
18	Maximum Temperature	The highest air temperature recorded in a particular location	21 - 37	Integer (°C)
19	Humidity	The quantity representing the amount of water vapor in the atmosphere	60 - 80	Integer (Percentage)

TABLE 1. (Continued.) List of dataset parameters for the proposed DRL framework.

20	Wind Speed	The rate at which the air moves in a particular area	40 - 50	Integer (Miles / Hour)
21	Minimum temperature	The lowest air temperature recorded in a particular location	16 - 20	Integer (°C)
22	K Soil	The amount of potassium present in the soil of the study area	≥ 42	Integer (ton / Hectare)
23	N Soil	The amount of nitrogen present in the soil of the study area	≥ 48	Integer (Kilogram / Hectare)
24	P Soil	The amount of phosphorus present in the soil of the study area	≥ 30	Integer (Kilogram / Hectare)
24	Aquifer Area Percentage	Percentage of area covered by aquifer. Aquifer is a body of permeable rock which can contain or transmit groundwater.	55 - 60	Integer (Percentage)
25	Aquifer Permeability	A measure of the rock property which determines how easily water and other fluids can flow through it. Permeability depends on the extent to which pores are interconnected.	25 - 30	Integer (Meter / Day)
26	Pre-Electrical Conductivity	Average Pre-monsoon Electrical conductivity of groundwater that determines the ability or power to conduct or transmit heat or electricity	55 - 60	Integer (Siemens / Meter)
27	Post-Electrical Conductivity	Average Post-monsoon Electrical conductivity of groundwater that determines the ability or power to conduct or transmit heat or electricity	65 - 70	Integer (Siemens / Meter)
28	Groundwater Pre-Magnesium	Average Pre-monsoon Magnesium level in groundwater	68 - 73	Integer (Milligram / Litres)
29	Groundwater Post-Magnesium	Average Post-monsoon Magnesium level in ground water	60 - 65	Integer (Milligram / Litres)
30	Groundwater Pre-Sodium	Average Pre-monsoon Sodium level in ground water	150-170	Integer (Milligram / Litres)
31	Groundwater Post-Sodium	Average Post-monsoon Sodium level in ground water	190-200	Integer (Milligram / Litres)
32	Groundwater Pre-Potassium	Average Pre-monsoon Potassium level in ground water	15-20	Integer (Milligram / Litres)
33	Groundwater Post-Potassium	Average Post-monsoon Potassium level in ground water	20 - 25	Integer (Milligram / Litres)
34	Groundwater Pre-Chloride	Average Pre-monsoon Chloride level in ground water	320 - 325	Integer (Milligram / Litres)
35	Groundwater Post-Chloride	Average Post-monsoon Chloride level in ground water	330 - 340	Integer (Milligram / Litres)
36	Yield Rice	Total amount of rice yield obtained	2.0 - 2.5	Integer (Ton)
37	Soil PH	A measure of the acidity and alkalinity in the soil.	6-7	Integer
38	Topsoil Depth	The outermost layer of the soil that is rich in organic matter and microorganisms	0.5 - 0.75	Integer (Meters)

A. PERFORMANCE ESTIMATION

During the construction of machine learning models, the dataset is arbitrarily split into training and test set, where the highest amount of data is taken as the training set. Even though the test dataset is small, there exist chances of leaving out some important information that may have enhanced the model. Also, there is a concern of high variance in the dataset. To handle this issue, K-fold cross-validation is utilized. It is a strategy that is utilized to assess the deep learning models by re-sampling the training information for enhancing the performance. Modeling and forecasting time series data are intricate and challenging. Randomly splitting a time series data for cross-validation does not hold well. It may lead to a temporal dependency problem as there is an implicit reliance on past observation and simultaneously, a leakage from the response variable to lag variables is bound to happen. This results in non-stationarity, which is the frequent changes in mean and variance in the information space. In such cases, cross-validation is performed in a forward-chaining manner.

For the proposed approach, five-fold forward chaining cross-validation is performed, which more precisely models the data prediction where the model is built on past data and predicts the forward-looking data. The results are tabulated in Table 2.

TABLE 2. Results of 5- fold forward chaining cross-validation for the deep reinforcement model M.

Model 'M' subset	Training Data (Years considered)	Test Data (Years Considered)	Correlation Measure	R ² Score
1	1996 - 2000	2001 – 2003	0.73	0.855
2	1996 - 2003	2004 – 2006	0.88	0.873
3	1996 - 2006	2007 – 2009	0.64	0.833
4	1996 - 2009	2010 – 2012	0.79	0.892
5	1996 - 2012	2013 - 2016	0.89	0.922

It is like starting with a small subset of data initially for training, predict for the following data and determining the exactness of the predicted data. The same forecasted data points are enclosed as a part of the next training data subset and the following data points are forecasted. Fig. 5 represents the deep reinforcement model plot before and after performing the cross-validation.

The cross-validation is performed using the python Scikit-Learn machine learning library. The dataset is preprocessed using the min-max scaling which normalized the dataset. From the Sklearn using the sub-library model_selection, the train_test_split function is imported to split the training and test sets. The hyperparameter tuning for the cross-validation in determining the best 'K' is obtained using the cross_val_score function of the Sklearn library. The data is split into 'K' subsets, in this case, it is five by setting the parameter n_splits as 5. The training and validation data set size is determined by the parameter test_size which is taken as 0.3 for the proposed work, indicating that 70% of the

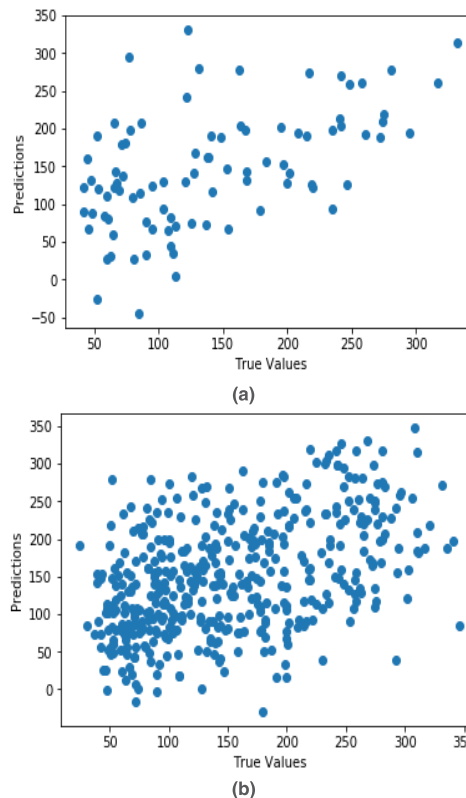


FIGURE 5. Data predictions of the deep reinforcement forecasting model against true values (a) Before cross-validation, (b) After cross-validation.

data is subjected to training and 30% of the data for testing. The model is trained throughout the K-fold forward chaining cross-validation process and the error metric is determined. The error metric is the r2 score, which is appended in every iteration and obtain the best value defining the overall model accuracy.

B. COMPARISION WITH OTHER MODELS

The proposed deep reinforcement model DRQN is explored and tested with other significant algorithms, namely Deep LSTM network, Artificial Neural networks (ANN), gradient boosting (GB), random forest (RF) and other deep learning based algorithms like Bernoulli Deep Belief Network (BDN), Bayesian Artificial Neural Networks (BAN), Rough Auto Encoders (RAE) and Interval Deep Generative Artificial Neural Networks (IDANN). An inducing aspect for the evaluation metric is their ability to discriminate between the results of different learning models.

1) EVALUATION METRICS

The proposed model's reasonability is determined by evaluating the model against different execution measures or specific evaluation metrics. The various evaluation metrics considered for the proposed work are the Determination Coefficient (R2), Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Median Absolute Error (MedAE), Mean Squared Logarithmic Error (MSLE), Mean

Absolute Percentage Error (MAPE) and Explained Variance Score (Exp. Var.). To assure a fair examination of the model error metrics, two sets of these four models for training and validation were constructed to forecast yield. The hyperparameter optimization for the proposed approach and the other models are observed through a manual selection approach for the respective models. The key objective of the manual hyperparameter selection is to tune the model’s capacity to match the target task complexity. The hyperparameters like the learning rate, the number of hidden units, optimizer, activation function and the dropout values are determined on the degree on which the training process and cost function reduce the test error. The DQRN based DRL model is constructed using an RNN network of one input layer, three hidden layers with each layer consisting of 8 neurons, a fully connected layer and an output layer presenting the crop yield value. The input layer consists of 30 neurons representing the crop dataset parameters. The RNN uses a ReLU activation function for the processing in the hidden layers. To attain the best performance accuracy without over-fitting, the agent learns by performing an action through 1000 epochs.

In the construction of the interval deep generative artificial neural networks [60] and the rough auto encoders [61], a rough set theory is introduced to the deep learning algorithm to deal with data ambiguity. The rough set theory is a mathematical function brought in by Pawlak [62] to handle uncertainty in learning. It is a proper theory obtained from the intrinsic research on logical characteristics of information systems. An information system S is identified as 4 tuple $S = \langle U, V, A, f \rangle$. Here U is the universe of primitive objects, A is the set of attributes, V is the domain set such that $V = \bigcup_{a \in A} V_a$. The mapping f is termed as the information or total function $f : U \times A \rightarrow V$. In concerning any attribute set $A_t \subseteq A$ and concept set such as $X \subseteq U$ rough set defines two approximations:

- $\overline{A_t X}$ represents the set of attributes in U which can possibly be members of X with respect to the attributes of A_t .
- $\underline{A_t X}$ represents set of all attributes in U that can be exactly identified as members of X with respect to the attributes of A_t .

The boundary region, $B = \overline{A_t X} - \underline{A_t X}$ defines the set of attributes that can’t certainly be identified to X only by considering set of attributes of A_t . If $B(X) = \emptyset$ then X is the crisp set with respect to A_t else it is defined as the rough set.

For the experimented rough autoencoder model a rough neuron based on the rough set theory is incorporated into the two layered stacked autoencoder model. The rough neurons are applied in the output and the hidden layers of the RAE model. The rough neuron used in this approach consists of an upper bound neuron $U = (w_1, b_1, \alpha)$ and a lower bound neuron $L = (w_2, b_2, \beta)$. w_1, b_1, w_2, b_2 are the weights and biases of the upper bound and lower bound neurons. The output coefficients α, β defines the contribution of the upper bound and lower bound outputs O_1, O_2 to the neuron’s output O . Beginning from the first layer of RAE, the rough

TABLE 3. Performance evaluation metrics of the proposed deep reinforcement model and other machine learning models.

Model	Training Dataset				Validation Dataset			
	MAE	MSE	RMSE	R ²	MAE	MSE	RMSE	R ²
Deep Reinforcement Learning	0.19	0.02	0.14	0.8	0.13	0.03	0.17	0.87
Deep Learning	0.19	0.07	0.25	0.61	0.18	0.05	0.22	0.68
Artificial Neural Network	0.28	0.09	0.35	0.68	0.22	0.09	0.37	0.62
Random forest	0.6	0.63	0.79	0.22	0.62	0.51	0.72	0.37
Gradient Boosting	0.48	0.31	0.64	0.49	0.41	0.28	0.53	0.54

TABLE 4. Performance evaluation metrics of the proposed deep reinforcement model and other machine learning models.

Model	Training Dataset			Validation Dataset		
	MSLE	Exp. Var	Med AE	MSLE	Exp. Var	Med AE
Deep Reinforcement Learning	0.002	0.83	0.19	0.002	0.875	0.11
Deep Learning	0.006	0.69	0.27	0.004	0.731	0.16
Artificial Neural Network	0.008	0.53	0.22	0.001	0.595	0.17
Random forest	0.074	0.29	0.47	0.06	0.221	0.52
Gradient Boosting	0.039	0.3	0.39	0.03	0.324	0.33

auto encoders are trained progressively using back propagation with stochastic gradient descent to determine the rough features for crop yield.

For the experimented IDANN, variational autoencoders with the rough set theory is incorporated to extract the data features. The variational autoencoder is a framework consisting of both encoder and decoder that is trained to reduce the reconstructed error between the generated and the actual data. The features are learned by means of stochastic generation of mean and standard deviation of the input samples. The initialization process is a regularization task; the randomly initialized parameters are moved to better latent space. The features are learned by maximizing the probabilities of the generative model-variational autoencoder to initialize the biases and weights of the multi-layered neural network. Naturally the mean vector oversees where the input encoding

TABLE 5. Performance evaluation metrics of other unsupervised deep machine learning models.

Evaluation metric	Model Training Dataset				Model Validation Dataset			
	BDN	BAN	ID ANN	RAE	BDB	BAN	ID ANN	RAE
MAE	0.23	0.27	0.28	0.25	0.21	0.23	0.26	0.22
MSE	0.07	0.079	0.09	0.11	0.05	0.07	0.07	0.09
RMSE	0.17	0.25	0.29	0.32	0.17	0.21	0.24	0.27
R ²	0.80	0.71	0.60	0.54	0.80	0.74	0.65	0.59
MSLE	0.007	0.005	0.008	0.01	0.004	0.004	0.006	0.01
Exp. Var	0.75	0.79	0.692	0.64	0.784	0.82	0.731	0.67
MedAE	0.19	0.157	0.192	0.21	0.13	0.137	0.15	0.19

TABLE 6. Accuracy and MAPE measure of the proposed deep reinforcement model and other machine learning models.

Model	Accuracy Measure (%)	MAPE (%)
DRL	93.7	17
BDN	92.1	20
BAN	91.7	27
IDANN	91	29
RAE	90.7	32
DL	91.85	28
ANN	90.5	38
RF	70.7	53
GB	81.2	41

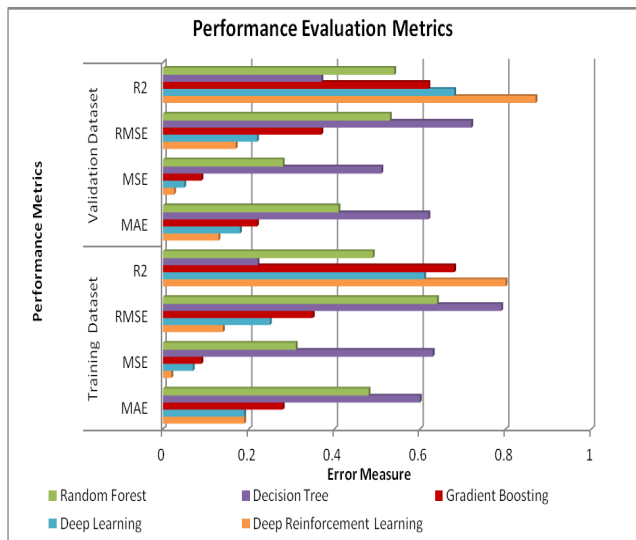


FIGURE 6. Model performance evaluation results using MAE, MSE, RMSE and R2 for the experimented models.

should be centered, and the standard deviation controls how much from the mean the encodings can vary. As the encodings are generated at random, the decoder understands that the sample refers to not a single point in latent space, but rather all nearby points.

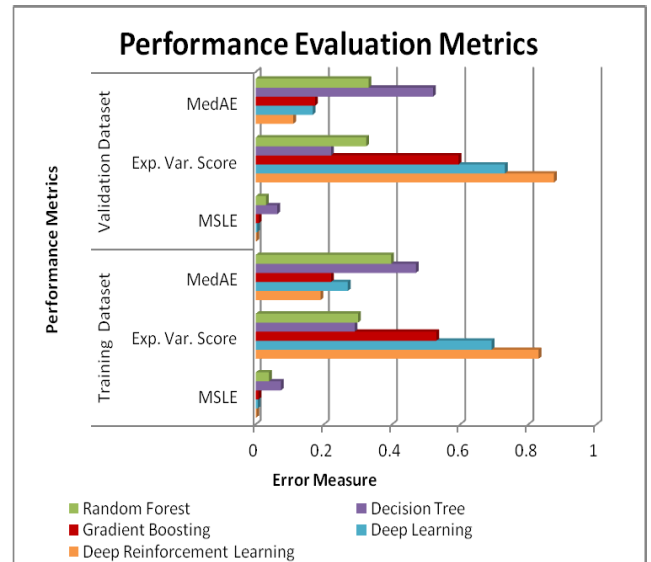
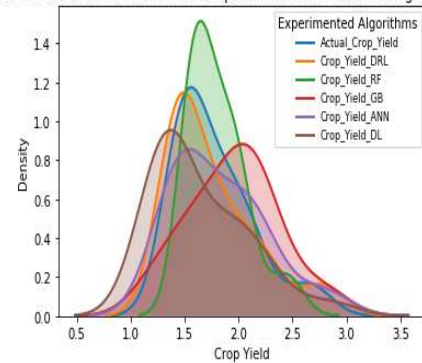


FIGURE 7. Model performance evaluation results using MSLE, Exp. Var. Score and MedAE for the experimented models.

Comparison of actual data distribution vs predicted observation using various algorithms



Comparison of actual data distribution vs predicted observation using various algorithms

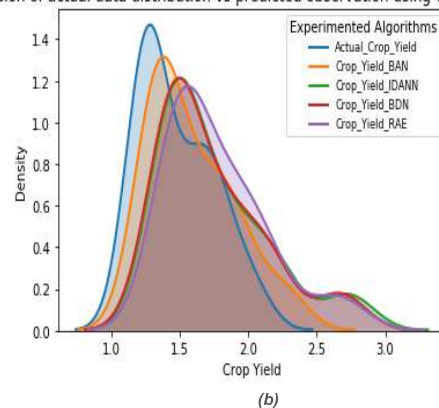


FIGURE 8. Comparison of actual data distribution with Predicted data distribution using experimented algorithms (a): Deep reinforcement learning, Random forest, Gradient boosting, artificial neural networks, Deep learning. (b): Bayesian artificial neural networks, interval deep generative artificial neural network, Bernoulli deep belief network, Rough autoencoders.

Tables 3 to 6 compare the performance of the machine learning models on both training and validation datasets to the metrics as mentioned earlier of evaluation.

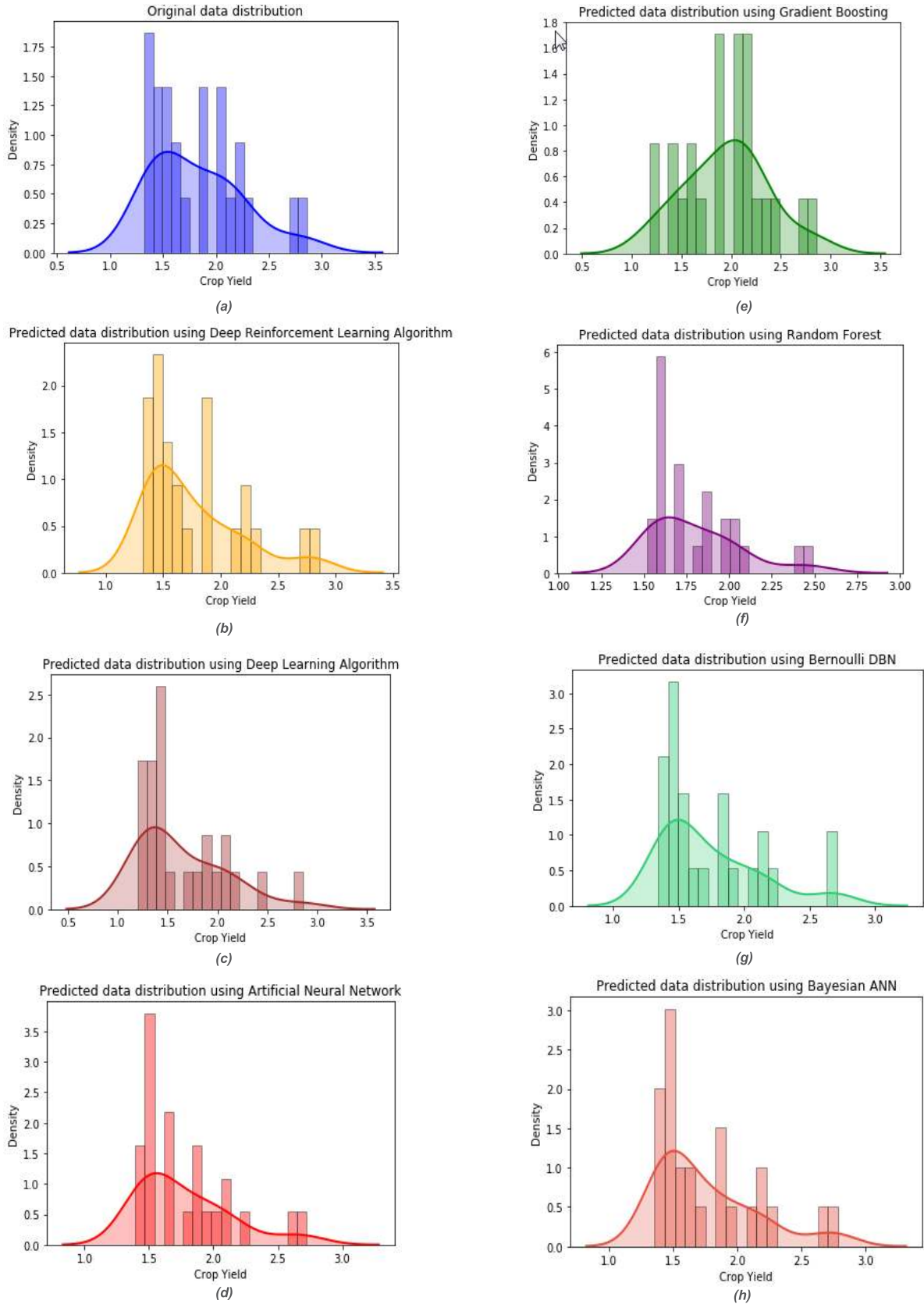


FIGURE 9. Probability density functions of the: (a) Original data and Predicted data using: (b) Deep reinforcement learning, (c) deep learning (d) Artificial neural network, (e) Gradient boosting, (f) Random forest, (g) Bernoulli DBN, (h) Bayesian ANN, (i) Interval deep generative ANN, (j) Rough autoencoder.

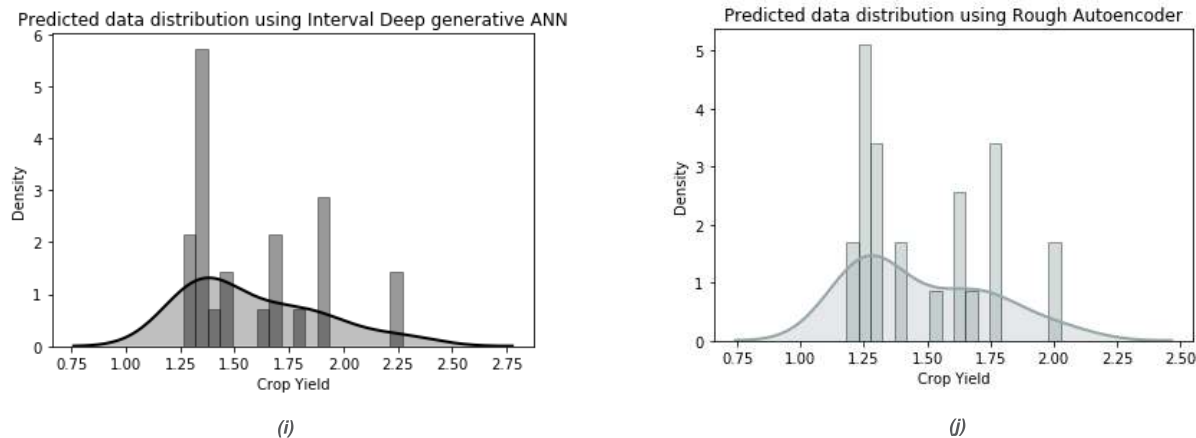


FIGURE 9. (Continued.) Probability density functions of the: (a) Original data and Predicted data using: (b) Deep reinforcement learning, (c) deep learning (d) Artificial neural network, (e) Gradient boosting, (f) Random forest, (g) Bernoulli DBN, (h) Bayesian ANN, (i) Interval deep generative ANN, (j) Rough autoencoder.

These models were enforced and implemented in python in the most effective aspect and tested under similar software and hardware conditions to assure reasonable comparisons. The error metric is used to define the performance degree during the execution of a model. The residuals obtained during the experiments, which are the difference between the actual and the predicted values are used to estimate the error measure. In other words, by observing the magnitude of the residual spread, the precision, as well as the efficiency of the model, is determined.

In terms of precision and efficiency, the proposed deep reinforcement model is observed to outperform the other machine learning models with an accuracy of 93.7% and improved error measures.

However, the performance of other deep learning models BDN, BAN, IDANN, RAE and Deep LSTN is reasonably close to the DRL approach. Fig. 6 and Fig. 7 explain the evaluated performance measures of the experimented models for the crop yield prediction.

2) DATA DISTRIBUTION PROPERTIES

In order to determine if the proposed DRQN model preserved the original distributional properties of the data, the probability density function of the actual crop yield data and the experimented models are observed. The Probability Density function (PDF) is an analytical expression that characterizes probability distribution for a continuous random variable against a discrete random variable. In graphically defining the PDF, the region under the curve will represent the interval where the predicted variable falls. The absolute area in the graph interval equates the probability of the continuous random variable occurrence. It enables us to calculate the probabilities of the range of outcomes.

The probability density functions of the actual crop yield and the predicted crop yield using the proposed deep reinforcement learning and the other machine learning algorithms are shown in Fig. 8. It is done to observe if the proposed model

and the other ML algorithms can preserve the distributional properties of the actual crop yield data.

Fig.9 defines the individual probability density plots of the original data, proposed DRL method and other experimented ML algorithms.

From Fig. 9, it is explicitly defined that the proposed deep reinforcement learning model can more approximately preserve the distribution properties of the actual crop yield data when compared to the other experimented machine learning algorithms.

3) MODEL ACCURACY MEASURES

Evaluation of the model accuracy is an integral part of the model development process. It enables in identifying the optimum model for the data representation and performance of the model for the future timestamps.

Accuracy refers to the ratio of predictions which the model has forecasted precisely. Accuracy reflects the closeness of the predicted value to the actual value or the true value. Fig.10 graphically represents the accuracy measure of the predicted data using the proposed deep reinforcement learning algorithm and the other experimented machine learning algorithms.

On observing the experimental values and results obtained for the paddy crop dataset, the deep reinforcement learning model is found to predict the data with better accuracy and precision of 93.7% over the other experimented algorithms.

Though the accuracy measures of other deep learning algorithms like BDN, BAN, IDANN, RAE and Deep LSTN are reasonably close to the proposed approach the computational cost and time complexity is more than the proposed model. The BDN and IDANN are asserted to be more suitable for predicting continuous data enabling greedy layer-by-layer learning efficiently by evaluating the parameters quickly. A critical disadvantage is that the approximation process is restricted to an individual bottom-up pass and the existing greedy process is very slow and inefficient.

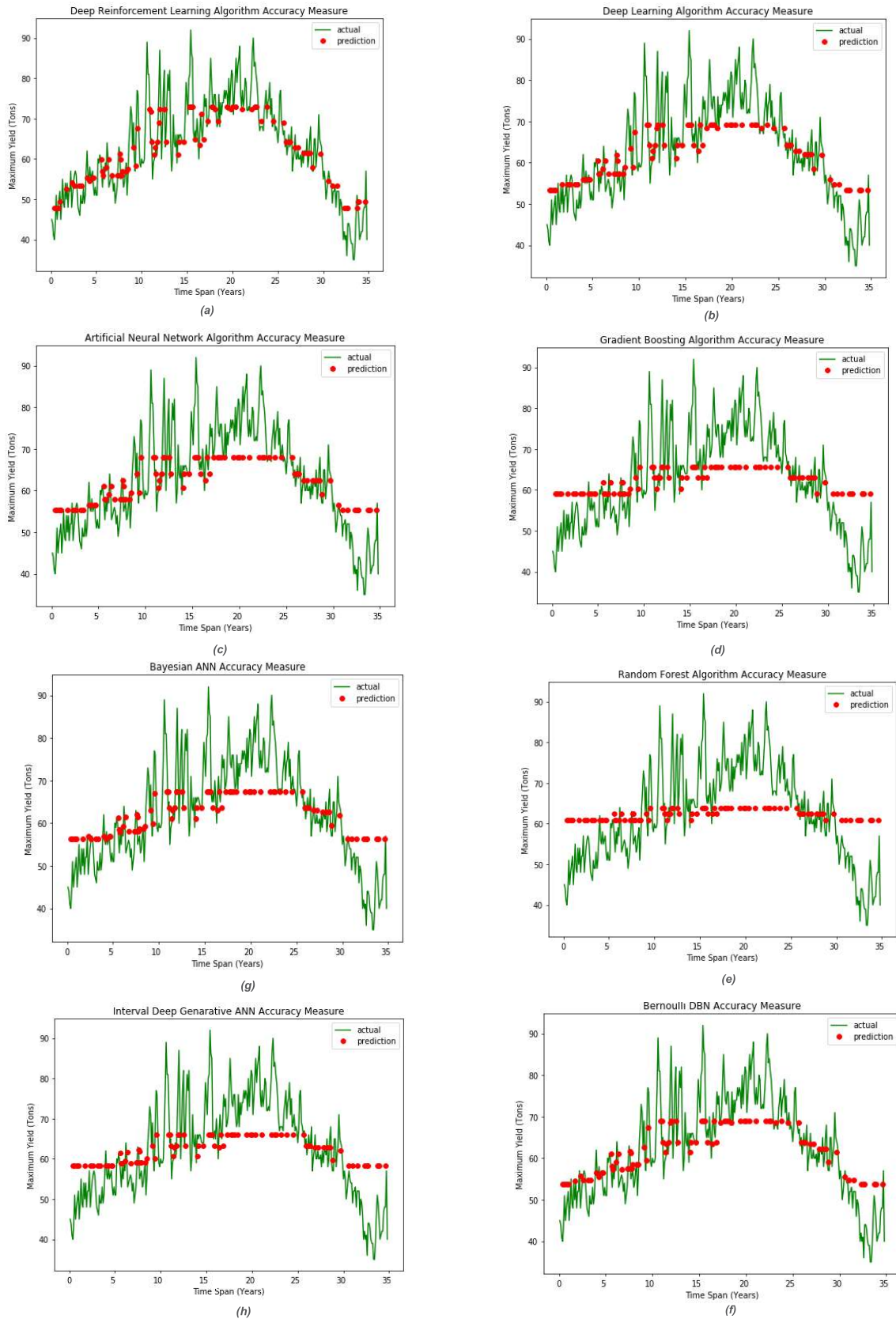


FIGURE 10. Prediction accuracy measure of: (a) Proposed DRL algorithm, (b) Deep learning algorithm, (c) Artificial neural network algorithm, (d) Gradient boosting algorithm and (e) Random forest algorithm, (f) Bernoulli DBN, (g) Bayesian ANN, (h) Interval deep generative ANN, (i) Rough autoencoders.

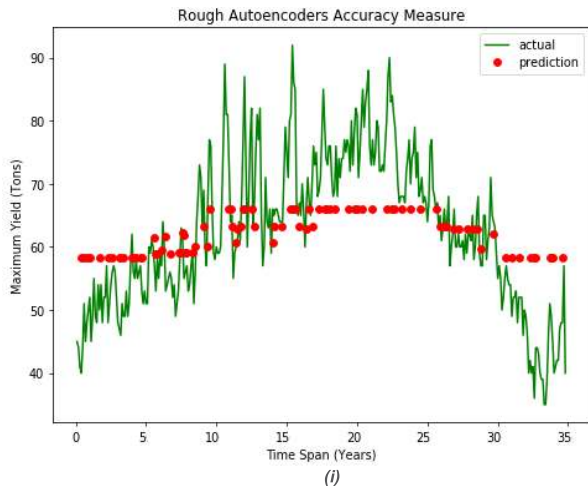


FIGURE 10. (Continued.) Prediction accuracy measure of: (a) Proposed DRL algorithm, (b) Deep learning algorithm, (c) Artificial neural network algorithm, (d) Gradient boosting algorithm and (e) Random forest algorithm, (f) Bernoulli DBN, (g) Bayesian ANN, (h) Interval deep generative ANN, (i) Rough autoencoders.

RAE automatically learns from the data samples, which is an essential feature, it is simple to train specific examples of the algorithm that will perform well on a particular kind of information. It doesn't require any new designing, just relevant training data. But the auto encoders' decompressed outcomes will be degraded on comparison to the actual inputs deviating from lossless arithmetic compression. Also in generalizing the model requires a large amount of training data. Though the RAE supports application in greedy layer-wise approach pertaining for deep networks, better random weight initialization schemes, batch normalization and residual learning could provide sufficient training for deep networks. BAN exposes few powerful insights and techniques to deep learning by automatically estimating errors associated with predictions but however, they are difficult to scale for large datasets. This is even evident by comparing the MAPE value obtained from the proposed approach and the BAN model.

In terms of error measures evaluation, the DRQN presented the lowest error values and almost preserved the original data distribution. It is evident from the results obtained that the proposed deep reinforcement learning DQRN model can solve the crop yield prediction problem by learning from the various dataset parameters through memory replaying and self-learning. Thus the predominance of the proposed method additionally enhances the system intelligence to predict the yield by minimizing the dependence on expert experience.

VI. CONCLUSION AND FUTURE WORKS

The evolution of DRL has raised the self-reliance and the intelligence of the Artificial Intelligence algorithms and motivates to propose a novel crop yield prediction system. The results observed from the precision and efficiency tests illustrate the effectiveness and versatility of the proposed Deep Recurrent Q-Network for yield prediction. By building a yield prediction environment, the proposed method makes

it feasible for the agent to identify and learn the crop yield prediction through self-exploration and experience replay. Through the dataset prediction results, it is evident that the yield prediction agent administers the process, suggesting that the proposed method can precisely define the characteristics for crop yield. The combination of RNN based feature processing and DQN based self experimental analysis is the key objective to attain favorable results. Unlike the supervised learning-based crop yield prediction process, DRQN based process provides a complete solution that independently mines the non-linear mapping between the crop yield and the climatic, soil and groundwater parameters. This advantage can definitely minimize expert dependency and prior knowledge for developing crop yield prediction models. Hence the proposed approach provides a perception of implementing a more generalized model for yield prediction. However, the RNN based DRL can cause the gradients to explode or disappear if the time series is very much longer. Experimenting data prediction through a wide range of ML predictive algorithms can be observed as a basis for decision making, but it is critical to interpret the statistical uncertainty related to these predictions. Hence there exist needs to design a framework that predicts both target and their prediction's uncertainty. Probabilistic predictive modeling strategies like information theory, probabilistic bias-variance decomposition, composite prediction strategies, probabilistic boosting and bagging approaches etc. can be considered to handle the uncertainty in statistical predictions that can be observed as a future extension of the current model. Another alternative approach to be considered is to use an LSTM based DRL. Exploration of more crop yield prediction parameters with respect to pest and infestations and crop damage can be included in the current framework to construct a more robust working model in the future. Further improvement in the computing efficiency of the training process is an intriguing option to be concentrated.

ACKNOWLEDGMENT

The authors would like to thank the India water portal for providing the meteorological data relevant to climatic factors from their MET data tool. The MET data tool provides district wise monthly and the annual mean of each metrological indicator values. They would also like to thank the Joint Director of Agriculture, Vellore, Tamil Nadu, India, for providing the details regarding the soil and groundwater properties for the respective village blocks.

REFERENCES

- [1] S. Li, S. Peng, W. Chen, and X. Lu, "INCOME: Practical land monitoring in precision agriculture with sensor networks," *Comput. Commun.*, vol. 36, no. 4, pp. 459–467, Feb. 2013.
- [2] A. D. Jones, F. M. Nguere, G. Peltó, and S. L. Young, "What are we assessing when we measure food security? A compendium and review of current metrics," *Adv. Nutrition*, vol. 4, no. 5, pp. 481–505, 2013.
- [3] G. E. O. Ogotu, W. H. P. Franssen, I. Supit, P. Omondi, and R. W. Hutjes, "Probabilistic maize yield prediction over East Africa using dynamic ensemble seasonal climate forecasts," *Agricult. Forest Meteorol.*, vols. 250–251, pp. 243–261, Mar. 2018.

- [4] M. E. Holzman, F. Carmona, R. Rivas, and R. Niclòs, "Early assessment of crop yield from remotely sensed water stress and solar radiation data," *ISPRS J. Photogramm. Remote Sens.*, vol. 145, pp. 297–308, Nov. 2018.
- [5] A. Singh, B. Ganapathysubramanian, A. K. Singh, and S. Sarkar, "Machine learning for high-throughput stress phenotyping in plants," *Trends Plant Sci.*, vol. 21, no. 2, pp. 110–124, 2016.
- [6] R. Whetton, Y. Zhao, S. Shaddad, and A. M. Mouazen, "Nonlinear parametric modelling to study how soil properties affect crop yields and NDVI," *Comput. Electron. Agricult.*, vol. 138, pp. 127–136, Jun. 2017.
- [7] Y. Dash, S. K. Mishra, and B. K. Panigrahi, "Rainfall prediction for the Kerala state of India using artificial intelligence approaches," *Comput. Elect. Eng.*, vol. 70, pp. 66–73, Aug. 2018.
- [8] W. Wieder, S. Shoop, L. Barna, T. Franz, and C. Finkenbinder, "Comparison of soil strength measurements of agricultural soils in Nebraska," *J. Terramech.*, vol. 77, pp. 31–48, Jun. 2018.
- [9] Y. Cai, K. Guan, J. Peng, S. Wang, C. Seifert, B. Wardlow, and Z. Li, "A high-performance and in-season classification system of field-level crop types using time-series Landsat data and a machine learning approach," *Remote Sens. Environ.*, vol. 210, pp. 35–47, Jun. 2018.
- [10] X. E. Pantazi, D. Moshou, T. Alexandridis, R. L. Whetton, and A. M. Mouazen, "Wheat yield prediction using machine learning and advanced sensing techniques," *Comput. Electron. Agricult.*, vol. 121, pp. 57–65, Feb. 2016.
- [11] T. U. Rehman, S. Mahmud, Y. K. Chang, J. Jin, and J. Shin, "Current and future applications of statistical machine learning algorithms for agricultural machine vision systems," *Comput. Electron. Agricult.*, vol. 156, pp. 585–605, Jan. 2019.
- [12] D. Elavarasan, D. R. Vincent, V. Sharma, A. Y. Zomaya, and K. Srinivasan, "Forecasting yield by integrating agrarian factors and machine learning models: A survey," *Comput. Electron. Agricult.*, vol. 155, pp. 257–282, Dec. 2018.
- [13] M. D. Johnson, W. W. Hsieh, A. J. Cannon, A. Davidson, and F. Bédard, "Crop yield forecasting on the Canadian Prairies by remotely sensed vegetation indices and machine learning methods," *Agricult. Forest Meteorol.*, vols. 218–219, pp. 74–84, Mar. 2016.
- [14] A. Kaya, A. S. Kececi, C. Catal, H. Y. Yalic, H. Temucin, and B. Tekinerdogan, "Analysis of transfer learning for deep neural network based plant classification models," *Comput. Electron. Agricult.*, vol. 158, pp. 20–29, Mar. 2019.
- [15] A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey," *Comput. Electron. Agricult.*, vol. 147, pp. 70–90, Apr. 2018.
- [16] I. M. Evans, "Reinforcement, principle," in *International Encyclopedia of the Social & Behavioral Sciences*, J. D. Wright, 2nd ed. Amsterdam, The Netherlands: Elsevier, 2015, pp. 207–210.
- [17] D. Vogiatzis and A. Stafylopatis, "Reinforcement learning for rule extraction from a labeled dataset," *Cognit. Syst. Res.*, vol. 3, no. 2, pp. 237–253, Jun. 2002.
- [18] D. A. Temesgene, M. Miozzo, and P. Dini, "Dynamic control of functional splits for energy harvesting virtual small cells: A distributed reinforcement learning approach," *Comput. Commun.*, vol. 148, pp. 48–61, Dec. 2019.
- [19] S. Wan, Z. Gu, and Q. Ni, "Cognitive computing and wireless communications on the edge for healthcare service robots," *Comput. Commun.*, vol. 149, pp. 99–106, Jan. 2020.
- [20] A. Tolba, O. Said, and Z. Al-Makhadmeh, "MDS: Multi-level decision system for patient behavior analysis based on wearable device information," *Comput. Commun.*, vol. 147, pp. 180–187, Nov. 2019.
- [21] V. Hassija, V. Saxena, and V. Chamola, "Scheduling drone charging for multi-drone network based on consensus time-stamp and game theory," *Comput. Commun.*, vol. 149, pp. 51–61, Jan. 2020.
- [22] S. Gheisari and E. Tahavori, "CCCLA: A cognitive approach for congestion control in Internet of Things using a game of learning automata," *Comput. Commun.*, vol. 147, pp. 40–49, Nov. 2019.
- [23] S. J. Shri and S. Jothilakshmi, "Crowd video event classification using convolutional neural network," *Comput. Commun.*, vol. 147, pp. 35–39, Nov. 2019.
- [24] M. Al-Ayyoub, A. Nuseir, K. Alsmearat, Y. Jararweh, and B. Gupta, "Deep learning for arabic NLP: A survey," *J. Comput. Sci.*, vol. 26, pp. 522–531, May 2018.
- [25] M. Usama, B. Ahmad, J. Yang, S. Qamar, P. Ahmad, Y. Zhang, J. Lv, and J. Guna, "Equipping recurrent neural network with CNN-style attention mechanisms for sentiment analysis of network reviews," *Comput. Commun.*, vol. 148, pp. 98–106, Dec. 2019.
- [26] Z. Liu, C. Yao, H. Yu, and T. Wu, "Deep reinforcement learning with its application for lung cancer detection in medical Internet of Things," *Future Gener. Comput. Syst.*, vol. 97, pp. 1–9, Aug. 2019.
- [27] H. Huang, M. Lin, and Q. Zhang, "Double-Q learning-based DVFS for multi-core real-time systems," in *Proc. IEEE Int. Conf. Internet Things (iThings) IEEE Green Comput. Commun. (GreenCom) IEEE Cyber, Phys. Social Comput. (CPSCom) IEEE Smart Data (SmartData)*, Jun. 2017, pp. 522–529.
- [28] H. Jahangir, M. A. Golkar, F. Alhameli, A. Mazouz, A. Ahmadian, and A. Elkamel, "Short-term wind speed forecasting framework based on stacked denoising auto-encoders with rough ANN," *Sustain. Energy Technol. Assessments*, vol. 38, Apr. 2020, Art. no. 100601.
- [29] K. Peng, R. Jiao, J. Dong, and Y. Pi, "A deep belief network based health indicator construction and remaining useful life prediction using improved particle filter," *Neurocomputing*, vol. 361, pp. 19–28, Oct. 2019.
- [30] J. Zhang, H. Wang, J. Chu, S. Huang, T. Li, and Q. Zhao, "Improved Gaussian-Bernoulli restricted Boltzmann machine for learning discriminative representations," *Knowl.-Based Syst.*, vol. 185, Dec. 2019, Art. no. 104911.
- [31] M. M. Rahman, D. Hagare, and B. Maheshwari, "Bayesian belief network analysis of soil salinity in a peri-urban agricultural field irrigated with recycled water," *Agricult. Water Manage.*, vol. 176, pp. 280–296, Oct. 2016.
- [32] F. Du, J. Zhang, J. Hu, and R. Fei, "Discriminative multi-modal deep generative models," *Knowl.-Based Syst.*, vol. 173, pp. 74–82, Jun. 2019.
- [33] A. V. Samsonovich, "Socially emotional brain-inspired cognitive architecture framework for artificial intelligence," *Cognit. Syst. Res.*, vol. 60, pp. 57–76, May 2020.
- [34] K. Ryan, P. Agrawal, and S. Franklin, "The pattern theory of self in artificial general intelligence: A theoretical framework for modeling self in biologically inspired cognitive architectures," *Cognit. Syst. Res.*, to be published.
- [35] R. Wason, "Deep learning: Evolution and expansion," *Cognit. Syst. Res.*, vol. 52, pp. 701–708, Dec. 2018.
- [36] X. Zhu, M. Zhu, and H. Ren, "Method of plant leaf recognition based on improved deep convolutional neural network," *Cognit. Syst. Res.*, vol. 52, pp. 223–233, Dec. 2018.
- [37] S. Zhang, W. Huang, and C. Zhang, "Three-channel convolutional neural networks for vegetable leaf disease recognition," *Cognit. Syst. Res.*, vol. 53, pp. 31–41, Jan. 2019.
- [38] P. Nevavuori, N. Narra, and T. Lipping, "Crop yield prediction with deep convolutional neural networks," *Comput. Electron. Agricult.*, vol. 163, Aug. 2019, Art. no. 104859.
- [39] K. Abrougui, K. Gabsi, B. Mercatoris, C. Khemis, R. Amami, and S. Chehaibi, "Prediction of organic potato yield using tillage systems and soil properties by artificial neural network (ANN) and multiple linear regressions (MLR)," *Soil Tillage Res.*, vol. 190, pp. 202–208, Jul. 2019.
- [40] A. Haghverdi, R. A. Washington-Allen, and B. G. Leib, "Prediction of cotton lint yield from phenology of crop indices using artificial neural networks," *Comput. Electron. Agricult.*, vol. 152, pp. 186–197, Sep. 2018.
- [41] J. Byakatonda, B. P. Parida, P. K. Kenabatho, and D. B. Moalafhi, "Influence of climate variability and length of rainy season on crop yields in semiarid Botswana," *Agricult. Forest Meteorol.*, vol. 248, pp. 130–144, Jan. 2018.
- [42] Q. Yang, L. Shi, J. Han, Y. Zha, and P. Zhu, "Deep convolutional neural networks for rice grain yield estimation at the ripening stage using UAV-based remotely sensed images," *Field Crops Res.*, vol. 235, pp. 142–153, Apr. 2019.
- [43] L. Zhong, L. Hu, H. Zhou, and X. Tao, "Deep learning based winter wheat mapping using statistical data as ground references in Kansas and northern Texas, US," *Remote Sens. Environ.*, vol. 233, Nov. 2019, Art. no. 111411.
- [44] S. Ramesh and D. Vydeki, "Recognition and classification of paddy leaf diseases using optimized deep neural network with Jaya algorithm," *Inf. Process. Agricult.*, to be published.
- [45] B. Saravi, A. P. Nejadhashemi, and B. Tang, "Quantitative model of irrigation effect on maize yield by deep neural network," *Neural Comput. Appl.*, to be published.
- [46] S. V. Desai, V. N. Balasubramanian, T. Fukatsu, S. Ninomiya, and W. Guo, "Automatic estimation of heading date of paddy rice using deep learning," *Plant Methods*, vol. 15, no. 1, p. 76, Dec. 2019.
- [47] A. Koirala, K. B. Walsh, and W. Z. McCarthy, "Deep learning for real-time fruit detection and orchard fruit load estimation: Benchmarking of 'MangoYOLO,'" *Precis. Agricult.*, vol. 20, no. 6, pp. 1107–1135, 2019.

- [48] K. Arulkumaran, M. P. Deisenroth, M. Brundage, and A. A. Bharath, "Deep reinforcement learning: A brief survey," *IEEE Signal Process. Mag.*, vol. 34, no. 6, pp. 26–38, Nov. 2017.
- [49] E. A. O. Diallo, A. Sugiyama, and T. Sugawara, "Coordinated behavior of cooperative agents using deep reinforcement learning," *Neurocomputing*, to be published.
- [50] F. Bu and X. Wang, "A smart agriculture IoT system based on deep reinforcement learning," *Future Gener. Comput. Syst.*, vol. 99, pp. 500–507, Oct. 2019.
- [51] Z. Liu, C. Yao, H. Yu, and T. Wu, "Deep reinforcement learning with its application for lung cancer detection in medical Internet of Things," *Future Gener. Comput. Syst.*, vol. 97, pp. 1–9, Aug. 2019.
- [52] Y. Liu, X. Guan, J. Li, D. Sun, T. Ohtsuki, M. M. Hassan, and A. Alelaiwi, "Evaluating smart grid renewable energy accommodation capability with uncertain generation using deep reinforcement learning," *Future Gener. Comput. Syst.*, to be published.
- [53] A. Plasencia, Y. Shichkina, I. Suárez, and Z. Ruiz, "Open source robotic simulators platforms for teaching deep reinforcement learning algorithms," *Procedia Comput. Sci.*, vol. 150, pp. 162–170, 2019.
- [54] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, and D. Hassabis, "Mastering the game of Go with deep neural networks and tree search," *Nature*, vol. 529, no. 7587, pp. 484–489, 2016.
- [55] M. A. Khan, S. Peters, D. Sahinel, F. D. Pozo-Pardo, and X. T. Dang, "Understanding autonomic network management: A look into the past, a solution for the future," *Computer Commun.*, vol. 122, pp. 93–117, Jun. 2018.
- [56] M. L. Littman, "Value-function reinforcement learning in Markov games," *Cognit. Syst. Res.*, vol. 2, no. 1, pp. 55–66, Apr. 2001.
- [57] P. Petersen and F. Voigtlaender, "Optimal approximation of piecewise smooth functions using deep ReLU neural networks," *Neural Netw.*, vol. 108, pp. 296–330, Dec. 2018.
- [58] X. Qian, H. Huang, X. Chen, and T. Huang, "Efficient construction of sparse radial basis function neural networks using L_1 -regularization," *Neural Netw.*, vol. 94, pp. 239–254, Oct. 2017.
- [59] X. Fan, X. Li, and J. Zhang, "Compressed sensing based loss tomography using weighted ℓ_1 minimization," *Comput. Commun.*, vol. 127, pp. 122–130, Sep. 2018.
- [60] M. Khodayar, J. Wang, and M. Manthouri, "Interval deep generative neural network for wind speed forecasting," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 3974–3989, Jul. 2019.
- [61] M. Khodayar, O. Kaynak, and M. E. Khodayar, "Rough deep neural architecture for short-term wind speed forecasting," *IEEE Trans. Ind. Informat.*, vol. 13, no. 6, pp. 2770–2779, Dec. 2017.
- [62] Z. Pawlak, "Rough sets," *Int. J. Comput. Inf. Sci.*, vol. 11, no. 5, pp. 341–356, 1982.



DHIVYA ELAVARASAN received the bachelor's degree in information technology from Anna University, India, and the master's degree in computer and communication technology from the Vellore Institute of Technology (VIT), India, where she is currently pursuing the Ph.D. degree in machine learning and analytics for agrarian frameworks with the School of Information Science and Technology. She worked as a Data Science Application Developer with IBM India Pvt., Ltd., for two years, and a Data Analyst with NetApp India Pvt., Ltd., for one year. She is currently doing her research for crop yield prediction system by developing machine learning models for an optimal solution to agricultural problems involving machine learning and deep learning technologies.



P. M. DURAIRAJ VINCENT (Member, IEEE) received the B.E. degree in electronics and the M.E. degree in computer science from Anna University, Chennai, India, and the Ph.D. degree from the Vellore Institute of Technology (VIT), Vellore, in 2015. He has more than 13 years of teaching and research experience. He has excellent industrial connectivity and also handled sessions for Wipro employees. He is currently working as an Associate Professor with the School of Information Technology and Engineering, VIT. He is a motivated researcher with more than 50 publications in Scopus Database. His current research interests include security, machine learning, the Internet of Things, and data analytics. He also acts as a doctoral committee member of a few universities. He has delivered a keynote in a few reputed conferences and also delivered invited talks in leading institutes, including Anna University. He has organized a few conferences, including one virtual conference. He is a Reviewer of a few reputed journals including the IEEE Access.

• • •