International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

COPY RIGHT

ELSEVIER
SSRN

2023IJIEMR.Personal use of this material is permitted. Permission from IJIEMR must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works. No Reprint should be done to this paper, all copy right is authenticated to Paper Authors

IJIEMR Transactions, online available on 07th Jan 2023. Link :http://www.ijiemr.org/downloads.php?vol=Volume-12&issue=ISSUE-1

DOI: 10.48047/IJIEMR/V12/ISSUE 01/46

Title CROP YIELD PREDECTION AND RECOMMENDATION

Volume 12, Issue 1, Pages: 514-523

Paper Authors
Mrs. M.Pratibha, K.Kavya, V.Prathyusha, O.Priya, P.Udaya Sri

USE THIS BARCODE TO ACCESS YOUR ONLINE PAPER

To Secure Your Paper As Per UGC Guidelines We Are Providing A Electronic Bar Code

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

CROP YIELD PREDECTION AND RECOMMENDATION

Mrs. M.Pratibha, Assistant Professor, Dept. of Information Technology, Sridevi Women's Engineering College, Hyd. madhunala.p@gmail.com

K.Kavya, B.Tech., Dept. of Information Technology, Sridevi Women's Engineering College, Hyd.

V.Prathyusha, B.Tech., Dept. of Information Technology, Sridevi Women's Engineering College, Hyd.

O.Priya, B.Tech., Dept. of Information Technology, Sridevi Women's Engineering College, Hyd.

P.Udaya Sri, B.Tech., Dept. of Information Technology, Sridevi Women's Engineering College, Hyd.

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%.

Keywords *– RNN, LSTM, Deep Q Network, RF and XGBoost Classifiers.*

1. 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

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

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. A country's policymaker depends on precise forecast, to make appropriate export and import assessments to reinforce national food security. Cultivators 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. On the other hand, crop yield forecasting is exceedingly challenging because of various complex aspects.

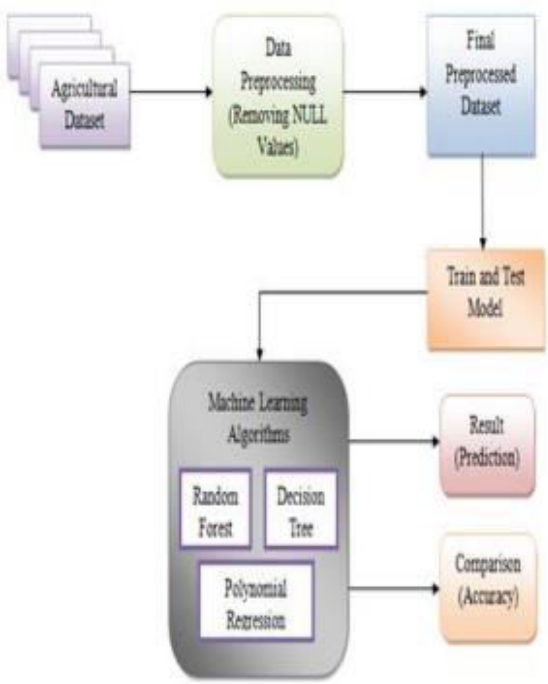

Fig.1: Example figure

Crop yield mainly depends upon climatic conditions, soil quality, landscapes, pest infestations, water quality and availability, genotype, planning of harvest activity and so on. The crop yield processes and strategies vary with time and they are profoundly non-linear in nature, and intricate due to the integration of a wide extent of correlated factors 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. 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. 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.

2. LITERATURE REVIEW

Probabilistic maize yield prediction over East Africa using dynamic ensemble seasonal climate forecasts

We tested the usefulness of seasonal climate predictions for impacts prediction in eastern Africa. In regions where these seasonal predictions showed skill we tested if the skill also translated into maize yield forecasting skills. Using European Centre for Medium-Range Weather Forecasts (ECMWF) system-4 ensemble seasonal climate hindcasts for the period 1981–2010 at different initialization dates before sowing, we generated a 15-member ensemble of yield predictions using the World Food Studies (WOFOST) crop model implemented for water-limited maize production and single season simulation. Maize yield predictions are validated against reference yield simulations using the WATCH Forcing Data ERA-Interim (WFDEI), focussing on the dominant sowing dates in the northern region (July), equatorial region (March-April) and in the southern region (December). These reference yields show good anomaly correlations compared to the official FAO and national reported statistics, but the average reference yield values are lower than those reported in Kenya and Ethiopia, but slightly higher in Tanzania. We use the ensemble mean, interannual variability, mean errors, Ranked Probability Skill Score (RPSS) and Relative Operating Curve skill Score (ROCSS) to assess regions of useful probabilistic prediction. Annual yield anomalies are predictable 2-months before sowing in most of the regions. Difference in interannual variability between the reference and predicted yields range from ±40%, but higher interannual variability in predicted yield dominates. Anomaly correlations between the reference and predicted yields are largely positive and range from +0.3 to +0.6. The ROCSS illustrate good pre-season probabilistic prediction of above-normal and below-normal yields with at least 2-months lead time. From the sample sowing dates considered, we concluded that, there is potential to use dynamical seasonal climate forecasts with a process based crop simulation model WOFOST to predict anomalous water-limited maize yields.

Deep reinforcement learning with its application for lung cancer detection in medical Internet of Things

Recently, deep reinforcement learning has achieved great success by integrating deep learning models into reinforcement learning algorithms in various applications such as computer games and robots. Specially, it is promising for computer-aided diagnosis and treatment to combine deep reinforcement learning with medical big data generated and collected from medical Internet of Things. In this paper, we focus on the potential of the deep reinforcement learning for lung cancer detection as many people are suffering from the lung tumor and about 1.8 million patients died from lung cancer in 2018. Early detection and diagnosis of lung tumor can significantly improve the treatment effect and prolong survival. In this work, we present several representative deep reinforcement learning models that are potential to use for lung cancer detection. Furthermore, we summarize the common types of lung cancer and the main characteristics of each type. Finally, we point out the open challenges and possible future research directions of applying deep reinforcement learning to lung cancer detection, which is expected to promote the evolution of smart medicine with medical Internet of Things.

Early assessment of crop yield from remotely sensed water stress and solar radiation data

Soil moisture (SM) available for evapotranspiration is crucial for food security, given the significant interannual yield variability of rainfed crops in large agricultural regions. Also, incoming solar radiation (Rs) influences the photosynthetic rate of vegetated surfaces and can affect productivity. The aim of this work is to evaluate the ability of crop water stress and Rs remotely sensed data to forecast yield at regional scale. Temperature Vegetation Dryness Index (TVDI) was computed as an indicator of crop water stress and soil moisture availability. TVDI during critical growth stage of crops was calculated from MODIS products: MODIS/AQUA 8-day composite LST at 1 km and 16-day composite vegetation index at 1 km. Rs data were obtained from Clouds and the Earth´s Radiant Energy System (CERES). The relationship between TVDI, Rs and yield of wheat, corn and soybean was analyzed. High R2 values (0.55-0.82, depending on crop and region) were found in different agro-climatic regions of Argentine Pampas. Validation results showed the suitability of the model: RMSE=330 kg ha-1-1300 kg ha-1, Relative Error= 13-34%. However, results were significantly improved considering the most important factor affecting yield. Rs proved to be important for winter crops in humid areas, where incoming radiation can be a limiting factor. In semi-arid regions, soils with low water retention capacity and summer crops, crop water stress showed the best results. Overall, results reflected that the proposed approach is suitable for crop yield forecasting at regional scale several weeks previous to harvest.

Machine learning for high-throughput stress phenotyping in plants

Advances in automated and high-throughput imaging technologies have resulted in a deluge of high-resolution images and sensor data of plants. However, extracting patterns and features from this large corpus of data requires the use of machine learning (ML) tools to enable data assimilation and feature identification for stress phenotyping. Four stages of the decision cycle in plant stress phenotyping and plant breeding activities where different ML approaches can be deployed are (i) identification, (ii) classification, (iii) quantification, and (iv) prediction (ICQP). We provide here a comprehensive overview and user-friendly taxonomy of ML tools to enable the plant community to correctly and easily

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

apply the appropriate ML tools and best-practice guidelines for various biotic and abiotic stress traits.

Rainfall prediction for the Kerala state of India using artificial intelligence approaches

Three artificial intelligence approaches - K-nearest neighbor (KNN), artificial neural network (ANN), and extreme learning machine (ELM) - are used for the seasonal forecasting of summer monsoon (June-September) and post-monsoon (October-December) rainfall from 2011 to 2016 for the Kerala state of India and performance of these techniques are evaluated against observations. All the aforesaid techniques have performed reasonably well and in comparison, ELM technique has shown better performance with minimal mean absolute percentage error scores for summer monsoon (3.075) and post-monsoon (3.149) respectively than KNN and ANN techniques. The prediction accuracy is highly influenced by the number of hidden nodes in the hidden layer and more accurate results are provided by the ELM architecture (8-15-1). This study reveals that the proposed artificial intelligence approaches have the potential of predicting both summer monsoon and post-monsoon of the Kerala state of India with minimal prediction error scores.

3. METHODOLOGY

In Existing System, 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.

Disadvantages:

- 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.

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

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

performed by minimizing the error and maximizing the forecast accuracy.

Advantages:

- The proposed model efficiently predicts the crop yield outperforming existing models by preserving the original data distribution with an accuracy of 93.7%.

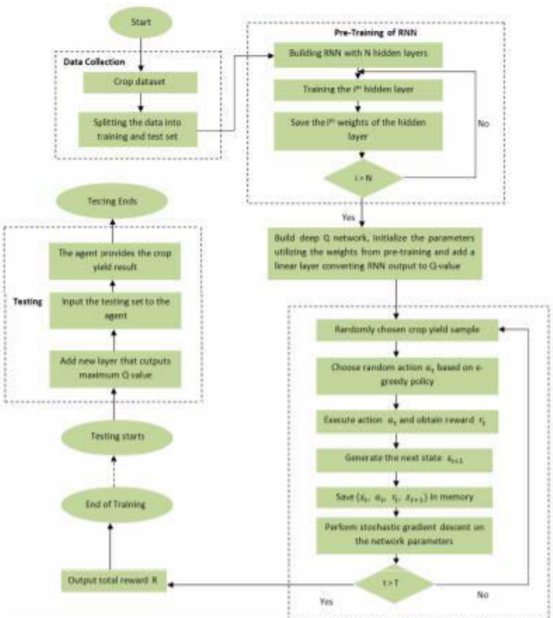

Fig.2: System architecture

MODULES:

To implement aforementioned project we have designed following modules

- Data exploration: using this module we will load data into system
- Processing: Using the module we will read data for processing
- Splitting data into train & test: using this module data will be divided into train & test
- Model generation: Build RNN,LSTM,Deep Q Network,RFAnd XGBoost ClassifiersAlgorithms accuracy calculated.
- User signup & login: Using this module will get registration and login
- User input: Using this module will give input for prediction
- Prediction: final predicted displayed

4. IMPLEMENTATION

Here in this project we are used the following algorithms

RNN:

A recurrent neural network is a type of artificial neural network commonly used in speech recognition and natural language processing. Recurrent neural networks recognize data's sequential characteristics and use patterns to predict the next likely scenario.

LSTM:

LSTM stands for long short-term memory networks, used in the field of Deep Learning. It is a variety of recurrent neural networks (RNNs) that are capable of learning long-term dependencies, especially in sequence prediction problems.

Deep Q Network:

Reinforcement Learning can broadly be separated into two groups: model free and model based RL algorithms. Model free RL algorithms don't learn a model of their environment's transition function to make predictions of future states and rewards. Q-Learning, Deep Q-Networks, and Policy Gradient methods are model-free algorithms because they don't create a model of the environment's transition function.

RF:

Random Forest is a powerful and versatile supervised machine learning algorithm that grows and combines multiple decision trees to create a "forest." It can be used for both classification and regression problems in R and Python.

XGBoost Classifiers:

XGBoost is an optimized distributed gradient boosting library designed to be highly efficient, flexible and portable. It implements machine learning algorithms under the Gradient Boosting framework.

5. EXPERIMENTAL RESULTS

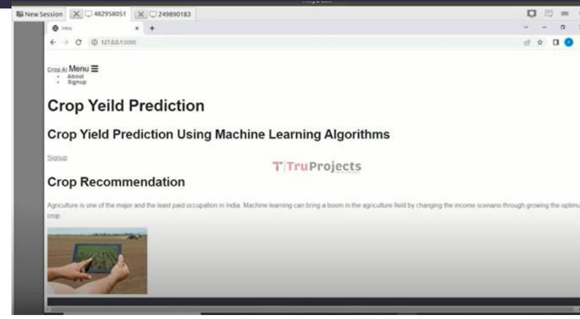

Fig.3: Home screen

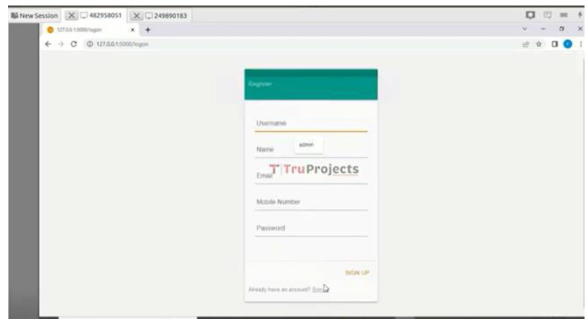

Fig.4: Registration

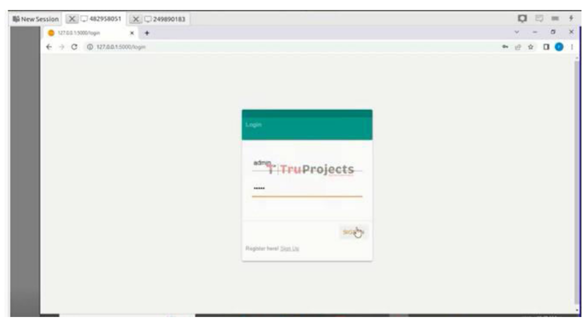

Fig.5: Login

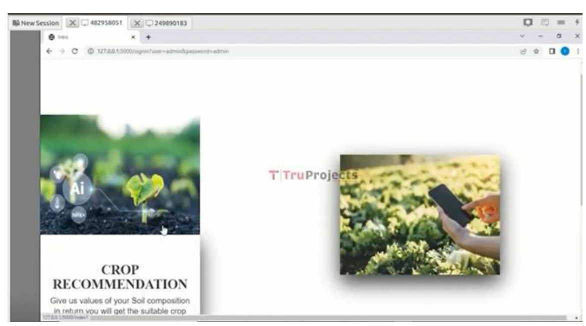

Fig.6: Crop recommendation home screen

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

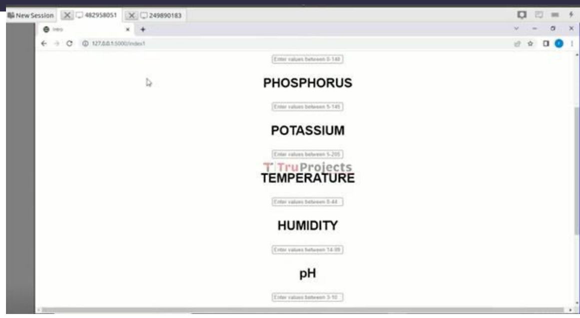

Fig.7: User input

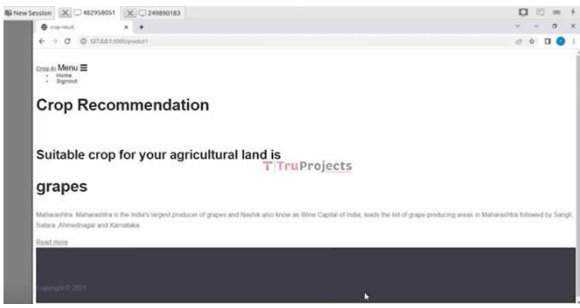

Fig.8: Prediction result

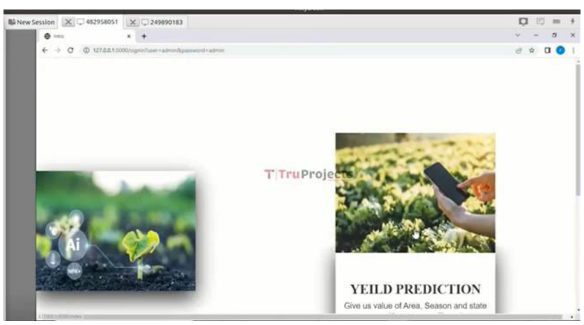

Fig.9: Yield prediction home screen

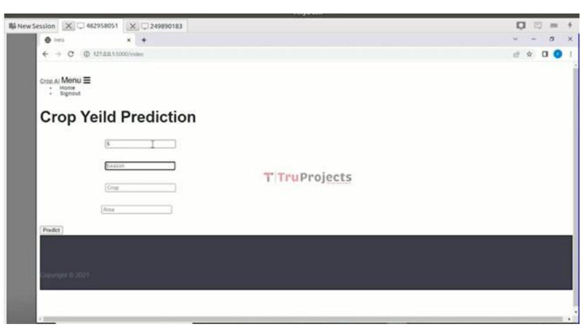

Fig.10: User input

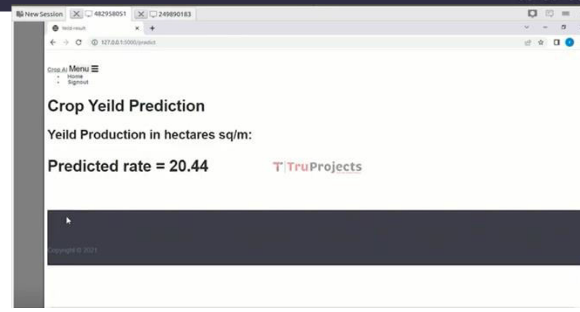

Fig.11: Prediction result

6. CONCLUSION

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.

7. FUTURE SCOPE

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.

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. Ngure, G. Pelto, 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. Ogutu, 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. Finkenbiner, ''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. Keceli, 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.