Educational Administration: Theory and Practice
2024,30(5), 1413-1420
ISSN:2148-2403
https://kuey.net/

Research Article

Applications Of Big Data Analytics In Cloud Computing And Smart Farming

P. Roopmathi[1*], Dr. J. Chockalingam[2], Dr.A.Shaik Abdul Khadir[3]

[1*]Research Scholar, Dept of Computer Science, Khadir Mohideen College, Adirampattinam, (Affiliated to Bharathidasan University, Trichy-24), Thanjavur, Tamilnadu, India.
[2]Research Supervisor, Associate Professor of Computer Science (Retd), Khadir Mohideen College, Adirampattinam, (Affiliated to Bharathidasan University, Trichy-24), Thanjavur, Tamilnadu, India
[3]Research Co-Supervisor, Head & Associate Professor of Computer Science, Khadir Mohideen College, Adirampattinam, (Affiliated to Bharathidasan University, Trichy-24), Thanjavur, Tamilnadu, India

Citation: P. Roopmathi et.al (2024), Applications Of Big Data Analytics In Cloud Computing And Smart Farming, *Educational Administration: Theory And Practice*, *30*(5), 1413-1420 Doi: 10.53555/kuey.v30i5.3101

ARTICLE INFO

ABSTRACT

Big Data technologies have attracted much attention among researchers due to their potential to handle large amounts of data. Big Data continues to become a hot topic of research in the agricultural field. Recent developments in the field of technology have led to a renewed interest in the field of smart agriculture. The current smart agricultural system produces and depends on large amounts of data, yet, it is hard to process the vast amounts of data using traditional data analysis systems Thus, for numerous possibilities and powerful data processing capabilities.The internet, sensor devices and data analytic paves the way for better performance, advanced resources planning and enhanced production in all industries. The agriculture is the most important and essential need for the world to invest and make research to strengthen the yield, crop disease control, farming community and advance agricultural production.In this paper, there is a brief description about big data and its techniques involved in agriculture.

I.INTRODUCTION

A. BIG DATA

Big Data A complex and massive collection of data which is hard to process by applying traditional data processing techniques or on-hand database management tools is referred as 'big data. These data are available on heterogeneity structures which are structured or unstructured or semi-structured **[28]**. Data which is having pre-defined format is called structured (banking information), data without predefined format is unstructured (text files, audio) and the semi-structured is data that are able to convert from unstructured to structured using available descriptions (xml document). Also big data is characterized by 3Vs to 4Vs which is meant by Volume, Variety, Velocity and Veracity **[29] [22]**. Each dimension has an opportunities to advance taking decisions as well as challenges for data management **[22].**

- Volume – amount of data generated
- Variety – variation of data with respect to the time.
- Velocity – how speed data is generated and processed.
- Veracity – availability and accountability

The characterizing of the big data as bad, good or undefined is depended on incompleteness, in-consistence, latency, deception, ambiguity and approximations.

B. BIG DATA ANALYTICS

The complexity of big data is very huge which is unable to process using commonly used software tools. Also data sizes are continuously increasing and changing from time to time. Hence, the used technique for analyzing data is called big data analytic. The process of collecting, keeping and analyzing the data with the aim of revealing hidden patterns, unknown correlations and other facts is called big data analysis **[8][4]**. Analyzing big data is needed analytical capabilities as well as optimal processing power **[3]**. In big data analysis, raw format data are converted into standard format with the support of tools **[12] [20]**. This is a process of information gathering, data analysis, visualization and scheduling.

C. BIG DATA ANALYTICS IN AGRICULTURE

Big data analysis is successfully adopted to industries like banking, insurance etc. Although agriculture did not adopt big data analysis for the past few years, recently it is used to use it. Agriculture produces countless types of data from economic models, crop productivity, crop diseases etc. **[18].** Smarter decisions can be obtained by using real time data on air, weather, soil, crop maturity and labor cost and even equipment **[3]**. Utilization of big data in agriculture diminishes the failures of farmers and recommend soil, water level etc. **[19].**The better service is given when the analysis is done continuously **[15].**
Big data analytic in agriculture can be studied under two major areas:

- Smart Farming
- Precision Agriculture (Pa) **[30].**

The idea of farming management that is including about measuring, observing and responding is called precision agriculture **[26].** In PA, it is needed collection of data, analyzing and processing the gathered information **[20]**. It uses big data tools for maximizing the productivity while using minimum number of resources **[30]**. Global Positioning System (GPS), Geographic Information System (GIS), and Variable-rate Technology (VRT) are some of the technologies used in Precision Agriculture **[19]**. Precision agriculture gives not only great challenges but also great opportunities for Computer Scientists specially who are working in the field of data analysis **[4]**. Smart farming explains the relationships among functions, variables, and concepts **[30].** This is concentrated on the big data analytic applications including agricultural value chain and business processes [30].

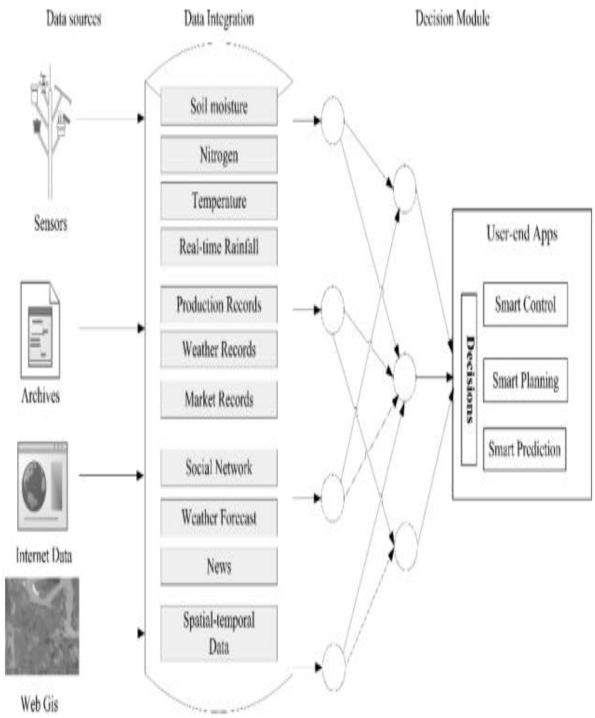

Fig.1. Big Data Analytic in Agriculture

(I)SMART FARMING

Smart farming can be defined as the multidimensional combination of several different technological implementations [cite]. As per the domain of cognitive science "a process is considered smart if it has the following six characteristics which are adapting, inferring, sensing, anticipating, learning and self-organizing" [cite]. In order to achieve smart farming and to develop solutions; integration of technologies is required **[26].** The smart farming model which consists of the integrated technologies such as IoT, Cloud and Mobile technology uses the concept of Map-reduce in terms of data analytics. This is easy to handle data which uses multiple nodes. The process is divided in two parts which is map and reduce. Map is used to filter and sort. Whereas, reduce function can perform summary operations. Hence, this specific analytic technique is used for predictive analysis. It can also be implemented to perform Data mining because of the huge amount of generated data by the IoT sensors. The flow is data obtained from the sensors is stored in the cloud platform. Following which pre-processing of data is done then it focuses towards categorization and then attribute selection, algorithm implementation and later evaluating with pattern prediction. Hence the output that is obtained can be visualized using Business analytic tool and can be used for decision making such as fertilizer requirement, crop sequence, weather patterns etc **[26].**

(i) Internet of Things(IoT):
Technologies such as Internet of things (IoT) in which it uses sensors placed in fields and farms that can gather data. Wireless sensors can also be used to analyze various agricultural parameters such as soil properties, temperature data etc.

(ii) Cloud Computing:
Similarly, the collected data needs to be accumulated at a common platform that should be easily accessible. This can be achieved by introducing cloud computing which provides sharing of resources at a reasonable cost. It can be majorly used to store agricultural data and can be implemented alongside IoT.

(iii) Mobile Computing:
Mobile computing which is used essentially in every domain including agriculture. It can be used to send daily or seasonal messages to agriculturists/farmers regarding weather, market or product details. Finally, these above-mentioned technologies are integrated with analytic techniques in order to find out new design, evaluation, concept and development.

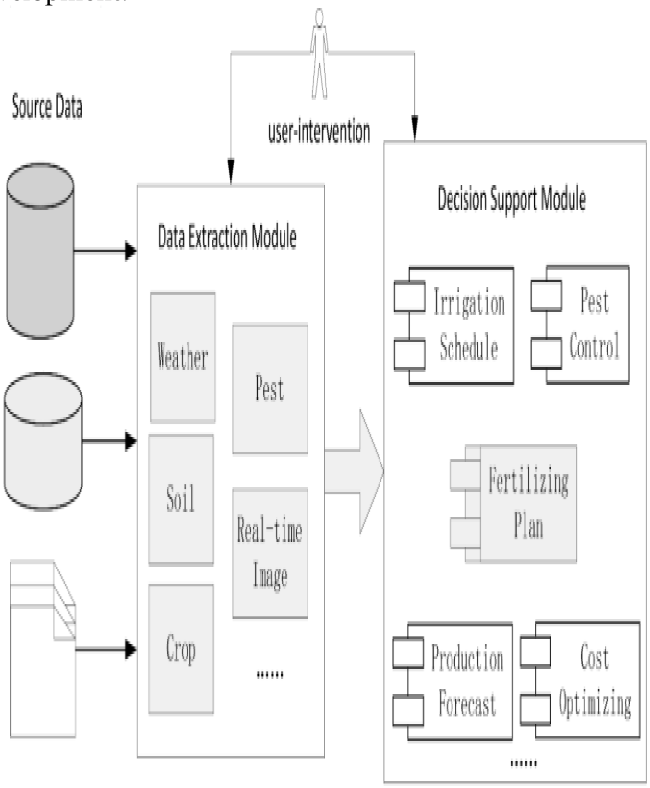

Fig.2. Big Data Analytic in Agriculture

D. BIG DATA ANALYTIC TOOLS AND TECHNIQUES
As discussed in earlier, handling and controlling big data by using traditional tools and techniques is very difficult. Therefore big data analytic framework has found several tools and techniques in order to manage data to take decisions. Efficiency of these techniques is highly important as it makes easier to take decisions. Clustering, Classification and Association are the techniques which are currently used in big data analytic **[3].** Map Reduce, Storm and Apache Spark are emerging tools in big data analytic **[29].** It is said that a lot of tools use stream processing, interactive analysis and batch processing **[29].** Most batch processing tools follow Apache Hadoop which is an open-source software which is reliable and very fast **[29] [27] [4]**. The framework Hadoop is having two major parts which are HDFS and Hadoop Map Reduce **[20] [24].** Apache Spark is also a computing framework for real time data as well as having memory oriented architecture with flexibility **[24]**. Apache Hive, Apache Mahout, Apache Pig and Apache HBase are some tools categorized under Hadoop **[29] [10].** Apache Hadoop and Apache Spark most popular among big data analytic applications **[24].** Google charts, Tableau, D3, Fusion chart, High chart, Canvas etc. are the some of the visualization tools in big data analytic **[14].**

II.METHODOLOGY

A. DATA INTEGRATION
Data integration is to centralize data from different sources and heterogeneous data in order to provide a complete data foundation for data analysis. Usually, extracting data from different databases or data tables is

to form a subject data warehouse relating to a certain field. Big data technology provides a new solution for processing large amounts of heterogeneous data and unstructured data **[9].** Hadoop is a distributed system infrastructure of the Apache open source organization. The core component, HDFS (Hadoop Distributed File System), provides unified storage support for massive, structured, and unstructured data. ETL (Extract-Transform-Load) tools can be used to integrate multiple source data and construct data warehouses to form thematic databases for various purposes **[10]**. Typical ETL activities include Extraction, Transformation, and Loading of data. Software tools are developed to take data out of a variety of sources, carry out necessary transformation and cleansing, and put into data marts for application analysis.

B. INTELLIGENT CROP RECOMMENDATION SYSTEM

This is an intelligent system which is a combination of Big data analytics and big data analytic and is used for decision making, recommending crops and making predictions. This system takes into consideration all the parameters such as rainfall, soil conditions, temperature and location. The recommendation system is generally split into two sub-systems which are interrelated to each other. Firstly, the sub system consists of

- Crop Predictor
- Rainfall Predictor **[27]**.

(i). Crop predictor:

This sub system is the primary function of the recommendation system which helps agriculturists recommend crops. Hence the initial part is acquiring of training data set because the accuracy of a Big data analytics algorithm is based on the selection of parameters and the correctness in its training data. The dataset generated has as schema such as soil type, aquifer thickness, soil PH, thickness of top-soil, precipitation, temperature and location. The next part of the crop predicting system is the prepossessing of data. This is further divided into two steps. Basically, there could be missing values in the original training dataset which has to be removed and replaced. The reason for removing and replacing the values is because the presence of these can cause deterioration to the values and can affect the Big data analytics algorithms performance. Following that, the second step is to generate class labels for the dataset before applying it to the algorithm.

The major reason behind labeling the data is because the recommend-er system is being trained using supervised learning. Since the system uses Multi-labeled classification. Therefore, the model has more than one class which is being assigned to a single instance for which the best Big data analytics algorithms would be K nearest neighbors (K-NN), Random forest, Decision trees and Artificial neural network **[27]**. Hence, after applying the appropriate Big data analytics algorithms. The trained models are used to recommend crops to the farmers; this can be done by saving the weights of the model. Henceforth, recommendation can be availed by just giving in raw inputs given initially which were soil type, aquifer thickness, soil PH, top-soil thickness, precipitation, temperature and and location.

(ii). Rainfall Prediction System:

This sub system as the name suggests predicts the occurrence of rainfall. As we know each crop has its individual rainfall requirement and if the requirement is not met then the yield may suffer. However, if the rainfall is surplus it may also cause adverse consequences. Hence, rainfall can be considered one of the critical parameters in the field of agricultural analytics. However, prediction of rainfall accurately during the sowing and harvesting season can be one of the major challenges. Therefore, this sub system predicts rainfall for each month across the year. The dataset is obtained from the governments meteorological official site which consists of month wise rainfall data. In millimeters (mm) **[27]**. Similarly, as the previous sub system data must be per-processed here the missing values are removed and are replaced with negative values in order to avoid deterioration. The Big data analytics algorithm that is implemented here is linear regression which is again a supervised learning approach where the parameters are considered as X (Location) and Y (precipitation values) which further can be used to predict the quantitative value of rainfall occurring in the region.

III.CURRENT ISSUES OF BIG DATA ANALYTICS IN AGRICULTURE

Even though big data analytics makes the path to interpret data in a better way, there can be found some limitations and challenges in this area. It is said that anything can be prevailed for a long time if it is applied in a correct way. However, dealing with big data is somewhat difficult as it wants proper storage, cleansing, processing and analyzing which are normally increased in an exponential way **[1].**

(i). Security and Privacy

Security and privacy of data is a highly important fact when considering big data analysis. Because of this, although big data analytics is having capability of using in agriculture, willingness of the farmers to share data on the agricultural products is a challenge. It is revealed that data privacy is an issue in big data analysis in agriculture because misusing of data **[6]**. Further, data access rights, transparency and usability, data ethics, data independence, data timeliness and barriers in data sharing are also a challenge in big data analysis in agriculture **[16] [14].**

(ii). Lack of technical knowledge

It is clearly noticed lack of knowledge on technologies and techniques as a challenge. Even, there is a well identified space between developed and developing countries because of unbalance access to the technologies and lack of skills **[6].**

(iii) Difficulty in Scalability and Visualizing

It is mentioned that the characterized 4Vs of the big data cause for many issues in agricultural big data. Since in each minute massive amount of data is generated and they are increased in exponentially and most of the generated data are semistructured or unstructured, dealing with the growth of big data is identified as an issue. With the growth of data, scalability **[9]** and data visualization are highly difficult tasks **[6].** Also noted that some computational techniques work well in a small set of data, it makes difficult when the data size is increasing **[7].**

(iv) Limited Data Storage

To store huge amount of data with 4Vs, there should be a high data storage. As well as data should be stored in an efficient manner **[16]** and there should be high throughput **[7].**

(v) Quality of Data

Since there is massive collection of data, the data accuracy and quality should be in a high position. When doing analysis extracting most accurate data is important for taking the better decision **[14] [23].**

IV.LITERATURE SURVEY

Xiong et.al. [1] demonstrated the Green leaves D1log(1/T)-CARS-PLS (Competitive adaptive reweighted sampling - Partial least squares) dataset with comparison, red-tip lettuce has stronger light absorption in the Visible Partial least squares and radial basis function neural network. Carried out to estimate the feasibility of using Visible/near-infrared spectroscopy to determine the potassium concentration and petioles of distinct variety and mixed lettuce leaves of two varieties Partial least squares offered R2 of 0.83, residual predictive deviations of 1.95, and RMSE (Root-mean-square error) of 39.07.

Abdipour et.al. [2] focused on analyzing the environmental constraints that impact the crop yield, namely area under cultivation, annual rainfall, and food price index. RA (Regression Analysis) analyzed the factors and groups them into explanatory and response variables that aids in attaining a decision.

Abbas .et.al. [3] presented a scalable, accurate, and inexpensive technique to forecast crop yields using accessible Remote Sensing statistics (open source). The proposed scheme improved the accuracy of the yield prediction pointedly along with a novel dimensional reduction technique.

Wang .et.al. [4] demonstrated the applications of the proposed algorithm to compute the probability of working days. – Performance criteria were considered, such as RMSE (Root-mean-square error) , MAPE (Mean Absolute Percentage Error), and R2 . Radial basis function offered the highest R2 compared with multiple linear regressions.

Khaki et.al. [5] evaluated five different Artificial Neural Network methods, namely generalized feed-forward, multilayer perceptron, Jordan/Elman, principal component analysis, and radial basis function. Among these models, multilayer perceptron offered the best prediction.

Esfandiarpour-Boroujeni et.al. [6] compared various artificial intelligence models to attain the most excellent crop yield prediction for the Midwestern United States (US). Notably, the Deep Neural Network model performed well, and its optimization process ensured the most acceptable configurations for the drop-out ratio, layer structure, cost function, and activation function.

Maya Gopal et.al. [7] used the enhanced vegetation index from MODIS (Moderate Resolution Imaging Spectroradiometer)and solar-induced chlorophyll fluorescence from GOME-2(Global Ozone Monitoring Experiment - 2) and SCIAMACHY(Scanning Imaging Absorption Spectrometer For Atmospheric Cartography)as metrics to predict crop production. The Machine Learning method offered the best yield prediction compared with the regression method.

Khosla et.al. [8] evaluated the performance of a proposed method to forecast apricot yield and identified significant factors affecting the yield. The proposed scheme offered relatively high accuracy of prediction (RMSE(Root-mean-square error) of 1.737 and 2.329 for training and testing data, respectively). Hybrid particle swarm optimization imperialist competitive algorithm, Support Vector Regression.

Kosari-Moghaddam et.al. [9] predicted the extent of monsoon rainfall using modular Artificial Neural Networks. Predicted the extent of chief crops yielded considering the rainfall data and area using Support Vector Regression.

Chlingaryan et.al. [10] carried out to estimate the feasibility of using Visible/near-infrared spectroscopy to determine the potassium concentration and petioles of distinct variety and mixed lettuce leaves of two

varieties. Partial least squares offered R2 of 0.83, residual predictive deviations of 1.95 and RMSE(Root-mean-square error) of 39.07.
Gu et.al. [11] surveyed the soil and crop variables potentially for variations in yield. – Four algorithms were demonstrated: linear regression, elastic net, k-nearest neighbor, and support vector regression to forecast potato yield from soil and crop data properties collected over proximal sensing.
You et.al. [12] discussed research growths accompanied within the last fifteen years on Machine Learning-based methods for accurate crop yield prediction and compared with Remote Sensing methods. Concluded that the fast developments in sensing tools and Machine Learning techniques could deliver cost-effective and wide-ranging resolutions for improved crop and decision making.
Marko et.al. [13] showed the design strategy for selecting soybean varieties to exploit maximum yield in the best season based on the knowledge attained from heterogeneous historical data. The outcomes with the existing regression algorithm proved that the proposed algorithm offered an optimal selection of seed varieties.
Feng et.al. [14] compared various artificial intelligence models to attain the most excellent crop yield prediction for the Midwestern United States (US). Notably, the Artificial Neural Network model performed well, and its optimization process ensured the most acceptable configurations for the drop-out ratio, layer structure, cost function, and activation function.
Mishra et.al. [15] used the agricultural dataset to contain 745 cases; 70% of statistics are randomly nominated to train the model and 30% are used for testing the model to evaluate the predictive capability. Among the four algorithms, random forest offered the best accuracy in prediction.
Deepak Murugan et.al. [16] has proposed "Precision agriculture monitoring at larger scales is critical for improved agricultural productivity and food management. Drones have been used for precision agricultural monitoring at smaller scales in recent years, and satellite data has been used for land cover classification and agriculture monitoring at higher scales for many decades".
Jochen Huster et.al. [17] has proposed In harvesting technology, cutting processes are among the most important crushing procedures. Trying to figure out when to regrind crop cutting blades is a common practise that rarely yields the desired results. "Continuous monitoring of harvesting Machine cutter bars may provide the best regrinding time to maintain cutting performance and, ideally, long maintenance intervals. Based on an analytical simulation, acoustic measurements, and statistical analysis, a method for real-time acoustic monitoring of the sharpness of crop cutting blades is demonstrated in this study.
Aruvansh Nigam et.al. [18] conducted experiments on Indian government dataset and it's been established that Random Forest Machine learning algorithm gives the best yield prediction accuracy. Sequential model that's Simple Recurrent Neural Network performs better on rainfall prediction while LSTM is good for temperature prediction. The paper puts factors like rainfall, temperature, season, area etc. together for yield prediction. Results reveals that Random Forest is the best classifier when all parameters are combined.
Balamurugan et.al. [19] have implemented crop yield prediction by using only the random forest classifier. Various features like rainfall, temperature and season were taken into account to predict the crop yield. Other Machine learning algorithms were not applied to the datasets. With the absence of other algorithms, comparison and quantification were missing thus unable to provide the apt algorithm.
Mishra et.al. [20] has theoretically described various Machine learning techniques that can be applied in various forecasting areas. However, their work fails to implement any algorithms and thus cannot provide a clear insight into the practicality of the proposed work.
Bhuyar et.al. [21] proposed an approach where different classification algorithms such as J48, Naïve Bayes, and Random forest algorithm were applied to soil data set to predict its fertility. J48 algorithm gave better result with an accuracy of 98.17% than other algorithms.
Rajeshwari et.al. [22] performed a comparative analysis of ML algorithms i.e. Naive Bayes, JRIP and J48 for prediction of soil types. The experiments were performed on soil data consisting of 110 samples using data analytics tool R. The experimental results predicted that JRIP algorithm performed better as it gave highest accuracy of 98.18% with kappa statistic of approximate 1.0.
Awasthi et.al. [23] performed comparative study on two data mining techniques namely Artificial Neural Network and Support Vector Machine with the help of data analytics tool R. ANN was implemented with 7 hidden nodes and this model trained for 73073 steps. It predicted accuracy of 55% with root mean square error is of 15. SVM implemented with Radial basis kernel and it achieved much better results with 74% of accuracy.
Rakesh et.al. [24] implemented Random Forest technique for crop yield prediction in Tamil Nadu state. The dataset for the study includes various parameters such as rainfall, temperature, crop production, etc. and experiments were performed using R Studio.
Renuka et.al. [25] applied Machine learning techniques namely KNN, SVM and Decision Tree for yield prediction of sugarcane crop. The study was carried out in Python platform. Decision Tree predicted highest accuracy of 99% with less mean square error.

IV.CONCLUSION

Currently, big data become more powerful with the development of the technology. As big data analytics is vastly adopted by many industries especially in agriculture, it has to face many challenges. As a result of these challenges, directions for the our research are revealed for further improvement.It is identified as a our research area to implement a mobile application which can be used by the farmers without language barrier in order to gather real time information **[12]**. Also, needed to implement an application in the agricultural input supply sector **[5].** Tools for better demand and yield prediction should be developed **[6].** Among the suggested our research directions, agricultural robots which are self-operated to identify weeds and remove them, tools with high performances, programming language for big data analysis can be identified **[6] [7].**Further, usage of IoT and cloud for storing and retrieving data is also a way that can be scaled up in our research

V.REFERENCE

1. Xiong, Y.; Ohashi, S.; Nakano, K.; Jiang, W.; Takizawa, K.; Iijima, K.; Maniwara, P. "Application of the radial basis function neural networks to improve the nondestructive Vis/NIR spectrophotometric analysis of potassium in fresh lettuces". Springer, Vol.110, pp. 298-317, 2021.
2. Abdipour, M.; Younessi-Hmazekhanlu, M.; Ramazani, S.H.R.; Omidi, A.H. "Artifificial neural networks and multiple linear regression as potential methods for modeling seed yield of safflflower (*Carthamus tinctorius* L.)". Springer. Vol.127, Issue no.2 pp.185–194, 2021.
3. Abbas, F.; Afzaal, H.; Farooque, A.A.; Tang, S. "Crop Yield Prediction through Proximal Sensing and Big data analytics Algorithms". Elsevier Vol.10, Issue no. pp.1046, 2020,
4. Wang, L.; Wang, P.; Liang, S.; Zhu, Y.; Khan, J.; Fang, S. "Monitoring maize growth on the North China Plain using a hybrid genetic algorithm-based back-propagation neural network model". Elsevier. , Vol.170, Issue no.2, pp.105-112, 2020
5. Khaki, S.; Wang, L. "Crop Yield Prediction Using Deep Neural Networks" IEEE,Vol.10, Issue no.12 pp.621-630, 2020.
6. Esfandiarpour-Boroujeni, I.; Karimi, E.; Shirani, H.; Esmaeilizadeh, M.; Mosleh, Z. "Yield prediction of apricot using a hybrid particle swarm optimization-imperialist competitive algorithm- support vector regression (PSO-ICA-SVR) method". IEEE. Vol.108, Issue no.6, pp.257, 2019
7. Maya Gopal, P.S.; Bhargavi, R. "Performance Evaluation of Best Feature Subsets for Crop Yield Prediction Using Big data analytics Algorithms". Springer, Vol.*33*, Issue no.12 pp.621–642,2019
8. Khosla, E.; Dharavath, R.; Priya, R. "Crop yield prediction using aggregated rainfall-based modular artificial neural networks and Support vector regression" Elsevier. , Vol.22, Issue no.14, pp. 5687–5708, 2019
9. Kosari-Moghaddam, A.; Rohani, A.; Kosari-Moghaddam, L.; Esmaeipour-Troujeni, M. "Developing a Radial Basis Function Neural Networks to Predict the Working Days for Tillage Operation in Crop Production". Springer. Vol.9, Issue no.10, pp.119–133, 2018.
10. Chlingaryan, A.; Sukkarieh, S.; Whelan, B. "Big data analytics approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review". IEEE, Vol.151,Issue no.9, pp. 61–69, 2018
11. Gu, J.; Yin, G.; Huang, P.; Guo, J.; Chen, L. "An improved back propagation neural network prediction model for subsurface dripirrigation system". IEEE, Vol.60, Issue no.8 pp.58–65, 2017.
12. You, J.; Li, X.; Low, M.; Lobell, D.; Ermon, S. "Deep Gaussian Process for Crop Yield Prediction Based on Remote Sensing Data". IEEE , Vol.16, Issue no.6, pp.4–9, 2017.
13. Marko, O.; Brdar, S.; Panic, M.; Lugonja, P.; Crnojevic, V. "Soybean varieties portfolio optimization based on yield prediction". IEEE Vol. 127, Issue no.8 pp.467–474, 2018.
14. Feng, P.; Wang, B.; Liu, D.L.; Waters, C.; Yu, Q. "Incorporating Big data analytics with biophysical model can improve the evaluation of climate extremes impacts on wheat yield in south-eastern Australia". Springer. Vol. 275, Issue no.19, pp.100–113, 2016.
15. Mishra, S.; Mishra, D.; Santra, G.H. "Applications of Big data analytics Techniques in Agricultural Crop Production: A Review Paper" Springer. Vol.9, Issue no7, pp.1–14, 2016
16. Deepak Murugan , D. J. Mulla, "Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps," Elsevier, vol. 114, Issue no.5, pp. 358–371, 2016
17. Jochen Huster, A. McBratney, B. Whelan, T. Ancev, and J. Bouma, "Future directions of precision agriculture," Elsevier, Vol. 6, Issue no.2, pp. 7–23, 2016.
18. Aruvansh Nigam, Saksham Garg, Archit Agrawal "Crop Yield Prediction using ML Algorithms" Elsevier, vol. 117, Issue no.pp. 378–384, 2019
19. Priya, P., Muthaiah, U., Balamurugan, M."Predicting Yield of the Crop Using Big data analytics Algorithm", Springer Vol.112, Issue no.12, pp. 645-656, 2015
20. Mishra, S., Mishra, D., Santra, G. H.,"Applications of Big data analytics techniques in agricultural crop production", Elsevier, vol. 114, Issue no.15, pp.402–414, 2016.

21. Bhuyar, V. “Comparative Analysis of Classification Techniques on Soil Data to Predict Fertility Rate for Aurangabad District”, Springer Vol.3, Issue no.1, pp.200-203, 2014.
22. Rajeshwari, V. & Arunesh, K. “Analyzing Soil Data using Data Mining Classification Techniques”, Springer Vol. 9, Issue no.4, pp.1-4, 2016.
23. Awasthi, N. & Bansal, A. “Application of Data Mining Classification Techniques on Soil Data using R”. IEEE Vol.4, Issue no.6, pp. 33-37, 2017.
24. Rakesh . “Performance analysis of rapeseed-mustard under different agro-climatic conditions of Jammu division of J&K state”,IEEE Vol.5, Issue no.9, pp. 464-470, 2018
25. Renuka & Terdal, S. “Evaluation of Big data analytics Algorithms for Crop Yield Prediction”. Springer, Vol.8, Issue no.6, pp.4082-4086, 2019