TEST Engineering & Management

January - February 2020
ISSN: 0193-4120 Page No. 16972 - 16979

A Review of Various Agriculture Systems Based on IoT, Data Mining and Cloud

Shital Jadhav[1], Dr.Bindu Garg [2]

[1,2]Department of Computer Engineering, BharatiVidyapeeth's College of Engineering, Pune,India

Article Info
Volume 82
Page Number: 16972 - 16979
Publication Issue:
January - February 2020

Article History
Article Received*: 18 May 2019*
Revised*: 14 July 2019*
Accepted*: 22 December 2019*
Publication*: 29 February 2020*

Abstract
Evolution in technology has opened new gates for agriculture. Various sensors, machines, Information and communication technologies are available and applied in agriculture. Agriculture is vast domain with multiple facets like yield of crop, crop diseases and application of irrigation, fertilizers and pesticides. Climate change and environment effects agriculture which leads in gap in demand and supply of agriculture commodity affecting prizes. To empower farmers need to explore various technologies at cheaper cost. The review of various research work carried by scientist based on IoT, Data mining and Cloud Computing carried out in this paper. This work excludes research papers on crop image and satellite images of farms which is other aspect of analyzing agriculture data.

Keywords – *Agriculture, IoT, Data mining, Cloud Computing*

I. INTRODUCTION

Increasing population is challenge due to limited resources available on earth. Food requirement of population must be satisfied. Agriculture and food security is of at most importance. In India agriculture land is 60.45 per cent of total land , fresh water utilized for agriculture is 80 per cent of total and 51per cent workforce working in agriculture sector. In Economic Survey 2018 estimated that percentage of agricultural workers of total work force would drop to 25.7 per cent by 2050 from 51 per cent in 20018.The Cost of production of crops are high due to intensive labor involvement. There is requirement in enhanced mechanization. Precision Agriculture and Management of Information in agriculture is field of research to satisfy new requirement. Combination of Internet technology and analysis of various farm data will be useful in decision making for Farmers and Government. Crop yield is affected by several factors like geographical location, soil type, environmental factors, and nutrition and seed quality. Considering this fact up till now there is lack of established concept in Agriculture. The comprehensive review of the application based on Internet of Things (IOT), Data Mining and Cloud will be provided with examples how these has been applied to various agriculture problems.

II. BACKGROUND

In precision Agriculture modern technology like remote sensing , Global positioning system (GPS) and Geographic Information System (GIS) used with respect to soil, weather and crop need which will improve productivity, quality and profitability in Agriculture. These inputs used for efficient application of pesticides, fertilizer and water to bring more crop yield and quality without environmental pollution .This section give brief introduction of technologies for modernization of Agriculture .

Wireless Sensor Network (WSN): WSN is multidisciplinary research area that draws on contribution from signal processing, network and protocols, databases, information management, distributed algorithm and embedded system and

Published by: The Mattingley Publishing Co., Inc.

January - February 2020
ISSN: 0193-4120 Page No. 16972 - 16979

architecture. Major building blocks are sensors, sensor node and network topology. Sensor are transducer that convert physical phenomenon such as heat, light, sound into electrical or other signals that may be further manipulated by other approaches. Sensor nodes is basic unit of sensor network with onboard sensor, processor, memory, wireless modem and power supply. Ii is often abbreviated as node. Network topology is connectivity graph where nodes sensor nodes and edges are communication link. In wireless network link represents one hop connection and neighbors of a node are those within radio range of node. Bluetooth, Zigbee, and Ultra Wide Band (UWB) are low cost wireless technology.

For large sensor network construction aspect of cost of sensor, networking and software platform to support system deployment is important. Each application of Wireless sensor network has its unique characteristics. In Agriculture sensors can be used to monitor condition of plants, air quality and track environmental pollutant or natural or manmade disaster [16].

Internet of Things (IoT) :IoT connect things such as devices, appliances and machines to Internet. IoT allows things to communicate and exchange data while executing meaningful application towards common user or machine goal. Sensor can send data directly to Internet in IoT. Conversely WSN has router or central node. IoT system can utilize WSN by communicating with router to gather data. Smart irrigation system uses IoT devices and soil moisture sensors to determine the amount of moisture in the soil and release the flow of water through irrigation pipes only when the moisture level goes below a predefined threshold. Data on moisture level is also collected in cloud where it is analyzed to plan watering schedules [17].

Smartphone Technology: Farmer have smart phone very easily available having mobility that matches farming. These devices now a day equipped with various sensors which makes them promising tool to assist farming task. Some dedicated sensor perform farming task can be replicated on smart phones with build in sensors. Three category of sensor embedded in smart phone are motion, environment and position sensors. Example of position sensor is accelerator, gravity and rotational vector sensor. The second category of environmental sensor provides measurement of surrounding environment. This ranges from ambient air temperature, from thermometer pressure from barometer and illumination from photometer. The third category of position sensor provides measurement of device physical position such as GPS and orientation sensor [19].

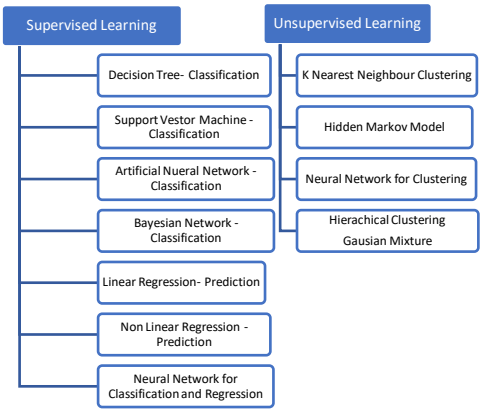
Data Mining (DM) : DM is category of algorithm that allows software application to become more accurate in predicting outcomes without explicitly programmed. In machine algorithms are build that receive input data and statistic logic predict output as new data become available . Machine learning focuses on development of computer programs that can access data and use it to learn for themselves [15]. DM categorized in two broad categories Supervised and Unsupervised learning. In supervised learning data are presented with example input and corresponding output. Supervised learning acquire expertise (trained model) is used to predict missing output label for test data. In unsupervised learning however there is no distinguish between training and test sets with data being unlabeled. In learner process input data with goal of discovery of hidden pattern. Neural Network evolved in Deep Learning models. Deep learning techniques require large amounts of labeled data and high computing power for parallel architecture. Training time required for deep learning can be reduced by integrating with cloud.

Cloud Computing (CC): CC is computing paradigm that involves delivering application and services over Internet. CC involves provisioning of computing, networking and storage resources on

Published by: The Mattingley Publishing Co., Inc.

January - February 2020
ISSN: 0193-4120 Page No. 16972 - 16979

demand and providing these resources as metered services to the users in "pay as you go" model.



Description of the Agricultural Systems based on IOT, Data Mining and Cloud

The paper [1] proposes system that helps farmer to use Android phones as remote control for his farm irrigation. The system proposed Arduino based automated irrigation system that collects data through sensors like light, moisture, temperature and humidity. Externally hosted cloud computing platform used to manage database, isolated server by across country and Android phones. The authors claim to propose first time integration of microcontroller, data mining, cloud and smart phone. Distance and range problem of previous systems are overcome and system is used in less water areas. User interface on Android phone provides graphical data used by Farmers to monitor water and crop yield. In proposed work there is no prominent use of data mining technique.

The paper [2] reports on design proposed for smart automated irrigation system with disease prediction. Various sensors data such as temperature, soil moisture and leaf wetness will be compared with predetermined threshold value of specific crop. Possible disease is predicted using markov model and observed results are transmitted to farmers. Android phones used to turn motor pumps on and off using Arduino based commands. Authors propose future work to detect pest and insect favorable condition in agriculture field.

The research work [3] carried out for monitoring plant needs at real time by Jojoba Israel by collecting soil moisture data with sensors. Additional sources of information like meteorological station and irrigation plant records for weather, irrigation and yield recorded used for analysis. Different regression and classification algorithms applied to data to predict weekly irrigation recommendation plan.

The authors of paper [4] implemented various data mining methods to find crop yield. For Analysis used Agriculture data of last 6 years for Karnataka state including 28 districts for various crops, seasons and area .Parameters such as production in tons , temperature , average rainfall ,soil PH value , major fertilizers, minimum and maximum rainfall are selected .Implementation of modified DBSCAN, PAM (Partition around medoids) and CLARA (Clustering Large Application) used for clustering Large amount of Data. The performance of each technique is evaluated and comparison between techniques carried out using quality metrics. Multiple linear regression used to estimate crop yield. In future this work can be extended for other soil and crop parameters under different climate to increase crop production.

In Paper[5] authors proposed model which will collect real time data from Wireless Sensor Network node with soil moisture, environmental temperature ,humidity , co2 and sunlight intensity sensors. Gateway node will transfer this data using WiFi to server computer for soil moisture content prediction. Scaled Conjugate Gradient (SCG) and Broyden Fletcher Goldfarb Shanno (BFGS) Quasi Network algorithm will use 11 sensor parameters to predict soil moisture content 1 hour before. Then after 1 hour actual soil moisture content are taken to calculate Mean Square Error (MSE) and Root Mean Square Error (RMSE) . Soil Moisture Content prediction supplied to fuzzy logic weather model for irrigation recommendation.

Published by: The Mattingley Publishing Co., Inc.

January - February 2020
ISSN: 0193-4120 Page No. 16972 - 16979

The Paper [6] proposed and implemented IOT framework to provide recommendation to farmers for irrigation timings, direction for fertilizer use etc. Sensor input like soil moisture , Temperature and Ph Sensor given as input to Wireless Sensor Network (WSN) through interface WSN sends data to base station which will be provided to central monitoring system . Soil moisture level is compared with already collected threshold value and take decision whether to apply irrigation or not. Proposed system collects PH Level of soil from sensors and uses available information base to identify needs of fertilizer. Android application provides this information to farmers.

The research work [7] analyzes gap in demand and supply by forecasting demand of agriculture commodity so that according to it cropping decision taken by farmers. Government data and Market data from year 2005 to 2016 collected for Karnataka state. Methods like simple linear regression, multiple linear regression and Holt Winters model used to forecast supply, demand and price specifically for tomato crop. Through analysis observations it is observed that gap between demand and supply creates price variation of agriculture commodity. Holt Winters model shows better results than rest two. Future work suggested in this paper is Intelligent Agriculture Analytics System with integration of forecast method suggested in paper with analytics that will be useful for Farmers to take better cropping decisions.

The paper [8] proposes framework to predict annual main crops in Egypt considering need of population, imports and production. This system would be useful in market decision in food security and countries productivity. Proposed Framework collect data from food security information center (FSIC) after data selection , information preprocessing and then applying data mining technique Artificial Neural Network via Weka using Multilayer Perceptron predict difference between available amount of crops and needed amount .Results visualizes data for wheat, rice and bean using graphs.

The paper [9] uses deep learning for prediction of agriculture environment data that is not explored much till now. Input data collected from R&D project grain cloud platform project for monitoring agriculture environment. This data is cleaned, feature are extracted and recursive neural network Long Short Term Memory (LSTM) used for prediction. Recursive neural network explore non linear relationship between data suitable for agriculture environment. Training data set applied to model. For optimization purpose 3 layered LSTM enhanced with additional layer GRU. This dropout layer prevents over fitting and uses 3 statistics RMSE, MSE and MAE for comparative analysis. Results obtained by this method are more accurate and helps in prior preparation in agriculture. This work can be extended in future with labeling system that will find normal or abnormal environment next movement based on current environment parameters.

The work carried out in paper [10] performs analysis of data for crops like millet , ground nut , pulses, sugarcane etc. Crops take water and nutrients from soil therefore soil parameters like depth, texture, PH, water holding etc. play crucial role in crop growth . Choosing correct crop considering soil parameter important . Research work analyses data collected from soil testing lab. Types of crop prediction carried out using majority voting popular ensemble model technique. Learner model like Random tree, Chaid, KNN, Naïve bayes predicts class label for each training data set. The class label that is predicted by majority of models is selected. Rules are generated by ensemble technique for cross validation, replace missing values and retrieve data set. This model gives 88% accuracy web based portal gives GUI for recommendation system with use of rule generated. In future this work can be extended for large number of attributes and total yield prediction.

Published by: The Mattingley Publishing Co., Inc.

January - February 2020
ISSN: 0193-4120 Page No. 16972 - 16979

The paper [11] uses data mining technique applied to crop disease and loss prediction which belongs to classification family. The various classifiers NN, KNN, DT, RF, SVM and GNB applied to data set of grass grab damage with 155 records and 8 feature and four labels. Accuracy, mean, F1 score, recall and precision are the measures on which classifier is evaluated. Ensemble model combines result of above algorithm to improve accuracy of weak classifier. Voting method for classifier used to predict grass grab damage. For data preprocessing chi-square method of selection is used. The classifier applied for both binary and actual data. In result binary data analysis give acceptable result compared to original data. In future local optima problem of data mining algorithms can be overcome by evolutionary deep learning technique of recurrent neural network.

The paper [12] presents research work on integration of agriculture data from different sources which give information on agriculture cultivation. Open Government data used for integration using first method which takes production and average yield of principal crops. Results of work shows that yield of Sugarcane is highest among all crops. In second method gross area under irrigation and production of principal crops analyzed based on results analyzed for percentage of high yield crop. Regression analysis curve used to estimate crop yield for given land area.

The paper [13] proposes system to classify soil based on fertility and predicting soil toxicity level. Based on level of fertility recommend fertilizer or based on toxicity action is taken. Various recommendations provided for crop including irrigation requirements. Data sets collected from sensors such as humidity, temperature, pH and Nutrients. Fertilizer data set organized with fertilizer name, crop name and date. Data stored on server ,pre processing carried out and Decision tree algorithm J48 used to classify soil based on fertility and toxicity. On client end results of analysis used to give recommendation of fertilizer, type of crop and water.

The paper [21] proposes to build cloud based agriculture information system for farmers and customers. Data collected from different sources like document, book, observation, feedback and media generated from different user of agriculture like scientist, policy maker, agriculture organization and agribusiness professional. Analysis carried out by using linear and multiple linear regressions. Data will be analyzed for forecasting crop production and pattern between different data set. Data and results stored on cloud would be safe and easily available any time anywhere and reliable since variety of authorized sources of data.

III. COMPARISON OF THE AGRICULTURAL SYSTEMS BASED ON IOT, DATA MINING AND CLOUD

Summary of 15 research paper by different scientist is presented in this paper. This section presents comparative analysis of various systems implemented or proposed in these research paper considered for review.

Agriculture Problem	Parameter/ Data Source type	Technique	Description/Result	Future Scope
Soil Analysis	Soil type, temperature, humidity, pH, nutrients	KNN,SVM,NN,LR and J48 DT	Prediction of soil moisture with more accuracy with KNN [20]. Classify soil based on fertility parameters accuracy of J48 decision tree algorithm is higher as compared to other classification algorithms[13].	New machine learning algorithms for different geographic location and meteorological station[20]

Published by: The Mattingley Publishing Co., Inc.

TEST Engineering & Management

January - February 2020
ISSN: 0193-4120 Page No. 16972 - 16979

Crop Diseases	Data sets with 155 records and 8 features for crop disease caused by Grass grub insect that cause diseases [11].	NN,DT,RF,KNN, SVM and GNB	Different data classifier applied using feature set to predict damage by Grass grab insect that increases profit for both customer and farmer [11].	To solve the crops related problem using various evolutionary and hybrid approaches of deep learning [11].
Agriculture commodity Prizes	10 year market prices for tomato crop	Holt Winter's forecasting model, LR, Multiple LR	Prediction models for agricultural crop prizes shows that compared to Simple Linear Regression and Multiple Linear Regression, the Holt Winter's forecasting model gives better results [7].	Prize prediction for all type of crops can be provided as service on cloud to farmers to alleviate the loss for both customers and farmers and hence improving the overall performance of agricultural system[7].
Crop Production and Information management	Agriculture Documents, reports, feedback and media from different sources like Agriculture organization, scientist and policy makers [21]. soil pH, Minimum and maximum rainfall[4].	SPSS software used to analyze data and regression used for prediction [21]. DBSCAN, PAM (Partition around medoids) and CLARA (Clustering Large Application) used for clustering Multiple Linear Regression[4].	Proposed model for building agricultural Information system by bridging gap between farmer and stockholders. Prediction of crop production performed by regression [21].Various soil and Environmental parameters used to estimate crop yield[4].	Multiple factors can be considered for more crops, vegetables and fruits[21]. Work can be extended for other soil and crop parameters under different climate to increase crop production.[4]
Irrigation automation and control	soil moisture, temperature , environmental humidity , temperature , co2 and sunlight intensity [5],	HMM and Arduino [2] ,Quasi network algorithm ,Fuzzy Logic [5]	Farmers can monitor water and crop yield using Android phones used to turn motor pumps on and off using Arduino Based Automated Irrigation System with Moisture Sensor [2].Prediction of soil moisture compared with actual moisture and root mean square error and mean square error calculated. Irrigation recommendation based on fuzzy logic.	Work to detect pest and insect favorable condition in Agriculture field[2].
Agriculture Environment	Data collected by R&D project for Agriculture Environment Monitoring[21]	Recursive neural network Long Short Term Memory (LSTM)	Deep learning method used for prediction of Agriculture environment data used in prior preparation in Agriculture[9].	Agriculture labeling system that will find normal or abnormal environment next movement based on current environment parameters[9].

IV. GAPS IN IMPLEMENTATION AGRICULTURAL SYSTEMS BASED ON IOT, DATA MINING AND CLOUD

Following gaps are observed in research work for which literature survey carried out :

Present work is carried out for particular crop or vegetable in future these system can be expanded for more fruits , crops and vegetables.

Integration of different analysis of soil and other factors related to crops to increase crop production under different climatic condition need to be explored.

Accuracy of sensor data affects results and use more number of sensors to cover large area of farm to give correct decisions.

Prediction Model can be applied to deal with multidimensional time series data and enhanced with more advanced algorithms.

Future work can be carried out by choosing a different and more efficient method for data mining with new and improved algorithm designed for executing, produces the least possible error in comparison to the other methods.

Need of exploring mathematical models devised by mathematicians in detail, so that some productive methods among those could be identified.

Agriculture environment parameters of labeling system can be implemented to realize input of current period environment parameters to obtain normal or abnormal environment variable of next moment.

TEST Engineering & Management

January - February 2020
ISSN: 0193-4120 Page No. 16972 - 16979

V. CONCLUSION

In this paper we reviewed various approaches analyzing agriculture data. Identification of paper based on IoT, Data mining and cloud helped in examining agriculture problem, solution provided, tools and technology used. The comparison of agriculture system helped in understanding relation between various farming processes. Correct decision about sowing right crop considering soil analysis and smart application of input like irrigation, water, nutrient helps in increasing production of crop. Environmental data analysis with soil and crop data helps to predict crop disease at early stage. Since processing of dependent and independent time variant values require efficient model. Parallel algorithm need to be explored for increasing data and computational need. Integrated agriculture analytical system can be offered as service to agriculture stakeholders.

REFERENCES

[1] SubhashreeGhosh ,SumaiyaSayyed ,KanchanWani ,MrunalMhatre, "Smart Irrigation: A Smart Drip Irrigation System Using Cloud, Android And Data Mining " IEEE International Conference on Advances in Electronics, Communication and Computer Technology (ICAECCT), 978-1-5090-3662-2/16/$31.00 , IEEE (2016) , pp236-239.

[2] Ms.Yashaswini L S,Ms.Sinchana H N, Ms.Vani H U,Mr. Nithin Kumar , "Smart Automated Irrigation System with Disease Prediction ", IEEE International Conference on Power, Control, Signals and Instrumentation Engineering , 978-1-5386-0814-2/17/$31.00, IEEE(2017) pp 422-427.

[3] Anat Goldstein, Lior Fink1 , AmitMeitin ,hiranBohadana , Oscar LutenbergGiladRavid "Applying machine learning on sensor data for irrigation recommendations: revealing the agronomist's tacit knowledge" , Springer Precision Agric (2018) , doi- 10.1007/s11119-017-9527-4 19, pp 421–444.

[4] JharnaMajumdar, SnehaNaraseeyappa ,ShilpaAnkalaki "Analysis of agriculture data using data mining techniques: application of big data", Springer Open Access Journal of Big Data (2017), doi-10.1186/s40537-017-0077-4, pp 1-15.

[5] Ambarish G. Mohapatra, Saroj Kumar Lenka, "Neural Network Pattern Classification and Weather Dependent Fuzzy Logic Model for Irrigation Control in WSN Based Precision Agriculture" , International Conference on Information Security & Privacy (ICISP), Elsevier, Science Direct, Procedia of Computer Science (2016). doi- 10.1016/j.procs.2016.02.094 pp 499 – 506.

[6] Rekha P, Maneesha V. Ramesh, Venkata P. Rangan, Nibi K.V. "High Yield Groundnut Agronomy: An IoT Based Precision Farming Framework", 978-1-5090-6046-7/17/$31.00 , IEEE (2017), , pp 1-5.

[7] BalajiPrabhu B V, M Dakshayini "Performance Analysis of the Regression and Time Series Predictive Models using Parallel Implementation for Agricultural Data", Science Direct Procedia Computer Science 132 (2018) doi-10.1016/j.procs.2018.05.187 ,pp 198–207.

[8] Ayman E. Khedr, Mona Kadry, GhadaWalid , "Proposed framework for implementing data mining techniques to enhance decisions in agriculture sector", Science Direct Procedia Computer Science 65 (2015) doi-10.1016/j.procs.2015.09.007, pp 633 – 642.

[9] ShuchangChen,Bingchan Li, JieCao,BoMao,China ,"Research on Agricultural Environmental Prediction using Deep Learning", Science Direct, Procedia Computer Science (2018), doi-10.1016/j.procs.2018.10.214, pp 33-40.

[10] S.Pudumalar, E.Ramanujam,R.HarineRajashree, C.Kavya, T.Kiruthika, J.Nisha, "Crop Recommendation System for Precision Agriculture", IEEE Eighth International Conference on Advanced Computing (ICoAC), IEEE (2016) , doi- 978-1-5090-5888-4/16 ,pp 32-36.

[11] UmairAyub, Syed AtifMoqurrab "Predicting Crop Diseases Using Data Mining Approaches: Classification" ,978-1-5386-5482-8/18/$31.00 , IEEE (2018) , pp 1-6.

[12] Purva Grover, Rahul Johari ,"PAID-Predictive Agriculture Analysis for Data Integration in

January - February 2020
ISSN: 0193-4120 Page No. 16972 - 16979

India", 978-9-3805-4421-2/16/$31.00, IEEE (2016) , pp 184-188.

[13]MayuriPawar ,GeethaChillarge, "Soil Toxicity Prediction and Recommendation System Using Data Mining In Precision Agriculture", 978-1-5386-4273-3/18/$31.00, IEEE (2018) , pp 1-5.

[14] Konstantinos G. Liakos ,PatriziaBusato , DimitriosMoshou , Simon Pearson and DionysisBochtis ,"Machine Learning in Agriculture: A Review", Sensors (2018), doi-10.3390/s18082674 ,pp 1-29.

[15] LokeshaHuchaiah , "The Economics of Applications of Artificial Intelligence and Machine Learning in Agriculture", International Journal of pure and applied Bioscience (2019), doi-10.18782/2320-7051.7324 , pp 296-305.

[16] Feng Zhao, Leonidas GuiBas , "Wireless Sensor Networks An Information Processing Approach ", Elsevier Publication.

[17] ArshdeepBahga, Vijay Madisetti, "Internet of Things A hands on Approach" , Universities Press India Pvt. Ltd.

[18] Christopher M. Bishop, "Pattern Recognition and Machine Learning" , Springer International Edition.

[19] SupornPongnumkul, PimwadeeChaovalit, and NavapornSurasvadi, " Applications of Smartphone-Based Sensors in Agriculture: A Systematic Review of Research", Hindawi Publishing Corporation, Journal of Sensors (2015) , doi- 10.1155/2015/195308, pp 1-18.

[20] OliviuMateli,TeodorRusu, Adrian Petrovan ,Gabriel Mihut, "A Data Mining System for Real Time Soil Moisture Prediction", Elsevier Procedia Engineering 181 (2017) doi-10.1016/j.proeng.2017.02.475, pp 837 – 844.

[21] Pallavi V. Jirapure , Prof. Prarthana A. Deshkar ,"Qualitative data analysis using Regression method for Agricultural data", IEEE Sponsored World Conference on Futuristic Trends in Research and Innovation for Social Welfare (WCFTR'16), 978-1-4673-9214-3/16/$31.00 , IEEE (2016), pp 1-6

Published by: The Mattingley Publishing Co., Inc.