International Conference on Recent Innovations in Science, Engineering & Management
R.R. INSTITUTE OF TECHNOLOGY, Bengaluru (India)
10th - 11th May 2019 www.conferenceworld.in ISBN : 978-93-87793-87-3

An Effective Big Data Framework based on Farmer's Agricultural Context

Prof. Manjunath R[1], Bindu S[2], HariKireeti b k[3], Irfan Pasha F[4], Kavya R[5]

[1]*Professor, Computer Science Department, RR Institute of Technology, Bengaluru, India*
[1234]*UG Students, Computer Science Department, RR Institute of Technology, Bengaluru,*

ABSTRACT-*In the recent years the Big Data technologies in agriculture presents a major challenge and also plays an important role in contributing effectively in many countries social and economic development. In this paper, we study the environmental data provided by precision agriculture information technologies which signifies a crucial source of data in essential of being wisely managed and analyzed with appropriate techniques and tools in order to extract the meaningful information. Also the main purpose of this work is to build an effective big data framework based of profile which assists the producers, consultants, researchers to make better decisions, enhance and monitor the agricultural productivity.*

Keywords: Big Data, Data Mining, Precision Farming, Prediction Analysis, Profiling System.

1. INRODUCTION

Big data is a term used to refer the data sets that are too large or complex for traditional data processing application software to adequately deal with Big data challenges.In the recent years, the huge volume of real time data in the agricultural sector and it need for an efficient and effective processing, stimulate the use of novel technologies and platform to acquire, store, process, analyze and visualize large data sets or future predictions and decision making. The stored data is analyzed and predicts the future condition in agriculture. It helps in order to meet the needs and requirements of various end users. Agricultural practices, in all their forms, are responsible of the considerable data quantity and they use a large number of external data to guide farmer decision.Farmer's predicting manually the demand of grains and vegetable so its effect to farmers economically to overcome this problem we are proposed machine learning auto demand and yield prediction process.Then the Admin will upload datasets to Hadoop Framework which contains crop data, crop demand data , crop yield data, Vegetable Demand data and Vegetable Yield data of the Previous 10 Years of Demands for the Vegetables and Grains in the Particular District. Next User will upload his PHANI Details of his Land Which Contains Aadhar number, Survey number, Total Land and Soil Type. After Comparing the Above Datasets with Testing Datasets using the Linear Regression Algorithm We Predict the Output Value for Predicting the Particular Crops to be Grown in the Specified .

1.1 Cloud Computing for Big Data Environment

Cloud computing is the ideal model for large data because of its endless scalability and resources used on-demand. It promises reliable services based on virtualized storage technologies. Resort to a cloud computing technology in agriculture can give considerable solutions to analysts and decision makers. Using the cloud, we can benefit from special bricks in Big Data management to collect and centralize data the maximum as possible regardless their sources, to make detailed analysis in order to obtain valuable data. This can be ensured by the use of new database models including mixed approaches between relational and non-relational (NoSQL)

databases, also by a distributed architecture at processing level of unstructured data with the aim to spread the load over a large number of servers (cluster) using a total abstraction of subjacent parallelizing mechanism.

2. LITERATURE REVIEW

[1] World agriculture in the 21st century will face three major challenges: how to feed a growing world population, how to contribute to reducing the still–high prevalence of rural poverty in the world, and how to respond to increased concerns about managing the natural resource base. [2] Agriculture traditionally maintained by families and villages where the passing on and sharing of knowledge is regarded as very important. The accumulation and sharing of knowledge has resulted in better overall efficiency and productivity. Cloud computing technology attracts more and more attentions of countries and enterprises with its powerful advantages and market potential, the feasibility and applicability of whose application are also exploring in various industries. This technology will bring greater opportunities to the agricultural development, and also be the inevitable choice to achieve modernization and informatization in agriculture. [3] Precision farming provides a new solution using a systems approach for today's agricultural issues, namely the need to balance productivity with environmental concerns. Based on advanced information technology, it includes describing and modeling variation in soils and plant species, and integrating agricultural practices to meet site specific requirements. Precision farming aims at increased economic returns, as well as reducing the energy input and the environmental impact of agriculture. The paper describes the principles and elements of Precision Farming. The paper also describes the scope and prospect of Precision Farming for small farm Agriculture in India. [4] Three major challenges are: To feed growing population, reduction of rural poverty and manage natural resources. Precision agriculture is an efficient method for crop production. Inputs are optimized leading to reduced cost and environmental impact. The advent of Big Data analytics is changing some of the current knowledge paradigms in Science as well in Industry. Although, the term and some of the core methodologies have been around for many years, the continuous price reduction of hardware and some services (e.g. cloud computing) are making more affordable the application of these methodologies to almost any Research Area being developed in Academic Institutions or Company Research Centers. [5] Many of agricultural-work pursuers are recording agricultural-work record on the paper basis. These situations based on paper recording has difficulty to check the past agricultural-work data and the cost control of agricultural products, compared to electrically stored data. Many farmhouses want to manage agricultural-work information, in order to grasp the cost price. Furthermore, the needs for the systematization of cultivated land management from many farming people in Japan. [6] The aim of this paper is to reveal the current situation regarding the use of mobile phones as a mean of information sharing by Cypriot farmers. In particular, a project at the Agricultural Research Institute is underway, to survey methods currently used for agricultural information and knowledge sharing, to determine the level of satisfaction of the farmers of the available sources of information, and to suggest how ICT tools can be applied to help in transferring agricultural knowledge to farmers who live and work in rural and remote areas. [7] The study examined the factors associated with the level of access of Agricultural Extension officers in Kwara State to Information and Communication Technology (ICT). It also identified the constraints to the usage of ICT for the purpose of sourcing agricultural information Apart from the general constraints to the use of ICT such as, high cost of personal computer, inadequate electricity supply and poor internet access, poor training and technological knowhow were also identified as specific constraints faced by the EAs.

3. PROBLEM STATEMENT

Due to the critical challenges facing in the agriculture sector, farmers feel more forced to adopt intensive farming practices and sustainable agricultural ones, in order to increase both economic and environmental costs.

Farmers predicting manually the demand of grains and vegetable so its effect to formers economically to overcome this problem we are proposed machine learning auto demand and yield prediction process.

A. Existing System

The manual prediction is difficult to farmers. Farmers are usually enabled to recall exactly when and for how much they have made investments and how long those investments were to be utilized. The amount of stored information has been enormously increasing day by day which is generally in the unstructured form and cannot be used for any processing to extract useful information. It requires to doing all the calculations manually which are complex, Critical, more time consuming and always prone to error. Disadvantage of Existing System • Less accuracy • Farmer will get loss economically

B. Proposed System

Propose an effective data mining technique based on profiling (PHANI), to improve their traditional decision-making process using Linear Regression algorithm we are suggesting the what type grains and vegetable former has to cultivate to get more profit.

In this module admin will upload the previous 10-year dataset like

- vegetable demand,
- crops demand
- vegetable yield dataset monthly wise
- vegetable demand dataset monthly wise
- crops yield dataset
- District wise Crop Demand data

Admin will upload all the above datasets using the Excel API and Calculates the Predicted Output Using Linear Regression Algorithm.

New Demand Calculation process

In this module we predicting the Yield from Yield Dataset Using the Linear Regression Technique Let NYLD and No acres former going to cultivate NACR and calculate the new demand.

NewDemand1=NYLD*NACR

Demand Prediction Process

In this module we predicting the demand from Demand Dataset Using the Linear Regression Technique Let NewDemand2.

Land Suggestion Process

We are comparing the NewDemand1 and NewDemand2 if NewDemand2 is greater than NewDemand1 admin will give permission to cultivate.

Advantages of Proposed System

- In this context, varieties of terminologies and techniques have been done to make agricultural practices more efficient.

- Using these advanced technologies to facilitate crop management, minimize losses and maximize yields.

4. SYSTEM DESIGN

The below figure 1.1 contains the System Architecture of the Project. That contains two phases as admin and the farmer phase. With the help of the internet we can access and predict the values of crops to specified field which predicts that which crop can be grown in the specified field.

In Farmer Interface the farmers can upload their Land details so they admin suggests the good crop production which get high yields and the suggestion report. In Main server the Aadhar and Phani details are maintained and by using this Details the admin suggests the good crop production using the necessary Data Mining Technique.

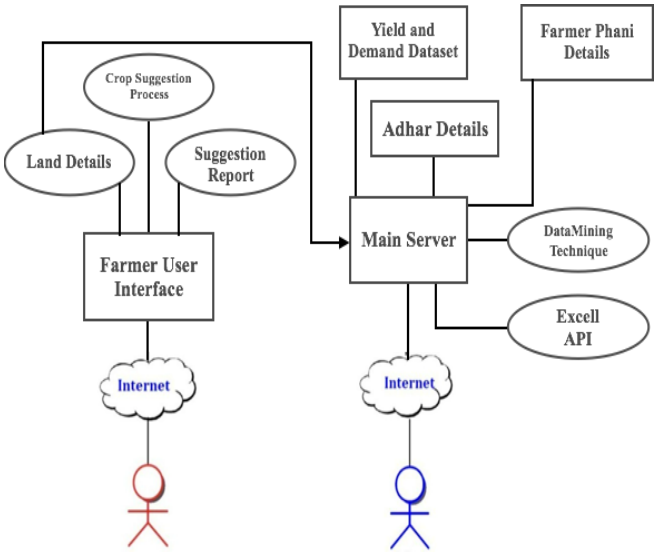

Figure 1.1 Big Data Framework

5. IMPLEMENTATION

5.1 Model View Controller or MVC

As it is popularly called as a software design pattern for developing web applications. The figure 1.2 below is a Model view Controller. Model–view–controller is an architectural pattern commonly used for developing user interfaces that divides an application into three interconnected parts. This is done to separate internal representations of information from the ways information is presented to and accepted from the user. The MVC design pattern decouples these major components allowing for efficient code reuse and parallel development.

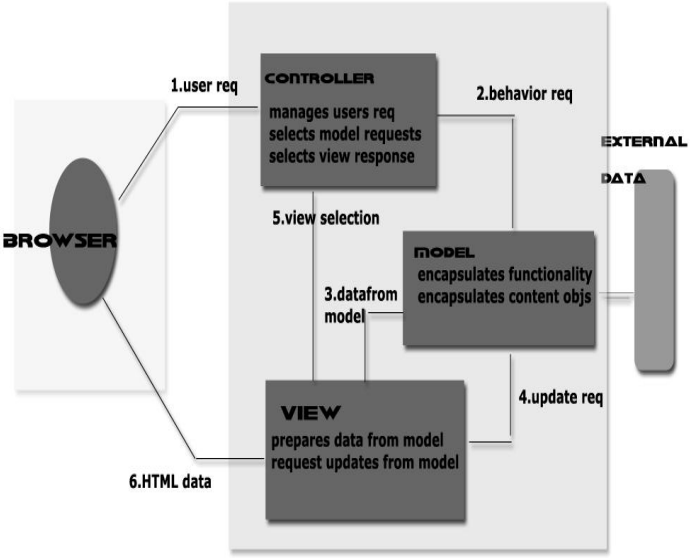

Figure 1.2. MVC Architecture

Traditionally used for desktop graphical user interfaces (GUIs), this architecture has become popular for designing web applications and even mobile, desktop and other clients. Popular programming languages like Java, C#, Ruby, PHP have MVC frameworks that are used in web application development straight out of the box. A Model View Controller pattern is made up of the following three parts –

i. Model − the lowest level of the pattern which is responsible for maintaining data.
ii. View − this is responsible for displaying all or a portion of the data to the user.
iii. Controller − Software Code that controls the interactions between the Model and View

5.2 Linear Regression Algorithm

Linear Regression is a machine learning algorithm based on supervised learning. It performs a regression task. Regression models a target prediction value based on independent variables. It is mostly used for finding out the relationship between variables and forecasting. Different regression models differ based on – the kind of relationship between dependent and independent variables, they are considering and the number of independent variables being used.

- Simple linear regression is a type of regression analysis where the number of independent variables is one and there is a linear relationship between the independent(x) and dependent(y) variable.
- Training Dataset and Testing Dataset gives a Linear Output of Predicted Values For the Crops.

Algorithm

Step1

Let
double[]year_data = {2008,2009,2010,2011,2012,2013,2014,2015,2016,2017};
double[]demand_data={18,24,32,40,58,72,80,93,102,118};

Step2

SumX= ∑ year_data

SumX2= SQRT(year_data);

SumY= ∑ demand_data

Step3

doublexbar = sumx / n;

doubleybar = sumy / n;

Where n=no of the year.

Step4

xxbar += (year_data[i] - xbar) * (year_data[i] - xbar);

yybar += (demand_data[i]-ybar) * (demand_data[i] - ybar);

xybar += (year_data[i] - xbar) * (demand_data[i] - ybar);

double beta1 = xybar / xxbar;

double beta0 = ybar - beta1 * xbar;

Step5

predicted_value=(beta1*(present_year))+beta0;

6. RESULTS AND DISCUSSION

The Admin maintains and manages all the required user data and profile as shown in figure 1.3 below.

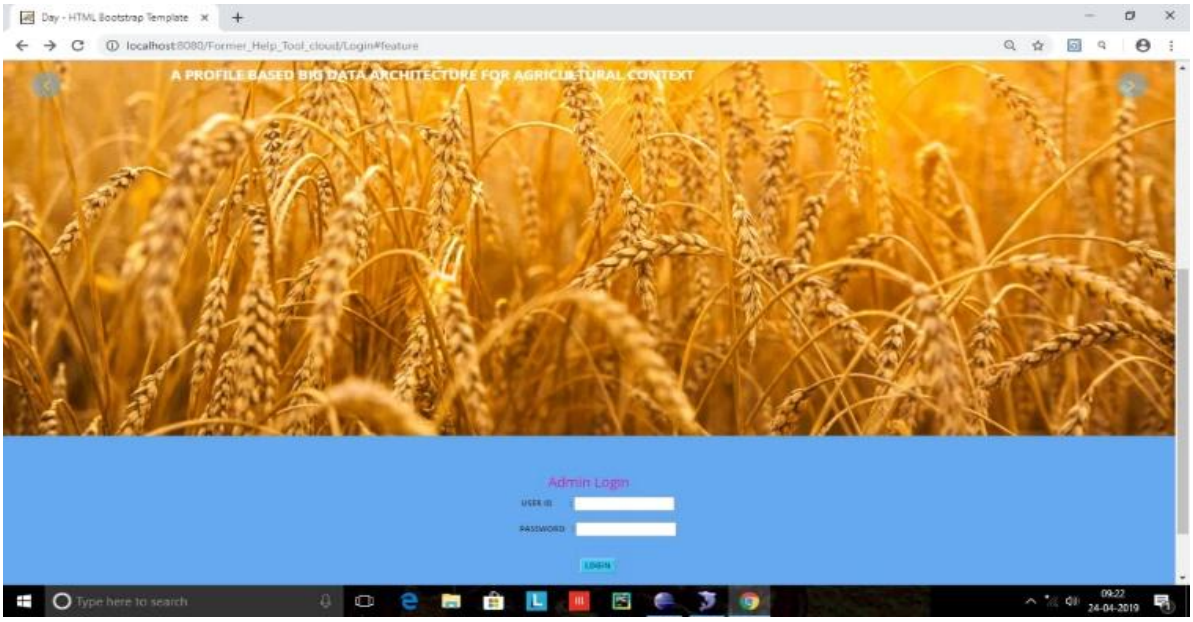

Figure 1.3 Admin page

Once user logged in, they can add their related land details to the system which contains detailed information of their land as shown in figure 1.4 below.

Figure 1.4 Land Details

Uploading the previous year data set to predict which crop yields better production is shown in below figure 1.5.

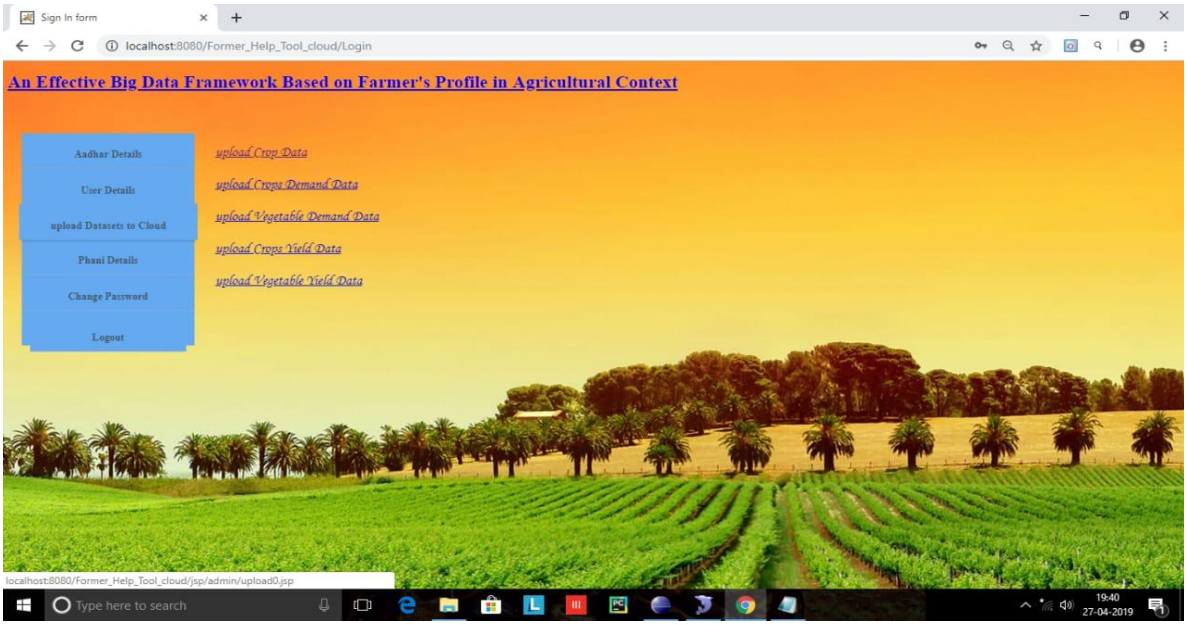

Figure 1.5 Uploading Data Sets

The farmer or user uploads their Phani details of their land such as Survey number, Total land and Soil type as shown in below figure 1.6.

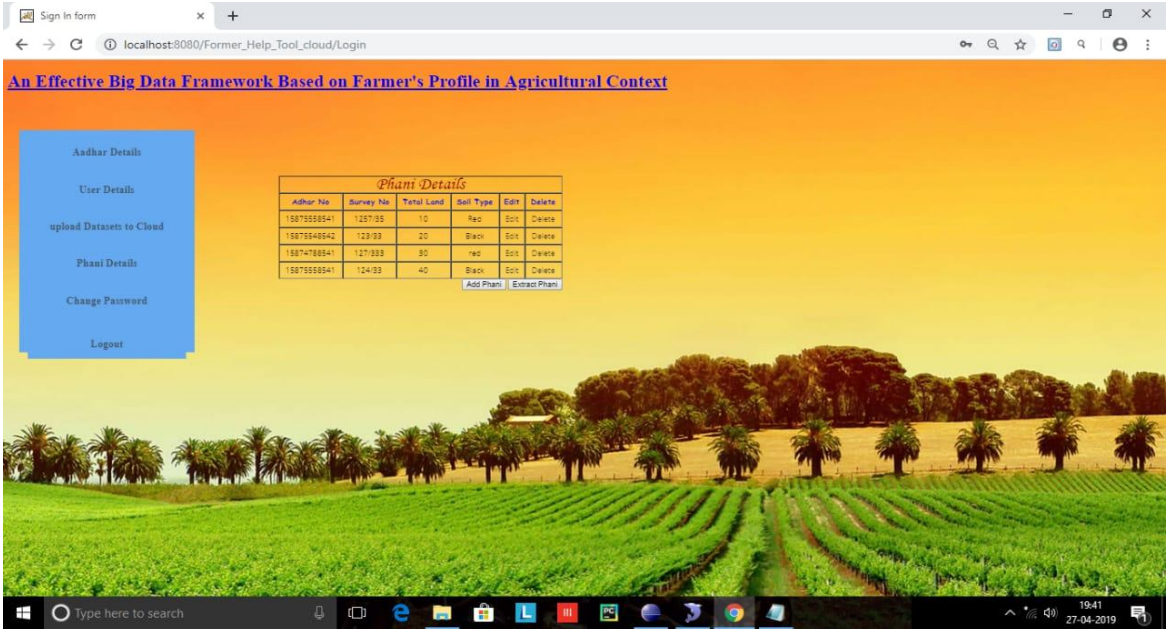

Figure 1.6 Phani Details

The list of Users who are registered to farmer profile which shows the user details such as username, email, address, audhaar details etc. as shown in figure 1.7 below.

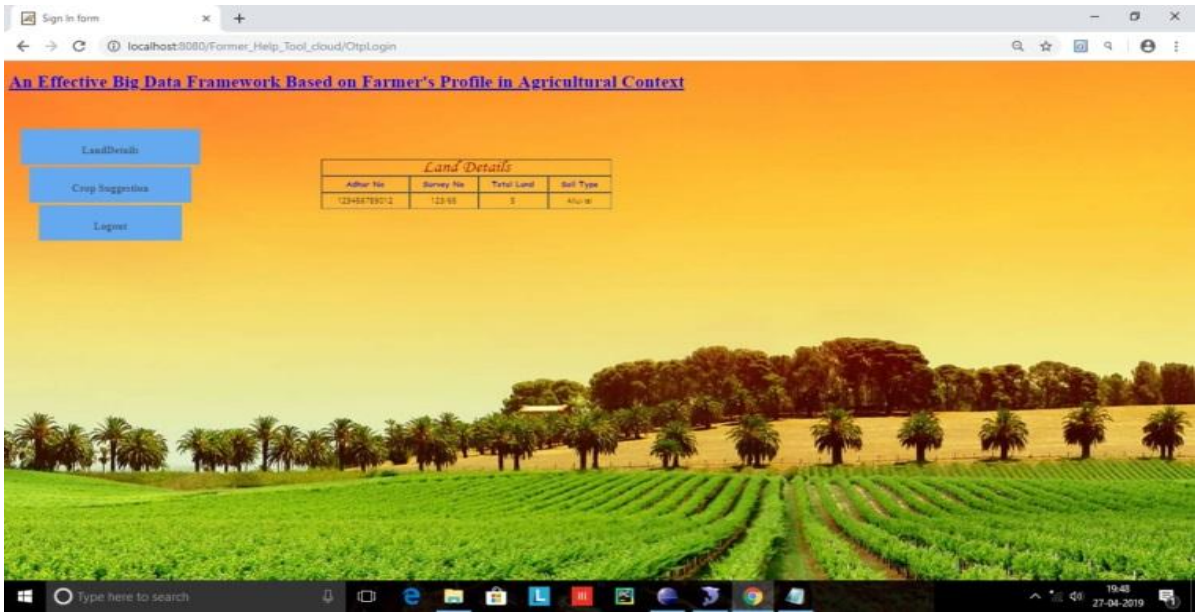

Figure 1.7 Farmer Profile

It suggests the related crops according to the predicted data set of previous years which is shown in figure 1.8below

Figure 1.8 Farmer Profile

7. CONCLUSION

In the present paper, we proposed an effective big data framework based on profile to manage agricultural data within a cloud computing architecture. The environmental data provided by precision agriculture information technologies which represents a crucial source of data in need of being wisely managed and analyzed with appropriate methods and tools in order to extract the meaningful information. Our approach guides the big data providers and the various agricultural actors to identify and select the best services adapted to their specific needs. By using our approach, the agricultural actors can easily integrate the big data customers or researchers to benefit from the advantages of big data technologies. Form this approach the farmers can also get the maximum profit.

ACKNOWLEDGEMENT

The authors of the paper would like to thank and express their gratitude to the anonymous reviewers for their valuable comments and suggestions which has greatly improved the quality of the paper and also we acknowledge the authors referred in the references.

REFERENCES

[1] N. Alexandratos, J. Bruinsma," World agriculture towards 2030/2050", the 2012 revision, ESA Work Pap, vol. 3, 2012.
[2] M. Amini, N. Sadat Safavi, S. Sohaei, and S. M. Noorbakhsh, Agricultural Development In IRAN Base On Cloud Computing Theory, Int. J. Eng. Res. Technol. IJERT, vol. 2, no. 6, pp. 796–801, 2013.
[3] R. D. Grisso, M. M. Alley, and G. E. Groover, Precision Farming Tools. GPS Navigation, 2009.
[4] JosephF.Coates, Agriculture in 21st century, 1995.
[5] SlametKristantoTirtoUtomo, Yukikazu Murakami, Keita Hosono, Kazuhiro Shigeta, "Proposed of cultivation management system with information and communication technology,2013
[6] Omotesho, K. F., Ogunlade, I. O., Muhammad Lawal,"Effectiveness of Knowledge Management Approaches on Agriculture,2015
[7] R. D. Ludena, A. Ahrary et al., Big Data approach in an ICT Agri-culture project, in Awareness Science and Technology and Ubi-MediaComputing (iCAST-UMEDIA), 2013 International Joint Conference on,2013, pp. 261–265.

[8] A. F. McCalla, Challenges to world agriculture in the 21st Century, Update Agric. Resour. Econ.Univ. Calif. Davis, vol. 4, no. 3, 2001.

[9] Aqeel-ur-Rehman, A. Z. Abbasi, N. Islam, and Z. A. Shaikh, A review of wireless sensors and networks' applications in agriculture, Comput. Stand. Interfaces, vol. 36, no. 2, pp. 263–270, Feb. 2014.

[10] S. Sonka and I. IFAMR, Big Data and the Ag sector: More than lots of numbers, Int. Food Agribus. Manag.Rev., vol. 17, no. 1, p. 1, 2014.

[11] S. Sakr, A. Liu, D. M. Batista, and M. Alomari, A Survey of Large Scale Data Management Approaches in Cloud Environments, IEEE Commun. Surv.Tutor., vol. 13, no. 3, pp. 311–336, 2011.

[12] A. Fernández, S. del Río, V. López, A. Bawakid, M. J. del Jesus, J. M. Benítez, and F. Herrera, Big Data with Cloud Computing: aninsight on the computing environment, MapReduce, and programming frameworks: Big Data with Cloud Computing, Wiley Interdiscip. Rev. Data Min. Knowl. Discov., vol. 4, no. 5, pp. 380–409, Sep. 2014.

[13] X. Li, Z. Liu, and H. Jifeng, A formal semantics of UML sequence diagram, in Software Engineering Conference, 2004. Proceedings. 2004 Australian, 2004, pp. 168–177.

[14] P. P. Jayaraman, D. Palmer,A. Zaslavsky, A. Salehi, and D. Geor-gakopoulos, Addressing Information Processing Needs of Digital Agri- culture with OpenIoT Platform, Interoperability and Open-Source Solutions for the Internet of Things. Springer International Publishing, 2015. pp. 137-152.