Development of Smart Crop Production System using Big Data- A Review

Shikha Ujjainia[1], Pratima Gautam[2], S.Veenadhari[3]

[1]*Research Scholar, AISECT University, Bhopal, Madhya Pradesh, India*
[2, 3] *Associate Professor, AISECT University, Bhopal, Madhya Pradesh, India*

Abstract **- The main source livelihood of many people is agriculture. Approximately 70 % of the people directly rely on agriculture as a mean of living. Development in agriculture may also increase savings. The rich farmers we see today started saving particularly after green revolution. This surplus quantity may be invested further in the agriculture sector to develop the sector. Green revolution began in India with an objective to give greater emphasis on Agriculture. The introduction of improved methods of agriculture and high yielding varieties (HYV) seeds, mainly wheat, had resulted into remarkable improvement in agricultural outputs. And all this was just because of technology. In this paper, we propose some novel method to find the most favorable conditions based on geographic factors for growing crops which helps in maximum crop production.**

I. INTRODUCTION

Everyone seems to be talking about big data, but what is big data really? " Big data is a term that describe massive data which can be structure and unstructured form ". Data is increasingly almost everywhere either it is online or offline. And this increased data we termed as " Big Data ". The term big data is also used to capture the opportunities and challenges facing all researchers in accessing, managing, analyzing and integrating database of diverse data type [1].

Due to advent of new technologies, devices and communication means like social networking sites, the amount of data produced by mankind is growing rapidly every year. The amount of data produced by us from beginning of the time till 2003 was 5 billion gigabyte. This rate is still growing enormously [2]. From 2005-2020, the amount of data is predicted to increase 300 times, from 130 Exabyte's to 40000 Exabyte's. There are some common source from where data are generated like mobile phones, audio, video, internet search, social network, document, scientific research, finance and business informatics, government, photography, click streams and so on.

It is not the particular data type which we want, but we can retrieve data by some computational tools like Hadoop, Cloudera, MongoDB etc. In order to access big data, we need program which extend multiple physical and/or virtual machines working together in order to process all the data in acceptable amount of time. Big data technologies are important in providing more accurate result resolutions, which gives detailed decision making results with grater operational efficiencies, cost and risk reduction for the business.

Sources of Big Data

To analysis big data we would require technologies that can store, manage and process huge volumes of structured, semi-structured and unstructured data. So looking into the technologies that can handle the big data, we categorized technologies into following-

- Operational big data
- Analytical big data

Operational big data - It can be defined as the tools which provides real time operational capabilities like searching, processing etc, where data is already captured and stored. It is mainly used by end users. MongoDB technology is comes under it which uses noSQL method to compute data efficiently and inexpensively. This makes operational big data easier to manage and faster to implement, with least coding.

Analytical big data - Big data analytics is slightly different from operational big data as it provide framework for using data mining techniques to analyse data, discover patterns and then enhance the performance of business processes by the analytical models.

It uses MapReduce method that provide new means of analyzing data that is reciprocal to the capabilities provided

by SQL. A system based on MapReduce that can be scaled up from single server to thousands of machine.

(i) Big Data Concept - Big data explicate in various ways. But there are three "V" features are there common characteristics. These are-

- Volume
- Velocity
- Variety

Volume - Of course volume of data is first and obvious important characteristics, because data is now more than text data. We can find this data in the format of videos, images, music etc**.**

Velocity - We can define velocity as data collected in real time. Time is moving so fast and everyone want to get update with latest news. So it depends how fast we receive data.

Variety - It shows different types of data and data can be stored in multiple format like- JPG, Excel, Word, Txt etc.

Farming system in India is strategically utilized, according to the locations where they are most suitable. The farming system that significantly contribute to the domestic GDP of India are subsistence farming, organic farming and industrial farming. Regions throughout India differ in types of farming they use. Smart Farming has a real potential to deliver a more productive and sustainable agricultural production, based on a more precise and resource-efficient approach.

At present, we are on the verge of a technological revolution, called the Internet of Things (IoT). The IoT term was coined by Ashton in 1999 [3]. The internet of things (IoT) is the system of interrelated computing devices, mechanical and digital machines that has ability to transfer data over a network without requiring human to human or human to computer interaction. The way to increase the quantity and quality of agricultural production using sensing technology is called "Smart Farming".

In the agriculture sector, ICT plays an important role to provide new technologies for data generation, transformation and management. It provides information to farmers through mobile app, Short Message Service(SMS) services, agriculture knowledge hub and new generation web applications.

Smart Farming applications do not target only large, conventional farming exploitations, but could also be new levers to boost other common or growing trends in agricultural exploitations, such as family farming (small or complex spaces, specific cultures and/or cattle, preservation of high quality or particular varieties,…), organic farming, and enhance a very respected and transparent farming according to society and market consciousness. Smart Farming can also provide great benefits in terms of environmental issues, for example, through more efficient use of water, or optimization of treatments and inputs.

The objective of this study is to reduce the problems faced by Indian farmers which is sporadic monsoon in India and unawareness about various technologies which will advantageous for our farming. To addressing the above problems, we are proposing a relationship between environmental variables like temperature, rainfall, humidity and soil moisture to-

(a) Developing smart farming system.

(b) Maximize crop production.

(c) Minimize overhead cost.

Typically we are developing wireless connection between devices which offer real time data for smart decision making.

II. RELATED WORK

In [4], the author designed a system which deals with environmental variable like temperature, humidity, water level and pH etc as well as desert specific challenges such as dust, sandy soil, constant wind etc to maintain the adequacy of the agriculture environment. This system is also providing distant access. The system is modeled with Phidget interface8/8/8/ (The phidget is a device which connect analog sensors, control switches, and check the push buttons. One board handles up to 24 devices.) which connects to the various sensors like temperature sensor, pH sensor and soil moisture sensor that will monitor the activities according to environment. Then software will analyze these various values and work will done according to that. So this system is proposed to maximize the productivity at low cost.

In [5], Arduino is an android software that works on two parts, hardware and software. The main focus of this software is on addressing the two important problems of watering plants i.e. the number of times to water the plants and quantity of water to be supplied. In hardware, they used soil hydrometer sensor which will detect soil moisture level and send to the Arduino software. And the software compare that moisture level value with the database value for a particular plant. The main purpose of this system is to provide better farming with the technology.

In precision agriculture, real time and historically generated data is collect in structured and unstructured datasets. As precision agriculture generates more data in the unstructured form and current research is trend is to find knowledgeable information from them[6].

The main goal of this study is to increase the accuracy of the forecasting using different weather parameter such as rainfall, temperature etc for the future precision agriculture. It uses ICT components to collect data and applied preprocessing tools to analyze and manage these data. Mainly they used the MapReduce programming technique for analysis of data and also minimize time of execution and produce as early as results for decision making in precision agriculture. Machine learning algorithms such as K-means,

decision tree, clustering, K-nearest neighbor search are used for development of predictive model on big amount of data.

In [7], after observing the problems of farmers worrying about their productivity this paper proposes an approach combining the emerging technologies such as IoT and Web Services in order to help the farmers using modernization techniques. This paper present the use of IoT device such as sensors to sense environmental data which will further analyze and use for decision making.

This paper introduced construction of sensors with Cloud-connected, wireless system in this crop yield maximization, which automates agriculture tasks and offer real time monitoring for smart decision making. On the other hand, there is the app for farmer which provide interface to access any information about the field.

In this paper, author mentioned about a connect farm, which aims to provide suitable environment for growing crops based on the IoT systems[3]. They connected all sensors with a gateway installed with a device software platform called *Cube*, whereas *Mobius* communicate with IoT service server.

The Cube is a device software platform which can be installed into IoT gateways, Whereas the Mobius is an IoT service platform complying with one M2M (Machine-to-Machine) standard[3].

The three main components used to design the connected system are –

IoT devices - It include monitoring sensor and controller use to record environmental variables.

Cube - It is IoT gateway.

Mobius - It is IoT service platform.

Working system of the connected farm as follows –

- Each device such as sensor and controller are used in the connected farm has to be registered into the Mobius using Cube.
- After successful registration, every device can have it virtual representation in the Mobius.
- Now in connected farm, sensor gather environmental data and send it to Cube.
- After that, Cube transmits the same data to the Mobius.
- Finally end user can monitor their connected farm data using smart phone or tablets.

So the aim of this system is to make connect farm using IoT devices that monitor farm activities which can be control by farmers with their personal devices like smart phones.

In [8], paper presents the prediction of yield using weather data such as air pressure, temperature, humidity etc. Author uses MapReduce method to process the data and apply nearest neighbor modeling to discover similar pattern.

III. EXPECTED OUTCOME

After studying the above review papers, the keen work is going on to make our agriculture system more contemporary and smart. Seeing this greatfull work, we are expecting to develop accurate relationship between various environmental variables which will give us possible conditions to increase crop production in different situations. So following idealize outcomes will be-

- Indentify the climate conditions for a particular crop
- Maintain the climate conditions for crops
- Maximize crop production.
- Reduce efforts to maintain the farming system
- Minimize the cost and energy of production.

REFERENCES

[1]. Hua Fang, Zhaoyang Zhang, Chanpaul Jin Wang, Mahmoud Daneshmand, Chonggang Wang and Honggang Wang " A Survey of Big Data Research " Journal of IEEE Network Volume:29, Issue: 5, September-October 2015.

[2]. Minwoo Ryu et al. " Design and Implementation of a Connected Farm for Smart Farming System " Journal of IEEE, SENSORS, pages:1-4,2015.

[3]. Aalaa Abdullah et al. " AgriSys: A Smart and Ubiquitous Controlled Environment Agriculture System " at 3rd MEC International Conference on Big Data and Smart City, pages: 1-6, 2016.

[4]. Ankita Patil et al. " Smart Farming using Arduino and Data Mining " at International Conference on Computing for Sustainable Global Development (INDIACom),pages: 1913-1917,2016.

[5]. M.R. Bendre et al. " Big Data in Precision Agriculture : Weather Forecasting For Future Farming " at 1st International Conference on Next Generation Computing Technologies (NGCT-2015),pages: 744-750, 2015.

[6]. M.K. Gayatri et al. "Providing Smart Agriculture Solutions to Farmers for better yielding using IoT " at IEEE International Conference on Technological Innovations in ICT for Agriculture and Rural Development (TIAR 2015), pages: 40-43, 2015.

[7]. Wu Fan[1] et al. " Prediction of crop yield using big data " Journal of 8th International Symposium on Computational Intelligence and Design, Volume: 1, Pages: 255-260, 2015.

[8]. Laizhong Cui Laizhong Cui, F. Richard Yu, and Qiao Yan " When Big Data Meets Software-Defined Networking : SDN for Big Data and Big Data for SDN " Journal of IEEE, Volume: 30, Issue: 1, January-February 2016.