# Big Data Analytics for Agricultural Development in India

Arjun S
Information Science and Engineering Department
AMC Engineering College
Bangalore, India

Anish Joshi
Information Science and Engineering Department
AMC Engineering College
Bangalore, India

H Pooja Das
Information Science and Engineering Department
AMC Engineering College
Bangalore, India

Amutha R
Information Science and Engineering Department
AMC Engineering College
Bangalore, India

***Abstract*— Agricultural production system is an outcome of a complex interaction of seed, soil, water and agro-chemicals (including fertilizers). In India, dominant parts of provincial inhabitants are reliant on agriculture for their occupation. Be that as it may, the current farming practices are not prudentially suitable neither naturally supportable and the yields of numerous agrarian items in India are basically low. Sooner rather than later, it will get to be the key for the nation to fabricate a high yielding, focused, and shifted agricultural part and speed up country, non-ranch business enterprise. This paper recognizes the culminations of conventional cultivating practices and delivers how to build the yield of the agrarian things by utilizing present day PC innovations. Further, it additionally recognizes the basic figuring and analytic capacity of Big Data in preparing colossal volumes of value-based information continuously circumstances. The target of this paper is to display the revisions in the rural area and supports the dialogs on how government can cultivate developments in the enormous information examination to enhance the rustic agricultural framework.**

***Keywords*— *Big Data Analytics, Rural agrarian systems, Precision Agriculture, Electronic farm records.***

## I. INTRODUCTION

India is a dauntless nation with more than billion or more individuals, furthermore one of the world's quickly thriving economies. Out of the tremendous populace, 58.4% are open agrarian collection. India's late achievements in harvest yields while being noteworthy, are still only 30% to 60% of the best product yields achievable in the ranches of created and in addition other creating nations [1].

Today, India positions second worldwide in ranch yield. Agriculture and united divisions represented 13% of the aggregate GDP in 2014; around half of the aggregate workforce. Agriculture is geographically the broadest area and assumes a typical part in the thorough financial habitation of India. The existent report states that the agriculture area keep on trailing for India [1]. The Financial enlargement of agriculture to India's GDP is in a roundabout way declining with the nation's monetary development growth.

*Precision agriculture (PA)* is a cultivating administration idea in light of watching, measuring and reacting to bury and intra-field variability in harvests. Crop variability regularly has both a spatial and transient segment which makes measurable/computational medications entirely included [1].

Some agribusiness specialists have seen the interest for sensor information joining in rural framework and realized how to conquer the reasonable suggestions including the government and its approaches. The information investigators or the information inspector, farming experts and others utilizing this innovation lay forward how to achieve better zenith at lower costs.

The widespread advancement of information over the previous decade has started a one of a kind domain in the space of data innovation and information science called Big Data. This sharp innovation is procuring heightening thought to overhaul the execution of farming frameworks by joining diverse frameworks by means of information and accord stage to decrease repetitive yield disappointments, upgrade and hurry agricultural overseeing, and allow section to all stratum of rural administrations for an extensive variety of circumstances. It accumulates all feasible product data created through electronic shrewd gadgets for a point by point region. These shrewd gadgets will produce enormous measures of information, incited by record keeping, consistence and administrative prerequisites, which are considered as Big Data.

According to the above exchanges, the e-Agriculture administration information can be considered as a Big Data due to its assortment of information with enormous volumes streaming with high speed. A portion of the answers for the e-Agriculture service big data include the predominant current technologies like HDFS, Map Reduce, Hadoop, STORM etc.

The accompanying are the key focuses that make the execution of farming frameworks better and build efficiency:

1. Measure, store and dissect the data to enhance yield quality.

2. Manage revenue costs by decreasing crop failure's probability.

3. Enhance preventive care and increase producer-consumer satisfaction

Adoption of big data in agriculture fundamentally diminishes the likelihood of crop failure and farmer's primary concerns and recommends the soil sensing and crop yield information to be stored in data centers.
This paper is organized as follows: in section II, we describe the present farming system in India, section III discusses about big data in general, section IV focuses on the problems in the existing agricultural system, section V tells the use of the big data analytics in agricultural system and section VI provides description about technologies for precision agriculture and
VII concludes the work.

## II. PRESENT FARMING SYSTEM IN INDIA

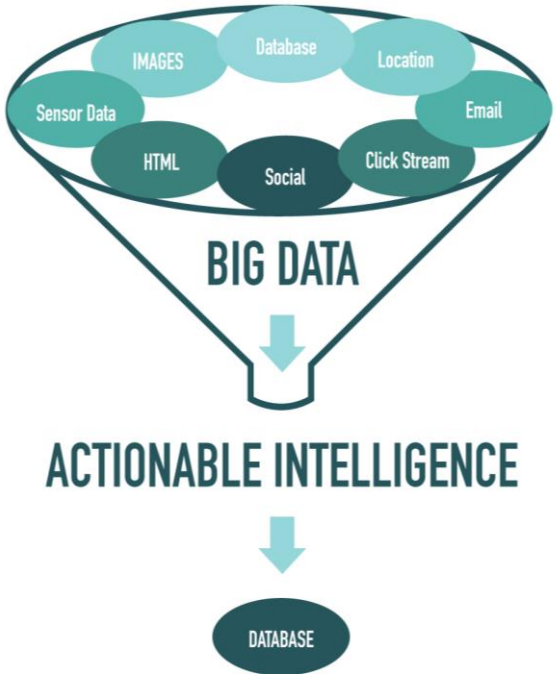
Farming Systems in India **is** adapted prudently, as per their most suited locations. That is harvests are developed as indicated by the ranch or the dirt conditions present over a specific zone or the area. The districts in India fluctuate as per the sorts of cultivating they utilize; some depend on agriculture, agro ranger service, and some more. India's geological site, cause diverse parts to encounter unmistakable atmospheres, which influences each district's agricultural profitability particularly. India holds the second position in agricultural generation on the planet at the present day. In 2007, agriculture and other industry made up more than 16% of India's GDP. In spite of the enduring decrease in farming's commitment to the nation's GDP, India agribusiness is the greatest business in the nation and assumes a key part in the financial development of the nation. It is additionally the second greatest gatherer of vegetables and organic product, speaking to 8.6% and 10.9% of general creation, individually. India likewise has the greatest number of domesticated animals on the planet, holding 281 million [1].
Around one-sixth of the area region experiences genuine harvest yielding issues, for example, disintegration, water logging, aridity, causticity, saltiness, and alkalinity. Around, soil protection measures are required for as expansive as 80 super hectare of developed territory. After the watering system practices were presented, inside of a couple of years, the issues of saltiness and water logging had stimulated. Obviously 7 Mega hectare of area are influenced by alkalinity and salinity [1]. According to the soil conditions of the particular farm area the crops to be cultivated are decided based on the moisture content, humidity, degree of nutrients present etc. So it is very important to keep records of all the soil quality properties (for which the big data can be used).

## III. BIG DATA IN GENERAL

Big data is an evolving term that describes any voluminous amount of structured, semi-structured and unstructured data that has the potential to be mined for information [3]. Big data is a set of techniques and technologies that require new forms of integration to uncover large hidden values from large datasets that are diverse, complex, and of a massive scale [2].

The processing of such data using common database management tools or conventional data processing applications is a very strenuous task. Everything around us is contributing to the generation of big data at every time instance. Every social media exchanges as well as digital processes involving systems, sensors, mobile devices and every other digital device present are transmitting it. To extract meaningful value from big data, you need optimal processing power, analytics capabilities and skills. Big data is creating a new culture in which business and IT leaders must join forces to realize value from all data. But for the proper understanding of this continuously growing data we need a new fundamental approach to architecture, tools as well as practices.

The boundless Indian agrarian framework should tackle agribusiness "Big Data" by deciphering an intricate arrangement of information, including electronic ranch records and sensor information. This empowers agriculturists to get to and examine horticulture's enormous information to find out quality, decide best practice, survey treatment methodologies and distinguish crops which are at danger.



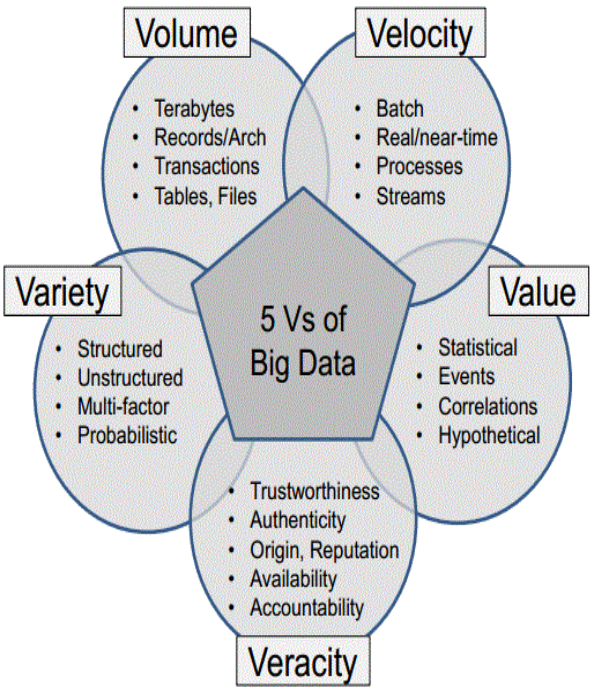

Despite the exponential development of information the changing client conduct and globalization is likewise in charge of coordinating Big Data. In this manner, numerous associations are looking for examining such models to upgrade their working.

The typical characteristics of the big data are:

*Volume:* The amount of data generated as Big Data ranges from Terabytes to Exabyte's and Zettabyte's of data. The volume has been increasing exponentially: up to 2.5 Exabyte

Of data is already generated and stored every day. This is expected to twofold by the end of 2015[4].

*Velocity*: Big data is growing rapidly, generating a bizarre of quantities needed to be stored, transmitted, and processed rapidly. It deals with the pace at which data flows in from sources like business processes, machines, networks and human interaction with things like social media sites, mobile devices, etc [9].
*Variety*: Variety refers to the many sources and types of data both structured and unstructured. We used to store data from sources like spreadsheets and databases. Now data comes in the form of emails, photos, videos, monitoring devices, PDFs, audio, etc [9].
*Veracity*: Veracity alludes to the predispositions, commotion and irregularity in information. Exactness of examination relies on upon the veracity of the source information. In contrast with Big Data's volume and speed, veracity is the most difficult trademark in information examination.
*Value*: if we are producing so much information, then what amount of it is really valuable? Understanding the estimation of this information is critical so that the procedure can be streamlined. Putting resources into Big Data arrangements is costly so it is vital to guarantee that the information being removed is of quality to the organization [9].




## IV. EXISTING ISSUES IN PRESENT AGRICULTURAL SYSTEM IN INDIA

More than 68% of India's aggregate populace dwells at the rustic zone, and around three-fourth of the inhabitants of provincial zones are subject to agriculture for their job [1].

Mr. Gopal Naik, a teacher in the region of financial matters and sociologies and Chairperson for the Center for Public Policy at Indian Institute of Management, Bangalore, in one of his meeting has brought up some basic issues. Right now, the issues that beset the Indian farming are the inadequacy of legitimate learning and framework in the rustic regions. Issues identified with watering system, market foundation and transport frameworks add huge expense to agriculturist's operations. Another issue is absence of conveyance systems. There are various plans pointed towards creating horticulture. We don't have successful conveyance systems that can make an interpretation of those into powerful help at the ground level, as far as expanding efficiency or diminishing cost or expanding value acknowledgment. Lacking government support worsens these issues.

Government failure is a noteworthy worry in farming in light of the fact that the high dangers included make help and assistance vital. Like some other business venture, horticulture is subjected to high dangers on account of the unpredictable
Way of the elements included. For example, climate is frequently an issue - you have dry spells in one year and overwhelming downpours in the following. In both cases, ranchers miss out; thus they need to search for a typical period to profit. Government, along these lines, needs to assume a noteworthy part in giving backing to agriculturists. This is genuine everywhere throughout the world and there is not really any nation where government mediation is not show. There may obviously be varieties in the degree of intercession; yet in the event that you check the circumstance in many nations or locales, including created ones like the US, Canada and the European Union, you see a significant mediation by the administration. Along these lines the government assistance is the key for sound agrarian advancement [1].

## V. USE OF BIG DATA ANALYTICS IN AGRICULTURAL SYSTEMS

We have gone into the time of huge information. Enormous information gives a ground to gathering, putting away and breaking down information to uncover the data not already known by astutely utilizing the ever-increasing amount of data available we could grow new vision by reexamination of the information or combining it with other accessible data. In agriculture this implies not simply mining crop records, precipitation maps, and symptomatic reports and so on. For experiences, judgments and choice bolster gadget, additionally persistent examination of the information streams (fully fledged records) created for and by the predetermined territory at each time moment.

The objective of this paper is to provide top notch Farming methods furthermore means to guarantee expanded efficiency of the yields to rustic individuals and defeat the issues in the rural frameworks like utilization of hurtful pesticides, exorbitant utilization of composts, giving appropriate watering system offices and extortion administration in the horticultural framework. The proposed idea empowers agriculturists, enormous information examiners and staff to have part constructs access to data in light of electronic homestead records.

In detail, the Big data in agriculture refers to the Electronic Farm Records (EFR) which incorporates soil temperatures maps and information, precipitation maps and information, electrical conductivity maps and information, dampness content information, air penetrability maps the supplement substance and pH level information, past development records, protection and yield related data and online networking posts including tweets, websites, new sustains and articles in agribusiness diary. The occupation of the huge information researcher is to mine the huge information and find the affiliations, comprehend examples and patterns to enhance the horticultural frameworks, build crop profitability and lower costs included by appropriate diagnosing the different variables.

## VI. TECHNOLOGIES FOR PRECISION FARMING

In order to collect and utilize information effectively, it is important for anyone considering precision farming to be familiar with the modern technological tools available.
The vast array of tools includes hardware, software and the best management practices. These are described briefly in the following paragraphs.

*Global Positioning System (GPS) receivers:* Global Positioning System satellites broadcast signals that allow GPS receivers to compute their location. This information is provided in real time, meaning that continuous position information is provided while in motion. Having precise location information at any time allows soil and crop measurements to be mapped. GPS receivers, either carried to the field or mounted on implements allow users to return to specific locations to sample or treat those areas. Uncorrected GPS signals have an accuracy of about 300 feet. To be useful in agriculture, the uncorrected GPS signals must be compared to a land-based or satellite-based signal that provides a position correction called a differential correction. The corrected position accuracy is typically 63-10 feet. When purchasing a GPS receiver, the type of differential correction and area coverage should be considered.

*Yield monitoring and mapping:* In highly mechanized systems, grain yield monitors continuously measure and record the flow of grain in the clean-grain elevator of a combine. When linked with a GPS receiver, yield monitors can provide data necessary for yield maps. Yield measurements are essential for making sound management decisions. However, soil, landscape and other environmental factors should also be weighed when interpreting a yield map. Used properly, yield information provides important feedback in determining the effects of managed inputs such as fertilizer amendments, seed, pesticides and cultural practices including tillage and irrigation. Since yield measurements from a single year may be heavily influenced by weather, it is always advisable to examine yield data of several years including data from extreme weather years that helps in pinpointing whether the observed yields are due to management or climate induced.

*Variable-rate Technology (VRT):* Variable rate technology refers to a technology that enables the variable rate application of materials in precision agriculture. Variable rate fertilizer application allows crop producers to apply different rates of fertilizer at each location across fields. The technology needed to accomplish variable rate fertilization includes an in-cab computer and software with a field zone application map, fertilizer equipment capable of changing rates during operation and the Global Positioning System (GPS). VRT describes any technology that enables the variable application of inputs. Therefore, VRT mounted on equipment permits input application rates to be varied across fields in an attempt to site-specifically manage field variability. This type of strategy can reduce input usage and environmental impacts along with increasing efficiency and providing economic benefits.

*Crop scouting***:** In-season observations of crop conditions may include: Weed patches (weed type and intensity); Insect or fungal infestation (species and intensity); Crop tissue nutrient status; Flooded and eroded areas using a GPS receiver on an all-terrain vehicle or in a backpack, a location can be associated with observations, making it easier to return to the same location for treatment. These observations also can be helpful later when explaining variations in yield maps.

*Geographic information systems (GIS):* Geographic information systems (GIS) are computer hardware and software that use feature attributes and location data to produce maps. An important function of an agricultural GIS is to store layers of information, such as yields, soil survey maps, remotely sensed data, crop scouting reports and soil nutrient levels. Geographically referenced data can be displayed in the GIS, adding a visual perspective for interpretation. In addition to data storage and display, the GIS can be used to evaluate present and alternative management by combining and manipulating data layers to produce an analysis of management scenarios.

*Information management:* The adoption of precision agriculture requires the joint development of management skills and pertinent information databases. Effectively using information requires a farmer to have a clear idea of the business' objectives and crucial information necessary to make decisions. Effective information management requires more than record-keeping analysis tools or a GIS.

It requires an entrepreneurial attitude toward experimentation and education.
A large portion of the investigation of the agrarian information is executed by yearly information amusement in social databases that create pre-handled reports. The examination of the information is should be done on the spot. Likewise, information vivification should be done continuously at least once in a month.
To offer better administrations to the general population, the agricultural framework needs to advance and innovate ceaselessly.

The information acquired by the big data analysis can be comprehensively used for precision agriculture. Some of them are listed as:

- The information ingested from the different frameworks can be assessed continuously to signal the fundamental values that assume a vital part underway basic leadership.
- The geological maps in this way produced are of high determination recognizing the variety of soil dampness. It would direct to the definitive utilization of watering system.
- The thorough focusing of controls will likewise be made conceivable with the assistance of unpredictable pictures of nuisance harm in the field.

These utilizations of enormous information can be tried, cleaned and upgraded quickly and financially and will altogether change conveyance and exploration in the rural area.

However, the enormous information investigation in horticulture assumes an urgent part to give better agrarian administrations, it give examination on the authentic information to reveal concealed data. The enormous information investigation has challenges like heterogeneity and fragmentation of information, scale, security and human joint effort [7].

## VII. CONCLUSION AND FUTURE WORK

.Precision agriculture gives farmers the capacity to utilize crop inputs more successfully including manures, pesticides, culturing and watering system. More viable utilization of inputs means more noteworthy harvest yield and/or quality, without dirtying the earth. Be that as it may, it has demonstrated hard to decide the money saving advantages of accuracy agribusiness administration. At present, a number of the innovations utilized are as a part of their early stages, and valuing of hardware and administrations is difficult to bind. This can put forth our current financial expressions around a specific innovation dated. Exactness agribusiness can address both monetary and natural issues that encompass creation farming today. Questions stay about cost-viability and the best approaches to utilize the mechanical apparatuses we now have, however the idea of "making the best choice in the correct spot at the ideal time" has a solid instinctive request.

The advancement of enormous information examination in the agrarian area is prompting the understanding of new enhancing results from these substantial information sets. Presently, the dirt and product detecting has changed the rural framework to wind up a great deal more proficient, less costly and accomplish preferable quality over some time recently. In the mean time, numerous new businesses are creating both physical and satellite-based sensors which are accelerating the move to 'connected' agriculture [8]. The spread of shrewd sensors with the compelling investigation of big data will lead us a step forward towards freeing farmers from the constraints of uncertain weather. The utilization and selection of Big Data, inside legislative procedures, is advantageous and permits efficiencies regarding cost, profitability and advancement. Be that as it may, this information examination frequently requires numerous parts of government (central and local) to work in cooperation and make new and inventive procedures to convey the fancied result [2]. Through the execution of this paper utilizing Hadoop HDFS and Map Reduce we could reveal the data lying in the huge agribusiness information sets. At last, the achievement of accuracy agribusiness depends to a great extent on how well and how rapidly the information expected to manage the new innovations can be found.

The future research is all about to get through the obstacles and use big data analytics in agriculture for unveiling the proficiency from the raw unstructured data.

## VIII. ACKNOWLEDGEMENT

Our sincere thanks to our mentor and guide Assistant Professor Bineet Kumar Jha who helped us in the successful completion of the paper .We are even grateful to our Head of The Department Mr.Selvin Paul Peter J for providing us all the required facilities.

## REFERENCES

[1] http://en.wikipedia.org/wiki/Agriculture_in_India
[2] http://en.wikipedia.org/wiki/Big_data
[3] http://www.ibm.com/big-data/us/en/
[4] Tekiner, F.; Keane, J.A., "Big Data Framework," Systems, Man, and Cybernetics (SMC), 2013 IEEE International Conference on , vol., no., pp.1494, 1499, 13-16 Oct. 2013
[5] http://www.patrickcheesman.com/wp-content/uploads/2014/08/Big-Data-5V.gif
[6] http://www.blogging4jobs.com/wp-content/uploads/2014/04/bigDataFunnel.jpg
[7] Venkatareddy Konasani, Mukul Biswas and Praveen Krishnan Koleth, "Fraud Management using Big Data Analytics", A Whitepaper by Trend wise Analytics.
[8] http://www.atelier.net/en/trends/articles/farmers-starting-capitalize-big-data-technology_424444
[9] http://insidebigdata.com/2013/09/12/beyond-volume-variety-velocity-issue-big-data-veracity/
[10] Rupika Yadav, Jhalak Rathod, Vaishnavi Nair, "Big Data Meets Small SensorsinPrecisionAgriculture", International Journal of Computer Applications (0975 – 8887) Applications of Computers and Electronics for the Welfare of Rural Masses (ACEWRM) 2015.
[11] Anil Kumar Singh "Crop Growth Simulation Models" Water Technology Center, I.A.R.I., Library Avenue, New Delhi - 11 0012
[12] http://cema-agri.org/page/precision-farming-key-technologies-concepts
[13] http://www.precisionag.com/
[14] V.Radha Krishna Murthy "Crop Growth Modeling and its Applications in Agricultural Methodology"