International Journal of Agricultural and Environmental Information Systems
Volume 11 • Issue 1 • January-March 2020

# Big Data Challenges and Opportunities in Agriculture

Maya Gopal P.S., VIT University, Chennai, India
Bhargavi Renta Chintala, School of Computing Science and Engineering, VIT University, Chennai, India

## ABSTRACT

This article reviews various aspects of research concerning the background and state-of-the-art of big data in agriculture. This article focuses on data generation, storage, analysis and visualization in big data. In every phase, technical challenges and the latest advancement are discussed, and these discussions aim to provide a comprehensive overview and complete picture of this exciting area. This survey is concluded with a discussion on the application of big data in precision agriculture and its future directions.

## KEYWORDS

Agriculture, Big Data, Data Acquisition, Data Analytics, Data Storage, Machine Learning

## 1. INTRODUCTION

The United Nations estimates that it is the need to increase the food production to 50 percent by the middle of the current century (FAO, 2009). Agricultural production tripled during the last decades as the world's population doubled (Kitzes et al., 2008). Agriculture remains essentially a primary source of food for the population and raw material for a large number of industries (da Silva et al., 2009). Population growth, climate change and bio-energy crops are worldwide trends that are increasing the importance of using science to improve agriculture (Tilman et al., 2011). With the need to produce more food using fewer inputs, agriculture is seeking new products, practices and technologies. Research activities centering on genomics, bioinformatics and computational biology of plants and animals enable the scientists and organizations to better feed the world and improve the quality of food and animal crops. Progress in agricultural growth can serve as a critical position for designing successful strategies to transform the economy and meet sustainable development (Christiaensen et al., 2010) and the investments in agricultural research play a key role to agricultural growth. Farmers will have the tools to get the most from every acre. The future of farming depends largely on adoption of cognitive solutions. While large scale research is still in progress and some applications are already available in the market, the industry is still highly underserved. When it comes to handling realistic challenges faced by farmers and using autonomous decision making and predictive solutions to solve them, farming is still at a budding stage (Jones et al., 2017). Research on new generation agricultural design models shows that the data is most important parameter for

DOI: 10.4018/IJAEIS.2020010103

on-farm decision support, research investment and policy decision making. The agricultural industry will be transformed by data science and artificial intelligence. Collecting reliable agriculture data for farm management decision making is important scenario. The developments in the concept of smart farming make agriculture more efficient and effective with the help of high-precision algorithms (Baseca et al., 2019). The mechanism used in smart farming is machine learning (ML), the scientific field that gives machines the ability to learn without much programming. It has emerged together with big data technologies and high-performance computing to create new opportunities to ease, quantify and understand data intensive processes in agricultural operational environments. The developments indicate that agriculture can benefit from machine learning at every stage like spices management, field management, crop management and livestock management. The artificial intelligence (AI) and machine learning are used in a number of agricultural applications today include the yield prediction algorithms based on weather and historical yield data, image recognition algorithms to detect pest and diseases in plants and robotics to harvest different types of specialty crops (Tibbetts 2018). This aspect needs an adaptive method to control the data sources and decision-making systems for better production and marketing with less waste of resource. Agriculture big data is playing important role by incorporating the AI and ML. The farmers are using data to calculate harvest yields, fertilizer demands, cost savings and even to identify optimization strategies for future crops as smart machines and sensors on farms and farm data grow in quantity and scope, farming processes will become increasingly data driven and data enabled.

In order to obtain better productivity, the people are using precision agriculture (De Rango et al., 2019), Somayeh et al. (2018) and Maes and Steppe (2019), automated irrigation scheduling (Li et al., 2018; Soulis & Elmaloglou, 2018), optimization of plant growth, farmland monitoring, greenhouse gases monitoring, production process management and security in crops (Yuan et al., 2018; Huang et al., 2018, Groenveld et al., 2019).

## 2. BIG DATA TECHNOLOGIES AND TOOLS

Recent technological development led to automation of several processes in various domains like agriculture, health care, fraud detection, etc., which in turn led to the generation of humungous data. McKinsey & Co. (Manyika et al., 2011)) foresees that the society is now facing a tremendous wave of innovation, productivity, and growth as well as new modes of competition and value capture - all driven by Big Data. The term Big Data is mainly used to describe massive, often unstructured, and heterogeneous digital content which is difficult to store and process using traditional data management tools and techniques Talia (2013), Stephen Kaisler et al. (2015), Rob Lokers et al. (2016)). Once these started gaining attention, the data analyst developed it further and currently big data can be described using the 10V model. Borne (2014) has listed the V's as challenges in deploying Big Data into any application. These V-based characterizations represent ten different challenges associated with the main tasks involving big data like- capture, cleaning, curation, integration, storage, processing, indexing, search, sharing, transfer, mining, analysis, and visualization. Big data can be described by the following 10 characteristics which are illustrated in Figure 1 and Table 1.

These covered most of the challenges in big data including collecting, storing, transferring, analyzing, and visualizing. Big data is primarily defined by the volume of a data set. The data sets are generally huge measuring tens of terabytes and some cases crossing the petabytes. The term big data was preceded by very large databases which were managed using database management systems. Currently, big data falls under three categories of data sets namely structured, unstructured and semi-structured. The structured data sets comprise of data which can be used in its original form to derive results. The unstructured data sets comprise of data are without proper formatting and alignment (Khan et al., 2014). Semi-Structured data sets are a combination of both structured and unstructured data. Big data processing requires a particular setup of physical and virtual machines to derive results. The processing is done simultaneously to achieve results as quickly as possible.

## Related Content

Doubly Fed Induction Generators: Overview and Intelligent Control Strategies for Wind Energy Conversion Systems

Vinod Kumar, Steven Kong, Yateendra Mishra, Z.Y. Dongand Ramesh C. Bansal (2010). *Intelligent Information Systems and Knowledge Management for Energy: Applications for Decision Support, Usage, and Environmental Protection (pp. 147-178).*

www.irma-international.org/chapter/doubly-fed-induction-generators/36966

Corporate Environmental Management Information Systems Influence of Green IT on IT Management and IT Controlling

Andreas Gadatsch (2011). *Green Technologies: Concepts, Methodologies, Tools and Applications (pp. 1408-1420).*

www.irma-international.org/chapter/corporate-environmental-management-information-systems/51769

IoT-Based Framework for Smart Agriculture

Jian Yang, Amit Sharmaand Rajeev Kumar (2021). *International Journal of Agricultural and Environmental Information Systems (pp. 1-14).*

www.irma-international.org/article/iot-based-framework-for-smart-agriculture/275239

Research on Measurement and Control System of Common Parameters of Agricultural Equipment Based on Wireless Transmission

Fan Wu, Aiqin Li, Saihua He, Mohammad Ikbaland Mohamed A. Sharaf Eldean (2021). *International Journal of Agricultural and Environmental Information Systems (pp. 73-86).*

www.irma-international.org/article/research-on-measurement-and-control-system-of-common-parameters-of-agricultural-equipment-based-on-wireless-transmission/275244

Modeling of current and future state of biodiversity in Central America using GLOBIO3 methodology

Denisse McLean R. (2011). *Land Use, Climate Change and Biodiversity Modeling: Perspectives and Applications (pp. 349-375).*

www.irma-international.org/chapter/modeling-current-future-state-biodiversity/53760