*International Journals of Advanced Research in Computer Science and Software Engineering*
*ISSN: 2277-128X (Volume-9, Issue-4)*

**Research Article** | **April 2019**

# Big Data in Agriculture

**Matthew N. O. Sadiku, Suxia Cui, Yonghui Wang**
*Roy G. Perry College of Engineering, Prairie View A&M University, Prairie View, TX 77446, United States*
Email: sadiku@ieee.org; sucui@pvamu.edu; yowang@pvamu.edu

***Abstract:*** ***We live in a big data age where huge volumes of information are generated, collected, and managed to add value to our daily lives. Today, farmers are deploying enormous amounts of data generated by interconnected devices to get an overall understanding of their farms. If tractors transformed farming in the middle ages, the evolution of big data is revolutionizing the agriculture industry in our time. The big data revolution is modernizing the farming industry at a pace we have never witnessed. This paper provides a brief introduction big data in agriculture.***

***Key Words:*** ***big data, agriculture, farming, food***

## I. INTRODUCTION

Agriculture is going through a digital revolution. Increasing agricultural production to feed a growing worldwide population in an environmentally sustainable way is a major challenge of the 21st century. The global population is growing exponentially, while the total land available for agriculture is reducing daily. There is widespread concern that agricultural inputs may have detrimental impacts on human health and environmental quality. For example, pesticides have adverse affect on human health, while fertilizers applied on farms can degrade environmental quality [1].

New technologies such as the Internet of things and cloud computing are supposed to help farming by introducing more robots and artificial intelligence in farming. This is encompassed by the phenomenon of big data, massive volumes of data that can be captured, analyzed, and used for decision-making [2]. Agricultural information is important to the World Food Organization, governments, food traders, and farm managers. Big data sources include weather, market data, and benchmarks with other farms.

## II. BIG DATA

Big data (BD) describes the large volume of data sets obtained from various sources such as social media, sensors, public data, transactions, and data warehouse appliances. Big data in agriculture often refers to the massive amount of data generated in agricultural practices. Modern farming practices are intensely data-driven. Data collected from commercial farmers is inherently huge, multidimensional, noisy, and heterogeneous. The amount of data can exceed PetaBytes. It requires computational intensive processing and management, which is a challenging task over traditional methodologies. This may require a new hardware and software platform with tools and techniques [3].

Originally, big data was defined to have only three basic properties [4]: volume, velocity and variety. This has been improved to include veracity and value. So we now characterize BD by the 5Vs: volume, velocity, variety, veracity, and value, as shown in Figure [5].

- *Volume*: This refers to the size of the data being generated both inside and outside organizations and is increasing annually. The sources may include text, audio, video, images, social networking, medical data, weather forecasting, etc. Big data refers to data sets of extreme size (e.g. exabytes, zettabytes) which are beyond the capability of the commonly used software tools.
- *Velocity*: This refers to the speed at which data is generated by Internet users, mobile users, social media, etc. Data are generated and processed in a fast way to extract useful, relevant information. Big data could be analyzed in real time, and it has movement and velocity.
- *Variety*: This refers to the fact that big data originates from heterogeneous sources and is in different formats (e.g., videos, text, logs). BD comprises of structured, semi-structured or unstructured data. Structured data (such as bank transactions) is easy to analyze and validate using structured Queries. Unstructured data (such as email, text, images, etc. originating from social networking sites) is difficult to handle. Most data is unstructured. Keep in mind that what is structured or unstructured is not the formatting but the content of the document.
- *Veracity*: By this, we mean the truthfulness of data, i.e. whether the data comes from a reputable, trustworthy, authentic, and accountable source. It suggests the inconsistency in the quality of different sources of big data.

 

The data may not be 100% correct. In other words, how certain are we about the data? How good and clean is the data for performing analysis? Is the data from a reputable and trustworthy origin? The more data we accumulate, the harder it becomes to keep everything correct and consistent.

- *Value*: This is the most important aspect of the big data. It is the desired outcome of big data processing. It refers to the process of discovering hidden values from large datasets. It denotes the value derived from the analysis of the existing data. If one cannot extract some business value from the data, there is no use managing and storing it.

Another relevant "V" could be visualization, indicating the need of presenting complex data structures in an easy-to-understand manner. To convert big data into actionable information requires big data analytics, which is the process of examining the data sets to uncover hidden patterns, unknown correlations, market trends, and other useful business information [6]. This will hopefully allow making decisions in a faster and more accurate manner.

## III. TYPES OF AGRICULTURE DATA

Typically precision agriculture dataset refers to the following data types [3]: (1) Historical Data: It includes, soil testing, weather conditions, GIS data, and labor data; (2) Agricultural equipment's data: It includes data collected from remote sensing devices; (3) Social and Web Based Data: It includes agricultural websites and blogs, social media groups;(4) Streamed Data: It includes data from drones, aircraft's, wireless sensors, smart phones, security surveillance's; and (5) Business, Industries and External Data: The data from billing and scheduling systems.

## IV. APPLICATIONS OF BD IN AGRICULTURE

Although big data analysis has been applied in various industries, it has not yet been widely applied in agriculture. Big data holds enormous promise for farmers. Its applications include crop identification, accurate crop predictions, automated farming or precision agriculture, and food security estimations [7]. Analytics looking forward and backward can assist in determining the best crops to plant. Weather forecasting can benefit from big data analytics.

## V. BENEFITS

Big data has been a major driver of progress in precise agriculture or smart farming.

The analysis of big data will enable farmers to extract value from it, improving their decision-making and productivity. Big data enables weather forecasting and provides information about parameters such as seed varieties and fertilizers. The insights offered from big data will provide farmers with effective, long lasting solutions and enable them harvest their yields at the right time. It enables them to see the big picture as they have never seen it before. It can help farmers make decisions that will increase yields and deliver safe, nutritious food to the society, which will enable all people to have sufficient access to safe and nutritious food.

Big data is the main underlying technology for precision agriculture, which uses GPS technology and enables a farmer to locate their vehicular positions on the field. The predictive analytic framework can be applied to evaluate big data to enhance decision making for livestock production, health, and welfare [8].

## VI. CHALLENGES

Although big data analysis has successful been applied in agriculture, it has created some problems too. One of the biggest challenges of big data is ensuring privacy and security. Another challenge is determining who retains ownership of the data generated by John Deere's agricultural tractors, which are proprietary [9]. As a data-driven industry, it is crucial to determine who owns the data, how can they use it, and what does that mean for competition. The collection, management, and analysis of agriculture big data are cumbersome and frequently resource intensive.

Some are concerned that data may fall into the wrong hands such competitors and intruders. Being a new trend, big data must be used according to the highest ethical standard in order to ensure user confidence. Consumers and farmers should be able to trust the data and how it is being collected should be transparent. Big data analytics is an expensive tool which only big farmers can afford.

## VII. CONCLUSION

Big data **(**BD**)** technology is a popular term that is commonly used in many fields such as social platform, e-commerce, financial analysis, and agriculture. Big data analytics provides a new insight that is needed to improve yield productivity and avoid unnecessary costs related to the use of pesticide and fertilizers. Big data in conjunction with newly emerging technologies such as the Internet of things, cloud computing, and data mining can revolutionize farming, reduce scarcity, and drastically increase food supply.

 

Big data will drastically affect many aspects of the agricultural industry. It is regarded as the key to the future of increased food production and sustainable agriculture [10]

## REFERENCES

[1] M.. H. Meisner, "Enhancing data-driven decision making in agriculture: A big data approach," *Doctoral Dissertation*, University of California at Davis, 2014.

[2] S. Wolfert et al., "Big data in smart farming – A review," *Agricultural Systems*, vol. 153, 2017, pp. 69–80.

[3] M. R. Bendre, R. C. Thool, and V. R. Thool, "Big data in precision agriculture: weather forecasting for future farming," *Proceedings of the 1st International Conference on Next Generation Computing Technologies*, Dehradun, India, September 2015, pp. 744-750.

[4] M. N.O. Sadiku, M. Tembely, and S.M. Musa, "Big Data: an Introduction for engineers," *Journal of Scientific and Engineering Research*, vol. 3, no. 2, 2016, pp. 106-108.

[5] H. Asri et al, "Big data in healthcare: challenges and opportunities," *Proceeding of International Conference in Cloud Technologies and Applications*, June 2015.

[6] H. Jain and R. Jain, "Big data in weather forecasting: Applications and challenges," *Proceedings of the International Conference on Big Data Analytics and Computational Intelligence*, 2017, pp. 138 – 142.

[7] A. Kamilaris, A. Kartakoullis, and F. X. Prenafeta-Boldú, "A review on the practice of big data analysis in agriculture," *Computers and Electronics in Agriculture*, vol. 143, 2017, pp. 23–37.

[8] B. J. White, D. E. Amrine, and R. L. Larson, "Big data analytics and precision animal agriculture," *Journal of Animal Science*, vol. 96, no. 4, April 2018, pp. 1531-1539.

[9] K. Bronson and I. Knezevic, "Big Data in food and agriculture," *Big Data & Society*, January-June 2016, pp. 1–5.

[10] M. Stubbs, "Big data in U.S. agriculture," January 2016, https://fas.org/sgp/crs/misc/R44331.pdf

## ABOUT AUTHORS

**Matthew N.O. Sadiku** is a professor in the Department of Electrical and Computer Engineering at Prairie View A&M University, Prairie View, Texas. He is the author of several books and papers. His areas of research interest include computational electromagnetics and computer networks. He is a fellow of IEEE.

**Suxia Cui** is an associate professor of Electrical and Computer Engineering Department at Prairie View A&M University. She has published journal and conference articles in the field of wavelets, image processing, and video coding. Her research interests include data compression, signal classification, image and video processing.

**Yonghui Wang** is currently an associate professor with the Department of Engineering Technology, Prairie View A&M University, Prairie View, TX. His research interests include digital signal processing, image and video coding, and wavelets.

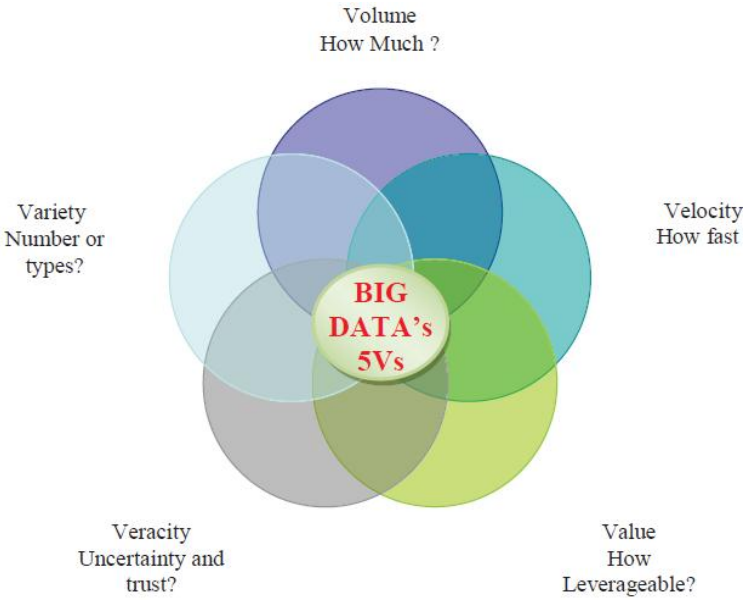


**Figure 1 The 5V of big data [5].**