**Usage of Big Data in Agriculture**

**Chiluveru Sanathkumar (LC88575)**

**Data science, University of Maryland Baltimore County**

[**DATA 603** – **Platforms**](https://blackboard.umbc.edu/webapps/blackboard/execute/courseMain?course_id=_69058_1) **for Big Data Processing**

**Dr. Najam Hassan**

**November 17, 2022**

**Usage of Big Data in Agriculture**

**Introduction:**

The rise of next-generation computer technology is strongly intertwined to the adoption of Internet of Things, cloud computing, plus artificial intelligence (AI). These technologies have resulted in a massive accumulation of data, resulting in the era of big data (Osinga et al., 2021). Big data technology has altered how people think about medical care, transportation, finance, and other industries, and it has also aided agricultural stakeholders in determining how to cope with agricultural information technology.

Agriculture is vital not only for food production, but also for raw material supply for textile industries, sugar, cotton, oil, and tobacco. Agriculture is a way of life for many people, not simply a profession. It is the foundation of a country's economy and plays an important role both in human and economic growth. It not only produces food and raw resources, but it also employs a sizable section of the people in most nations. Agriculture was originally supposed to be restricted to the production of essential foods, but it now includes forest, dairy, fruits, poultry, and other industries. In conventional data warehousing, predefined analysis is implemented to abstracted data, whereas non-predicted analysis is utilized in big data. Sensor information, location information, clickstream logs, and other types of data cannot be organized or queried using traditional technology. However, in today's world, we must deal with such details. It is challenging to conduct fast information analysis and decision-making due to the large volume of data.

Agricultural data is often distinguished by at least three features: huge data volume, diverse data sources, and complicated structure. Precision agriculture has emerged as an important development path in agricultural industrialization as next era information technology has gained popularity. This research offered a big data application platform for precision agriculture. Data sources, data integration, as well as data analytics are all part of the concept. Precision agriculture benefits from the utilization of big data applications. In the agriculture business, big data focuses on the use of data, technology, and analytics to provide relevant data that farmers can use. Big data may be utilized to give information for the whole agricultural business, or it can help individual segments or places improve their efficiency.

**Big Data:**

Big data is a combination of organized and unstructured data that is rapidly developing. Big data is described as information assets with a huge quantity, high speed, and high variety that demand cost-effective, creative data processing for greater insight and decision making. It's difficult to comprehend modern life without data. Globally, a growing amount of data is generated and consumed. However, having the ability to analyze and handle large data is critical for the effective functioning of an agricultural firm (*Role of big data in agriculture)*. Corporate farms are now becoming bigger and more diversified, resulting in increasing amounts of complicated data that must be maintained on a continuous basis. External sources from social networks, and sensor data from the ground are all included. This transforms the agriculture sector, creating new opportunities.

There are four steps in agriculture: 1. data gathering, 2. data storage, 3. data transformation, 4. data analysis, and 5. data marketing. Data acquisition is the process of obtaining huge amounts of diverse data from several sources, such as temperature readings, remote mapping, crop production, agricultural characteristics, and other Department of Agriculture information. Data storage is a key challenge in large data-driven systems, and NoSQL solutions are gaining popularity. The act of integrating disparate data sources that describe the same variable into a single coherent representation is known as data transformation. Data analytics derives value from acquired data and evaluates the outcomes, converting them to information.

**Big Data sources in Agriculture**

**Sensor Data:**

Internet of Things connects real-world things to digital entities. The sensors quantify the characterization information and deliver to the information receiving end after measuring the features of the natural environment (iopscience.iop.org). For instance, sensors placed on fields may detect the temperature, precipitation, humidity, and PH of the soil in real time. The secret to system control might be found in these data recordings. Most agricultural big data comes from sensor data, which is also the primary source for data mining analysis.

**Statistical Data:**

Historical data are the results of the long-term accumulation of details on agricultural production operations. Logs of farms, sowing, fertilization, pesticide use, preservation, breeding, etc. are included in the historical data, together with statistics from many companies and institutes. Agricultural cross-cutting data, certain incomplete historical data, and publications are also included in the category of historical data. There are several possibilities for gathering agricultural documentation in the era of information. On the website, Internet provides text, images, videos, and audios. Numerous policies, market, agricultural technology, and agricultural economic information sources may be found on respected agricultural websites.

**Geographical Data:**

Spatial data seems to have a unique place among the various sources of farm big data. Remote sensing satellite data, mapping data, environmental survey data, agriculture zoning data, and other types of spatial data all include a significant quantity of spatial information. In the current state of precision agricultural development, the use of GIS technology may successfully meet precision agriculture development goals. GIS assists in the organization and management of agricultural geographical data. Using maps as a framework to visually show agricultural statistics data and real-time data is a significant way to increase the usability of big data research.

With all the data obtained from various sources is analyzed and used to aid with optimum agricultural practices. Let us now look at some of the uses of bigdata in agriculture in the following categories.

**Big data applications in Agriculture**

Precision agricultural methods help farmers save money while also opening new economic prospects. Here are the primary benefits of using big data in agriculture.



**Controlling pesticides and fertilizers using big data:**

Due to lack of information, fertilizer and pesticide use has increased. Its negative impacts are sometimes seen in people's present health problems. Overuse of these harmful compounds would harm both human and animal health. Because of a lack of study, many growers are unaware about their harmful repercussions (*Big Data and agriculture: A complete guide*.). It is difficult to remove pests and insects from farms in the current environment, but their usage may be minimized by leveraging data obtained from a number of sources, such crowdsourcing, resources on the internet, and contributions from agriculture agencies and farmer groups.

with all the information, farmers can make more precise and dependable decisions, thus increasing farm output. Pesticide application optimization Pesticide usage is regarded as a problem owing to the negative impacts on the environment. Big data provides prospects for smart and accurate pesticide application, allowing farmers to make more informed decisions about which pesticides to use, when, and where. This type of monitoring assists food manufacturers in avoiding the abuse of chemicals. Furthermore, it boosts farmer earnings by reducing the usage of needless pesticides. Farm equipment administration large farms benefit from remote control of agricultural gear.

**Big data intelligence for crop selection:**

The most fundamental requirement for increased agricultural production primarily based on geographic data such as soil health and crops that are suitable for those specific locations. As a result, for years, scientists and growers have already been able to discern these patterns. Several crops, however, have been deemed inappropriate for long-term cultivation owing to global climate change.  The big data and Internet of Things are used to monitor agricultural soil and the conditions to provide farmers with proper crop selections and to track their farming operations and environmental issues. Prescriptive strategies use real-time and statistical data provide the timely suggestions for procedure advancement, such as seed, soil conditioner, and other farm inputs application rates, soil testing, and localize weather and disease/pest reports. Localized Explanatory plans use big data to offer better and more sophisticated perspective of an activity, whereas descriptive proposals use big data to make better and more sophisticated view of an operation.

**Meeting the need for food:**

The greatest method to meet the rising food demand without utilizing new resources and land is to use the present agriculture more effectively with the goal of increasing yields (impact-of-big-data-in-agriculture). In what ways does big data play a role in this? Most significantly, it informs farmers of variations in the weather, rainfall, soil moisture, and other elements that influence crop productivity.

**Food security:**

Every year, millions of individuals are affected by food-related illnesses and disorders. One of the responsibilities of contemporary agribusiness is to enable the identification of bacteria and symptoms of contamination in real time. To determine the condition of a growing plant, information on weather, humidity, and chemicals can be collected. So, how do farmers "harvest" all of the benefits that big data provides? The most convenient approach to gain access to previously studied agricultural data is through various web tools. They gather data from a variety of sources, including satellite photography and ground sensors, and typically analyse it in real time. Crop Surveillance by EOS is one such solution, able to process and analyzing massive amounts of data in record time.

**Efficient use of Agricultural machinery:**

Agricultural machinery is a key differentiator in the agriculture sector since it saves money and time. In the field, the use of machines was efficient and capable of lowering labor costs. Even now, we are still relying on outmoded equipment to do jobs. These older machines require more fuel, which is environmentally unfriendly, and if they break down, it will influence the industry or lower the yield achieved. Big data is used when it comes to identifying these potential issues with the arrival of GPS-equipped machinery and farmers can know wherever their machinery is at all times using these equipped machines. Users may also set machines and water systems to function in specific regions based on data insights. For example, localized weather information can assist you determine which crops require water and which do not.

**Big Data Problems:**

**Privacy and security:**

The Cloud Security Alliance's big data working group identifies the top privacy and security issues that must be addressed to make big data technology and infrastructure safer. Most of these challenges are associated with large data storage and computing (*Nebula.wsimg.com*). Secure data storage is one of the issues. Various data safety and privacy problems are covered, including data theft, data integrity, data availability, and data backup.

**Dynamic Provision:**

A cloud computing service is infrastructure as a service, which delivers compute resources as needed. Many cloud computing firms are applying this concept and making it simple for clients to utilize these services. Dynamic provisioning is not a feature of current frameworks. The problem is that compute resources may be insufficient for the submitted work; some processes may necessitate extra resources. Another difficulty is that present algorithms do not take scheduling and protection into account.

**Procedures:**

Organizations were giving papers based on key terms extracted from abstracts and titles. It was tough to analyze the science by hand. After that, program analyst completed the task. Algorithms are used to do this operation. These algorithms may differ from one another. This discrepancy might impair the final result's efficiency and dependability. Improvements in data management will result in improved technologies, but there will be numerous challenges.

**Big Data Misuse:**

Because information is power, there will be challenges, including the possible exploitation of big data. The kind of data that humans will generate in the future remain uncertain. To overcome these obstacles, we must strengthen and expand our aim and capabilities.

**Applying machine learning to big data analysis**

The benefit of large data in decision making may be exploited through the application of machine learning techniques. Machine learning is a discipline that combines computer science, statistics and artificial intelligence (*Challenges to use machine learning in Agricultural Big Data)*. Agriculture, from sowing to harvesting, is a natural activity that requires a wide range of information to comprehend the fundamental science and construct decision models. Recently, data analysts have been using machine learning to harness the knowledge buried in big data by finding correlations and comprehending the trends and patterns of the data gathered. A massive quantity of data produced on a regular basis in agriculture must be processed and understood using machine learning algorithms. Data interpretation must provide farmers with decision-making tools. Using machine learning, researchers are attempting to create a large-scale data analytics tool. Agricultural information from many repositories that is diverse might be quite valuable for a learning assignment. Apache Mahout is a machine learning algorithm library that may be used on large data systems. The collected data highlights several challenges that must be addressed before applying machine learning directly. To solve challenges such as redundant information, noisy data, inconsistency, and data imbalance, data transformations are required.

**Future scope of Big Data and Cloud in Agriculture**

Farming success used to be heavily dependent on beneficial natural factors, but that is no longer the case. The combination of cloud computing as well as big data have provided farmers with enough data points to make wise decisions (*Big Data in smart farming)*. As data centers as well as storage have become available on a 'pay-as-you-go' basis, cloud computing is revolutionized the availability of massive computing power. This has enabled the gathering of knowledge repositories including information such as temperature, irrigation procedures, plant nutrition requirements, and a variety of other farming approaches. Cloud-based software may advise farmers on how to modify their output in response to market demand, as well as how to increase yield and profitability.

**Conclusion**

Big data is becoming popular in precision agriculture. The ultimate objective of building a big computer network for agriculture field is to improve agricultural economic optimization by offering more effective assistance and instruments (Wolfert et al., 2017). We created a big data analytic system for precision agriculture based on agricultural data. We also went through the specifics of the computational pipeline. The adoption of this big data strategy has several guiding implications for the integration, mining, and use of enormous volumes of data created in crop yields and applications.

**Reference:**

Osinga, S. A., Paudel, D., Mouzakitis, S. A., & Athanasiadis, I. N. (2021, October 21). *Big Data in agriculture: Between opportunity and solution*. Agricultural Systems. Retrieved November 17, 2022, from https://www.sciencedirect.com/science/article/pii/S0308521X21002511#s0020

*Role of big data in agriculture - researchgate.net*. (n.d.). Retrieved November 17, 2022, from https://www.researchgate.net/profile/Madhuri-Javaregowda/publication/341879849\_Role\_of\_Big\_Data\_in\_Agriculture/links/60543bbfa6fdccbfeaedb6c1/Role-of-Big-Data-in-Agriculture.pdf

You are being redirected... (n.d.). Retrieved November 17, 2022, from https://www.analyticsinsight.net/the-impact-of-big-data-in-agriculture/

<https://iopscience.iop.org/article/10.1088/1742-6596/1757/1/012107/pdf>

*Big Data and agriculture: A complete guide*. Talend. (n.d.). Retrieved November 17, 2022, from https://www.talend.com/resources/big-data-agriculture/#:~:text=Big%20data%20provides%20farmers%20granular,profitability%20and%20when%20to%20harvest

*Nebula.wsimg.com*. (n.d.). Retrieved November 17, 2022, from https://nebula.wsimg.com/f541730d81c447b6de4e976f313791d1?AccessKeyId=4CF7FAE11697F99C9E6B&disposition=0&alloworigin=1

*Challenges to use machine learning in Agricultural Big Data: A ...* (n.d.). Retrieved November 17, 2022, from https://www.researchgate.net/publication/359386300\_Challenges\_to\_Use\_Machine\_Learning\_in\_Agricultural\_Big\_Data\_A\_Systematic\_Literature\_Review

*Big Data in smart farming - a review - wur.nl*. (n.d.). Retrieved November 17, 2022, from https://www.wur.nl/upload\_mm/6/c/9/4bbd5425-7889-42ee-9da9-e52750269904\_Wolfert%20et%20al%20Big%20Data%20in%20Smart%20Farming.pdf

Wolfert, S., Ge, L., Verdouw, C., & Bogaardt, M.-J. (2017, February 7). *Big Data in smart farming – a review*. Agricultural Systems. Retrieved November 17, 2022, from https://www.sciencedirect.com/science/article/pii/S0308521X16303754#s0105