ISBN: 978-93-48091-98-7

4. Big Data Analytics in Agriculture

Prashanth J.

Ph.D Scholar, Department of Agricultural Economics,
University of Agricultural Sciences, Dharwad, Karnataka, India.

Rakesh B. L.

M.Sc. Student, Department of Agricultural Economics,
University of Agricultural Sciences, Dharwad, Karnataka, India.

Hanumanthappa R.

Ph.D Scholar, Department of Agricultural Economics
University of Agricultural Science Bengaluru, GKVK.

Abstract:

Big data analytics in agriculture is completely changing the way farming is carried out and pushing the sector toward higher productivity, sustainability, and efficiency. Integration of large and heterogeneous data sets-such as weather patterns, soil conditions, crop health, and market trends-can allow data-driven decisions to be taken by farmers. Advanced analytics approaches to agricultural interventions range from the anticipation of pest outbreaks to the optimization of irrigation and fertilization schedules by use of machine learning, predictive modelling, and artificial intelligence. This approach saves costs and conserves resources while improving crop management. Big data analytics also fuels progress in precision agriculture by analysing data from sensors, drones, and satellite images to create insights and recommendations one could act on immediately. Some of the possible benefits it holds for increased crop yield, resource management, and resilience of the agriculture industry to adjustments in the market and climate change. However, for the proper application of big data analytics in farming, problems related to data privacy and technology infrastructure, along with a gap in terms of skills among farmers, need to be attended to. The present research focuses on an analysis of the applications, benefits, and challenges of big data analytics in agriculture, underlining important technical developments but also possible directions for further research and exploitation.

Keywords:

Big data, Analytics, Agriculture, forecasting

4.1 Introduction:

A country's economy is largely based on its agricultural sector, which is regarded as one of its most important industries. It is also among the oldest professions in human history. Farmers mostly base their decisions on human intellect and knowledge gained through experience. It experiences a great deal of unforeseen natural disasters as a result, including droughts, floods, absence of or premature monsoon, and climate change. In addition, there are institutional and governmental issues that do not provide agriculture with improved agricultural schemes, credit facilities, etc.

That being said, technology has to take the lead in this shift and find a solution. The agricultural industry is experiencing significant advancements and innovation due to the rise in "Big Data Analytics," which involves analysing large and diverse data sets to uncover information such as hidden patterns, unknown correlations, market trends, and customer preferences. This field has impacted almost every industry worldwide and is expected to continue to grow in the future.

Agriculturalists are rapidly shifting from traditional approaches to more standard tools and procedures for decision making and sophisticated analytical modelling, which would further aid in boosting yield and revenue, based on scientific understanding. Farmers are aware that they have access to a vast quantity of data. So, in order to meet their needs, farmers and agricultural businesses deal with enormous volumes of both organised and unstructured data.

Big Data analytic techniques, when integrated with farm data and information from multiple sources, such as corporate agronomical data, satellite-drone imagery, public data (data from government/governing bodies), Emerging data (consumer-based data), field-level sensors, weather stations, and various historic data obtained from a variety of growers and growing conditions, can offer farmers valuable insights and information. These decisions can help farmers increase their yields and minimise financial losses resulting from unpredictable fluctuations.

In order to make wiser judgements, people today desire and require not only the collection of data but also their understanding and determination of the set's significance. It is challenging to process and manage big data using conventional methods and instruments since it is a sizable collection of datasets with high velocity, volume, and variety (Elgendy and Elragal, 2014). It is possible for it to be semi structured, unstructured, or structured. Big data analytics is a sophisticated analytical method that may be used to examine large amounts of data in order to find unknown, hidden, and valuable patterns. Big data is therefore crucial to the decision-making process.

The huge role that information technology plays today in almost every sector, like health care and education, is a result of the rapid economic growth of countries. Another major sector that contributes to the economy of all countries is agriculture. Due to the rapid growth of IT, collection, storage, and analysis of data will become easier in order to receive information. Technology can be applied in agriculture directly to enhance productivity and indirectly to incentives farmers toward better decision making, (Patel and Sayyed, 2014). Information Technology is mainly applied in precision farming, which is prevalent in developed countries, to enhance productivity.

While big data analytics is widely used in other industries, it is not used in agriculture. In 2050, there will be more people on the planet than 9 billion, according to research. Consequently, it is crucial to produce and transport goods in an effective manner.

Big data analytics application in agriculture yields extremely advanced benefits, such as innovative resource-minimization solutions with harvests of promise similar to precision agriculture. Analysing real-time data from a variety of sources, including soil, weather, air, equipment, availability, and others, is crucial to precision agriculture (Surya *et al*., 2016).

The field of smart farming focusses mostly on leveraging big data analytics to improve commercial processes. Furthermore, data is growing faster than computation speed. Therefore, it is imperative to reduce the amount of time required for data analysis. Therefore, timely information may be gained by employing big data analytic tools and methodologies to help policy makers and farmers make better decisions about product import or export and how to get a good harvest (Ojugo., 2017).

4.2 Overview of Big Data, Big Data Analytics and Big Data Analytics in Agriculture:

The productivity of farmers in rural areas is too low because most of them still use traditional farming practices and lack knowledge of contemporary technologies. Every day, 2.5 quintillion bytes of data are produced, according to research. Furthermore, it's shown that by the end of 2015, the expansion of big data had doubled. It is quite difficult to manage and govern these data using traditional tools and approaches. Consequently, knowledge of big data and its connected fields is required (Sadiku *et al.*, 2020).

4.2.1 Big Data:

"Big data" refers to a complicated and large-scale collection of data that is difficult to handle using conventional data processing methods or readily available database administration tools (Himesh *et al.,* 2018). These data are accessible on semi structured, structured, and unstructured heterogeneous structures. Information with a predetermined format is referred to as structured (banking information), unstructured (text files, audio), and semi structured (data that may be transformed from unstructured to structured using existing descriptions, such as xml documents). Also, big data is characterized by 4Vs

V's of Big data

a) Volume: The 'volume' refers to the huge quantity of data that businesses generate and maintain. It covers the exponential growth of data repositories from terabytes to petabytes and beyond. This exponential rise in data volume from the Internet of Things devices, from social media sites, and from various online transactions, demands scalable infrastructure and advanced analytics tools for handling such huge data sets and extracting meaningful insights from them.

b) Velocity: Velocity refers to the speed at which real-time or almost real-time data is generated, processed, and analyzed. This reflects the dynamic nature of streams of data defined by fast informational inflows from many sources. The velocity of data poses problems related to intake, lagging in processing, and responsiveness to actionable insights in such diverse areas as social media feeds, sensor networks, financial transactions, and online clicks.

c) Variety: Variety includes the wide array of sources, formats, and types that compose a Big Data environment. It has structured and semi-structured data besides text documents, images, videos, sensor readings, and log files. Extraction of insight from heterogeneous information requires flexible data structures and sophisticated techniques for data wrangling due to proliferation of diversity, which presents challenges in terms of data integration, interoperability, and analysis.

d). Veracity: Veracity in Big Data refers to the reliability, accuracy, and trust of data. It summarizes noise, biases, and natural uncertainty that large datasets inherently possess because of poor quality data, biased sampling, or inaccurate observations. Veracity makes it much harder for techniques of data cleansing and anomaly detection and maintaining integrity of the insights obtained from probably noisy or untrustworthy data sources.

4.2.2 Importance and Implications:

The four V's of Big Data hold profound implications for organizations across diverse sectors, offering both opportunities and challenges in leveraging data as a strategic asset:

Opportunities:

Harnessing insights: Businesses can find actionable insights, trends, and patterns that spur innovation, improve customer experiences, and streamline business operations by utilising the volume, velocity, and variety of data available.

Real-time decision-making: Enterprises can react quickly to market developments, customer preferences, and emerging opportunities by using the velocity of data to inform their decisions.

Agility and creativity: The diversity of data encourages creativity and agility, enabling businesses to test out new models, technologies, and data sources in order to obtain a competitive advantage in the digital economy.

Challenges:

Infrastructure and Scalability: Processing frameworks that can handle large datasets and real-time data streams, as well as scalable storage systems, are necessary to manage the amount and velocity of data.

Data Integration and Interoperability: To guarantee consistency and coherence across various data sources, strong data management techniques and standards are required. The variety of data presents issues in terms of data integration, interoperability, and governance.

Data Integrity and Trustworthiness: Ensuring the accuracy of data is crucial since biases, errors, and inconsistencies can compromise the integrity of conclusions and judgements drawn from Big Data analytics.

4.2.3 Big Data Analytics:

As it was mentioned above, big data has an extremely high rate of complexity that cannot be processed with regular software tools. Besides, sizes of data are increasing day by day and sometimes even changing. Accordingly, big data analytics is terminology addressed to the method of data analysis. Big data analysis is the procedure for collecting, storing, and analyzing data to recognize hidden patterns, unknown relationships, and other useful information. Big data analytics requires sufficient processing capacity and analytical capability. Tools are there for turning data in raw format into a standard format for big data analysis, which involves information gathering and analysis and its visualization, and also scheduling (Kumar and Menakadevi., 2017).

4.2.4 Big Data Analytics in Agriculture:

Industries like banking and insurance have successfully embraced big data analysis. Despite not using big data analysis in agriculture for a while, it has recently been put to use. Numerous forms of data are generated by agriculture, including data from economic models, agricultural yield, crop diseases, *etc.* Real-time data on the air, weather, soil, crop maturity, labour costs, and even equipment can help make smarter decisions. Big data utilisation in agriculture reduces farmer failures and provides recommendations for soil, water level, etc. The more frequent the analysis, the better the service (Bendre *et al.*, 2015). There are two main areas of study for big data analytics in agriculture: precision agriculture (PA) and smart farming.

4.3 Potential Applications of Big Data Science in Agriculture:

The following are some specific agricultural industries where big data analytics can be very helpful:

Faster Technology Adoption and Development: Technology creation time can be shortened by applying the right analytical tools in conjunction with sensors and automated technologies for data collection. Collaboration between professionals in plant breeding, agronomy, engineering, statistics, modelling, and computer science is needed for this. Extension specialists can encourage quicker adoption and get input for technology evaluation and improvement using social media (Masih and Rajasekaran, 2020)

Plant Breeding: From the first crossings to the commercial release of crop varieties, it typically takes 7-8 years, including 3–4 years for field testing. For tree crops and perennials, the duration is significantly greater. Making sure the material is appropriate for a given area and provides farmers with an economically feasible answer is the problem here. To forecast how well crosses would do in various weather scenarios, data from individual fields, plants, and leaves are being gathered. Machine learning and deep learning techniques are then being used.

High Throughput Field Phenotyping (HTFP): This technique replaces the laborious and arbitrary ratings employed by plant breeders by quickly scanning a large-scale field evaluation of crosses thanks to big data analytics.

This can be used to gather spatial and temporal data via sensors, drones, or satellite imagery in field studies as well as breeding nurseries. Evaluating crop growth throughout the course of a season yields useful data. The HTFP can provide fresh perspectives on genotypic-environment interactions when combined with meteorological data and management techniques.

Bioinformatics: Additional uses of Big Data include genome assembly, proteomics, metagenomics, and high throughput sequencing. Important plants, animals, and bacteria are the subject of ongoing genetic data generation by agricultural researchers. Building blocks, rules for sharing data, parallel and grid computing, and de novo assembly development are needed to identify the genes underlying various features. ICAR has started some research on database construction and prediction through the use of machine learning algorithms to mine genome and proteome sequences. It is necessary to pursue the development of genomic data storage, workflow parallelisation, SNP resources for animal species, genomic resources for diverse crops and animals, and other transcriptome resources.

Precision Farming: It is the application of site-specific management techniques to reduce input use and production costs by analysing intra- and inter-site variability using geotagged data. Developed nations are pursuing precision farming. Farmers can receive personalised advice from the software solutions, which also offer cloud storage, satellite imagery, and remote management. Harvesting machines are another sector of agriculture where Big Data can have a big impact. Machines for harvesting rice and wheat are widely used.

A harvesting machine's sensors can offer useful data on crop production, soil temperature, and moisture. Since all of the data are geotagged, it is simple to analyse them to determine the inter- and intra-plot variability at the plant level.

Crop Planning: Seen through a different lens, crop planning has always been a traditional big data problem requiring the analysis of several data sets, such as soil, water, weather, demand, and price forecasts. Additionally, useful for Mega Environment Mapping of Crops is big data analytics.

Quality Assessment: Another possible use case for big data analytics is price discovery for farmers at the market level based on quality utilising picture recognition of various crops.

Grading and Marketing: Through mobile applications, technologies that use picture analysis are now accessible to evaluate the quality of fruits, vegetables, and spices. These methods may aid producers in realising better prices by bringing openness and uniformity to the quality assessment process during grading.

Customized farm Advisory Services: Plans, management techniques, and disease and insect identification are all being addressed in the pipeline for customised farm advises. Before planting, for instance, a farmer would like to know what the best yield in his or her field or town is, how it differs from the best in the state, the country, and the world, and why.

Farmers will be able to comprehend and prepare for higher yields thanks to this gap analysis. The creation of planting schedules and backup plans, the forecasting of phenology and the suggestion of agronomic measures, crop-specific soil test-based fertiliser recommendations, and agro-met advisories will be beneficial for advisories particular to a given site.

Traceability: The product's journey from farm to plate can be tracked using blockchain technology. Food safety will be guaranteed as a result. Furthermore, e-ledger based transactions will make e-market interventions less speculative and more transparent.

Sustainable Resource Management: When it comes to marine fisheries resources, sustainability and optimality represent the two extremes of the managerial spectrum and frequently result in trade-offs between them. Real-time data utilisation will aid in closing the profit gap between capture volume and value, and a new way of thinking will optimise fishermen's earnings while staying within the sustainable bounds of fisheries. An in-depth analytical analysis on the dynamics of fisheries will be relevant and timely, as numerous international authorities are pressing for the certification of catch and the attestation of marine ecosystems.

4.4 Specific Applications of Big Data Analysis in Agriculture:

Big data analysis is used in agriculture in a number of areas to boost productivity and spur innovation:

4.4.1 Precision Agriculture:

Precision agriculture entails adjusting farming techniques based on in-depth data analysis to particular field circumstances. Important uses consist of:

- **Crop monitoring:** Examining sensor and satellite data to identify illnesses, control pests, and evaluate crop health.
- **Soil Management:** Utilising soil data to maximise land usage, fertilisation, and irrigation is known as soil management.
- **Yield Prediction**: Accurately predicting yields requires combining historical data, weather projections, and present crop conditions.

4.4.2 Livestock Management:

Livestock management is improved by big data analysis by:

- **Health Monitoring:** Using information from animal wearables to track health, identify illnesses, and improve reproduction.
- **Feed Optimisation:** It is the process of fine-tuning feeding procedures to increase effectiveness and lower costs by utilising data on animal growth and health.
- **Traceability:** Ensuring the traceability of livestock products from the farm to the table in order to improve the quality and safety of food.

4.4.3 Supply Chain Optimization:

Using big data analysis in the agricultural supply chain helps with:

- **Demand Forecasting:** Estimating consumer demand for different livestock and crop products in order to maximise output and minimise waste.
- **Logistics Management**: It is the process of maximising logistics, cutting expenses, and minimising spoiling by analysing data from storage and transit.
- **Market Analysis**: It is the process of using data from the market to identify trends, determine prices, and plan marketing strategies.

4.5 Current Big Data Analytics Issue in Agriculture:

While big data analytics opens up new avenues for improved data interpretation, there are still certain obstacles to overcome. Anything can be successful for a long period if implemented correctly, so the saying goes. Managing large amounts of data may be challenging, though, as it requires appropriate handling, cleaning, processing, and analysis—all of which often expand exponentially (Kamilaris *et al.*, 2017).

a) Security and Privacy:

Data security and privacy are crucial factors to consider while doing big data research. As a result, even if big data analytics has applications in agriculture, farmers' willingness to share information on their products can be difficult to come by.

It is discovered that due of data misuse, data privacy is a problem for big data analysis in agriculture. Big data analysis in agriculture also faces challenges related to data access rights, openness and usability, data ethics, data independence, data timeliness, and hurdles in data sharing.

b) Lack of Technical Knowledge:

It is evident that one of the challenges is a lack of information about technology and approaches. Even Nevertheless, there is a clear divide between industrialised and developing nations due to disparities in technology access and skill levels.

c) Difficulty in Scalability and Visualizing:

It is stated that many problems with agricultural big data are caused by the big data's defined 4Vs. Managing the rise of big data is recognised as a challenge since enormous amounts of data are generated every minute, grow quickly, and are primarily semi structured or unstructured.

Scalability and data visualisation become increasingly challenging as data grows. It should be emphasised that while many computational algorithms perform well with small data sets, they become more challenging to use with larger data sets.

d) Limited Capacity for Data Storage:

High data storage capacity is necessary to store massive volumes of data using 4Vs. In addition, there should be a high throughput and effective data storage.

e) Data Quality:

Since a great deal of data is being collected, accuracy and quality of the data should be prioritised. Finding the most accurate data to extract during analysis is crucial to making the best choice.

4.6 Way Forward:

The application of big data analytics to agricultural research and development can be very beneficial. It's time for everyone to take the lead and create a plan for using big data analytics in Indian agriculture. Here are some recommendations to help in this direction:

Bringing Data Together: To accomplish effective data governance, it would be necessary to dismantle silos and establish identification and security controls. Policies for data integration, data retention, and data sharing must be followed throughout the domain.

Multidisciplinary Group: Big data problems necessitate the involvement of experts in a variety of fields, including computer science, engineering, statistics, and domain knowledge experts for ML/AI, ML/Graph analytics, NLP, augmented analytics, blockchain, recurrent neural networks, deep learning, etc. Therefore, working as a multidisciplinary team to tackle a problem is essential.

Partnerships & Collaborations: Working together with the commercial sector and other academic institutions is crucial. Collaborations with other countries will also be necessary.
Infrastructure: Create or improve the platforms for data management (data fabric, permanent memory servers, forgetting insights) and big data analytics and warehousing.
Moving to the cloud: One possibility is cloud-based computing, where services can be hired according to need. When compared to the costs of setup and upkeep, computation fees could be lower. The organisations need to consider a hybrid model that combines cloud software as a service with on-site resources.

Standard Community of Practices (CoPs): All across the system, Standard Community of Practices (CoPs) are necessary. They interact with ongoing efforts and define data standards, interoperability protocols, security, and privacy concerns.

Showcase Applications: It will be necessary to identify prospective fields in which big data science might be used. These include the creation of data management platforms for big data analytics and data warehouses, as well as web applications for location-specific, timely, and data-driven information and advisories, crop planning, mega-environment mapping, and phenology prediction, sowing schedule creation, and artificial intelligence and machine learning-based contingency planning.

Building Capacity: Given that the field of big data analytics is still developing, it is imperative that scientific and technical staff members receive training in order to increase their capacity. International cooperation might be a part of it.

4.7 Conclusion:

Data analytics have transformed and continue to do so by providing deep insights which is critical for the decision making in Agriculture. Using the information, you can glean from many sources is how farming operations will be optimized and profitability, sustainability and production are only going to increase. Precision agriculture has become much better through technical advances in machine learning, artificial intelligence and predictive modelling. This allows the farms to track and handle resources as well as crop at a real-time.

If utilized, major benefits of big data analytics for agriculture include higher crop yields, more responsible resource management and greater resilience against environmental crisis market vagaries. Many challenges remain in order to harness the full advantages of these technologies, ranging from limited expertise among farmers and strong but optional IT infrastructure, to data privacy issues. Together, legislators and technology suppliers must cooperate with farmers to produce such solutions.

To sum up, big data analytics has the ability to greatly improve agriculture's efficiency, sustainability, and adaptability. The agriculture sector may achieve more innovation and resilience by tackling current difficulties and pushing the integration of big data technology forward, which will ultimately contribute to environmental sustainability and global food security.

4.8 References:

1. Bendre MR, Thool RC and Thool VR. 2015. Big data in precision agriculture: Weather forecasting for future farming. *In 2015 1st international conference on next generation computing technologies (NGCT)* (pp. 744-750). IEEE.
2. Elgendy N and Elragal A. 2014. Big data analytics: A literature review paper. In *Advances in Data Mining. Applications and Theoretical Aspects: 14th Industrial Conference, ICDM 2014, St. Petersburg, Russia, July 16-20, 2014. Proceedings 14* (pp. 214-227). Springer International Publishing.
3. Himesh S, Rao EP, Gouda KC, Ramesh KV, Rakesh V, Mohapatra GN and Ajilesh P. 2018. Digital revolution and Big Data: a new revolution in agriculture. *CABI Reviews*, 1-7.
4. Kamilaris A, Kartakoullis A and Prenafeta-Boldú FX. 2017. A review on the practice of big data analysis in agriculture. *Computers and Electronics in Agriculture*, *143*, 23-37.
5. Kumar H and Menakadevi T. 2017. A review on big data analytics in the field of agriculture. *International Journal of Latest Transactions in Engineering and Science*, *1*(4), 1-10.
6. Masih J and Rajasekaran R. 2020. Integrating big data practices in agriculture. *IoT and Analytics for Agriculture*, 1-26.

7. Omo-Ojugo E. 2018. Relevance of big data analytics in agriculture: Focus on Nigeria agricultural sector. *Int. J. Sci. Res*, *7*: 1-11
8. Patel S and Sayyed IU. 2014. Impact of Information Technology in Agricultural Sector, *International Journal of Food, Agriculture and Veterinary Sciences*, *4*(2): 569-578
9. Sadiku MN, Ashaolu TJ and Musa SM. 2020. Big data in agriculture. *Int. J. Sci. Adv*, *1*(1): 44-48.
10. Surya P, Aroquiaraj IL and Kumar MA. 2016. The role of big data analytics in agriculture sector: a survey. *Int J Adv Res Biol Eng Sci Technol*, *2*(10): 830-838.