Data science and analytics in agricultural development

Sidhu, K. S. ¹, Singh, R.², Singh, S.³, Singh, G.¹

Received: 06.11.2020  Revised: 06.01.2021  Accepted: 2.01.2021

Abstract
The consistent advancement of innovation has implied information and data being created at a rate, not at all like ever previously, and it's just on the ascent. The world makes an extra 2.5 quintillion bytes of information every year. The demand for individuals talented in investigating, deciphering, and utilizing this information is now high and is set to become exponential over the coming years. The total populace is relied upon to arrive at 9.7 billion by 2050 from the current population of 7.8 billion. The Food and Agriculture Organization (FAO) has predicted that the development of farming must be expanded by 70% to provide for the extended interest. Data-driven agriculture choices can be a potent technology to manage the needs of this much high population, as this technology gives higher efficiency, rehearses support-ability, and even assists with giving straightforwardness to purchasers and consumers needing to find out about their food as reported in the studies. The current and future interests will require more data researchers, data engineers, data specialists, and chief data Officers. This paper tries to examine the need, use, role, and issues faced by data science and data analytics to improve the quality as well as quantity of Agricultural produce thereby leading to an increase in production, a decrease in costs, and overall sustainability.

Key Words: Agriculture, data analytics, data collection, data science, precision farming

Introduction
The world is moving towards untold challenges, and the population is the primary cause behind unsustainable development (Mittal, 2013). To manage the resources for mankind under some fixed requirements is not a simple job. Therefore the use of artificial intelligence for the agriculture sector management can be fruitful in several aspects. The need is high, and production is not compatible as per the demand. Moreover, artificial intelligence in the farming sector is no less than a wonderful gift for humanity. By installing modern technology in agriculture, the 24-hour working can be attained based on different equipment and programming software. Humans are always curious about learning new things, which is very needful to accept the importance of technology. Data Science is the most potent weapon in the 21st century; it has the capability to manage, manipulate and generate the data to face unexpected consequences and challenges. Fig. 1, generated by Google Trends, shows the popularity of "Data Science" search term over the past five years. As the world entered the new period of big data, its stockpiling requirement likewise developed. It was one of the principal challenges and worries for endeavour businesses till 2010. The main focus was to build a system and find solutions to store information. Today, systems like Hadoop have constructively tackled and solved the issue of storage and capacity, and hence the centre of interest has moved to the handling of this information. Data science is a multi-disciplinary field, which utilises cycles, logical techniques, frameworks, and calculations, to separate information and data from numerous structured and unstructured data (Dhar, 2013; Leek, 2013). Data science may be defined as an "idea to bind together measurements, information investigation, and their connected strategies" in order to "comprehend and examine real marvels" with data (Hayashi, 1998).

Author’s Address
¹Department of Computer Science and Engineering
²Department of Mechanical Engineering
³Department of Agricultural Engineering
Baba Banda Singh Bahadur Engineering-140407, Punjab, India
E-mail: karamveersidhu14@gmail.com

© ASEA
This work is licensed under Attribution-Non Commercial 4.0 International (CC BY-NC 4.0)
It uses methods and hypotheses drawn from numerous fields of mathematics and statistics, space information, software engineering, and data science. Turing grant victor, Jim Gray visualised data science as a "fourth worldview" of science (experimental, computational, hypothetical, and now information-driven) and announced that "everything concerning science is changing as a result of the effect of data innovation" and the information deluge (Hey, 2009; Bell, 2009). The cutting edge origination of data science as an autonomous control is once in a while ascribed to William S. Cleveland. In 2001, there was an extension of insights past hypothesis into specialised territories. Since this would fundamentally transform the field, it justified another name (Gupta, 2015). "Data science" turned out to be all the more broadly utilised in the following years. In the year 2002, the Committee on Data for Science and Technology dispatched the Data Science Journal. In the year 2003, Columbia University dispatched The Journal of Data Science (Press, 2014). In the year 2014, the American Statistical Association's Section on Statistical Learning and Data Mining chose a new name and was called the Section on Statistical Learning and Data Science, mirroring data science's ascendant prominence (Talley, 2016). The expert title of "Data Scientist" has been credited to Jeff Hammerbacher and DJ Patil in the year 2008 (Davenport, 2012). Though it was utilised by the National Science Board in the report of 2005, "Long-Lived Digital Data Collections: Enabling Research and Education in the 21st Century," it alluded comprehensively to any essential function in dealing with computerised information collection (Anonymous, 2005). Nevertheless, there is no agreement on the meaning of data science, and a lot of people consider it to be a buzzword (Press, 2013).

Data science is transforming how farmers and agrarian experts have been undertaking decisions (Matthews, 2019). Current innovation has made it conceivable to gather information of water, soil, and minerals from fields and farms, and store it in a unified framework, famously known as the “Internet of Things (IoT)”. IoT helps transmit the data by using sensors and microcontrollers (Sain et al., 2020), in lesser time unlike manual operations, which are time-consuming and include human errors (Singh et al., 2020). IoT alludes to associating interrelated gadgets to the Internet and enable them to share and trade information autonomously (Clark, 2016). This type of information can be joined with information from outer sources, for example, satellites, climate stations, and even information from adjacent homesteads to frame a greater volume. Data Analytics may be utilised in the aggregated mass to acquire information that can be utilised by farmers or ranchers to upgrade their cultivating. Farmers would thus be able to settle on savvy cultivating choices utilising that data all through the production cycle; from decision making and planning, plantation, reaping, and harvesting, right to its marketability (Maru, 2018).
The Need of innovation in Agriculture

Farming has prevailed since the beginning of human development as the basis of human survival. It has consistently been a labour-intensive and hard work concentrated task. Cultivating has gone through a ton of changes concerning techniques, devices, and apparatus. A great deal of innovative work has been done in order to carry farming to the current stage, and it is in a period of constant improvement. One of the components driving advancement is farmers' reliance just on their instinct to settle on farming choices (Anonymous, 2016). It is a danger; if the farmer commits an error, there might be no creation in that season at all. In this manner, there is a requirement for the farmer to ease such weight of danger and settle on sure and practical choices. The use of data mining and unique applications to operate the different systems with respect to the changing environmental conditions and parameters can enable farmers to make better and informed decisions towards profitable and productive farming. Besides, the rising field of nanotechnology along with the use of different sorts of sensors makes it possible to accumulate data of large volume, amount, and exactness from the fields and farms. Likewise, there is a colossal measure of advanced information accessible on the web. All this data and information can be used adequately to enhance farming and agriculture (Es and Woodard, 2017). It likewise has been seen that the conduct of customers has changed in recent times. They are more interested in eating great food and want information about where and how their food has been delivered, bundled, altered, and dispersed (Brown, 2017).

As shown in Fig. 2, the total populace is relied upon to arrive at 10.9 billion by 2100 from the current population of 7.8 billion. Fig. 3 shows the projected populace by region (Roser, 2014). As shown in Fig. 3, the world's maximum population will be living in Asia and Africa. These are the countries with a large number of people practising agriculture. This increase in population clearly indicates the need to increase agricultural production to feed the extended population. Hence, the most significant factor to drive advancement in the field of agribusiness is to make food accessible to the whole human population. Subsequently, there is a massive need to improve and follow better approaches to increment rural creation and agricultural production by utilising fewer assets (Anonymous, 2016; Bhutiani and Ahamad, 2019).

Figure 2. World population growth from 1700 to 2100 (adapted from Roser, 2014)
How data collection can improve agriculture

Farmers continually shuffle between a lot of factors before settling on agricultural choices. To raise the assortment of yield, they have to plan and arrange for what they will develop, where, and when. At that point, they ought to choose a water system, manures, and pesticides that must be utilised. It is trailed by selecting the time to procure, gather, and dispatch the items to the market (Maru, 2018). This cultivating strategy is a discretionary science, and hence it is crucial to get all of the factors right in order to get the maximum profit. Luckily, at this time, farmers can utilise data and information to help them in settling on those intricate choices. Farmers can accumulate data from various sources and use it alongside data analytics to pick up experiences on their homestead and yields. The data got by utilising sensors placed inside the fields, for example, that of soil supplements, water content, air penetrability, and so on (additionally named as localised data) can be utilised independently, or with information from outer sources like temperature or precipitation to obtain various sorts of information. All this data may be combined to evaluate and make changes persistently (Maru, 2018). Moreover, the information accessible from new advances like spectrosopes may be utilised in fields to recognise the nature of the soil along with the availability and nature of organic products delivered in the homestead. The spectroscopy standard is passing light of different frequencies through an object with the goal that multiple properties, for example, temperature, glow, mass, and structure, can be determined (Cuffari, 2018). In this way, the data obtained by spectroscopy can be utilised to computerise a few cycles in the field too, for example, the opening and shutting off water system sprinklers as per the dampness content of the soil. The data may also be utilised in the creation of prediction models that can help in estimating production (Es and Woodard, 2017). Data is able to empower consumers or shoppers to separate the components present in their food right to its atomic arrangement. The data collected with the help of sensors and spectrosopes, which are able to detect substance organisation and chemical composition...
noticeable in the soil, air, and yields can be made accessible on the web, which can be seen by the customers utilising the smart gadgets used by them (Brown, 2017). It will assist with building trust among concerned clients, along with an increment in client adherence. Accessibility of localised and outside data and information can assist with practising precision agriculture. The possibility of precision agriculture rotates around supportability and utilising assets strictly at the amount required, without squandering them. Precision agriculture enables human beings with the technology to evaluate plants and exclusively tend to their requirements of minerals, composts, and water. Also, those plants that lack sufficient assets can be provided with whatever they need. Along these lines, a ton of assets may be spared using this technique, and at last, the expense of production will automatically be diminished (Creutzberg, 2014). In horticulture also, accuracy can be acquired by appending Radio Frequency Identification (RFID) chips to the domestic animals. These RFID chips are used to distinguish and follow livestock. When a debilitated animal is recognised, the farmer can continue the treatment for the well being of the animal (Kshetri, 2016). It is accepted that computerised agriculture rehearses by both, smallholding and enormous scope farmers can calm the world of craving and hunger.

**Using data science to improve Agriculture**

To help small farmers

Since Data Scientists are equipped with instruments to examine and measure colossal amounts of data productively, organisations are working towards deciding the use of that data to help little scope farmers participate in the fight to unravel overall food deficiencies. In the year 2018, an alliance dispatched a project that aims to collect data from about 500 million farmers of devastated territories in 50 nations (Tollefson, 2018). The members completing the project believe that the data about seed assortments, salary, and the farmers' mechanical limit will show whether the investments done in agriculture by the organisations and leaders of different nations are paying off or not and will help in creating strategies for the farmers. For an immense scope, this task is similar to the Sustainable Development Goals of the United Nations that aim to twofold the agricultural productivity and salaries of farmers in non-industrial countries and helps them diminish world hunger.

---

Figure 4. A new initiative expects to collect data of about 500 million farmers, for example, the small scale farmers in Uganda. Credit: Geovien So/SOPA Images via ZUMA. Retrieved from [https://www.nature.com/articles/d41586-018-06800-8](https://www.nature.com/articles/d41586-018-06800-8)
An important test locale is Uganda that started utilising the survey of the World Bank in the year 2009. The results of the survey demonstrated that little farmers were experiencing difficulty getting to veterinary administrations for their domesticated animals. Presently, the government of Uganda is upgrading its program for supporting country farmers so that it is simpler for individuals to get clinical consideration for their animals. The outcome will be a data set able to be represented statistically with exact geographic data about country farms. It will help to track trends and change government outreach endeavours towards the best possible result. MyCrop, an insightful and self-learning continuous framework which considers every farmer's area, climate, and crop information, utilises big data, AI, and cell phone innovation to convey information, skill, and assets to small scale farmers (Trendov, Varas, and Zeng, 2019).

**Overseeding crop diseases and pests**

Pests can immediately and drastically reduce the benefits of a farmer. However, large use of pesticides may adversely affect plants, individuals, and other living creatures. Luckily, a few organisations enlist data scientists and create client confronting platforms that examine and guide farmers about when to apply pesticides, what kind of pesticides to apply, and the amount to utilise. One such organisation is Agrosmart, and it is a Brazilian organisation. Its innovation is dependent on Artificial Intelligence and the Internet of Things (IoT) sensors to decide the sort and amount of bugs present on a crop. Hence, farmers get a related report and may utilise it to design their pest management approaches. The objective is to enable farmers to control pests with minimum cost investment and with a limited environmental effect (FAPESP, 2018). An Israeli startup, Saillog has created an application for mobile devices named Agrio, which illuminates farmers regarding the diseases that are presently influencing their crops, or the ones found on farms that are nearby (Anonymous, 2018).

![Figure 5: Agrio, the mobile application that "detects" diseases in plants through images. Retrieved from https://blueberriesconsulting.com/en/agrio-la-aplicacion-movil-que-detecta-enfermedades-en-plantas-a-traves-de-imagenes/](https://blueberriesconsulting.com/en/agrio-la-aplicacion-movil-que-detecta-enfermedades-en-plantas-a-traves-de-imagenes/)

**To examine agricultural niches**

Data Scientists are able to utilise tools which can help in identifying patterns and connections that have a high probability of somehow staying covered up. They also have the ability to make determinations which enable agrarian science to be pushed forward by assessing the explicit elements. For instance, analysts realise that animals and poultry's metabolic elements are decidedly influenced by trace minerals, while the expansion...
of quality and sustenance of egg yolk is done by carotenoids (Anonymous, 2019). At TH Milk office, Vietnam milk production in cows is being controlled by utilising data science where each dairy animal is labelled with an RFID chip. The process of milking or the draining cycle is done automatically by using sensors in the suckers which are able to identify irritation inside the cow's mammary organs. The draining cycle is also stopped automatically, and the cow is then stamped and checked. An organisation, AfiMilk, uses a comparable chip, which is joined to a goat's legs that track the goat's movement. If a goat doesn't show movement for an extensive stretch or is demonstrating vague resting designs, it will be checked for disease. (Kshetri, 2016). Very invaluable seeming factors in agrarian cycles can help in achieving considerable changes. Various discoveries found out after searching through information bases have proved this. Animal feed brands, farmers, and everyone else involved in the agricultural sector can benefit by exploiting such discoveries by data scientists who can help in changing their activities to improve results.

To predict crop production
International Business Machines (IBM) has a developed platform, which can predict corn yields a few months ahead of time, reducing undesirable shocks for agricultural experts and farmers (Mello, 2019). Additionally, scientists at the University of Illinois rely upon seasonal forecasts and satellite data to make end-of-season predictions earlier than expected (Anonymous, 2020). Lab results indicate that this new methodology is dramatically more reliable than the United States Department of agriculture's real-time data.

Coping with climate change
Climate change is an approaching and alerting issue that has just influenced the agriculture area. Data Scientists are diligently working towards sorting out the problem of climate change. One project includes providing IoT sensors to the rice farmers of Taiwan to gather critical data regarding their crops. The data will be stored in a database and will be used to assist farmers with enhancing their production cycles, in any event, when environmental change makes that task particularly challenging (Anonymous, 2018).
Following the traditional cultivating schedule at this point is not adequate in view of the changes taking place in the climate and environment. However, data analysis could everlastingly transform the eventual fate of farming.

Soils add to the changes taking place in the environment by delivering ozone-harming substances. Hence, scientists are also examining agricultural soil data to improve their comprehension of just how soil data may help adjust to climate change (Barbieri, 2018). Gathering this type of data is dubious. However, scientists believe that it could fill in knowledge gaps related to climate change and soil connection.

Figure 7: A more far-reaching comprehension of agricultural soils is expected to support sustainable development. Retrieved from https://bigdata.cgiar.org/grow-more-food-and-mitigate-climate-change-agricultural-soil-data-could-help/

Role of data scientists in providing pictures of land and crops

Some agricultural businesses oversee such large acres of land that it becomes difficult to receive fast updates and cautions about possible issues without the help of innovation and technology. Canadian organisation Farmers Edge captures everyday satellite pictures of homesteads and consolidates them with important external information, including data from more than 4,000 weather stations. Besides, some nations, like Ghana and Ireland, depend on satellite-based crop observation to assess regions faster than traditional techniques. Data Scientists assume an important part in facilitating the utilisation of satellite imagery. Pilot ventures and several new platforms could enable such data scientists to choose the techniques most appropriate for handling the needs of the future and the present inside the agriculture sector.

Issues to the implementation of data science

The issue of the agriculture sector not promptly inviting change is an important issue looked at in the execution and implementation of Data Science. Farmers are exceptionally distrustful about transforming their cultivating strategies since it is exorbitant to them in case things turn out badly. Moving to advanced cultivating strategies likewise requires an enormous capital, which just large scale and rich farmers can manage. Huge organisations
can get returns generally quickly when contrasted with smallholder farmers (Anonymous, 2016). Uninformed and small scale poor farmers will most likely be unable to actualise and implement advanced farming. Furthermore, they might be unequipped for observing information and data that might be provided to them. Thus, another worry for the execution of data and information-driven solutions is that it might just profit educated and large scale farmers (Kshetri, 2016).

The subsequent issue accompanies any data-based solution: collecting data, cleaning of that data, storing of that data and its circulation through secure channels. Agricultural data is also the same. A farmer ought to incur the expense of buying and installing sensors and connecting them to a centralised server to store and use the local data. The agricultural data comes in different forms and time intervals. For instance, the data provided by a sensor that measures the content of water and pH estimation of soil can be taken at any occasion of time, however those, for example, checking fruits and vegetables for ripeness, which should be gathered physically can't be obtained so promptly. Thus, it is important to change such information or data at a point in time or aggregate it so that it can be compared with standard information or data of comparative nature and kind. The data must be assessed by its reusability, interoperability, handiness, relevance, propriety, and adequacy with the goal that it very well may be utilised in analytics. A standard must be defined, and that must be utilised for a specific zone and kind of plant as a pattern or a baseline for examination. Likewise, farmers ought to impart their local data to different farmers to achieve a significant volume of data to perform analytics. This data should be imparted to trustworthy people only as it might be abused, adapted by another gathering, or might present little rivalry or unfair competition (Maru, 2016).

**Conclusion**

Many farmers, especially those in the mechanical and industrial nations, utilise modern innovation and technology for mechanisation and automation of processes in their homesteads and farms. They have transformed their homesteads and farms into a manufacturing plant and have computerised all cycles imaginable with a substantial computing measure. Data Science is playing a significant part in providing those computerised solutions with high accuracy. With the world's development towards computerised farming, a great deal of investment has been made in this regard. Presently, considerably more than ever, agricultural experts can delve into information and use it to settle on profoundly educated choices. Various research and development projects are proceeding to amplify the proficiency of farms. Fusing innovation will heighten the yields of both large and small-scale farmers. New patterns in cultivating by including data science and analytics will reform the agriculture sector; delivering a better quality of food in a more considerable amount reasonably to achieve the worldwide objective of expanding food creation by 70% will be met by 2050. The future would need more data scientists and agricultural experts to make even better decisions and make the computerised processes highly accurate and affordable so that not just the large and educated but small farmers can also utilise the latest technology for their benefit as well as for the benefit of humanity.

**References**

Anonymous, 2005. Long-lived digital data collections enabling research and education in the 21st century. Retrieved from https://www.nsf.gov/pubs/2005/nsb0540/

Anonymous, 2016. The future of agriculture. Retrieved from https://www.economist.com/technology-quarterly/2016-06-09/factory-fresh

Anonymous, 2018. Taiwan’s rice farmers use big data to cope with climate change. Retrieved from https://www.ft.com/content/9f5438fa-ee2d-11e8-89e8-d36339d855e0

Anonymous, 2019. Animal nutrition solutions. Retrieved from https://www.kemin.com/na/en-us/markets/animal/nutrition

Anonymous, 2018. Startup uses AI to identify crop diseases with superb accuracy – NVIDIA developer news center.

https://www.sciencedaily.com/releases/2018/09/180927145320.htm
Hey, T. 2009. The fourth paradigm: Data-intensive scientific discovery. Microsoft Press.

Kshetri, N. 2016. Big Data’s Big Potential in Developing Economies. Retrieved from https://books.google.com.au/books?id=4fl-DQAAQBAJ.

Leek, J. 2013. The key word in "Data science" is not data, it is science - Simply statistics. Retrieved from https://simplystatistics.org/2013/12/12/the-key-word-in-data-science-is-not-data-it-is-science/

Lohr, S. 2015. The Internet of Things and the Future of Farming. Retrieved October 21, 2020, from https://bits.blogs.nytimes.com/2015/08/03/the-internet-of-things-and-the-future-of-farming/

Matthews, K. 2019. 6 Ways the Agricultural Industry Is Benefiting From Data Scientists. Retrieved from https://towardsdatascience.com/6-ways-the-agricultural-industry-is-benefiting-from-data-scientists-b778d83f61db

Mello, U. and Raghavan, S. 2019 Smarter farms: Watson decision platform for agriculture. Retrieved from https://www.ibm.com/blogs/research/2018/09/smarter-farms-agriculture/

Mittal, R. 2013. Impact of population explosion on environment. WeSchool "Knowledge Builder" - The National Journal vol.1(01)

Press, G. 2013. Data science: What's the half-life of a buzzword? Retrieved from https://www.forbes.com/sites/gilpress/2013/08/19/data-science-whats-the-half-life-of-a-buzzword/

Press, G. 2014. A very short history of data science. Retrieved from https://www.forbes.com/sites/gilpress/2013/05/28/a-very-short-history-of-data-science/

Sain, M., Singh, R. and Kaur, A. 2020. Robotic Automation in Dairy and Meat Processing Sector for Hygienic Processing and Enhanced Production. Journal of Community Mobilization and Sustainable Development, 15(3), 543-550.

Singh, R., Sain, M., Singh, B., Nagi, H. S. and Bala, N. 2020. Development of a Cost Effective Beverage and Food-Serving Robot for Hygienically Outcomes and Human Comfort. International Journal of Current Microbiology And Applied Science 9(5), 247-257. doi: https://doi.org/10.20546/ijcmas.2020.905.028
Data science and analytics in agricultural development

Roser, M. 2014. Future population growth. Retrieved from https://ourworldindata.org/future-population-growth

Talley, J. 2016. ASA Expands Scope, Outreach to Foster Growth, Collaboration in Data Science. Amstat News. American Statistical Association.

Tollefson, J. 2018. Big-data project aims to transform farming in world’s poorest countries. Retrieved from https://doi.org/10.1038/d41586-018-06800-8

Trendov, N. M., Varas, S. and Zeng, M. 2019. Digital Technologies in Agriculture and Rural Areas. Retrieved from http://www.fao.org/3/ca4887en/ca4887en.pdf