# Smart Farming: Big Data and Digital Agriculture

**Manish Kumar*, Dr. (Prof.) Deva Prakash****

*(Research Scholar, Department of Mathematics & Computer Science, Magadh University, Bodhgaya, Bihar
**(Associate Professor, Head of Department, Department of Mathematics, S.M.D. College, Punpun, Bihar)

*Abstract:* **The last several years have seen robust expansion in the agriculture industry. The industry, which employs the most people, contributed a significant 18.8% (2021–2022) of the country's Gross Value Added (GVA), growing by 3.6% in 2020–2021 and 3.9% in 2021–2022. The main factors influencing the sector's overall growth have been expansion in related industries including cattle, dairying, and fishing. The Indian economy's cornerstone is its agricultural sector, which accounts for 15% of the country's Gross Domestic Product (GDP). More crucially, over half of India's population depends entirely or heavily on agriculture and related industries for a living.**

**There are already over 7 billion people on the earth, and by 2050, that number is projected to increase to 9.6 billion.[1] By 2030, there may be 5 billion people in the middle class, who typically have more money to spend on food, driving up demand.[2] The present world population is 7.7 billion, and the United Nations predicts that number to increase to 8.5 billion by 2030 and 9.7 billion by 2050. [3] If these figures are accurate, the world's food supply will need to double in a relatively short amount of time to keep up with demand and feed everyone on the planet.[4] The good news is that cutting-edge digital technologies now make it feasible to gather and use enormous volumes of important data for little or no expense, making a farm's field activities more insight-driven and possibly more effective and productive.**

**The ecosystem for agriculture is already beginning to invest in these digital technologies. Between 2014 and 2020, the market for digital services, also referred to as "precision agriculture," is anticipated to reach $4.55 billion, growing at a CAGR of 12.2 percent.[5] Utilization of precision agricultural services is essential to meeting the growing population's food needs as well as enhancing a farm's financial performance.**

**Although an advance over conventional approaches, the impact of precision agricultural solutions, such as variable-rate spraying, has, up until recently, been constrained by the granularity and timeliness of the data they use and the absence of operational decision support on a day-to-day basis.[5] Big Data, which affects the entire supply chain, is anticipated to have a significant impact on smart farming. Smart sensors and devices provide enormous volumes of data (Big Data) that enable unprecedented decision-making capabilities. Significant changes in power dynamics between traditional and non-traditional players are anticipated as a result of big data. Big data necessitates the use of a variety of novel integration-capable techniques and technologies in order to extract insights from extremely large, diverse, and complicated information[6]. Big Data refers to information assets with such large volume, velocity, and variety that their conversion into value requires specialized technology and analytical techniques[7].**

*Keywords:* **Agriculture, Agriculture Industry, Precision Agriculture, Gross Domestic Product, Precision Agriculture, IOT, Smart Farming, Big Data, Big Data Analytics**

## 1. INTRODUCTION

### 1.1 Overview

The purpose of this paper is to propose an IOT-based smart farming system that will give farmers access to real-time information on soil moisture and environmental temperature at a very low cost, allowing for real-time monitoring.

India is the second most populated nation in the world, with 1.27 billion people. With 3.288 million square kilometers of land, it is the seventh-largest nation in the world. Over 7,500 kilometers of shoreline stretch along it. India is a multicultural nation with 415 dialects and more than 22 major languages. The country is home to a wide Agro-ecological diversity due to its proximity to the largest mountain range in the world, the Himalayas, the Thar Desert, the Gangetic Delta, the East, and the Deccan Plateau in the South. India is the world's second-largest producer of rice, wheat, sugarcane, groundnuts, vegetables, fruit, and jute. It is also the world's top producer of milk, pulses, and jute. Furthermore, it is a major producer of plantation crops, fish, poultry, and spices. After the US and China, India has the third-largest economy in the world, with a GDP of $ 2.1 trillion.

India has a wide variety of habitats and climates, ranging from humid and dry tropical in the south to temperate alpine in the northern regions. Four of the 34 biodiversity hotspots on the planet and 15 of the WWF's top 200 ecoregions are entirely or partially located in India. India is home to over 8% of all species known to exist, including more than 45,000 plant and 91,000 animal species, but having only 2.4 percent of the global land area. The strengthening of agriculture will benefit both farmers and a sizable portion of the rural poor who are either directly involved in agriculture as workers or indirectly as customers. A climate more conducive to the growth of the national economy as a whole and of the rural population in particular would be created in the nation through efficient methods of production, stabilized pricing, and higher income from agriculture.

Small-scale sugarcane farmer Sanjeev Bharti (real name changed) lives in western Maharashtra. He and his family are the landowners of roughly 3 hectares. Bharti has two boys who both have degrees and are employed in Pune. He said that farming is labor-intensive, unprofitable, and challenging to find workers when I questioned him about why he did not raise his boys as farmers. In addition, he believes that his sons live better lives in Pune and that farmer's sons are unmarriable commodities.

As he ages, Bharti is considering selling his land to the highest bidder. He might even travel to Pune to reside with his boys. Now a days, this is the common narrative among Indian farmers, from village to village and state to state. They want to leave farming

and sell their land. Agriculture in India is under trouble. No matter how developed or wealthy we grow, we must all eat. Software and hardware are off limits to us!
India still has a number of pressing challenges, though. From 1951 to 2011, the contribution of agriculture to India's GDP decreased steadily as the country's economy diversified and expanded. India continues to have a quarter of the world's hungry people and more than 190 million undernourished people despite having reached food self-sufficiency in production. At this time, the poverty rate is estimated to be close to 30%. India ranks 114th out of 132 nations for under-5 stunting, 120th out of 130 for under-5 wasting, and 170th out of 185 for anemia prevalence, according to the Global Nutrition Report (2016). In the nation, anemia continues to impact 60% of kids and 50% of women, including pregnant women.
The production is resource-intensive, cereal-centric, and regionally biased even though Indian agriculture has attained grain self-sufficiency. Serious sustainability challenges have also been brought up by the resource-intensive methods used in Indian agriculture. A restructuring and reconsideration of policy would be necessary given the country's increasing demand on its water supplies. Agriculture in the nation is also seriously threatened by desertification and land degradation.
Agriculture-related societal issues have also seen shifting patterns. The increase in male rural-urban migration, the rise of families headed by women, and the expansion of the production of labor-intensive cash crops are the main causes of the rising feminization of agriculture. Women contribute significantly to both agricultural and non-farm operations, and their involvement in the sector is growing. However, their employment is seen as an extension of their household duties, which adds to the weight of domestic duties.
India also has to make numerous improvements in the way it manages agricultural operations. Although there is a tenuous connection between improvements in agricultural productivity and nutrition, the agriculture sector can still make progress in this area by, among other things, raising the incomes of farming households, diversifying crop production, empowering women, enhancing agricultural diversity and productivity, and developing thoughtful price and subsidy policies that should promote the cultivation and consumption of nutrient-rich crops. The diversification of agricultural livelihoods through Agri-allied industries including animal husbandry, forestry, and fisheries has increased chances for livelihood, boosted resilience, and resulted in a significant rise in labor force participation in the sector.

## 1.2 IOT Technology and Agriculture

### 1.2.1 IOT: Concept and Definition

The acronym IOT stands for Internet of Things. Things in the IOT are different IOT devices with distinct identities and the ability to do remote sensing, actuating, and live monitoring of specific types of data. IOT devices can also share data in real time with other connected devices and applications, either directly or indirectly. They can also gather data from other devices, process it, and send it to different servers. The alternative definition of the term "internet" is a global communication network that links trillions of computers worldwide and permits the sharing of knowledge.
The definition of the term "internet" is a worldwide network of computers that allows for the sharing of information. Consequently, the Internet of Things (IOT) can be defined as "a dynamic Global Network Infrastructure with self-configuring capabilities based on standard and inter operable communication to protocol where physical and virtual things have identities, physical attributes, and virtual personalities and use intelligent interfaces and are seamlessly integrated into the information network, often communicating data associated with user and their environment."
An optimal IOT device has a variety of interfaces for connecting to other devices, both wirelessly an over wired network.
Each IOT-based gadget has the following parts:

- Sensor I/O interface.
- interface used to access the internet.
- Memory and storage interface
- Audio/video interface

Wearable sensors, smart watches, IOT smart home monitoring, IOT intelligent transport systems, and other IOT devices are just a few examples of the many different types of IOT gadgets.

### 1.2.2 Technologies that enable IOT

The Wireless Sensor Networks, Cloud Computing, Big Data, Embedded Systems, Security Protocols and Architectures, Protocols facilitating communication, Web Services, Internet, and Search Engines form the foundation of the Internet of Things.
A ***wireless sensor network (WSN)*** is made up of a variety of sensors and nodes that work together to track different types of data.
***Cloud computing***, commonly referred to as on-demand computing, is a category of Internet-based computing that makes pooled processing resources and data available instantly to computers and other devices. It can take many different forms, including IaaS, PaaS, SaaS, and DaaS.
***Big Data Analytics:*** Analyzing huge data sets with a variety of data kinds, or "Big Data," is the process of finding undiscovered patterns, correlations, market trends, consumer preferences, and other pertinent business information.
***Communication protocols:*** These are the building blocks of IOT systems that enable connectivity and coupling to applications. By enabling data exchange formats, data encoding, and data addressing, communication protocols also make it easier to share data over networks. Embedded systems are a type of computer system that combine hardware and software to carry out particular functions. Microprocessor/microcontroller, RAM/ROM, networking elements, I/O units, and storage devices are all included.

### 1.2.3 Applications of IOT in Agriculture

There is a great deal of potential to make everything intelligent and smart as IOT is used in a variety of settings, including industry, homes, and even cities. These days, even the agricultural industry is adopting IOT technology, which has resulted in the creation of "AGRICULTURAL Internet of Things (IOT)".

When we discuss IOT-based smart farming, we are referring to a system created to keep an eye on the agricultural field using sensors. These sensors manage the irrigation system and monitor every factor necessary for crop production, including soil moisture, humidity, light, and temperature. With the help of this system, farmers can keep an eye on their fields from any location. Comparing IOT-based farming to traditional farming, it is far too efficient.

In addition to helping to modernize traditional farming practices, IOT-based smart farming also focuses on organic farming, family farming (complex or constrained spaces, specific cattle and/or cultures, preservation of particular or high-quality varieties, etc.), and farming that is highly transparent. Smart farming based on the Internet of Things is also advantageous for environmental concerns. It can assist farmers in using water effectively and maximizing inputs and treatments.

### 1.3 Key IOT-based smart farming solutions that are transforming the agricultural industry

1. Precision Farming
2. Agricultural Drones
3. Livestock Monitoring
4. Smart Greenhouses
5. Monitor Climate Conditions
6. Remote sensing
    - Crop Assessment
    - Weather conditions
    - Soil quality
7. Computer imaging
    - Quality control
    - Sorting and grading
    - Irrigation Monitoring
8. Data Analytics

### 1.4 Smart Farming: Future of Agriculture

#### 1.4.1 Concept and Definition:

An emerging idea known as "smart farming" refers to managing farms utilizing IOT, robotics, drones, and AI to improve product quality and quantity while eliminating the need for human labor.

Nearly every sector imaginable has seen improvements because to the Internet of Things (IOT). IOT has not only made it possible to complete tasks in agriculture that were previously time-consuming and tiresome, but it is also fundamentally altering the way we view the industry. A smart farm, though, what smart farm is exactly one? What smart farming is and how it's transforming agriculture are described below.

Modern information and communication technology are used to manage farms in a way that increases production while minimizing the need for laborers. This technique is known as "smart farming."

**The following technologies are accessible to modern farmers:**

Sensors for managing soil, water, light, humidity, and temperature

Software: tailored software solutions for particular farm kinds or IOT platforms that are not application-specific

Cellular, wireless LoRa

Location: Satellite and GPS

Robotics: Self-driving tractors and processing plants

***Data analytics:*** Stand-alone analytical programs and data pipelines for following-up programs, with such technologies, farmers can keep an eye on the state of their fields and decide how best to proceed without ever having to set foot in the field, whether it be for their entire operation or just one plant. IOT, which connects devices and sensors installed on farms to make farming activities data-driven and automated, is the driving force behind smart farming.

To increase agricultural production, new technologies are being deployed. Using the most recent technological advancements, such as cloud computing[8], the Internet of Things (IOT) [8], big data [9], data mining [10], and artificial intelligence, aids in improving and modifying agricultural activities. Furthermore, using these technology supports farmers in taking deliberate decisions and actions to improve farming practices.

A management idea called "smart farming" aims to equip the agricultural sector with the infrastructure needed to take advantage of cutting-edge technologies, such as big data, the cloud, and the internet of things (IOT), for tracking, monitoring, automating, and analyzing processes. Most farmers are not familiar with the specifics of smart farming because it is a relatively new concept.

The use of technologies like the Internet of Things, sensors, navigational aids, robotics, and artificial intelligence on your farm is referred to as "smart agriculture." The ultimate objective is to maximize the use of human labor while raising crop quality and yield.

- Precision Irrigation and Precise Plant Nutrition
- Climate Management and control in Greenhouses
- Sensors – for the soil, water, light, moisture, for temperature management
- Agriculture Drones – for taking pictures, spraying the pesticides
- Software Platforms -
- Location based systems – GPS, Satellite, Etc
- Communication Systems – Based on mobile connection, LoraWan
- Robots

- Analytics and Optimization Platforms

The data we can extract from devices and send over the internet is the basis of the Internet of Things (IOT). IOT devices put on a farm should collect and process data in a repeated cycle that enables farmers to respond rapidly to developing challenges and changes in ambient conditions. This will optimize the agricultural process.

The cycle of smart farming looks something like this:

***(1) Supervision*** - Sensors capture observational data from the atmosphere, soil, livestock, and crops.

***(2) Assessment***- The sensor data is sent to an IOT platform housed in the cloud that has predefined decision rules and models (also known as "business logic") that determine the status of the investigated device and pinpoint any needs or shortcomings.

***(3) Making the decisions***- The user and/or machine learning-driven IOT platform components analyze issues to decide whether location-specific therapy is required and, if so, which type.

***(4) Action*** -The cycle repeats from the beginning following end-user assessment and action.

### 1.4.2 IOT impacted the agricultural industry

Sensor uses in agricultural operations is now a thing of the past. However, this conventional approach to sensor technology has the drawback of not providing real-time data. These sensors were previously used to store data in the associated memory and then use that data.

Modern sensors are now used in agriculture thanks to the emergence of industrial IOT. The cellular/satellite network connects these sensors to the cloud. With the aid of this system, we may collect real-time data and make smart decisions.

Farmers have benefited from the use of IOT in several ways, including tank water level monitoring. The fact that everything is done in real time makes the entire irrigation process more effective. The ability to track seed-growth is another development made available by IOT technology. Farmers can now monitor resource use and the length of time it takes a seed to fully develop into a plant. A follow-up to the Green Revolution was the adoption of IOT in agriculture. The farmers have benefited from IOT in two ways. With the aid of precise data collected via IOT, they can now do the same number of chores in less time and also boost agricultural yields.

### 1.5 Big Data

Modern agriculture systems include sophisticated tools and sensors. By using sensing techniques, these devices are connected to soil and crops to gather all the information present in the actual world and convert it into digital data. The internet of things is what this is (IOT). Big data is the term used to describe the enormous volumes of data that these smart gadgets will produce. Big data is defined by the authors of[11] using the following 5Vs:

***Volume:*** Big data can be gathered using digital farming technologies for irrigation from a variety of sources, including manual acquisition, weather data (rainfall, temperature, and humidity), crop data (crop height and surface resistance), soil data (texture, structure, and depth), and aerial systems (drones and aircraft).

***Velocity:*** information gathered almost instantly. One minute might be the lowest possible periodicity.

***Variety:*** the data collection methods may include photographs, excel files, and graphs[12].

***Veracity:*** In order to apply the analytics method, agricultural data must be exact and correct.

The merging and interaction of internal and external agricultural operations is represented by agriculture 4.0, which, like industry 4.0, enables the use of precise digital data to inform choices throughout the agricultural value chain[13]. Moreover, in this instance, the term "farm 4.0," sometimes known as a "smart farm," refers to a land that has been outfitted with tools, procedures, and personnel specifically geared toward digital agriculture[14]. Agriculture 4.0 includes as one of its primary goals the adaptation of production systems and the development of efficiency[15], in addition to the introduction of new methods and instruments. Increased productivity, which is the capacity to produce more, enables the latter. Remotely gather, utilize, and communicate data in real time[16]. Automation and intelligent decision-making so become crucial for achieving these goals[17].

Smart agriculture is advancing thanks to the growth of information and communication technologies, IOT cloud computing, and big data[18]. Digital agriculture is developing as a result of the ability to utilize the enormous amount of agricultural data generated by the activities performed on farms through the use of technologies that combine cloud computing and IOT[19]. Agricultural operations are using big data technologies to give predictive information, enabling real-time operational choices and revamping business processes. The use of these technologies is crucial to the creation of a hypothetical situation in which humans are exclusively responsible for high-level intelligence analysis and action planning[18].

With the explosion of information in recent years, the idea of big data, particularly with the IOT, cloud computing, and other new technologies, continues to grow, giving people a greater understanding of the notion. First, big data describes a lot of data sets, big data is not only very massive but also followed by other indications.

A few fundamental components of enormous data are

(a) containing a specified volume of data;

(b) the composition of the data is somewhat complex;

(c) a hidden value worth mining and additional investigation is present in the data collection;

(d) must be possible to mine the data set's value within a specific time frame and cannot increase indefinitely.

## 2. PROPOSED TECHNIQUE

### 2.1 IOT system for agriculture

Internet- and mobile-based IOT technology, a vast communication network can accommodate all needs. Several domains as science and technology have advanced technology. However, given that agriculture is affected by the growth of IOT technology in agriculture is delayed due to numerous measurable variables.

Agriculture is one of the most essential and important major IOT technology development areas.

A significant number of sensor nodes will create various networks for agricultural monitoring through Using a variety of sensors to gather information, IOT technology aids farmers in locating the locate the problems precisely and locate problems in good time. One use for IOT in agriculture IOT system that integrates manufacturing, processing, and distribution of agriculture.

Utilizing a lot of automated, intelligent, remote-controlled production machinery as part of a software-centered production paradigm. To maximize the revenues of agricultural output, operators, producers, or managers can modify their planting plans in accordance with varied data from the agricultural IOT system. Technology for IOT A hierarchical system is used to implement agricultural at the degree of business. It has five layers, from the bottom to the top.

### 2.2 BIG DATA ANALYTICS AND MACHINE LEARNING FOR SMART FARMING

Big data analytics is a methodology that makes use of internet of things (IOT)-connected devices for data collection and sharing. Internet-connected devices enable measurements that are concerned with soil and crop conditions with sensors put directly on the plant and soil[20]. These huge data will be transmitted via wireless sensor networks along with weather data (rain, temperature), and will subsequently be stored and analyzed using cutting-edge big data analytics technologies to yield practical insights[21]. By fusing the farmer's knowledge with the outcomes of the analysis of the acquired data, big data analytics attempts to make the task easier for the farmer enormous data.

By analyzing historical climate data that are daily updated during the irrigation season, calculating the daily soil water balance, analyzing the soil water data, determining the impact of the weather on the plant, and identifying the driest areas that exist, he can estimate the daily crop irrigation need to be watered first[22].

A data scientist, who is an analyst with a thorough understanding of analytical tools and has training in statistics and mathematics, is also a key player who can help to the solution of the irrigation scheduling problem.

To help the business make strategic or operational decisions, he wants to make sense of the data that is being modified and extract value from it. The data scientist may be in charge of ingestion—the gathering of agricultural data from various sources and sensors—storage—of that data, preparation—the cleaning and filtering of that data (processing), advanced processing (analyzing data using sophisticated tools and analytical methods like deep learning and machine learning), and visualization (returning the result in an interactive way)[23].

Machine learning is one of the technologies that helps big data analytics evolve[24]. It is an algorithm that can learn from information. Through experience, it becomes better at some things.

## 3. RESEARCH METHOD

### 3.1 Apache Spark and databricks platform

The data used in the irrigation industry is divided and stored in the databricks file system (DBFS), a distributed and redundant parallel file system. Additionally, the support vector machine classifier uses machine learning to assess the present dataset. This section describes the planned research technique and describes the technical context of this effort, which is based mostly on the following technologies.

A rapid engine for handling lots of data is Apache Spark. It is a unified framework and open-source tool for large data analytics. It makes it possible to use techniques like graph processing, machine learning, statistics, and structured query language (SQL) to acquire insightful information. Spark is renowned for its distributed parallel processing on affordable hardware[25]. It allows rapid speeds due to in-memory caching and directed acyclic graph (DAG) processing engine. It comes to correct the constraints of Hadoop. For in-memory computations, Spark is in fact 100 times faster than MapReduce in Hadoop. Additionally, it may sort100 TB of data (1 trillion records) can be processed three times faster and with ten times fewer machines than Hadoop's MapReduce.[26]

Apache Spark established the big data processing platform called Databricks. It was developed as a MapReduce replacement for data scientists, engineers, and analysts and offers a just-in-time cloud-based platform for massive data processing. Automated cluster management was offered by the open-source distributed computing platform Databricks. Python, Scala, R, SQL, and other computer languages are supported, in addition to data science frameworks[27]. In order to execute all forms of analytics workloads, it developed its own file format, the data bricks file system (DBFS), which is distributed and immediately mounted on the unified platform[28]. Additionally, it enables the deployment of sophisticated big data analytics via its virtual analytics platform.

### 3.2 Data science and machine learning models

We will use an existing dataset containing data gathered from sensors for the example (temperature and moisture). The platform's databricks will house the used dataset. For uploading data from files and databases, the latter offers a straightforward and user-friendly interface. Once the data have been stored, we will build a cluster and a databricks notebook that is connected to it, as shown in Figure 1. The machine learning technology is used as the study methodology to analyze this data and then reply to the use case for irrigation. Additionally, it is suggested that the neural network and support vector machines be used to represent the state of the pumping motor utilizing the Spark engine. Depending on the temperature and moisture levels, ON or OFF. Before presenting the results, we outline the dataset utilized and go over how a data science model was used in this section [597–60524]. As shown in Figure 2, big data analytics can be performed utilizing machine learning techniques by doing the following:

Finding relevant data that can be kept in databases and comma separated values (CSV) files is known as data intake. To gather data, we can make use of big data solutions like Apache Flume and Apache Sqoop. After that, combine it into one table. Data transformation involves cleaning and preparing the data. This process is for getting the data ready for analysis. Tools like Apache Spark, Hive, or Pig are available.

Prepare a dataset from which we can learn by using training data. Training is the procedure of reviewing the practice data and creating a model. Prepare a data collection for testing to see if our model is accurate. The dataset for the experiment with 200 records in it. The NN and SVM classifiers are trained using 80% of the records, and 20% are employed in testing. The handled data is divided up and kept in the databricks file system (DBFS).

***Data Science Model:*** A rational and sequential series of actions to address a specific issue is a data science model. Once we have a well-structured table, we may model a number of algorithms based on analytical methods and then select the algorithm that produces the best results for deployment based on a survey.

***Prediction:*** Employing SVM and neural network techniques applied to the current dataset, we can first determine whether the soil is dry or not by using a function that predicts target values of the test data supplied in a parameter. A text report with the classification metrics of accuracy, precision, recall (sensitivity), and f1-score should then be used to assess the prediction.

To anticipate the status of the pumping motor based on moisture and temperature, we will use support vector machines and neural network approaches in the demonstration. The notebook we have already built will be used to apply the suggested models.

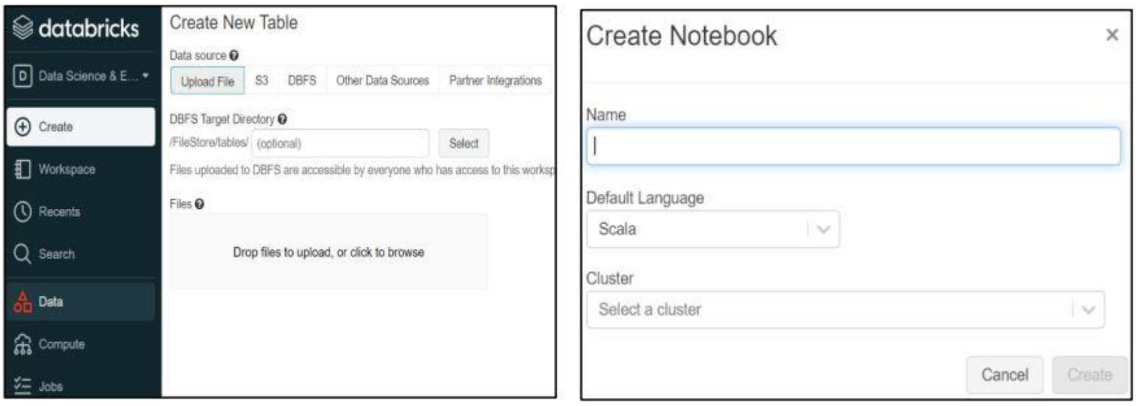


**Figure 1. Interface of databricks**

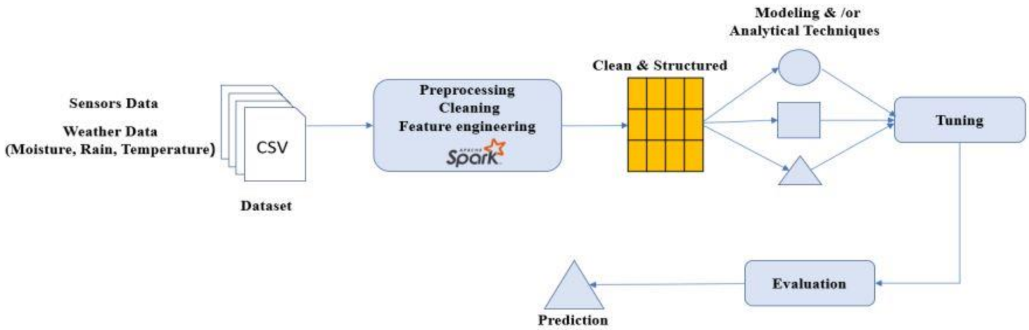


**Figure 2. Data advanced processing using machine learning**

SVMs are a group of machine learning algorithms that may be used to create, train, and classify models utilizing cell record data. Support vector machines categorize data points even when they are not otherwise linearly separable by mapping the data to a high-dimensional features space (this gets done by Kernel function of SVM classifier). The data is modified as a separator between the categories is discovered so that the separator could be drawn as a hyperplane. Equation[29] describes how the SVM model classifies data.

$$Y = sgn(\vec{x} \cdot \vec{w} - b)$$

where: x: is the feature vector
w: is the model weights vector
b: is the bias value

Whether the pump will be ON or OFF is something that neural network (NN) attempts to predict. In fact, temperature and moisture data are present in the input layer. Additionally, the output will be a number 0 or 1, indicating whether the pump is more likely to be ON or OFF. The input and output layers will be separated by a third hidden layer. The output's parameters (neurons) are determined by those layers. All of the levels will be fully interconnected. The network's last (output) layer will then provide the prediction[26].

## 4. RESULTS AND DISCUSSION

The experiment's findings are presented in this section. We demonstrate the class "pump" prediction rate for the used algorithm. Because we have a labeled dataset in this study that is made up of a collection of existing data that includes the target values, we applied supervised learning. The targets may be one of two results (pump ON or OFF) (pump ON or OFF). We employed one of the most well-liked data mining algorithms in this simulation:
NN and SVM.

The used dataset is kept in distributed file system on the databricks platform (15,3 GB memory, 2 cores, 1 DBU, databricks file system), and the dataset contains a significant amount of data relating to moisture, temperature, and other variables temperature, and the position of the pumping motor. If the pump field has a value of 1, pumping is active and the soil is dry; if it has a value of 0, pumping is inactive and the soil is wet.
A cluster is first constructed to act as a database, and then a table is created and attached to the cluster in order to examine the dataset in databricks. Then, a CSV file containing data will be imported, and this table will be filled with our dataset. The CSV file has three columns; as shown in Figure 3, the first two columns include the parameters (moisture and temperature), and the third column denotes the matching class (pump).

| Moisture | temp | pump |
|---|---|---|
| 540 | 18 | 1 |
| 685 | 12 | 1 |
| 721 | 16 | 1 |
| 745 | 22 | 1 |
| 680 | 31 | 1 |
| 544 | 27 | 1 |
| 675 | 24 | 1 |
| 650 | 13 | 1 |

**Figure 3. Example of the dataset used in the simulation**

## 5. CONCLUSION

Big data analytics for smart irrigation was the latest relevant issue we covered in this study. Irrigation scheduling is getting more exact in front of intelligent agriculture, which embodies the widespread global use of internet of thing (IOT) analytic solutions. It gives the farmer access to new technological skills, less hardship on the farm, water resources for future generations, decreased losses, higher yields, and more competitive products.
Based on moisture and temperature parameters, NN and SVM algorithms are used to analyze the current irrigation information. The big data platform "databricks" was used to execute the simulation uses the spark framework as its large data processing engine. SVM and neural network results for the models were good (SVM 96.60%, NN 98.10%). The ISO 19440 enterprise metamodel will be expanded in upcoming work, adding specific constructs related to cutting-edge technologies (big data analytics tools, IOT, AI, cloud, and machine learning), and reflecting a new metamodel that will likely be projected on a smart farm as part of an agriculture 4.0.
Therefore, ranchers and farmers are now able to gather useful data thanks to IOT agricultural applications. Large landowners and small farmers must be aware of the IOT market's potential for agriculture and implement smart technology to boost their productions' sustainability and competitiveness. If ranchers and small farmers successfully use agricultural IOT solutions, the need can be satisfied despite the population's rapid growth.
According to FAO predictions, the world would need to produce 70% more food by 2050 as a result of the exponential rise in human population. As a result, agricultural fields would be reduced in size and limited natural resources will be used up. Therefore, it becomes imperative to boost crop production. IOT may therefore play a significant role in this process.
Farmers and ranchers can now practice smart farming thanks to the Internet of Things. a high-tech, capital-intensive process. Smart farming offers two advantages: it allows farmers to spend less time in the fields while yet increasing crop yields. Numerous uses for the IOT-based ecosystem exist in the agricultural industry. The applications have been covered in great detail. We can draw a conclusion from the fact that IOT apps enable farmers to get useful data that is used to improve efficiency. The promise of IOT-based smart farming must be understood by large landowners and small farmers, and they must successfully apply IOT solutions.

**References:**


1. World population projected to reach 9.6 billion by 2050 with most growth in developing regions, especially Africa says UN. © 2013 United Nations. Reprinted with the permission of the United Nations. http://esa.un.org/wpp/documentation/pdf/wpp2012_press_release.pdf
2. 2 OECD (2012), OECD Yearbook 2012: Better Policies for Better Lives, OECD Publishing. http://www.oecd-ilibrary.org/economics/oecd-observer/volume-2011/issue-5_observer-v2011-5-en
3. [1] United Nations, World Population Prospects 2019. Department of Economic and Social Affairs. Population Division (2019).
4. 3 Food Production Must Double by 2050 to Meet Demand from World's Growing Population, Innovative Strategies Needed to Combat Hunger, Experts Tell Second Committee. General Assembly, Meetings Coverage. © 2009 United Nations. Reprinted with the permission of the United Nations. http://www.un.org/press/en/2009/gaef3242.doc.htm
5. Precision Farming Market by Technology (GPS/GNSS, GIS, Remote Sensing & VRT), Components (Automation & Control, Sensors, FMS), Application (Yield Monitoring, VRA, Mapping, Soil Monitoring, Scouting) and Geography - Global Forecasts to 2020", MarketsandMarkets, October 2014. http://www.marketsandmarkets.com/Market-Reports/precision-farming-market-1243.html

6. Hashem et al., 2015 I.A.T. Hashem, I. Yaqoob, N.B. Anuar, S. Mokhtar, A. Gani, S. Ullah KhanThe rise of "Big Data" on cloud computing: Review and open research issues
7. De Mauro et al., 2016 A. De Mauro, M. Greco, M. Grimaldi A formal definition of Big Data based on its essential features Libr. Rev., 65 (2016), pp. 122-135
8. N. Ahmed, D. De, and I. Hussain, "Internet of Things (IOT) for Smart Precision Agriculture and Farming in Rural Areas," IEEE Internet Things J., vol. 5, no. 6, pp. 4890–4899, 2018, doi: 10.1109/JIOT.2018.2879579.
9. X. Pham and M. Stack, "How data analytics is transforming agriculture," Bus. Horiz., vol. 61, no. 1, pp. 125–133, 2018, doi: 10.1016/j.bushor.2017.09.011.
10. J. Muangprathub, N. Boonnam, S. Kajornkasirat, N. Lekbangpong, A. Wanichsombat, and P. Nillaor, "IOT and agriculture data analysis for smart farm," Comput. Electron. Agric., vol. 156, no. January 2019, pp. 467–474, 2019, doi: 10.1016/j.compag.2018.12.011.
11. L. Rabhi, N. Falih, A. Afraites, and B. Bouikhalene, "Big Data Approach and its applications in Various Fields: Review," Procedia Comput. Sci., vol. 155, pp. 599–605, 2019, doi: 10.1016/j.procs.2019.08.084.
12. H. Cadavid, W. Garzón, A. Pérez, G. López, C. Mendivelso, and C. Ramírez, "Towards a smart farming platform: From IOT-based crop sensing to data analytics," Communications in Computer and Information Science, vol. 885, pp. 237–251, 2018, doi: 10.1007/978-3-319-98998-3_19.
13. I. Zambon, M. Cecchini, G. Egidi, M. G. Saporito, A. Colantoni, Revolution 4.0: Industry vs. agriculture in a future development for SMEs, Processes, 2019, Vol. 7, No. 1, p. 36.
14. M. J. O'Grady e G. M. P. O'Hare, Modelling the smart farm, Inf. Process. Agric., 2017, Vol. 4, No 3, pp. 179–187.
15. D. Andritoiu, L. Bazavan, F. Besnea, H. Roibu, N. Bizdoaca, Agriculture autonomous monitoring and decisional mechatronic system, in 2018 19th International Carpathian Control Conference (ICCC), 2018, p. 241–246.
16. V. Bonneau, B. Copigneaux, L. Probst, B. Pedersen, Industry 4.0 in Agriculture: Focus on IOT Aspects, European Commission, 2017.
17. D. Yane, "Research and Analysis about System of Digital Agriculture Based on a Network Platform", in International Conference on Computer and Computing Technologies in Agriculture, 2010, p. 274– 282.
18. S. Wolfert, L. Ge, C. Verdouw, e M.-J. Bogaardt, "Big Data in Smart Farming – A review", Agric. Syst., vol. 153, p. 69–80, 2017.
19. S. Rotz et al., "The Politics of Digital Agricultural Technologies: A Preliminary Review", Sociol. Ruralis, vol. 59, no 2, p. 203–229, 2019.
20. K. Singh, S. Jain, V. Andhra, and S. Sharma, "IOT based approach for smart irrigation system suited to multiple crop cultivation," Int. J. Eng. Res. Technol., vol. 12, no. 3, pp. 357–363, 2019.
21. Y. Shekhar, E. Dagur, S. Mishra, R. J. Tom, M. Veeramanikandan, and S. Sankaranarayanan, "Intelligent IOT based automated irrigation system," Int. J. Appl. Eng. Res., vol. 12, no. 18, pp. 7306–7320, 2017.
22. V. Jinubala, R. Lawrance, and P. Jeyakumar, "Classification of Rice Pest Data Using Decision Tree Algorithm," Int. J. Res. Advent Technol., vol. 7, no. 5S, pp. 148–154, 2019.
23. M. N. Islam Sarker, M. Wu, B. Chanthamith, S. Yusufzada, D. Li, and J. Zhang, "Big Data Driven Smart Agriculture: Pathway for Sustainable Development," 2019 2nd International Conference on Artificial Intelligence and Big Data (ICAIBD), 2019, pp. 60-65, doi: 10.1109/ICAIBD.2019.8836982.
24. U. Moorthy and U. D. Gandhi, "A Survey of Big Data Analytics Using Machine Learning Algorithms," pp. 95–123, 2017, doi: 10.4018/978-1-5225-2863-0.ch005.
25. C. -Y. Lin, C. -H. Tsai, Ching-Pei Lee, and C. -J. Lin, "Large-scale logistic regression and linear support vector machines using spark," 2014 IEEE International Conference on Big Data (Big Data), 2014, pp. 519-528, doi: 10.1109/BigData.2014.7004269.
26. K. Jin and S. Bagui, "Apache Spark SVM for Predicting Obstructive Sleep Apnea," Big Data Cogn. Comput., vol. 4, no. 4, 2020, doi: /10.3390/bdcc4040025.
27. R. Singh and J. Woo, "Applications of machine learning models on Yelp data," Asia Pacific Journal of Information Systems, vol. 29, no. 1, pp. 65–49, 2019, doi: 10.14329/apjis.2019.29.1.35.
28. Y. O. Sayad, H. Mousannif, and H. Al Moatassime, "Predictive modeling of wildfires: A new dataset and machine learning approach," Fire Safety Journal, vol. 104, pp. 130–146, 2019, doi: 10.1016/j.firesaf.2019.01.006.
29. L. Wang, Y. Yu, L. Deng, and H. Pang, "A two-stage agriculture environmental anomaly detection method," Commun. Comput. Inf. Sci., vol. 763, pp. 779–789, 2017, doi: 10.1007/978-981-10-6364-0_77.