Journal of Advances in Science and Technology
Vol. 20, Issue No. 2, September-2023, ISSN 2230-9659

Data mining for environmental analysis and effect of weather on agriculture

Namrata Kumari*

Research Scholar (Computer Science), Magadh University, Bodh Gaya, Bihar, India

Email- namratakrishnasinha@gmail.com

Abstract- The present study deals with "Data mining for environmental analysis and effect of weather on agriculture." A government platform's cluster largest sample on the shifting climate toward agriculture has been gathered. The question of whether agriculture has been impacted by the changing climate was examined. In practically every economy, the use of data mining techniques for environmental analysis in agriculture is important. Thus, it can be concluded from the statistics that agriculture generates enormous amounts of data per second at a high pace and variety. It is quite challenging to analyze these data and make judgments based on the data using conventional tools and procedures.

As a result, a significant quantity of goods as well as labor, fertilizers, and other resources are squandered. To study the data, big data analysis is the most effective method. The government's statistics clearly demonstrates the huge impact that data mining techniques have on weather and agriculture. The primary goal of this paper is to present the discussion of big data analytics and how its tools are used in the agricultural sector. The analysis revealed that the algorithm (data base technology) is a major factor in the weather information provided by agricultural platforms. The idea is supported by the findings.

Keywords: Data Analysis, Agriculture, Tools and Techniques, effect of weather.

- x -

INTRODUCTION

Farmers and other stakeholders must make critical decisions due to the expanding demands of the agriculture industry, which are influenced by a number of variables including soil type, pollution level, humidity, temperature, rainfall, and geographic features. The paper discusses the different data mining methods for examining environmental influences on agricultural indicators. These methods provide answers to a range of decision-making issues that the agriculture industry currently faces. The primary goal is to maximize the impact of weather on agriculture by employing a variety of strategies.

Data mining, an interdisciplinary subfield of computer science, is the computational pathway of determining outlines in large data sets involving ways at the intersection of artificial intelligence, machine learning, statistics, and database systems. It is the superior scrutiny step of the "familiarity find in Databases" progression, or KDD. The overarching goal of data mining development is to extract information sequentially from a data set and transform it into a structure that can be easily understood for further applications. In addition to the first stage of study, it includes database and data management, pre-processing of the data, considerations of copy and speculation, interestingness metrics, and complexity, as well as post-processing of the configurations that are found, visualization, and online updating.

The term is an exhortation that is frequently misused to refer to any type of large-scale data or sequence processing (anthology, extraction, warehousing, breakdown, and information). However, it can also refer to any class of computer decision support system, including business intelligence, artificial intelligence, and machine learning. The essential phrase in correctly using the word is uncovering, which is generally understood to mean "detecting a bit new."

The definitive task of data mining is the iterative or partially automated examination of vast amounts of data in order to eliminate previously undefinable intriguing patterns, such as dependencies (association rule mining), strange records (irregularity detection), and clusters of data proofs (cluster analysis). Spatial indexes and other database systems are often used in this. These summaries can therefore be viewed as a type of summary of the incoming data and can be applied to additional research or, for instance, extrapolative analytics and machine learning. A decision support system can leverage the data mining step power, for instance, to find many groups in the data and use that information to provide predictions that are more accurate.

In order to make better judgments, people today want to gather data, but they also need to comprehend it and recognize the significance of the data collection. Big data is defined as an extensive dataset collection with high velocity, volume, and variety that is challenging to handle and process using conventional methods and tools. It is possible for it to be semi-structured, unstructured, or structured. Big data analytics is an advanced analytical method that may be used to analyze large amounts of data in order to find patterns that are helpful, hidden, and unknown. Big data is therefore crucial to the decision-making process. Due to nations' rapid economic expansion, information technology is now widely used in practically every industry.

Additionally, agriculture is a significant contributor to the economies of all nations. The swift advancement of information technology will facilitate the gathering, archiving, and analysis of data to extract valuable insights. Technology has two uses in agriculture: it can be used directly to boost productivity and indirectly to motivate farmers to make wiser choices. IT is primarily used in precision farming, which is common in developed nations, to boost productivity. Despite being widely used in other industries, big data analytics is not used in agriculture. It has been determined that by 2050, there will be more people on the planet than nine billion.

Consequently, it is crucial to produce and transport goods in an effective manner. Utilizing big data analytics in agriculture yields extremely advanced benefits, such as creative ways to reduce resource usage while producing promising results similar to precision farming. Accuracy Finding knowledge from databases is called data mining. The computer process of sifting through data analysis and then deriving its significance is known as knowledge discovery.

Data mining techniques are useful for providing answers to business questions that took too little time in the past. Data mining tools are useful for different purposes. They go through data sets to uncover facts that experts would overlook because it deviates from their preconceived notions. Businesses may be proactive and create predictions about behavior, finances, future trends, knowledge-driven decisions, and bioinformatics with the use of data mining tools and techniques.

Retail, manufacturing, transportation, healthcare, and medical science are just a few of the industries that find data mining to be advantageous. Data mining tools and techniques are already widely used in finance and aerospace to leverage historical data. Data mining analysts identify important facts, patterns, and relationship exceptions by sifting through warehoused information using model thanks technology and statistical and mathematical methodologies.

Data mining is used for business purposes to find patterns and relationships in data to assist make better business decisions. Data mining can be used to identify sales trends, create more intelligent marketing strategies, and anticipate consumer loyalty.

The concept and physical paradigm of data mining have been around for centuries. The 1700s saw the development of the Bayes theorem and regression analysis, two early techniques for finding patterns in data. The expansion, pervasiveness, and increasing capability of computer technology have enhanced data grouping, strategy gifting, thinking, and storage. Concepts related to data mining expand the scope of study. With the growth of data sets, indirect mechanical data processing has been used more and more. This has been made possible by other advancements in computer science and information technology, including neural networks, cluster analysis, and genetic algorithms, which were discovered in the 1950s.

Decision trees for residential data mining study (1960s). Support vector machines were also urbanized by data mining researchers in the 1990s. Using data mining is beneficial when using these techniques. These techniques reveal hidden patterns in large-scale datasets. Artificial intelligence, bioinformatics, and applied statistics are all connected through data mining. Through exploitation, data mining offers the mathematical foundation for database administration. Data mining is the storage and indexing of data in databases to enhance the efficiency of specific learning and sensing algorithms.

These techniques can be used to a wide range of data sets thanks to data mining. The field of data mining is ideally interdisciplinary. There are various ways to explain data mining. The terms "knowledge mining from database" and "knowledge discovery from data set" are more akin to what data mining actually is. Concepts and data mining are often used interchangeably. It's for a different widely used word, and others see data mining as just a necessary phase in the data mining process.

Data mining as a step in the process of knowledge discovery-

- **Data Cleaning-** The procedure data cleaning method removes noise and inconsistent data.
- **Data Integration**- The of data integration where compound data sources may be shared.
- **Data selection-** In data variety where data germane to the testing duty are retrieved from the data base.

- **Data transformation-** The process of data makeover where data significant to the breakdown mission are retrieved.

- **Data mining-** Data mining techniques are the awareness ruling form catalogue. Many types of concepts apply for data mining.

- **Pattern evaluation-** Pattern costing is to conclude the fascinating patterns indicating knowledge.

- **Knowledge presentation-** The process of knowledge presentation where image. And knowledge production system presents the knowledge discovery from data base.

- **Therefore, the analysis of real-time data—** which arrives in vast quantities and from a variety of sources, including soil, weather, air, equipment, availability, etc.—is crucial to agriculture. Smart farming focuses mostly on how to improve business operations through the use of big data analytics. Furthermore, data is growing faster than computing power. 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.

The remainder is set up as follows. In section two, an overview of big data, big data analytics, and the tools used to analyze large data are covered. The main studies on big data analytics in agriculture will be covered in section three. Big data analytics challenges are covered in part four, while applications for big data analytics in agriculture are covered in section five. The conclusion is the main topic of section six. Section seven will conclude with a discussion of future directions.

Data Analytics in Agriculture

Data analysis is effectively being used in sectors such as finance and insurance. 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, labor costs, and even equipment can be used to make decisions. Big data utilization in agriculture reduces farmer failures and makes recommendations for soil, water level, etc. Better services are provided when the analysis is done on a regular basis. Smart farming and precision agriculture (PA) are the two main categories under which data analytics in agriculture can be investigated.

Precision agriculture refers to a farming management approach that involves measuring, observing, and responding. PA requires data gathering, analysis, and processing of the information obtained. It makes use of big data techniques to maximize production while consuming the fewest resources possible. Among the technologies utilized in precision agriculture are the Global Positioning System (GPS), Geographic Information System (GIS), and Variable-rate Technology (VRT). Computer scientists, particularly those in the data analysis industry, have enormous opportunities as well as obstacles with precision agriculture. The connections between variables, concepts, and functions are explained by smart farming.

This focuses on the business processes and agricultural value chain applications of big data analytics. between variables, concepts, and functions. The focus is on big data analytics applications, such as corporate processes and the agricultural value chain.

Decision trees for residential data mining study (1960s). Support vector machines were also urbanized by data mining researchers in the 1990s. Using data mining is beneficial when using these techniques. These techniques reveal hidden patterns in large-scale datasets. Artificial intelligence, bioinformatics, and applied statistics are all connected through data mining. Through exploitation, data mining offers the mathematical foundation for database administration. Data mining is the storage and indexing of data in databases to enhance the efficiency of specific learning and sensing algorithms. These techniques can be used to a wide range of data sets thanks to data mining.

Data Mining: The field of data mining is ideally interdisciplinary. There are various ways to explain data mining. The terms "knowledge mining from database" and "knowledge discovery from data set" are more akin to what data mining actually is. Concepts and data mining are often used interchangeably. It's for a different widely used word, and others see data mining as just a necessary phase in the data mining process. Figure depicts the knowledge discovery process as an iterative run through the subsequent steps. This figure illustrates the process of data mining.

Environmental Data Mining- The process of obtaining knowledge from massive volumes of environmental data is known as environmental data mining. Environmental information management systems, recommender systems, decision support systems, environmental data analytics, and other related fields are all part of this multidisciplinary field that combines computer science with environmental sciences. An ensemble of learners is trained successively on the dataset in the relatively new field of environmental data mining based on ensemble learning in order to better analyze and comprehend environmental processes and systems. Nevertheless, how ensemble methods might be applied to enhance the performance of a particular method is still poorly understood.

HYPOTHESIS

sWhether data mining techniques on environmental analysis will effect of weather on agriculture.

METHODOLOGY

In order to find state-of-the-art in algorithms, technique, and problem domains, data mining For environmental analysis was conducted. The WEKA toolbox has been used to evaluate the method for data gathering and storage. The information gathered from the Open Data Government portal. Information Agriculture is becoming increasingly vulnerable to climate change due to the mining of vast amounts of crops, the soil used for their cultivation, shifting climate data, and non-experimental data optimization. A predictive data mining technique will be used to forecast how crops, fertilizers, and pesticides will be used in the future, as well as how much money will be needed for crop development and function in agriculture.

RESULTS

Smart agriculture, or the application of technology to improve and optimize agricultural processes, heavily relies on data analytics. Farmers and agribusinesses can obtain important insights into their operations and make data-driven decisions to raise yields, save expenses, and boost efficiency by utilizing data analytics tools and methodologies. The following are some applications of data analytics in agriculture:

- **Predictive analytics:** Data analytics can be used by farmers to predict insect infestations, weather trends, and agricultural yields. Farmers can anticipate crop loss and take preventative measures by analyzing sensor data in real time and previous data.

- **Precision farming:** Precision farming is the application of data analytics to agricultural methods that are customized to particular farm regions in order to maximize crop yield. Farmers may optimize crop yields and save waste by adjusting irrigation, fertilizer application, and seed planting strategies through the analysis of soil and moisture data.

- **Livestock management:** Data analytics can be used by livestock farmers to track feed consumption and growth rates, as well as to keep an eye on animal behavior and health. Farmers can use this information to make better-informed decisions on veterinarian care, feeding, and breeding.

- **Supply chain management:** Using data analytics, the agricultural supply chain—from seed production to distribution—can be made more efficient. Agribusinesses may reduce waste and guarantee on-time product delivery by analyzing data on demand, transportation, and inventory levels. All things considered, data analytics can assist agribusinesses and farmers in increasing productivity, cutting expenses, and making better judgments. Data analytics will probably play an even bigger part in smart agriculture as technology develops further. Processing, analyzing, and interpreting data using a variety of approaches and methodologies is the focus of the quickly developing discipline of data analytics.

- **Real-time estimates:** These estimates give decision-makers in many different businesses the ability to make well-informed judgments based on current facts. As such, they are a valuable tool. In agriculture, real-time estimations can refer to the employment of several tools and methods to collect and process data instantly in order to increase farming productivity and efficiency. Real-time estimations in agriculture include the following examples:

- **Weather and climate monitoring:** Farmers can use real-time climate and weather data to inform their planting, irrigation, and harvesting decisions. This can be accomplished by using a variety of sensors and weather stations that gather information and give farmers updates in real time.

- **Soil monitoring:** Farmers that use real-time soil monitoring can make more educated decisions regarding fertilization and other farming methods by having a better understanding of the health and nutrient levels of their soil.

- **Crop monitoring:** Drones, satellites, and other sensors can be used in crop monitoring to get current data on the health, growth, and production of the crop. This can assist farmers in seeing issues early on and fixing them before it's too late.

- **Livestock monitoring:** Farmers may discover possible issues and take action before they worsen by keeping an eye on the health and behavior of their animals through real-time monitoring.

Technologies used for real-time soil monitoring include:

- **Soil sensors:** You can install soil sensors in the field to monitor temperature, moisture content, and other attributes.

This information can be transmitted in real-time to a central system for analysis.

- **Wireless networks:** Data from soil sensors and other monitoring systems can be sent across wireless networks to a central system for analysis.

- **Satellite imagery:** Satellite images can be used to track variations in soil moisture content and other properties over time, giving farmers important information on the fertility and health of their land.

- **Automated irrigation systems:** With the ability to program irrigation systems based on real-time soil moisture data, farmers can ensure that their crops receive the right amount of water without wasting any resources.

- **Weather stations:** Data on temperature, humidity, wind speed, and other meteorological factors can be gathered using weather stations. Wireless transmission of this data to a central system for analysis is possible.

- **Satellites:** Large-scale weather patterns can be observed using satellites, which can also provide real-time information on precipitation, cloud cover, and other meteorological factors.

- **Doppler radar:** Farmers and emergency personnel can take prompt action by using Doppler radar to detect and track severe weather occurrences like thunderstorms and tornadoes.

- **Climate models:** Based on past data and present trends, climate models can be used to simulate and forecast future weather patterns.

CONCLUSION

"Data mining for environmental analysis and effect of weather on agriculture" was the study's analysis. By merging time series data from previous and current years, data analytics can assist increase productivity in the future by extracting insights from the data using the right algorithms. Additionally, data analytics helps farmers understand environmental trends better, which helps them plan ahead and seize opportunities without wasting resources. It was too late to take any action because farmers were unable to anticipate the early warning indications of failed crops prior to the development of data analytics.

Scientists can use data analytics to forecast future harvests and search for signs of disease in plants to find out if crops are susceptible to disease. Agriculture in the future will be greatly dependent on technologies like as GPS, aerial photography, robotics, and sensors. Precision farming, contemporary machinery, and robotic systems will increase the efficiency, safety, and environmental friendliness of agriculture. Data mining for weather and environmental analysis has a legitimate impact on agricultural practices. Since the weather and environment are entirely up to the farmer, the government has supplied precise data figures for farming in the future, making it a useful tool for agriculture and farming.

REFERENCES

1. Ramesh, D., and B. Vishnu Vardhan. "Data mining techniques and applications to agricultural yield data." International journal of advanced research in computer and communication engineering 2.9 (2013): 3477-3480.

2. S. A. Yadav, B. M. Sahoo, S. Sharma and L. Das, "An Analysis of Data Mining Techniques to Analyze the Effect of Weather on Agriculture," 2020 International Conference on Intelligent Engineering and Management (ICIEM), 2020, pp. 29-32, doi: 10.1109/ICIEM48762.2020.9160110.

3. Deepak Sharma, Priti Sharma "Rain Fall Prediction using Data Mining Techniques with Modernistic Schemes and Well-Formed Ideas" International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-9 Issue-1, November 2019

4. P.Kalaiyarasi [1], Mrs.A.Kalaiselvi [2] "Data Mining Techniques Using To Weather Prediction" International Journal of Computer Science Trends and Technology (IJCST) – Volume 6 Issue 3, May - June 2018

5. Hira, Swati, and P. S. Deshpande. "Data analysis using multidimensional modeling, statistical analysis and data mining on agriculture parameters." Procedia Computer Science 54 (2015): 431-439.

6. Sneha, N & Sushma, K & Muzumdar, Surekha. (2019). Precision Agriculture using Data Mining Techniques and IOT. 376-381. 10.1109/ICAIT47043.2019.8987333.

7. Majumdar, J., Naraseeyappa, S. & Ankalaki, S. Analysis of agriculture data using data mining techniques: application of big data. J Big Data 4, 20 (2017).

8. Tek B. Sapkota a,⁎, Sylvia H. Vetter b,M.L. Jat a, Smita Sirohi c, Paresh B. Shirsath d, Rajbir Singh e, HanumanS. Jat f, Pete Smith b, Jon Hillier g, Clare M. Stirling h "Costeffective opportunities for climate change mitigation in Indian agriculture" Science of the Total Environment 655 (2019) 1342–1354

9. Jeetendra Prakash Aryal, Tek Bahadur Sapkota, Dil Bahadur Rahut, M.L. Jat

"Agricultural sustainability under emerging climatic variability: the role of climatesmart agriculture and relevant policies in India" International Journal of Innovation and Sustainable Development · January 2019 DOI: 10.1504/IJISD.2019.10020869

10. EREIRA, G.C. Data Mining for environmental analisys and diagnostic. Ph.D. dissertation, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil. (In portuguese). 2005.

Corresponding Author

Namrata Kumari*

Research Scholar (Computer Science), Magadh University, Bodh Gaya, Bihar, India

Email- namratakrishnasinha@gmail.com