Cyber-Physical System Based Data Mining and Processing Toward Autonomous Agricultural Systems

Abdulrahman M. Abdulghani
Jackson State University
Jackson, Mississippi
abdulrahman.m.abdulghani@gmail.com

Mokhles M. Abdulghani
Jackson State University
Jackson, Mississippi
mokhles.m.abdulghani@students.jsums.edu

Wilbur L. Walters
Jackson State University
Jackson, Mississippi
wilbur.l.walters@jsums.edu

Khalid H. Abed
Jackson State University
Jackson, Mississippi
khalid.h.abed@jsums.edu

***Abstract*—Recent advancements in the internet and the proliferation of sensing devices have made it possible to deploy and communicate heterogeneous and bridge data in a variety of systems that link physical objects to the real world. This breakthrough also provides a lot of benefits for the farming industry, including improved resource leadership and human workforce. With objective evidence gleaned by sensors with the goal of increasing production and durability, principal benefits emerge. This type of automated and data-driven farm management relies on data to boost efficiency while reducing resource waste and environmental contamination. Smart farming, along with automated alternatives that use Artificial Intelligence (AI) approaches, lays the foundation for future food production. This paper proposed an Autonomous Agricultural Cyber-Physical System (AA-CPS) as a framework to predict the precise crop that fits the farm based on the soil and weather data. The data is collected by sensory technology, associated data mining techniques, and autonomous tractors in the field leveraging the best recommendation about what crop to be grown on any farm around the world.**

Keywords— *Cyber-Physical System, Precision Agriculture, Data Mining, Artificial Intelligence (AI), Autonomous Tractors, Crop Recommendation*.

I. INTRODUCTION

Agriculture is undergoing a revolution driven by modern knowledge, which appears to be highly capable meanwhile it will permit this key division to clamp a new territory of agricultural revenue and profit margins. Precision farming, which consists of implementing inputs (what is needed) where and when they are required, has become the third phase of the advanced agricultural revolution ("the first was mechanization and the second was the green revolution with its genetic modification"), and it is already being improved with the accessibility of high volumes of data. Improved agricultural technology, according to the United States Department of Agriculture (USDA), boosted net returns and operational profitability in October 2016 [1].

Nevertheless, when it approaches to nature, novel representations are currently being used alongside farms to ensure the long-term viability of agricultural produce. However, the use of these technologies is uptight with risks and trade-offs. As per a market study, variables that might help farmers embrace sustainable farming methods include increased farmer support and learning, information exchange, simple wealth accumulation, and growing consumer desire for organic food [2]. Moreover, when using these emerging innovations to get data from plants, the problem is to produce something logical and meaningful since data is nothing more than numbers or pictures. Farms that choose to be innovative in some way benefit in a variety of ways, including reducing the cost and time, increasing productivity or lowering expenses with less effort, and providing high-quality food with more ecologically friendly techniques. However, bringing these benefits to the farmland will be contingent on farmers' desire to employ new technology in their fields and on each farm's capacity for economies of scale and scope since profit margins rise with farm size [3].

Lastly, the focus of this research is to show how using today's current data-based agriculture environment (weather and soil) to make decisions may contribute to successful and lucrative actions that feed people while decreasing the negative environmental impact. This paper explores the major stages of a data agribusiness and concentrates on data processing by examining current finds that are suitable for each critical stage, from data capture in farmlands to the implementation of work with floating interest devices, in order to determine how contemporary agriculture can help in a self-sustaining crop choice procedure. Figure 1 shows some major problems of precision farming.

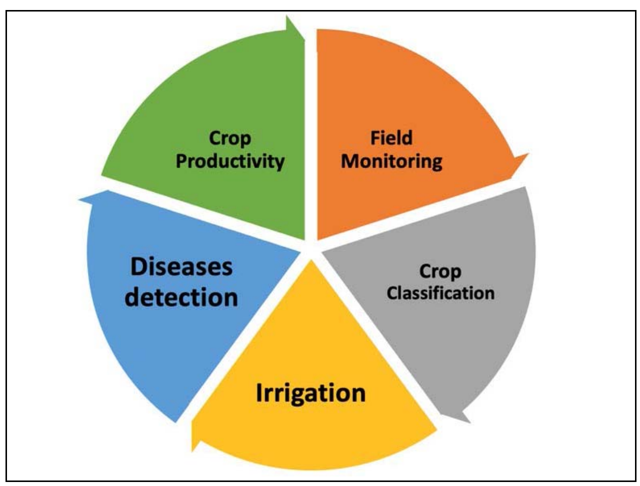

Figure 1: Major Problems of Precision Farming.

979-8-3503-2028-2/22/$31.00 ©2022 IEEE
DOI 10.1109/CSCI58124.2022.00214

This paper aimed to employ soil and weather data as the most important factors in growing a crop. The existing works have many limitations that focus on predicting what crop is perfect for a specific environment. Moreover, Cyber-Physical System (CPS) technologies find a way to manage the farms efficiently with ease the roles of humans and machines toward precision agriculture [4]. Additionally, we may conclude the problem of this research as follow:

- Agricultural CPS focuses on the physical parts (sensors) more than the cyber techniques such as Big Data Analysis, Data Mining, and Machine Learning to get better results for potential solutions that may be extracted from the collected data [5].
- Most of the existing works studied one kind of crop in a specific geographical spot around the planet without consideration for the climate changes that may affect other kinds of crops in the same location [6].

The remainder of the paper is organized as follows. Section II discusses the related work. Section III presents the proposed framework. Section IV discusses the results of this paper. In Section V, we provide a conclusion.

II. RELATED WORKS

The "US National Science Foundation" created the term "Cyber-Physical System" (CPS) in 2006. CPS has both cyber (software administration) and physical (hardware) aspects, as the name implies. Cyber-Physical systems (CPSs) are complex schemes that are meant to communicate with the networking of technological and physical elements in a consistent way. These technologies will serve as the backbone of our vital infrastructure, enhancing the superiority of being in a diversity of ways. CPSs and associated systems ("such as the Internet of Things and the Industrial Internet") have the potential to have a global influence on numerous areas of the economy [7]. CPS can also be found in medicine, autos, electrical energy grids, infrastructure needs, industry, aerospace, autonomous systems, and construction methods, among other places.

CPSs are made up of a collection of decentralized hardware, software, and networking components that are integrated into complex processes and surroundings. The most significant function is played by software, which comprises all software applications for processing, filtering, and storing data. Through networks, CPSs communicate with the physical system. Distributed real-time, scalability, and dependability are all important features of CPS. Real-time operations like real-time surveillance, actual power, and real-time prediction are supported by the majority of CPSs [8] and [9].

Primary observations of critical crop descriptions obligation to be quickly treated so that facts or images can be decoded into valuable material without haziness. Harvest control obtained from arena statistics had now matured by the time Exactitude Farming has introduced 30 years ago, but the present numerical evidence time has unquestionably revolutionized it. Field administration has continuously consisted of visually assessing the progress of crops to come to an assessment with which producers make judgments and activate applying alternative medications to their crops in places where automation has not yet been reached. This method is based on field experience and data viewed from the perspective of farmers [10]. Nevertheless, connected farmers might also follow the advice of cooperative technicians or architects engaged by the society to which they belong.

This technique, which is founded on accurate ground observations and sound decision-making, starts with the plant being handled, taking advantage of both space-time climate variations [5]. The Agricultural Cyber Physical System (ACPS) paradigm to agricultural production, which is founded on factual data obtained and rational decision, starts with the crop to be controlled, utilizing both space-time internal diversity. The sensors are the particular variables that gather real information, whereas the platforms are the mechanistic explanation that receives information. Data refers to the data that is directly acquired from the crop, soil, or ambient properties. Values from the sensors may be obtained in a variety of ways, including by inserting a pen drive into the device. Data from the sensors may be retrieved in a variety of methods, ranging from plugging a pen drive into a USB connection to data retrieval through software programs that are synced with the Internet [8]. Filtering procedures and AI algorithms are used at the nexus here between analysis-driven decision stages to ensure that only the proper data is obtained and that the farmer makes the best judgments possible. Finally, actuation is defined as the physical implementation of a decision framework action, which is often accomplished by sophisticated equipment that may accept commands from a computerized control unit.

In [9], the study's goals were to (i) style and create new field-based connectivity, as well as test and verify the homogeneity of HNT stress application on a scaled-up version of the prototype; (ii) enhance and expand a more advanced cyber-physical system to detects and enforce post-anthesis HNT stress homogeneously through physiological maturity inside the scaled-up tents. Moreover, the focus was on the physical part more than the cyber part as their limitation.

In [10], the goal was to use computer modelling to examine the interaction between a harvester rod and a coffee branch. Additionally, plagiotropic coffee branch samples were evaluated. In addition, simulations generated 3.14 % greater accelerations than tests, simulated separations were 23.2 points lower than observations, and calculated Von Mises stress was close to values obtained. They considered taking the interaction between equipment and plants.

In [11], utilized (WSN) in greenhouses for monitoring roses plantation with a prediction model that is based on Data Mining with the purpose to provide the environmental factors recorded by wireless sensors that use Zigbee technology to transfer the data represented by environmental temperature, relative humidity, soil moisture and light intensity. Finally, there is a central node that transfers information to a central Big Data database and a tracking and prediction client programme using Arduino and Raspberry Pi cards. In order to improve strategies to obtain an understanding of these natural conditions and to predict behavioural trends within these climate factors, information from the WSN is analyzed. However, their system lacks to be integration with the climate control API unit to reach excellent results with less effort.

In [12], farmers can use the suggested system to track the evolution of particular metrics of interest and make necessary decisions to boost agricultural yield. Creating such a strategy is not simple, and it involves a thorough understanding of the crop, management tactics, data processing, and other factors in order to be practical. Finally, the notion given in this work will serve as a starting point for future precision agricultural researchers.

In [13], researchers proposed weed detection using Artificial Neural Network (ANN) to detect a pattern for the plant shape feature, and they investigated four species of common weeds in sugar beet fields. The processing was done using the MATLAB image processing toolbox and the segmentation to distinguish and classify each plant shape and using a Support Vector Machine (SVM) as a pattern recognition method to detect the weed changes accurately. In general, their system shows success in distinguishing the plants depending on image processing but as an input, we need to focus on the aspect of plantation in earlier stages to avoid any other unwanted plants in our field instead of all these expenses and efforts that based on pictures and could be manipulated easily based on the parameters of the image.

Finally, [14] discussed that CPS is possible to function more productively, jointly, and enthusiastically if AI is used. Crop management in agribusiness is a method of ensuring that the crop is protected as much as possible against various pests. Infestations are any living species that can affect an agricultural crop, such as insects, mammals, grasses, infections, and other creatures. Every year, pests and other illnesses cause a 20-30 % global potential loss of agricultural yield. To decrease these losses, crop protection is essential. Moreover, crop selection, watering automation, farming techniques, plant tracking, and picking are all part of the CPS precision agricultural process.

We conclude that smart farming has grown in popularity as a result of recent technical developments in CPS. Farmers are beginning to adopt sensor technologies in order to increase agricultural productivity and efficiency. The preceding solutions enable agriculturists to areasonable amount of the ever-changing climate factor of their crops in real-time as well as their past data and how to get the most out of such data in order to make appropriate decisions

III. AA-CPS FRAMEWORK

This section will present the proposed method that extracted and fulfilled the gaps from the previous studies. As we noticed the limitations in the literature, we propose the Autonomous Agricultural Cyber-Physical System (AA-CPS) shown in Figure 2 that consists of the following three steps. The AA-CPS framework is shown in Figure 3.

1. Sensors & Communication (Physical)

In this part, we need to collect the data from the farming field as weather data (climate reads station) and the soil data from multi purposes sensors network. An energy-efficient wireless sensor network (WSN) was employed to construct the following two components [15].

In our experiment, we used a data set that was collected within different weather and soil reads and attributes with considering the seasons within the year. The dataset sources from the Kaggle website: *https://www.kaggle.com/code/chitrakumari25/smart-farming-optimizing-engine/data*.

In addition, it consists of (2200 Row, 8 Columns), the columns represented as follow:

- Soil: N = Nitrogen, P = phosphorous. K = Potassium, Ph = A scale used to identify acidity or basicity nature; (Acid Nature- Ph<7; Neutral- Ph=7; Base Nature-P>7).
- Weather: Temperature=The average weather temperatures from 50 to 75F, Humidity= weather Reads.
- Crop Label (Class) = Types of Crops ("Rice, Maize, Chickpea; Kidney beans; pigeonpeas; moth beans; mung bean; black gram; lentil; pomegranate; banana; mango; grapes; watermelon; muskmelon; apple; orange; papaya; coconut; cotton; jute; coffee").

Table 1 tabulated a sample of the data to show the attributes of the data:

TABLE 1: DATASET SAMPLE

N	P	K	Temp	Humidity	ph.	Rainfall	Crop
90	42	43	20.879	82.002	6.5	202.93	Rice
82	38	41	24.667	80.045	7.8	253.703	Banana
67	46	39	26.301	68.411	7.0	199.310	Cotton
83	49	42	28.223	77.790	7.3	185.471	Lentil

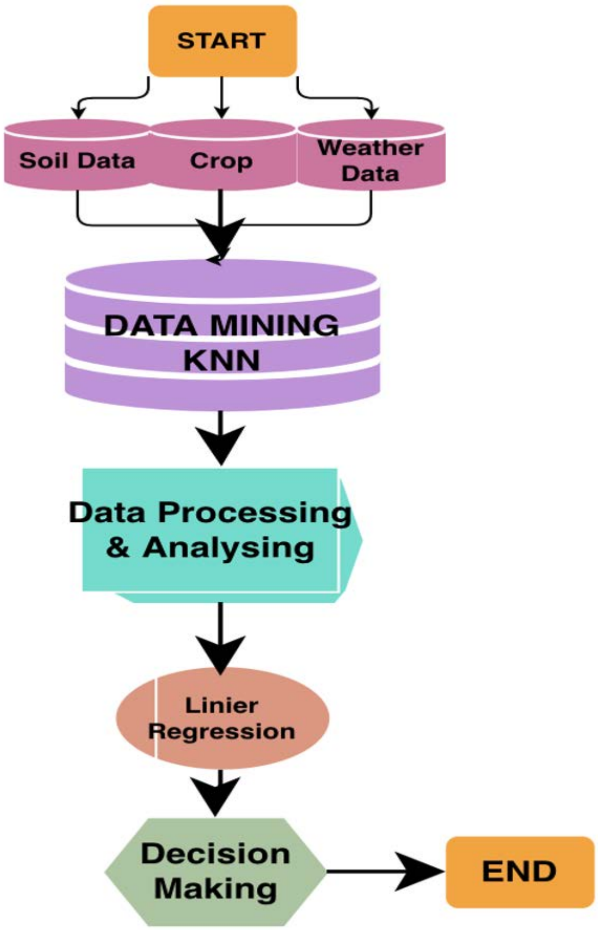

Figure 2: Autonomous Agricultural Cyber-Physical System Flowchart.

2. Data Processing & Decision Making (Cyber)

After collecting the data from the physical components (sensors), the data will transfer to the local server to be analyzed and filtered using K-Means Algorithm to classify the data and cluster to Training and Testing parts. In addition, it employs data mining techniques and a Machine learning algorithm to update the recommendation status for the kind of crop that fit the data that has been collected from the wanted environment.

Moreover, utilizing Machine Learning models such as "Linear Regression", helps to choose what crop fit to be grown in the same field that we collect the data from. Linear Regression is a method of constructing a model that best fits the pieces of data just on the chart such that we could use this to anticipate correct output for sources that are not in the given data we have, assuming that all those outcomes would fall mostly on-trend.

3. Autonomous Plantation

From the previous steps, after soil and climate data were collected, analyzed and processed, the type of crop suitable for the chosen environment was deduced. You can add this decision that was taken in the field of tractors, automated farming for the purpose of planting seeds or monitoring the health of the soil in case its needs fertilizers, and so on.

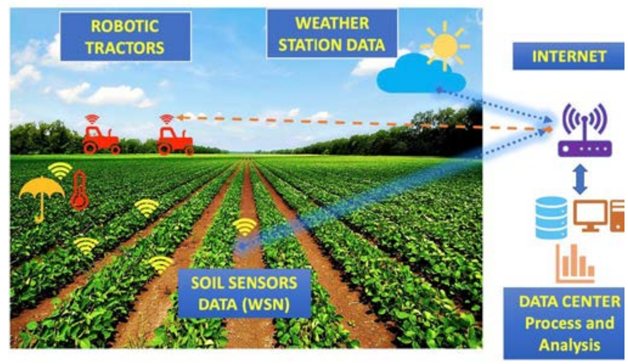

Figure 3: Autonomous Agricultural Cyber-Physical System Framework.

IV. RESULTS AND DISCUSSION

The proposed system is considered an ideal solution to discover the land and environment before starting farming it. Additionally, the decision-making of the cyber system will provide the benefit to acquire a specific crop and reduce the risk of spending the resources. Using the Internet of Things (IoT) techniques centric on CPSs in the agricultural field consisting of sensory devices and WSN that will collect the required data gathered in one repository to be analyzed, and calculate the crucial parameters with data mining techniques in order to deliver the results in an easy, fast, and smart way and also, accurate prediction results. Using the specific system, the end user (farmer) may simply interact with the system remotely.
As a result, it may be used on a wide range of devices, including any desktop browser. A user-friendly GUI with basic administration functions is required in the given solution. It gives you a simple way to keep track of the climate, soil data and status, as well as the option to program new crop recommendations.

The results showed that our model gives an accurate result that exceeds our expectations, which results in a 96 % accuracy ratio in the classification report, with precision and weighted output as we evaluate the results using the confusion matrix.

We implement the experiment in the system of 2.4 Corei5 CPU, 4GBs of RAM, 512 GBs HDD, based on Windows 10, Anaconda Tool/ Jupyter Lap 3.0.1, and Python 3.9 programming language.

Nevertheless, to test our model, we used the input of 50 for Nitrogen, 45 for Phosphorous, 10 for potassium, 20 for temperature, 80 for humidity, 7 for ph, and 200 for Rainfall, resulting in the recommendation for (Orange) crop as close to the real data to compare with.

V. CONCLUSION

Agricultural Cyber Physical System (ACPS) is considered a vital role in modern farming techniques that may impact human needs indirectly. Moreover, crop recommendations required deep research and experiments for the surrounding environment, in both influencing aspects of weather and soil. In this research, we employ a sensors network as a physical part, and data gathering, Data Mining techniques on other hand as cyber system to find out the parameters that are observed through such a system in order to make a decision for the farmers to know what crop my fit their land. The K-Means clustering and Logistic Regression Algorithms run the data which consists of 2200 Rows, and 8 Columns in order to obtain the results represented by 96% in terms of accuracy and precision to the Crop type recommendations. For future works, we recommend using Deep Learning Algorithms to enhance the accuracy of the prediction, and there is a variety of parameters that may be considered in such a study; for instance, tropical, rainy, cold zones and finally, many other attributes belong to the soil also may be considered for the next step of research work. Autonomous systems will play a great role in future agriculture, so future work will also examine autonomous systems that automate or augment operations such as pruning, thinning, and harvesting, as well as mowing, spraying, and weed removal.

REFERENCES

[1] Schimmelpfennig, D. Farm Profits and Adoption of Precision Agriculture. USDA 2016, 217, 1–46.

[2] Saiz-Rubio, V., & Rovira-Más, F. From smart farming towards agriculture 5.0: A review on crop data management. Agronomy, 2020, 10(2), 207.

[3] Selmani, A., Oubehar, H., Outanoute, M., Ed-Dahhak, A., Guerbaoui, M., Lachhab, A., & Bouchikhi, B. Agricultural cyber-physical system enabled for remote management of solar-powered precision irrigation. Biosystems Engineering, 2019, 177, 18-30.

[4] Malek, M., Dhiraj, B., Upadhyaya, D., & Patel, D. A Review of Precision Agriculture Methodologies, Challenges, and Applications. Emerging Technologies for Computing, Communication and Smart Cities, 2022, 329-346.

[5] Udutalapally, V., Mohanty, S., Pallagani, V., & Khandelwal, V. sCrop: A Novel Device for Sustainable Automatic Disease Prediction, Crop Selection, and Irrigation in Internet-of-Agro-Things for Smart Agriculture. IEEE Sensors Journal, 2021, 21(16), 17525-17538.

[6] Souza, V., Dias, G., Santos, A., Costa, A., Santos, F., & Magalhães, R. Evaluation of the interaction between a harvester rod and a coffee branch based on finite element analysis. Computers and Electronics in Agriculture, 2018, 150, 476-483

[7] Slysheva, D. Precision Farming Technologies and Their Role in on-Farm Land Management Projects Based on Adaptive Landscape Farming Systems. Stolypin Annals, 2022, 4(1).

[8] Sadiku, M., Wang, Y., Cui, S., & Musa, S. Cyber-Physical Systems: A Literature Review. European Scientific Journal, ESJ, 2017, 13(36), 52.

[9] Hein, N., Bheemanahalli, R., Wagner, D., Vennapusa, A., Bustamante, C., & Ostmeyer, T. et al. Improved cyber-physical system captured post-flowering high night temperature impact on yield and quality of field grown wheat. Scientific Reports, 2020, 10(1).

[10] Rad, C., Hancu, O., Takacs, I., & Olteanu, G. Smart Monitoring of Potato Crop: A Cyber-Physical System Architecture Model in the Field of Precision Agriculture. Agriculture And Agricultural Science Procedia, 2015, 6, 73-79

[11] Rodríguez, S., Gualotuña, T., & Grilo, C. A System for the Monitoring and Predicting of Data in Precision Agriculture in a Rose Greenhouse Based on Wireless Sensor Networks. Procedia Computer Science, 2017, 121, 306-313.

[12] Bakhshipour, A., & Jafari, A. Evaluation of support vector machine and artificial neural networks in weed detection using shape features. Computers And Electronics In Agriculture, 2018, 145, 153-160.

[13] Choubisa, M., & Doshi, R. Crop Protection Using Cyber-Physical Systems and Machine Learning for Smart Agriculture. Real-Time Applications Of Machine Learning In Cyber-Physical Systems, 2022, 134-147.

[14] Domínguez-Niño, J., Oliver-Manera, J., Girona, J., & Casadesús, J. Differential irrigation scheduling by an automated algorithm of water balance tuned by capacitance-type soil moisture sensors. Agricultural Water Management, 2020, 228, 105880.