Agri-PAD: a scalable framework for smart agriculture

Tehreem Qamar, Narmeen Zakaria Bawany

Center for Computing Research, Department of Computer Science and Software Engineering, Faculty of Science,
Jinnah University for Women, Karachi, Pakistan

Article Info

ABSTRACT

More recently, big data tools and technologies have been applied in the agriculture sector leading to major benefits. Many frameworks have been proposed that employ big data technologies in the field of agriculture, however, such existing frameworks are focused on a particular aspect of agriculture and do not consider multiple stakeholders and applications. The objective of this research is to develop a holistic framework named Agri-PAD that encompasses almost all aspects of agriculture including crop selection, crop monitoring, soil monitoring, weather conditions, precision farming, and market demand. The Agri-PAD framework includes three major categories of machine learning based agriculture applications that is precision, recommendation, and enterprise applications. The Agri-PAD framework is capable of providing remote sensing of fields, precision farming, effective supply chain, and support informed decision making leading to enhanced productivity. To validate the efficacy of the proposed framework, the two most prominent agricultural applications, crop production forecasting and crop harvesting recommendation have been investigated and accuracy of 99% has been achieved. We believe that the Agri-PAD framework enables all stakeholders in the agriculture cycle to connect and apply big data analytics at every step leading to a more efficient and smarter agriculture ecosystem.

Corresponding Author:

Tehreem Qamar
Center for Computing Research, Department of Computer Science and Software Engineering
Faculty of Science, Jinnah University for Women
Adress: 5C, Nazimabad, Karachi, Pakistan
Email: tehreem.qamar@juw.edu.pk

1. INTRODUCTION

Agriculture has been considered as the primary means of reducing poverty and improving food security of the world's 80% of impoverished people who live in rural areas as per the World Bank reports. Moreover, agriculture contributes to around 25% of the gross domestic product (GDP) in some developing countries [1]. Globally, the agricultural industry has been adversely affected by increased droughts, floods, erratic precipitation patterns, and heat waves brought on by climate change. Additionally, the high population growth rate and the existing climate-related impacts on agricultural lands, agricultural areas are under more pressure to provide enough food [2]. To satisfy the increasing food demand, available agriculture land must be utilized effectively such that sustainable and healthy crops are produced. However, traditional intensive agricultural practices are causing land degradation and lead to relatively lower yield [3]. Furthermore, the farmers are unaware about climate changes and market fluctuations in real-time ending up with inappropriate supply of required crops. Cost of per yield production is increasing due to inefficient utilization of resources [4]. For example, the fertilizers and pesticides are used extensively without the precise requirement.

The shortcomings are addressed by the application of information and communication technologies (ICT) to agriculture, widely known as smart agriculture [5]. Smart agriculture comprises extensive ICT infrastructure that generates enormous amounts of data that can be utilized to improve the traditional agriculture practices [6]. Smart agriculture employing big data, ICT and internet of things (IoT) has transformed the traditional experience-based agriculture and has gained much attention from researchers. Numerous studies are presented in the literature that provide specific application of smart agriculture including crop forecasting [7], [8], pest identification [9], [10], precision farming [11], [12], irrigation monitoring [13], [14], smart greenhouses [15], [16], and supply-chain [17], [18]. But relatively few studies cover all stakeholders of agriculture.

This research is an attempt to present an effective, complete and scalable framework for smart agriculture; Agri-PAD. The layered framework focuses on three major components of big data agriculture applications. The first layer is responsible for sensing real-time data from sensors and is called the perception layer. The second layer, called the data procession layer, stores and processes the data. The third, and the final layer comprises various applications which are available to all stakeholders to maximize the agri-productivity. The major contributions of this reseach are summarized below:

− We propose a comprehensive big data analytics framework; Agri-PAD that facilitates agriculture activities by offering three broad categories of applications that is precision, recommendation and enterprise.
− We demonstrate the implementation of Agri-PAD framework for crop prediction and crop recommendation.
− We discuss major challenges that hinder adoption of smart agriculture widely.

The rest of the paper is organized as follows: section 2 discusses big data role in agriculture, section 3 describes the related work, section 4 explains the Agri-PAD framework, section 5 presents potential use cases of Agri-PAD framework. Section 6 discusses open challenges in the adoption of big data in agriculture while section 7 concludes the paper.

2.   BIG DATA IN AGRICULTURE

Big data is termed as a technological paradigm for data. It refers to massive, heterogeneous, structured, semi-structured and unstructured datasets that includes text, images, videos, and audios [19]. Big data is considered as the data generated at high velocity, high volume, and with high variety that require advanced technologies and algorithms for processing [20]. Big data definitions have evolved rapidly; the following five dimensions have emerged as common criteria to illustrate big data:

− Volume: the enormous amounts of digital data generated and collected every second from a billion of devices and applications.
− Velocity: the speed of data creation and generation.
− Variety: the diversity of the data types, sources and their format (e.g., videos, documents, comments, and logs).
− Veracity: The truthfulness, reliability and accuracy of the data and their sources.
− Value: the usefulness of the collected data i.e., the insights that can be extracted from the big data.

Big data is more inclined towards the capability to search, aggregate, visualize and cross-reference large datasets in reasonable time to extract information and insights which was previously not feasible, both economically and technically [21]. Big data applications are utilized in almost every field; engineering, mathematics, computer science, business, management and accounting, biochemistry, genetics, physics and astronomy to name a few [19].

Big data in agriculture has immensely evolved the traditional agricultural practices [5]. Major sources of big data in agriculture include:

− On-field sensors which provides real time insights about farm (biosensors and weather stations).
− Airborne sensors, data captured from drones and satellites.
− Data collected by governmental agencies and third-party organizations such as yearly statistical reports, rules and regulations.
− Data from web available via online repositories, web services, and social media feeds.

It is evident that the data produced by above-mentioned sources are heterogeneous in nature with varied level of volume, velocity and variety. However, efficient storage, processing and analysis of such diverse data for extracting valuable information such as precision farming is guaranteed by the big data proponents [22]. Big data technologies that can be used for such applications include Hadoop distributed file system (HDFS), HBase, Cassandra, Spark, Hive, tensorflow and many others [23].

3.   RELATED WORK

Climate change and growing population are turned out as serious threats to food security and major challenges for agriculture [24]. The total agricultural land utilized for food production has experienced a decline and with time the gap between demand and supply has become more significant [25]. To meet this rising demand, farmers need to produce approximately 70% more food by 2050, according to the UN food and agriculture organization (FAO) [26]. Consequently, the agriculture industry adopted technological trends and transformed into smart agriculture. Smart agriculture comprises of IoT, cloud computing, big data, and artificial intelligence improves crop yield, reduces waste, and optimizes supply chains [27].

Various studies have been conducted for improving productivity in agriculture using smart agriculture solutions. Sarker *et al.* [5] proposed a conceptual framework incorporating big data technologies to facilitate farmers at field level. Their proposed solution is a seven-stage model that encompasses extracting data from the sensors and explains how that big data might assist in accomplishing sustainable agriculture. Liu *et al.* [28] proposed an experimental framework to monitor different agricultural aspects such as water, fertilization, temperature using IoT and cloud computing. Their framework is capable of acquiring, persisting and processing sensor data which would help in satisfying farm needs as and when required. A smart agriculture monitoring system for the detection of soil temperature and moisture was presented in Mekala and Viswanathan [29]. Analyzing the data with a stated accuracy of 94%, the proposed cloud-based solution may enable efficient monitoring of specific crop comfort levels and might be an accurate and useful decision tool for farmers. Cicioğlu and Çalhan [30] introduced an IoT enabled corn production monitoring system. The proposed solution monitored the cornfield for soil and air attributes via different sensors and provides crucial information through graphical interfaces about growth of corn, requirement of water and the actions to be taken for disease risks. Vincent *et al.* [31] targets the land suitability assessment and proposed a recommendation model that classifies the land suitability by employing machine learning models on the big data produced by sensors. The system assists farmers in classifying the land for cultivation in more suitable, suitable, moderately suitable, and unsuitable classes. To increase crop production and to control the agricultural cost, Rajeswari *et al.* [32] presented a smart agricultural model that predicts the crop yield and choose the best crop sequence based on previous crop sequences on the same farmland and soil nutrient data.

Apart from crop monitoring, researchers have employed advanced technologies to especially improve the irrigation processes. Tseng *et al.* [33] proposed a big data analysis technique to assist farmers in crop selection. They used a three-dimensional correlation analysis to analyze the irrigation cycle and determine the farmer's irrigation techniques. The soil moisture content was then computed to identify irrigation and determine whether the farmer had used pesticides or fertilizers. Nawandar and Satpute [34] proposed an intelligent system for smart irrigation intending to preserve water resources. Bu and Wang [35] presented a deep reinforcement learning based smart agriculture system which aims to reduce extraneous water consumption. Further, Kamienski *et al.* [36] proposed smart water management architecture for an efficient water irrigation system. These irrigation systems track the water requirements of crops based on collected data and actuate water flow in line with expected demands without any human involvement.

The application of smart agriculture is not limited to automating the traditional practices and reducing human involvement, but also includes fully automated farms that can be operated without any human intervention [37], [38]. These systems exploit agriculture big data in a robust and effective way. However, adoption of these unmanned farms is still in its infancy and a number of issues such as governmental and data security legislations are yet to be addressed.

Despite the challenges, big data remains the driving force behind smart agriculture. The data collected via sensors and other IoT devices provides detailed insights to the farmer leading to better decision making and effective utilization of resources. The aforementioned solutions prove that big data technologies have the potential to revolutionize the agriculture industry. However, each of these solutions has limited scope and is focused on a single aspect of agricultural activity. To fully utilize the potential of smart agriculture, there is a need for a platform that integrates all agricultural activities and provides a holistic view to all stakeholders. Keeping in view the limitations of existing solutions we propose a framework-Agri-PAD that includes the complete lifecycle of agriculture activities and enables all stakeholders in making informed decisions. The following sections describe the details of Agri-PAD framework.

4.   Agri-PAD THE PROPOSED FRAMEWORK

This section presents Agri-PAD, a smart agriculture framework that encompasses major agriculture activities. The integration of various services in a single framework ensures that all stakeholders are facilitated and that a broad view of the system is presented, which eventually helps in better decision making. The Agri-PAD framework offers a systematic classification of services and applications leveraging big data analytics. The Agri-PAD framework consists of three layers i) perception layer, ii) data processing layer and iii) application layer as depicted in Figure 1. The Agri-PAD framework is scalable and can handle batch and

real-time data streams. It incorporates a wide variety of techniques and technologies to aggregate, manipulate, analyze, and visualize big data. The details about each layer are described in the successive subsections.
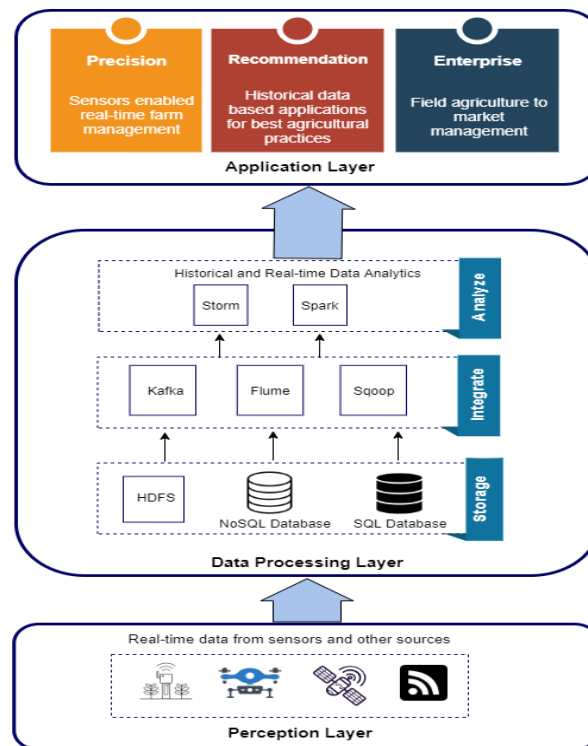

Figure 1. Agri-PAD-the scalable framework for smart agriculture

4.1. Perception layer
The perception layer is the first layer of Agri-PAD framework. This layer is horizontally scalable and includes all data sources as mentioned in section 2. Big data sources in agriculture can be broadly categorized into IoT devices (ground sensors, weather stations, and airborne sensors), social media posts, and web-based data i.e., the data stored in databases. The data will be acquired from each source and make available to the next layer for further processing.

4.2. Data processing layer
The data processing layer is the fundamental layer that acts as a data processing engine. It handles everything from data storage to knowledge extraction. It first transforms the data into an analysis-ready format. Then, apply big data tools and technologies to extract relevant insights. Data will be processed in Apache Hadoop [19], [39] ecosystem as it offers scalable platform for storing, managing and processing big data. Table 1 summarizes the applications provided by Hadoop. Moreover, it incorporates any programming language that offers processing of artificial intelligence algorithms.

4.3. Application layer
Application layer is the topmost layer of Agri-PAD framework. This layer is the key layer that interacts with the users and provides services to the farmers and other stakeholders. The layer consists of multiple applications that are classified as precision, recommendation and enterprise. Figure 2 provides the potential applications against each category.

4.3.1. Precision
Precision applications process the real-time data of various sensors and offers monitoring applications. These applications are providing real-time statistics which help in revealing accurately and precisely the needs of a farm hence, facilitating farmers in monitoring their farms remotely. Soil, weather, and crop conditions can be observed without being present on the farm. The requirement of pesticides and fertilizers can be precisely determined. Farmers will also receive notifications and alerts about the particular need of their farm. With these

applications, it will be possible for the farmers to perform the precise intervention, at the exact location, at the right time, responding to the specific demands of individual crops and individual areas of land.

4.3.2. Recommendation

Recommendation applications are predicting the best agriculture practice from sowing to market the harvested crop. This category of applications is based on the historical data and helps in effective decision making to all stakeholders of agriculture. Identify crops to study crop rotation, detect expansion and intensification of row crop agriculture, classify crop types and crop related land use practices, find areas that are most likely to be affected by damaged crops, estimate food availability are few of these applications.

4.3.3. Enterprise

Enterprise applications will assist in the administration of post-harvest processing to food processing, and its marketing. Field agriculture to human resources management, inventory, logistics, machinery, profit in buying and selling can be improved by using such applications. Understanding how management choices affect sustainability and operational efficiency, improving smallholder farmers' productivity and income, providing fair and attractive insurance and finance for farmers, facilitating farmers' financing and easier payments, and connecting smallholders directly with markets are just a few examples of these applications.

Table 1. Apache Hadoop data processing layers

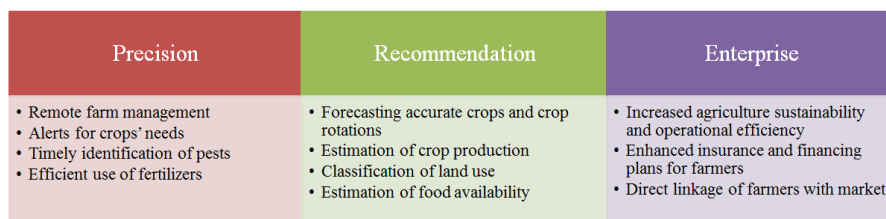| Application	Description
Storage	
HDFS	The native Hadoop data management system. HDFS is scalable, distributed, fault-tolerant and a high-performance reliable data storage. HDFS is meant to span large clusters of commodity servers and manage large volumes data files.
NoSQL databases	To handle huge volumes of semi-structured and unstructured data properly at which traditional relational databases are not designed for these types of data.
Kudu	Apache Kudu is an open-source distributed data storage engine that enables fast analytics on real-time data.
Integrate	
Apache Flume	Flume is designed to collect, aggregate and transfer data from external machines to HDFS. It streams data form high volume sources and provides real-time analysis.
Apache Kafka	Kafka is a publish–subscribe message streaming platform and is distributed in nature. It is used for high-performance real-time data pipelines, data integration and streaming analytics.
Apache Sqoop	Sqoop is a tool dedicated for transferring bulk data between relational databases (e.g. MySQL, SQL and Oracle) into HDFS and vice versa.
Analyze	
Apache Storm	Storm provides distributed real-time processing of high velocity with large variety of data. It can also perform micro-batch processing.
Apache Spark	Spark is a batch in-memory computing framework and an efficient alternative to Hadoop MapReduce programming framework. It offers unified analytics engine comprising SQL, machine learning and graph processing.
Impala	The data warehouse for Hadoop. It structures data at rest on a columnar data format which allows handling interactive and real-time analysis on big data.

Precision	Recommendation	Enterprise
• Remote farm management • Alerts for crops' needs • Timely identification of pests • Efficient use of fertilizers	• Forecasting accurate crops and crop rotations • Estimation of crop production • Classification of land use • Estimation of food availability	• Increased agriculture sustainability and operational efficiency • Enhanced insurance and financing plans for farmers • Direct linkage of farmers with market

Figure 2. Big data enabled applications in Agri-PAD framework

5.   DISCUSSION AND POTENTIAL USE CASES

Agri-PAD framework aims at providing a real-time view of major agriculture activities along with valuable insights and recommendations. The shared platform will provide big data applications to the farmers, crop experts, market owners and other appropriate decision makers. The farmers can monitor their farms and can communicate data with crop experts and other consultants in real time. These consultants can evaluate the data and give expert advice. Market owners can effectively forecast the needs and utilize the available resources efficiently while decision makers can make more productive long-term decisions. The following sub-sections provide use-cases employing big data analytics on the popular agriculture activities.

5.1. Crops production forecasting

Crops production forecasting application is an enterprise application type of the Agri-PAD framework. Accurate and timely forecasting of crop yields is significant for food security and planning of agricultural markets. Historical big data can be used in predicting the actual need for the years to come. To evaluate the efficacy of Agri-PAD framework, we have used the production data of crops cultivated in Pakistan. The statistical production data of all four provinces of Pakistan have been collected by agriculture marketing information service (AMIS), directorate of agriculture (economics and marketing) Lahore, Punjab. We have used 7 major crops data which includes wheat, cotton, rice, sugarcane, maize, ground nut and barley. The dataset includes the production of the crops and area available for crop cultivation in the respective province.

The selected dataset was divided in training and testing sets with 80:20 ratio. The training set was used to establish model to predict production of crops while the test set was used to test the quality of the model. Random forest algorithm was applied having 1097.2 RMSE. Figure 3 illustrates the accuracy of the model which clearly indicates that the model outperformed with nearly equal actual and predicted values for each crop. Crop production forecasting is essential for policy makers to make timely decisions and Agri-PAD framework has the tendency to incorporate historical data from different resources and generate effective crop production estimates that can be used as agricultural production warnings in food-insecure regions. Thus, empowering decision makers to plan expeditiously.
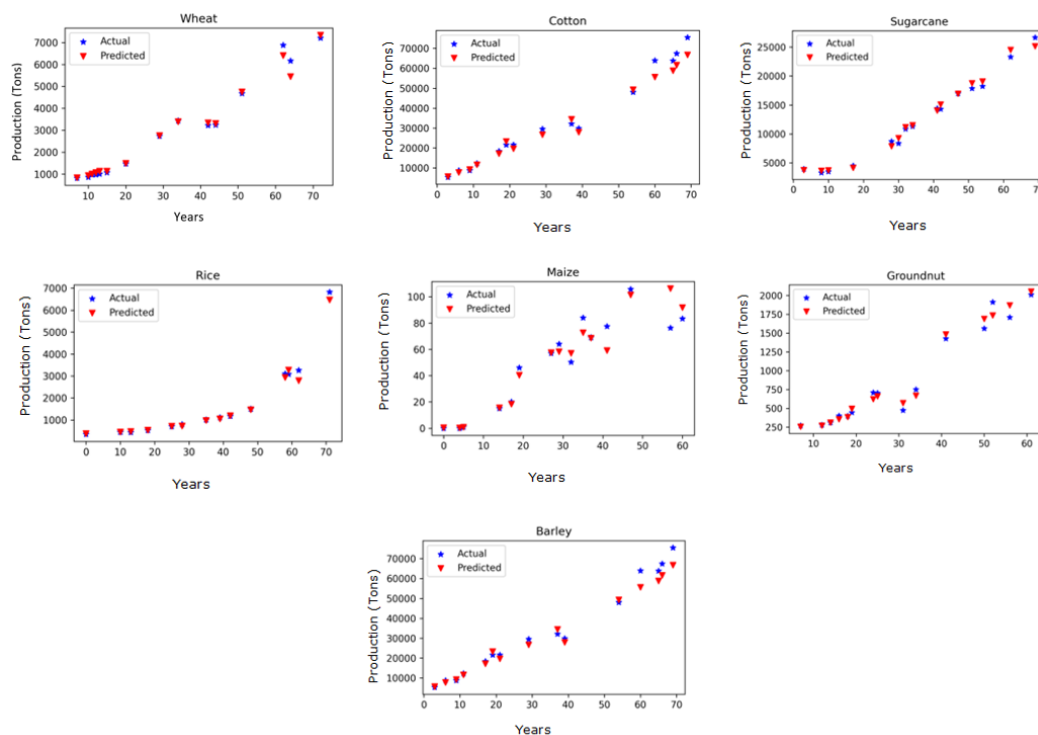

Figure 3. Actual production vs. predicted production of crops

5.2. Crop recommendation

Crop harvesting recommendation prior to its cultivation is an effective tool for farmers to increase productivity with no capital loss. This is a type of recommendation application in Agri-PAD framework. This application will help the farmers to make an informed decision by recommending the best suitable crop according to the climate condition before cultivation. To prove its applicability, we have used crop recommendation dataset available at Kaggle. The dataset comprised of 7 independent and 1 dependent features with no missing values. The dataset was balanced for 22 crops with features including temperature, humidity, pH value, rainfall, nitrogen, phosphorous, and potassium.

Naïve Bayes algorithm was used for training the model which attained the accuracy of 99%. The model was able to recommend the crop against the climatic parameters. This model is part of the data processing layer of Agri-PAD framework that process real-time data via its perception layer, thus enabling farmers to make appropriate decision instantly.

6.   OPEN CHALLENGES IN ADOPTION OF BIG DATA TECHNOLOGIES IN AGRICULTURE

Big data analytics undoubtedly provides tremendous insights for making better decisions in the agricultural domain but has not undergone its wide adoption in production agriculture [40]. The open challenges hindering its way can be broadly categorized into technical and operational challenges. Technical challenges deal with the installation of technical devices, their management, and the technical expertise required in disseminating the smooth transfer of data between participating entities. Researchers [41], [42] asserted that the availability of limited infrastructure and lack of technical human resources to collect, process and analyze the huge amount of diverse data is the potential barrier in smart agriculture adoption. The collected data must be consistent, clean, complete and in compliance with the protocols that allow to be pooled into a centralized server for processing. Hence, the consistency, accuracy, and veracity of big data and its analytics are critical challenges [43]. Nonetheless, security and privacy remain an inherent concern for all stakeholders [44]. Another barrier that inhibits smart agriculture from wide adoption is the insufficient internet connectivity. Transferring data from IoT devices to server needs high speed internet connectivity, but the availability of internet is still critical in rural areas of most developing countries [45].

Operational challenges involve the investment and management of the field and post-agricultural activities. Transforming the traditional agriculture system into smart agriculture requires huge amount of investment both in terms of financial cost and time. Further, smart agriculture demands proper maintenance of the infrastructure and includes certainly high costs for the smooth running of its operations and up-gradation. Additionally, proper training of farmers is also required to get the maximum benefit. Proper governance and structure of agricultural big data are still missing and require policy frameworks.

Apart from technical and operational challenges, acceptance of smart agriculture is also hindered by the lack of trust among its stakeholders. The primary stakeholder, farmer, mostly belongs to the older age group are less likely to utilize smart agriculture infrastructure as they are more habitual of traditional agriculture activities. Furthermore, the direct link of farmer to the consumer will end the middleman inclusion and dominance which would create chaos between them that may result in strong opposition in the adoption of smart agriculture.

7.   CONCLUSION

Agriculture and climate change are inextricably linked. The rapid pace of climate change will have far-reaching consequences on the agricultural ecosystem. Therefore, to ensure food security, smart agriculture is the prospective solution. Smart agriculture utilizes big data analytical techniques which have been proved to be outstandingly beneficial in providing valuable insights leading to better decision-making. In this research, we discussed the applications of big data analytics in agriculture industry and proposed a framework; Agri-PAD. The framework consists of three layers i.e., perception layer, data processing layer, and application layer. Agri-PAD incorporates big data analytics by providing three distinguished categories of applications namely precision, recommendation, and enterprise applications. Real-time data analytics applications that facilitate farmers in managing their farms and crops remotely fall under the precision category. Recommendation applications, based on historical data collected from various sources, provide insights that lead to informed decision-making. While enterprise applications assist farmers in reaching the market directly, eliminating the middleman dependence and improving the efficiency of the agriculture supply chain. Further, we present two use cases of smart agriculture employing machine learning. We also highlight the open challenges that are serving as barriers to the adoption of smart agriculture widely. Realizing the importance of security and privacy, we intend to enhance this framework by including a layer to handle data security in the future. In addition, the work on an integrated testbed to further validate the effectiveness of the Agri-PAD framework can be undertaken.

REFERENCES

[1]   "Agriculture overview." https://www.worldbank.org/en/topic/agriculture/overview#1 (accessed Aug. 15, 2021).
[2]   N. K. Arora, "Impact of climate change on agriculture production and its sustainable solutions," *Environmental Sustainability*, vol. 2, no. 2, pp. 95–96, Jun. 2019, doi: 10.1007/s42398-019-00078-w.
[3]   P. M. Kopittke, N. W. Menzies, P. Wang, B. A. McKenna, and E. Lombi, "Soil and the intensification of agriculture for global food security," *Environment International*, vol. 132, p. 105078, Nov. 2019, doi: 10.1016/j.envint.2019.105078.
[4]   A. Alengebawy, S. T. Abdelkhalek, S. R. Qureshi, and M.-Q. Wang, "Heavy metals and pesticides toxicity in agricultural soil and plants: ecological risks and human health implications," *Toxics*, vol. 9, no. 3, p. 42, Feb. 2021, doi: 10.3390/toxics9030042.
[5]   M. N. I. Sarker, M. Wu, B. Chanthamith, S. Yusufzada, D. Li, and J. Zhang, "Big data driven smart agriculture: pathway for sustainable development," in *2019 2nd International Conference on Artificial Intelligence and Big Data (ICAIBD)*, May 2019, pp. 60–65, doi: 10.1109/ICAIBD.2019.8836982.
[6]   A. Cravero and S. Sepúlveda, "Use and adaptations of machine learning in big data—applications in real cases in agriculture," *Electronics*, vol. 10, no. 5, p. 552, Feb. 2021, doi: 10.3390/electronics10050552.

[7] A. Ceglar, A. Toreti, C. Prodhomme, M. Zampieri, M. Turco, and F. J. Doblas-Reyes, "Land-surface initialisation improves seasonal climate prediction skill for maize yield forecast," *Scientific Reports*, vol. 8, no. 1, p. 1322, Dec. 2018, doi: 10.1038/s41598-018-19586-6.

[8] T. Iizumi, Y. Shin, W. Kim, M. Kim, and J. Choi, "Global crop yield forecasting using seasonal climate information from a multi-model ensemble," *Climate Services*, vol. 11, pp. 13–23, Aug. 2018, doi: 10.1016/j.cliser.2018.06.003.

[9] B. Jabir and N. Falih, "Deep learning-based decision support system for weeds detection in wheat fields," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 12, no. 1, pp. 816–825, Feb. 2022, doi: 10.11591/ijece.v12i1.pp816-825.

[10] P. Vasavi, A. Punitha, and T. V. N. Rao, "Crop leaf disease detection and classification using machine learning and deep learning algorithms by visual symptoms: a review," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 12, no. 2, pp. 2079–2086, Apr. 2022, doi: 10.11591/ijece.v12i2.pp2079-2086.

[11] E. F. I. Raj, M. Appadurai, and K. Athiappan, "Precision farming in modern agriculture," in *Smart Agriculture Automation Using Advanced Technologies. Transactions on Computer Systems and Networks*, Singapore: Springer, 2021, pp. 61–87.

[12] P. Lottes, J. Behley, A. Milioto, and C. Stachniss, "Fully convolutional networks with sequential information for robust crop and weed detection in precision farming," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 2870–2877, Oct. 2018, doi: 10.1109/LRA.2018.2846289.

[13] A. Benbatouche and B. Kadri, "Design and realization of low-cost solenoid valve remotely controlled, application in irrigation network," *Bulletin of Electrical Engineering and Informatics*, vol. 11, no. 3, pp. 1779–1788, Jun. 2022, doi: 10.11591/eei.v11i3.4123.

[14] L. García, L. Parra, J. M. Jimenez, J. Lloret, and P. Lorenz, "IoT-based smart irrigation systems: an overview on the recent trends on sensors and IoT systems for irrigation in precision agriculture," *Sensors*, vol. 20, no. 4, p. 1042, Feb. 2020, doi: 10.3390/s20041042.

[15] R. Rayhana, G. Xiao, and Z. Liu, "Internet of things empowered smart greenhouse farming," *IEEE Journal of Radio Frequency Identification*, vol. 4, no. 3, pp. 195–211, Sep. 2020, doi: 10.1109/JRFID.2020.2984391.

[16] I. Ardiansah, N. Bafdal, E. Suryadi, and A. Bono, "Greenhouse monitoring and automation using arduino: a review on precision farming and internet of things (IoT)," *International Journal on Advanced Science, Engineering and Information Technology*, vol. 10, no. 2, p. 703, Apr. 2020, doi: 10.18517/ijaseit.10.2.10249.

[17] C. Jin *et al.*, "Big data in food safety- a review," *Current Opinion in Food Science*, vol. 36, pp. 24–32, Dec. 2020, doi: 10.1016/j.cofs.2020.11.006.

[18] N. N. Misra, Y. Dixit, A. Al-Mallahi, M. S. Bhullar, R. Upadhyay, and A. Martynenko, "IoT, big data, and artificial intelligence in agriculture and food industry," *IEEE Internet of Things Journal*, vol. 9, no. 9, pp. 6305–6324, May 2022, doi: 10.1109/JIOT.2020.2998584.

[19] A. Oussous, F.-Z. Benjelloun, A. Ait Lahcen, and S. Belfkih, "Big data technologies: a survey," *Journal of King Saud University - Computer and Information Sciences*, vol. 30, no. 4, pp. 431–448, Oct. 2018, doi: 10.1016/j.jksuci.2017.06.001.

[20] I. Lee, "Big data: Dimensions, evolution, impacts, and challenges," *Business Horizons*, vol. 60, no. 3, pp. 293–303, May 2017, doi: 10.1016/j.bushor.2017.01.004.

[21] A. Kamilaris, A. Kartakoullis, and F. X. Prenafeta-Boldú, "A review on the practice of big data analysis in agriculture," *Computers and Electronics in Agriculture*, vol. 143, pp. 23–37, Dec. 2017, doi: 10.1016/j.compag.2017.09.037.

[22] B. Turapov and P. Siegfried, "Analyzing industry 4.0 with an example from agricultural manufacturing," *Accurate: Journal of Mechanical Engineering and Science*, vol. 2, no. 2, pp. 17–22, 2021, doi: 10.35970/accurate.v2i2.1079.

[23] J. A. Shamsi and M. A. Khojaye, *Big data systems*. Chapman and Hall/CRC, 2021.

[24] J. Iaksch, E. Fernandes, and M. Borsato, "Digitalization and Big data in smart farming – a review," *Journal of Management Analytics*, vol. 8, no. 2, pp. 333–349, Apr. 2021, doi: 10.1080/23270012.2021.1897957.

[25] M. Ayaz, M. Ammad-Uddin, Z. Sharif, A. Mansour, and E.-H. M. Aggoune, "Internet-of-things (IoT)-based smart agriculture: toward making the fields talk," *IEEE Access*, vol. 7, pp. 129551–129583, 2019, doi: 10.1109/ACCESS.2019.2932609.

[26] "UN: farmers must produce 70% more food by 2050 to feed population." https://www.theguardian.com/environment/2011/nov/28/un-farmers-produce-food-population (accessed Dec. 03, 2021).

[27] O. F. El-Gayar and M. Q. Ofori, "Disrupting agriculture: the status and prospects for AI and big data in smart agriculture," in *Research Anthology on Food Waste Reduction and Alternative Diets for Food and Nutrition Security*, IGI Global, 2021, pp. 771–812.

[28] S. Liu, L. Guo, H. Webb, X. Ya, and X. Chang, "Internet of things monitoring system of modern eco-agriculture based on cloud computing," *IEEE Access*, vol. 7, pp. 37050–37058, 2019, doi: 10.1109/ACCESS.2019.2903720.

[29] M. S. Mekala and P. Viswanathan, "CLAY-MIST: IoT-cloud enabled CMM index for smart agriculture monitoring system," *Measurement*, vol. 134, pp. 236–244, Feb. 2019, doi: 10.1016/j.measurement.2018.10.072.

[30] M. Cicioğlu and A. Çalhan, "Smart agriculture with internet of things in cornfields," *Computers & Electrical Engineering*, vol. 90, p. 106982, Mar. 2021, doi: 10.1016/j.compeleceng.2021.106982.

[31] D. R. Vincent, N. Deepa, D. Elavarasan, K. Srinivasan, S. H. Chauhdary, and C. Iwendi, "Sensors driven AI-based agriculture recommendation model for assessing land suitability," *Sensors*, vol. 19, no. 17, p. 3667, Aug. 2019, doi: 10.3390/s19173667.

[32] S. Rajeswari, K. Suthendran, and K. Rajakumar, "A smart agricultural model by integrating IoT, mobile and cloud-based big data analytics," in *2017 International Conference on Intelligent Computing and Control (I2C2)*, Jun. 2017, pp. 1–5, doi: 10.1109/I2C2.2017.8321902.

[33] F.-H. Tseng, H.-H. Cho, and H.-T. Wu, "Applying big data for intelligent agriculture-based crop selection analysis," *IEEE Access*, vol. 7, pp. 116965–116974, 2019, doi: 10.1109/ACCESS.2019.2935564.

[34] N. K. Nawandar and V. R. Satpute, "IoT based low cost and intelligent module for smart irrigation system," *Computers and Electronics in Agriculture*, vol. 162, pp. 979–990, Jul. 2019, doi: 10.1016/j.compag.2019.05.027.

[35] F. Bu and X. Wang, "A smart agriculture IoT system based on deep reinforcement learning," *Future Generation Computer Systems*, vol. 99, pp. 500–507, Oct. 2019, doi: 10.1016/j.future.2019.04.041.

[36] C. Kamienski *et al.*, "Smart water management platform: IoT-based precision irrigation for agriculture," *Sensors*, vol. 19, no. 2, p. 276, Jan. 2019, doi: 10.3390/s19020276.

[37] T. Wang, X. Xu, C. Wang, Z. Li, and D. Li, "From smart farming towards unmanned farms: a new mode of agricultural production," *Agriculture*, vol. 11, no. 2, p. 145, Feb. 2021, doi: 10.3390/agriculture11020145.

[38] P. K. R. Maddikunta *et al.*, "Unmanned aerial vehicles in smart agriculture: applications, requirements, and challenges," *IEEE Sensors Journal*, vol. 21, no. 16, pp. 17608–17619, Aug. 2021, doi: 10.1109/JSEN.2021.3049471.

[39] A. M. S. Osman, "A novel big data analytics framework for smart cities," *Future Generation Computer Systems*, vol. 91, pp. 620–633, Feb. 2019, doi: 10.1016/j.future.2018.06.046.

[40] E. L. White *et al.*, "Report from the conference, 'identifying obstacles to applying big data in agriculture,'" *Precision Agriculture*, vol. 22, no. 1, pp. 306–315, Feb. 2021, doi: 10.1007/s11119-020-09738-y.
[41] P. S. M. Gopal and B. R. Chintala, "Big data challenges and opportunities in agriculture," *International Journal of Agricultural and Environmental Information Systems*, vol. 11, no. 1, pp. 48–66, Jan. 2020, doi: 10.4018/IJAEIS.2020010103.
[42] N. Ramankutty *et al.*, "Trends in global agricultural land use: implications for environmental health and food security," *Annual Review of Plant Biology*, vol. 69, no. 1, pp. 789–815, Apr. 2018, doi: 10.1146/annurev-arplant-042817-040256.
[43] S. van der Burg, M.-J. Bogaardt, and S. Wolfert, "Ethics of smart farming: current questions and directions for responsible innovation towards the future," *NJAS: Wageningen Journal of Life Sciences*, vol. 90–91, no. 1, pp. 1–10, Dec. 2019, doi: 10.1016/j.njas.2019.01.001.
[44] K. A. Shastry and H. A. Sanjay, "Data analysis and prediction using big data analytics in agriculture," in *Internet of Things and Analytics for Agriculture*, Singapore: Springer, 2020, pp. 201–224.
[45] A. Weersink, E. Fraser, D. Pannell, E. Duncan, and S. Rotz, "Opportunities and challenges for big data in agricultural and environmental analysis," *Annual Review of Resource Economics*, vol. 10, no. 1, pp. 19–37, Oct. 2018, doi: 10.1146/annurev-resource-100516-053654.

BIOGRAPHIES OF AUTHORS

Tehreem Qamar 🆔 📇 sc 🔾 is currently pursuing the Ph.D. degree in Computer Science at Jinnah University for Women, Karachi, Pakistan. She has received her MS degree in Computer Science and Information Technology from N.E.D. University of Engineering and Technology, Pakistan and BS degree in Computer Science from Jinnah University for Women, Pakistan. She is working as lecturer in the department of Computer Science and Software Engineering, Jinnah University for Women. Her research interests include Machine Learning, Semantic Web and Human Computer Interaction. She can be contacted at email: tehreem.qamar@juw.edu.pk.

Narmeen Zakaria Bawany 🆔 📇 sc 🔾 has completed her Ph.D. in Computer Science from National University of Computer and Emerging Sciences, Karachi, Pakistan. She has over 20 years of experience in teaching, research, and administration. She is currently working as a Dean, Faculty of Science and is also the Head of the department of Computer Science and Software Engineering at Jinnah University for Women, Karachi, Pakistan. She has supervised many projects and has also received funding from Ignite (National Technology Fund, Pakistan) for her projects. Her research areas include Human Computer Interaction, Cloud Computing, Blockchain technology, Machine Learning, Cyber security, and Software Defined Networking. She has more than 30 publications in journals and conferences. She can be contacted at email: narmeen.bawany@juw.edu.pk.