



# Machine Learning Prediction Analysis using IoT for Smart Farming

Abdul Rehman<sup>1</sup>, Jian Liu<sup>2</sup>, Li Keqiu<sup>3</sup>, Ahmed Mateen<sup>4\*</sup> and Muhammad Qasim Yasin<sup>5</sup>

<sup>1,2,3</sup>College of Intelligence and Computing, Tianjin University, Tianjin Key Laboratory of Advance Networking, Tianjin, China.

<sup>4</sup>Department of Computer Science, University of Agriculture Faisalabad, Pakistan, 38000

Email: ahmedbuttar@uaf.edu.pk

<sup>5</sup>College of Intelligence and Computing, Tianjin University, Tianjin Key Laboratory of Cognitive Computing and Application, Tianjin, China

## ABSTRACT

Commonly, the implementation of agricultural practices (e.g., ploughing, sowing, watering, pests control, and harvesting.) purely depends on climate change, recommendations from previously experienced rules, and Governmental policies. For fulfilling Term smart farming, i), we employed real-time applications over sensors to capture climate changes of soil and atmosphere. ii) we defined agriculture practice rules by applying machine learning techniques over the last five years data iii) By federations of real-time data from the field sensors and rules, we define the time for implementation of the practice. This federation eliminates many malfunctions in old ways of smart farming for precision agriculture.

**Keywords:** Agriculture, IoT, Sensors, Smart Farming, Raspberry Pi3.

## 1. INTRODUCTION.

Pakistan is an agro-based country, 47 % of the total land area is for agriculture, and 61% of rural community make their earnings directly or indirectly from agriculture [1]. It is playing an essential contribution in the economy.

With the regular expansion of industrializations and urbanizations, the agricultural area is continuously reducing. It is boosting the demand for food due to the high growth rate. They are still practicing traditional agriculture with the old traditional methodology of cultivation. Which are needed to be replaced with new emerging technologies, e.g., field sensors implementations [2]. data analytics, experienced bases decision systems. Agri-data of the last fifteen years about pest attacks, agricultural practices regarding different zones is available. Smart farming is the demand of the era to meet the food demand by boosting crop production. The authors [3] implemented big data analytics to formulate agriculture data. The stored agri-data is needed to be cleaned and stored to be used for valuable decision support systems. In this manuscript, we import data in heterogeneous from heterogeneous systems. For example, the old agri-data is

available in CSV, TXT, and relational database format. Secondly, we collect real-time data with IoT sensors from

agriculture. This real-time data stream is composed of soil temperature, atmospheric temperature, and humidity. The practices plan is formulated with the information extracted from the following sources e.g.

- 1) Real-time data streams from Sensors
- 2) Information about the sensor locality
- 3) Information extraction from five years old data
- 4) Climate forecast by the government
- 5) Watering System

The Agricultural Big source of heterogenous Data Because the Agricultural Data is also being collected from different sources like public data, private data, industrial data, and Governmental data. So, data privacy and security are also required for private data and industrial data [4]. For such [3] a significant and diversified data, we need to develop a large cloud environment for data management and analytics [5]. The apache-spark [6] based cloud environment is constructed for data cleaning, collecting, storing, and data analytic & Machin learning.

## 2. BACKGROUND.

### 2.1 Precision farming

Precision farming is the term that refers to the effective and efficient use of limited inputs to get more outputs. It is a new way to employ digital technology to optimize agriculture practices. Different trends of technologies are changing the parameters and shapes of precision agriculture. The major trends are IoT, weather forecast, and big data technologies. The Internet of things has a significant impact in all fields; it gives new automation [31]. It leads to new directions for scientists to implement their research. Specifically, it improves the efficiency of operational work. In agriculture, it applications of agriculture [7]. For example, it made it possible to attain real-time field data about soils and atmospheric temperature. For large form, aerial imagery and drones, and satellites are implemented as a tool of precision agriculture. To make the precision agriculture more sustainable, reliable, optimization and productive we employed machine learning and data analytics to formulate the quick decision.

## 2.2 IoT in Agriculture

The main factors of sustainability and reliability can be ensured with implementation of the smart farming concepts. We can optimization agricultural practices by defining rules by application of Machine learning and by recording the real time rainfall, temperature, humidity and pest activity and prediction. For predicting soil drying, soil conditions, soil temperature, and soil moisture KNN ANN /MLP and RBF are used [8]. The diseases identifications [9,28] is no more unresolved. The agri-fields are monitored by satellites, sensors and drones. The data is collected a for crop pest management and irrigation system. The real time data we need field sensors for soil temperature, humidity and atmospheric temperature and humidity.

The remote sensing system is effective for detection of plant disease and it is very cheap as compare to traditional methods of pest scouting. In modern word, area which can be cultivated is rapidly reducing. Sensor technologists and agriculture scientists are working in collaboration to increase productivity in agriculture field [10]. Crop yield predication techniques have an important role in improving crop productivity because they help the farmers in timely decisions about production, storage and making of crops as well as risk management [11-13]. The efficient and sustainable production methods can be recognized to build the reliable agriculture strategies.

Machine learning is used for crop prediction in many techniques. Intelligent techniques have been suggested by various researchers and have given accurate results as well [14]. These tests have been done on small number of crops. Gonzalez-Sanchez et al.'s study [15] included data from a large number of crops. Accuracy of different learning methods was tested in this analysis and some techniques were proposed to predict crop growth [16]. The Support vector technique is more accurate and precise as compared to ARIMA. It is very easy to implement as compared to statistical approaches [17],[30].

In [18] the author shows that factors like temperature of pacific and Indian ocean, monsoon rainfall and pressure of sea level have direct impact on growth of agriculture products in Pakistan. Moreover, the results prove the production rate and amount of monsoon rainfall, all over Pakistan remain constant, except few cases. Machine learning algorithm U ChooBoost is used to accommodate idea of PA [19]. Lots of digital information obtained from farm sensors needs to be manipulated. Knowledge mining uses this supervised learning algorithm (UChooBoost) [20]. UChoo classifier ensembled over machine learning. PA votes are assigned with specific weights and highest weighted votes are combined to check the performance. Which boost up the performance. The extensive experimentation is carried out over extended data expression.

Results of many assumptions can also be tested with its help which will help in improvement of algorithm performance.

### Advantages of proposed system

- The temperature and humidity of a particular place to a particular depth for the specific crop can be record automatically with adjustment.
- The raspberry pi 3 requires very small processing space.

Information which is collected by farms sensors is used by artificial neural networks for predication of production rate of crops. Parameters like temperature, soil, rainfall, pressure and humidity in this information. These parameters and their effects on crop growth are discussed and the results evaluated [21].

In [22] the authors develop a system to detect animal growth in the livestock. It remains successful and very effective for reducing cost, efficient source of energy, and robustness. The Mobile Monitoring System based on RFID to hold the cattle efficiently by extracting dynamic information about their locality and behavior by wireless network. Their behavior is analyzed. The used IoT sensors

with embedded system. The measure many parameters for further analysis through this IOT-based system in a smart way. In [23] cattle are monitored in the fields. The system is developed over Zigbee and WSN. They effectively used WSN for identifying the locality of the cattle.

A team of researchers [24] develop disease detection system with “infrared thermography”. They efficaciously detect disease over foot and mouth of livestock. They used infrared thermography for screening the disease.

Scientist [25] recorded the heartbeat of the cattle. For the purpose they used polar sport tester (PST) monitoring technique for cattle. They monitor animal for their temperature and heartbeat to identify the diseases.

In [26] the authors developed a real time system by using the technology of smart mobiles and Bluetooth system with IOT. They recorded the parameters like temperature and heartbeat rate continuously.

The health monitor devices are not available in Pakistan market veterinary doctors monitor it manually. [27],[29]. But sensor for soil temperature and humidity at different levels are available which are accessible for implementation and testing for our framework.

## 3. PROPOSED SYSTEM

In this manuscript we proposed novel framework for smart framing with the federation of Machine learning prediction system and real-time data collections by sensors, Raspberry pi3 and IOT technology. We take agriculture practices as use case. The data is collected in DHT11. Raspberry pi is stored on cluster by employed the Dstream and txt splitting and cleaning techniques. Once the data is stored, with federated queries it can be accessed for planning and implementation of the specific agriculture practice.

The temperature and humidity are main features sensor DHT11 to ensure sustainability. The data being processed by Raspberry pi will be updated continuously and passed to streamed cluster. By data analytics the facts and figure can be presented to the user. The recommendation is also made for a specific condition.

- It is very useful for predicting and forecasting the upcoming climate risks
- Accessible from anywhere, efficient, user friendly.

### 3.1 Framework

The monitoring node is raspberry pi3 model b monitor each node and collect real-time data in the form streams which are converted to Discretized Stream (DStream) which is continuous form of Resilient Distributed Datasets (RDDs). With spark core functionality and SparkSQL we become enable to answer the federated query which enables the system to make some recommendations.

For getting the continuous and sustainable real-time data the sensor needed to be connected to the raspberry pi3 model b with jumper wires as in figure no 3.

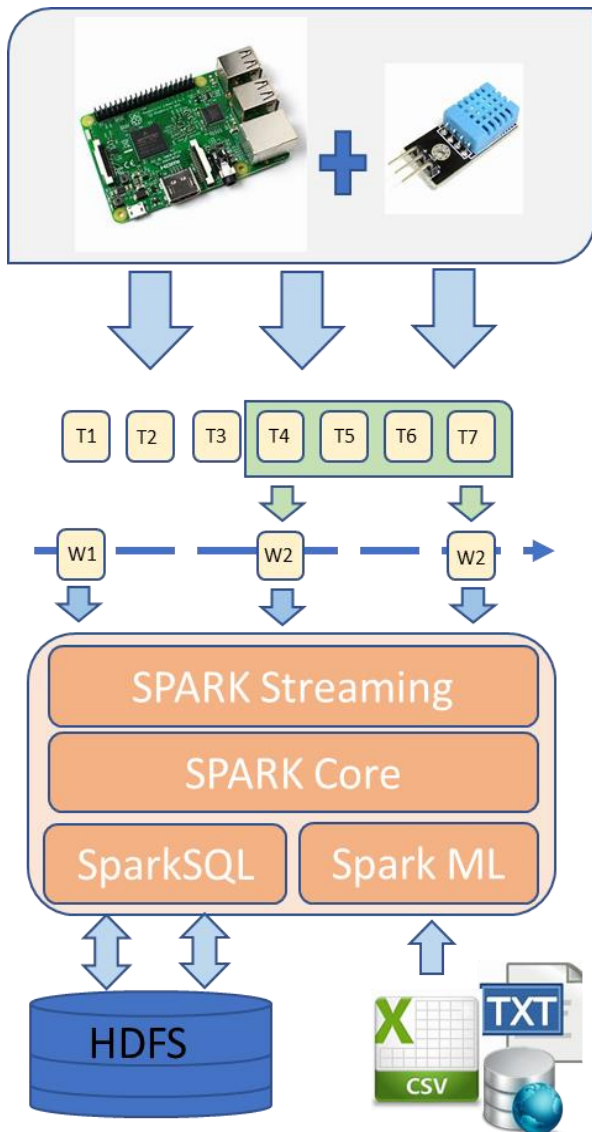


Figure 1: Framework of smart Farming

T1,T2,T3,T4,T5,T6,T7 are the series of times when the data is broadcasted. And W1, W2, W3 are the dstream windows. Which can be stored and can queried by SparKSQL. SparkML is the module of the spark which is used to implement different classifier of machine learning to extract the continuous and discrete, diversified agri-data. The data is available in heterogenous formats like text, web data, and CSV etc. the

required information is extracted and build rule of farming for recommendations as shown in figure 1. When real information is matched, and it is uneven it makes the recommendations.

### 3.2 Real-time DATA

The data is captured as mentioned in fig 1. For sensors in same way it also can be collected from satellites and other machines etc. In this paper, we present a simple framework for smart farming to highlight the new research directions. The data collect about the climate change and we made predictions accordingly with machine learning what it will have effects over plants.

### 3.3 Connectivity

As first step we can access raspberry pi interface on laptop by following steps

- Wireless WIFI connection is published and shared
- Through SSH the GUI of raspberry pi is accessible by USB /ethernet cable consoles.
- For identify the default IP address advanced IP scanner can be used.
- For GUI Xming server and for ssh Xshell can be used to view the SSH and GUI interface.

When it became accessible it shows the login, the default login and password, which are mentioned on it. We can connect to the terminal by giving the default password as shown in figure 2.



Figure 2: Password Screen

Now the commands ssh with -X on XShell or Xmanger. We can visualize the GUI interface Connections. The way of

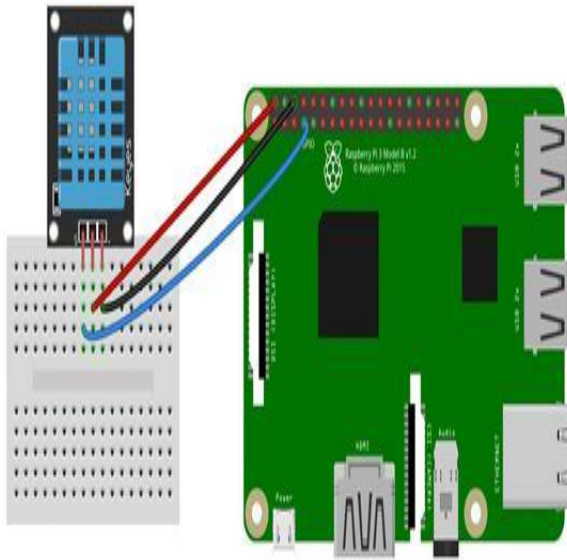


Figure 3: Outer Interface Connection Setting

Let take an example of sugarcane which has different growth stages such as

**Germination** of duration: 15 - 45 days,

**Tillering** of duration: 92 days,

**vegetative** of duration: 93 days

**maturing** normally duration: 60-75 days.

The rule for temperature is as followed for different stages.

Table 1: Temperature from IoT Devices

Stages	Optimal Temperature
Sowing	25-30 C
Germination	28 -30 C
Tillering	< 30 C
Vegetative	28-38C
Maturing	10C

For the rule generated from the data by machine learning, when the value of temperature form real-time system become worst, the system generate serious alert and also made recommendation that how, the temperature can be lessened or increased to get the optimal value. Is the irrigation required. When the temperature is greater or less then the optimal value but not to worry, the recommendation system sends warnings.

connection of Raspberry pi with DHT11 sensor using project board and jumper wires are shown in figure 3.

The data has been taken from the sensors using IoT setup and shown in table 1 and table2. Here we show the sample readings for temperature and humidity. We can see that Temperature and humidity are changes day to day in surrounding environment. From these three different samples calculate minimum, maximum and average values of both temperature and humidity. Here the DStream windows is the day when the day reading has conflict with the rule for the case sugarcane, the system generates alerts and made recommendations with rules and suggestions

From the table 1 and table2, we take readings of temperature and humidity for three samples.

Table 2: Humidity from IoT Device

Date	Sample 1	Sample 2	Sample 3	Ma x	Mi n	GM
29.04.2020	32	34	32	34	33	32.653
30.04.2020	28	29	33	33	28	29.924
31.04.2020	26	30	33	33	26	29.525

#### 4. CONCLUSION

In the manuscript we elaborate the novel framework of recommendations system for smart and precision farming. We try to ensure its sustainability be federating different heterogeneous data from diversified resources. We conduct analysis and made rules for making recommendations and suggestions. We also predict the upcoming risks on the basis of available data. The data includes the various parameters including weather forecast, soil conditions, atmospheric conditions, irrigation plan, cropping pattern information, pest scouting and control policies by Government, fertilizers doses and applications for various crops and crop yields etc. we used the partial for this prototype in future we use the whole data for big and complete solution of smart farming. The implementations of real-time prototype by using DHT11 sensor show the new way of research to be get informed about the field about the atmospheric and soil conditions of data on time. We effectively used the big data technologies like Hadoop ecosystem with spark to predict and suggest the plan for the agricultural practices.

## REFERENCES

- [1] C. Abrar and A. Nadia. **Climate-Smart Agriculture in Pakistan**, CIAT / World Bank <https://cgspace.cgiar.org/handle/10568/83340>, 2017
- M. Kaur, G. Heena and H. Kundra. **Data Mining in Agriculture on Crop Price Prediction Techniques and Applications**, *International Journal of Computer Applications*, Vol. 99, no.12, pp. 1-3, August 2014.
- [2] S. Hussain, F. Kabeer, W. Ali and K. Jamshed. **A New Framework to Integrate Wireless Sensor Networks with Cloud Computing**, in *Proc. of IEEE Aerospace Conference*, USA, 2013, pp. 2-9.
- [4] F. X. Diebold. **On the Origin(s) and Development of the Term “Big Data”**, *Penn Institute for Economic Research*, Pennsylvania, 2012, pp.12-37.
- [5] D. Laney. **3-D Data Management: Controlling Data Volume, Velocity and Variety**, *META Group Research Note*, Vol. 6, no. 70, pp.1, 2001.
- [6] Apache Spark: <https://spark.apache.org/>
- [7] V. C. Patil, K. A. A1-Gaadi, D. P. Biradar and M. Rangaswamy. **Internet Of Things (IoT) and Cloud Computing for Agriculture: An Overview**, in *Proc. of agro-informatics and precision agriculture (AIPA)*, India, 2012, pp. 292-296.
- [8] E. Khandakar, A. Unayes and M. A. Gregory. **Integrating Wireless Sensor Networks with Cloud Computing**, in *Proc. of 7th International Conference on Mobile Ad-hoc and Sensor Networks (MSN)*, Beijing, 2011, pp.16-18.
- [9] D. S. Sawaitul, K.P. Wagh and P. N. Chatur. **Classification and Prediction of Future Weather by using Back Propagation Algorithm- An Approach**, *International Journal of Emerging Technology and Advanced Engineering*, Vol. 2, no. 1, pp. 110-113, January 2012.
- [10] I. Jagielska, C. Mathehews and T. Whitfort. **An investigation into the application of neural networks, fuzzy logic, genetic algorithms, and rough sets to automated knowledge acquisition for classification problems**, *Neuro computing*, Vol. 24, no. 1-3, pp.37-54, 2012.
- [11] D. Ramesh and B. V. Vardhan. **Analysis of Crop Yield Prediction Using Data Mining Techniques**, *International Journal of Research in Engineering and Technology*, Vol. 4, no. 1, pp. 47-473, January 2015.
- [12] S. Veenadhari, B. Misra and C. D. Singh. **Data Mining Techniques for Predicting Crop Productivity – A Review Article**, *International Journal of Computer Science and Technology IJCSST*, Vol. 2, no. 1, pp. 90-100, March 2011.
- [13] W. Duncan, K. Rabah. **Environmental Conditions’ Big Data Management and Cloud Computing Analytics for Sustainable Agriculture**, *World Journal of Computer Application and Technology*, Vol.2, no.3, pp. 73-81, 2014.
- [14] R. Piyare, S. Park, S. Y. Maeng, S. H. Chan, S. G. Choi, H. S. Choi and S. R. Lee. **Integrating Wireless Sensor Network into Cloud Services for Real-time Data Collection**, in *Proc. of International Conference on ICT Convergence*, Jeju, pp. 752-756, 2013.
- [15] V.Rajesh, J. M. Gnanasekar, R. S. Ponmagal and P. Anbalagan. **Integration of Wireless Sensor Network with Cloud**, in *Proc. of IEEE International Conference on Recent Trends in Information, Telecommunication and Computing*, Kochi, pp. 321-323, 2010.
- [16] R. Kumar, M. P. Singh, P. Kumar and J.P. Singh. **Crop Selection Method to Maximize Crop Yield Rate using Machine Learning Technique**. In *Proc. of IEEE International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials ICSTM*, India, 2015, pp. 138-145.
- [17] Z. Peng, Y. Zheng and S. Hamlin. **A Novel Architecture Based on Cloud Computing for Wireless Sensor Network**, in *Proc. of 2nd International Conference on Computer Science and Electronics Engineering ICCSEE*, pp 472-475, 2013.
- [18] R. Sudarsan, V. Rao, N. Satyanarayana and V. L. Prasanna. **Generic Integrated Secured WSN-Cloud Computing for U-Life Care**, *International Journal of Engineering Science and Advanced Technology*, Vol. 2, no. 4, pp 897-907, 2011.
- [19] M. Firdhous, O. Ghazali and S. Hassan. **Trust Management in Cloud Computing: A Critical Review**, *International Journal on Advances in ICT for Emerging Regions*, Vol. 4, no. 2, pp. 24-36, 2011.
- [20] V. Rajesh, O. Pandithurai and S. Magesh. **Wireless sensor data on cloud**, in *Proc. of IEEE International Conference on Communication, Control and Computing Technologies*, Ramanathapuram, 2010, pp. 476-481.
- [21] Y. Rupika, J. Rathod and N. Vaishnavi. **Big data meets small sensors in precision agriculture**, *International Journal of Computer Applications*, Vol. 975, pp. 8887, 2015.
- [22] A. Lounis, H. Abdelkrim, A. Bouabdallah and Y. Challal. **Secure and Scalable Cloud-based Architecture for e-Health Wireless sensor networks**, in *Proc. of 21<sup>st</sup> International Conference on Computer Communications and Networks (ICCCN)*, Munich, 2012, pp 1-7.
- [23] P. Langendoerfer, K. Piotrowski, M. Diaz and B. Rubio, **Distributed Shared Memory as an Approach for Integrating WSNs and Cloud Computing**, in *Proc. of 5th International Conference on New Technologies, Mobility and Security*, Istanbul, 2012, pp 1-6.
- [24] C. Srimathi, P. Soo-Hyun, N. Rajesh. **Proposed framework for underwater sensor Cloud for environmental Monitoring**, in *Proc. of 5th International Conference on Ubiquitous and Future Networks*, Da Nang, 2013, pp. 104-109.
- [25] F. Tongrang, Z. Xuan and G. Feng. **Cloud Storage Solution for WSN Based on Internet Innovation Union**, *International Journal of Database Theory and Application*, Vol. 22, pp. 164-169, 2013.
- [26] S. Rajeswari, K. Suthendran and K. Rajakumar. **A Smart Agricultural Model by Integrating IoT, Mobile and Cloud-based Big Data Analytics**, *International Journal of Pure and Applied Mathematics*, Vol. 118, no. 8, pp. 365-370, 2018.
- [27] S. Rajeswari, K. Suthendran, K. Rajakumar and S. Arumugam. **An Overview of the MapReduce Model**, *Lecture Notes in Computer Science*, Vol. 10398, pp. 312-317, 2017.

[28] M. Ongayev, Z. Sultanova, S. Denizbayev, G. Ozhanov and S. Abisheva. **Engineering and Process Infrastructure of the Agro-Industrial Complex**, *International Journal of Emerging Trends in Engineering Research*, Vol. 7, no. 12, pp. 879-885, 2019.

[29] S. Aleksander, A. Konev, T. Kosachenko and D. Dudkin. **Threat Model for IoT Systems on the Example of OpenUNB Protocol**, *International Journal of Emerging Trends in Engineering Research*, Vol. 7, no. 9, pp. 834-290, 2019.

[30] L. Boggula, B. Navyasri. **Energy Efficient Routing Mechanism for Harsh Environment in Wireless Sensor Networks**, *International Journal of Emerging Trends in Engineering Research*, Vol. 7, no. 9, pp. 234-238, 2019.

[31] Apache kafka: <https://kafka.apache.org/>