1. INTRODUCTION

Agricultural sector plays a strategic role in the process of economic development of a country. It has already made a significant contribution to the economic property of advanced countries and its role in the economic development of less developed countries is of vital importance. The agriculture sector is the backbone of an economy which provides the basic ingredients to mankind and now raw material for industrialization.

Our country plays a key role in Agriculture for the crop production with the foremost food staples like Chilli, Ground nut, Paddy, Jowar, Bengal gram, etc. In the next few decades farming industry will become more important than before. According to the UN Food and Agriculture Organization "the world needs to produce 70% more food in 2050 than 2006 in order to feed the growing population of the Earth". To produce that amount of food farmer has to change the techniques and methods that are being used now. Farming, in India, is still mundane and doesn't see the usage of modern techniques that were employed in many other developed countries. The reason may be that most of the farmers are uneducated and many of them are small and marginal with just 1-2 hectares of land. But, the best out of the available land can be made in a low budget by effective use of IoT devices and sensors. Moreover, there are several factors that take stand when it comes to the prolificacy of a farm like monsoons, soil type, available water resources, soil micronutrients, temperature and humidity of that area, acidity, diseases that the crop could be prone to, and so on.

IoT, a sensation in the modern-day technology, has a major impact on the rapidly growing technological aspects. The Internet of Things is a network of physical things with built-in sensors, software, and other technologies for the purpose of connecting & integrating data with other devices and systems over the internet. It's making people's life easier and enabling us to do things that were previously seen as miracles. From smart appliances to smart security, smart meters to smart city technologies, IoT is revolutionizing the way the world works. We are in a hyper-connected world where things connect with each other with their "smartness" and make life easy for us. It helps in solving many complex real-time problems. One such major applications in the field of agriculture can turn out to be productive and profitable. This paper presents the application of various techniques infusing IoT in agriculture, that helps in productive and predictive results in that field, thereby leading towards precision agriculture. A low-cost power supple and definitive farming system using IOT with the cloud computing platform.

IoT is an intelligent technology which includes identification, sensing and intelligence. Life and even intelligence of life itself can also be regarded as part of IoT technology. It is used in pattern identification fields like measurement and computing as well as computer and communication fields like sensing, communication, information collection and processing. The definition of IoT changes as the time of cloud computing comes. It is now defined as IoT=cloud computing + ubiquitous network + intelligent sensing network. Cloud computing management platform is the “brain” of cloud computing and relevant data. It involves management of accession of cloud computing customization application by users of this IoT, computing and processing what is involved in customization service, organizing and coordinating service nodes in the data center. IoT is a technology which aims to extend internet to large number of distributed devices by defining standard, interoperable communication protocols”. Smart farming and precision agriculture involve the integration of advanced technologies into existing farming practices in order to increase production efficiency and the quality of agricultural products. As an added benefit, they also improve the quality of life for farm workers by reducing heavy labor and tedious tasks.

While world agriculture is undergoing industrialization, it is important to develop agricultural information at the same time. Agricultural information has become the trend of development for world agriculture. In recent years, we have been focusing on agricultural information service and infrastructure development. After years of hard efforts, remarkable results have been seen in agricultural infrastructure development, like “Every Village" project of Ministry of Industry and Information, "Golden Agriculture project” and “Three Dian Project" (computer, TV and telephone network coverage in rural area) of Ministry of Agriculture. This infrastructure provided foundation for agricultural information service. Moreover, information is not sufficiently used by farmers and the effect of information on agriculture, famers and rural area is not that notable.

1.1 Problem Definition including significance and objective

In the modern era where technology is taking steps towards automation leading to control everything from our fingertips. IoT triumphs over a vast number of upcoming technologies as it conquers life by making it easier and gluing people to it in almost all realm. IoT basically infuses smartness into the environment. In the field of agriculture where farmers' mechanical work plays a vital role the Internet of Things is capable of making smart and informed decisions as of what can be done to improvise the growth and outcome, improving the quality and quantity, minimizing the risk and reducing the effort to maintain the crops. The central theme revolves around the availability of required farm data anytime and anywhere. In key terms it is supposed to be centralizing the data necessarily.

The current project is essentially built to be focused upon selection of a certain type of crop suitable to a farm based on the conditions prevailing in the farm. Many sensors, methods and technologies have been used to collect the data from the field and this data is used for picking the appropriate crop. Moreover, the inclusion of the cloud brings a remote monitoring advantage where the requirements and circs can be reviewed every so often. Also, the data that has been accumulated is displayed in various graphs showing comparison from past times to recently recorded moments such that the trends in the farm can be judged.

1.2 Introduction to Domain Terminology

Table1.2: Key terminologies related to smart farm

IOT	The Internet of Things refers to "the ever-growing network of physical objects that feature an IP address for internet connectivity, and the communication that occurs between these objects and other Internet-enabled devices and systems".
Data Analysis	Analysis of historic and current data that is being constantly generated from the farm to draw statistics from them and represent them to the user in a much effective manner.

Cloud Computing	Cloud computing refers to “a network of remote servers hosted on the Internet to store, manage, and process data, rather than a local server or a personal computer”.
Sensor	“A sensor is a device that detects and responds to some type of input from the physical environment”. The output is generally a signal that is converted to human-readable display at the sensor location or transmitted electronically over a network for reading or further processing.
Wireless-sensor network	“A wireless sensor network is a group of specialized transducers with a communications infrastructure for monitoring and recording conditions at diverse locations”.
Automation	“Automation is the technology by which a process or procedure is performed with minimal human assistance”.
Machine Learning	Machine Learning is a method of data analysis that automates analytical model building. It is a branch of artificial intelligence based on the idea that systems can learn from data, identify patterns and make decisions with minimal human intervention.

1.3 Methodologies

Ever since the origin of agriculture, many studies have been made and several propositions were developed as to explain the historic emergence of farming. A myriad of hypotheses to explain the historical inception of agriculture have been developed. A transition from hunter-gatherer to agricultural societies indicate an antecedent period of intensification and increasing sedentism. To a large extent, the rate of technology development and the degree of innovation in future technologies will greatly influence the stability, and certainly the productivity, of agriculture. Technology, in the classical sense, includes the development and use of nutrients, pest control products, crop cultivars, and farm equipment; but it also includes the vision of genetically modified crops providing greater nutritional efficiency, manipulation of natural pest control agents, and use of farm management techniques that focus on whole-farm productivity over time, not just annual production per hectare. This is an era where technology is evolving rapidly, and there by required improvements are to be injected in every sector. Adopting modern techniques in the farm has become desideratum.

1.3.1 Existing System

- There are many systems that decide the farm's growth and productivity. Using existing systems individual factors can be measured but major decision as to decide what type of crop to cultivate is brutally abandoned.
- The system work upon already decided crops by farmers which typically ignored various factors that play a major role in retaining soil fertility and enhancing crop growth.

1.3.2 Proposed System

- It collects the prevailing conditions of temperature and humidity continuously.
- It collects the pH data of the soil to determine the acidity, alkalinity of the soil.
- The moisture levels of the soil at all times is collected to know if that serves the requirement of a certain crop type.
- The collected data is sent into a cloud for the purpose of centralizing the data such that it can be accessed anywhere.
- The data from temperature, humidity, pH, soil moisture sensors are aggregated to predict the type of crop that suits a farm.

The proposed system extends the existing systems that were able to innovate the farms to some extent, by supplementing it with new capabilities to predict the crop type and centralize farm data. The hardware components include Raspberry Pi 3, soil moisture sensor, pH sensor, Rain sensor, DHT11 sensor. The sensors are plugged at a place in the field by connecting to Raspberry Pi for data acquisition. From those sensors, temperature, humidity, acidity, moisture levels of the soil are recorded.

The acquired data is moved to the cloud instantly as they are noted. From that data using all those parameters as the basis for prediction, the k-means algorithm is applied to know the suitable crop. This can be reviewed wherever needed through a web-app in your smart device or Pcs. Also, the result of prediction is sent in the form of a mail to the farmer.

1.4 Outline of the results

The results of the system when deployed were consistent. The test cases that were considered are the standard types of soils found in nearby fields of Hyderabad and from certain districts of Telangana. The prediction of the crop type gave different results considering various parameters that were recorded like the soil moisture, temperature, humidity, rainfall, pH values. All these values are considered as key components in the prediction of the type of crop to be cultivated in that area. The results depend on the k value of KNN algorithm. In KNN algorithm, an object is classified by a majority vote of its neighbors, with the object being assigned to the class most common among its k nearest neighbors. For k=1, result will be the class of single nearest neighbor. Increase in the value of k increases nearest neighbors. The k value influences the result. In addition to the collection of data, it was made real-time such that more data was collected in order to give greater accuracy in results.

1.5 Scope of the project

IoT (Internet of Things) has become an inseparable concept in the world. Automation became a culture in the tribe of humanity, where everyone desires a comfortable lifestyle at their fingertips. This technology environment brings a paradigm shift in our professional and personal life. As a connected environment, IoT adds customer value and loyalty. This project is a real-time work on how IoT is used in the field of Agriculture for crop prediction using sensor values and monitoring the farm.

The project was initiated with a major goal to fill the gap between the farmers and their field. As the farmer can't stay at the site all the time to look after, this is designed to remotely monitor the crop. The major deliverables include a prediction module that uses K-Nearest neighbors to predict the crop type. The hardware used includes Raspberry Pi, DHT11 sensor, soil moisture sensor, pH sensor and mail API to send predicted crop to the given mail. A major milestone for this system is to reach out the farmers to the maximum extent such that it serves their purpose to increase the productivity in a cost-effective manner.

The key constraints include the changing level of acidity in the soil and sudden changes in temperature or unexpected rainfall that might go out of hand and is an extreme condition to the current system. One of the issues with implementation is the availability of data to predict accurately. The system used soils from Telangana and Andhra for training and prediction. The scope excludes other factors like the availability of minerals and components in the soil inclusion of which would make it a complex system. This is designed to work with continuous supply of power as the Raspberry Pi is installed in the farm. And a Wi-Fi module is integrated to the system to transmit real time data to the cloud from where they can be accessed through a Web Application.

In the coming years, IOT and ML techniques gone a play a major role in removing the manual work for monitoring and predicting the factors influencing productivity of farm. The awareness and knowledge about newer agricultural technology are yet to spread extensively, especially in emerging countries.

The scope of the project varies from small individual house gardens to the large agricultural farms for monitoring the necessary conditions that can be reviewed by the guardian/owner of the farm.

1.5 Organization of Report

The introduction section is followed by Literature Survey where the existing systems are explained with various references to the farm. It introduces domain specific terminology which form the background to understand the current system. It discusses in depth about existing solutions' core aspect which forms the basis for many other solutions. The section also discusses the drawbacks in all the solutions exhaustively. The literature survey section is followed by design of the Proposed System section. This section discusses the evolution and design of the proposed solution. Next section discusses about the implementation of the design discussed in the previous section. The Data Flow Diagrams and Flowcharts are discussed in this section of the project. The algorithm is also discussed about in this section. The data set being used, the features of the data set, and their significance are mentioned. The testing process is also included. The next section deals with the result analysis. The system is executed over the test cases and the results are analyzed and discussed. The final section deals with limitations and recommendations. The references showing the base papers used in this project are then mentioned.

2.LITERATURE SURVEY

2.1 Introduction to Problem Domain Technology

Internet of Things (IoT) is the new world where everyday devices talk to one another and everything you do is monitored real time. It used to be the stuff of science fiction machines communicating with each other to take care of tasks automatically. It is the network of physical objects that contain embedded technology to communicate and sense or interact with their internal states or the external environment. IoT is the put together of sensors in electricity grid, railway, bridge, tunnel, road, building, water supply system, dam, oil and gas pipes, appliance, etc., and connect the internet, to operate certain programs and realize remote control. The central computer can realize concentrated management and control of machine, equipment and personnel based on the internet and improve production and life through more detailed and dynamic means. This is useful for integration and harmony between human society and the physical world and is regarded as the third wave of information industry development following computer and internet. Major IoT technologies include radio frequency identification technology, sensor technology, sensor network technology and internetwork communication, all of which have been involved in the four links of IoT industrial chain, namely, identification, sensing, processing and information delivery. IoT can be considered a markable milestone as it contributes towards a success driven technological world where every technology sets to defeat others by proving its importance to the people. IoT is a trend-setting innovation in which all the data from sensors is stored in the cloud where it can be easily accessed from the cloud. Sensors and actuators for gathering the data and sending across the internet are also included in this advancement. A sensor is a device, module, machine, or subsystem whose purpose is to detect events or changes in its environment and send the information to other electronics, frequently a computer processor. Sensors are always used with other types of electronics. Sensors are used in everyday objects such as touch-sensitive elevator buttons and lamps which dim or brighten by touching the base, besides innumerable applications that contribute to the comfort and luxury of mankind. The output of sensors is generally a signal that is converted to human-readable display at the sensor location or transmitted electronically over a network for reading or further processing.

The data from the sensors serves a deciding factor in any aspect where IoT is employed. Sensors are connected to a microprocessor or microcontroller which enables it to collect data. In here, we use the Raspberry Pi as an aid for serving our purpose.

The Raspberry Pi is a series of small single-board computers developed in the United Kingdom by the Raspberry Pi Foundation to promote teaching of basic computer science in schools and in developing countries. It is a low cost, credit-card sized computer that plugs into a computer monitor or TV and uses a standard keyboard and mouse. It is a capable little device that enables people of all ages to explore computing, and to learn how to program in languages like Scratch and Python. An SD card inserted into the slot on the board acts as the hard drive for the Raspberry Pi**.** It is powered by USB and the video output can be hooked up to a traditional RCA TV set, a more modern monitor, or even a TV using the HDMI port. The Raspberry Pi is 40 times faster than an Arduino when it comes to clock speed. The original model became far more popular than anticipated, selling outside its target market for uses such as robotics. It now is widely used even in research projects, such as for weather monitoring, because of its low-cost and portability. It does not include peripherals or cases. However, some accessories have been included in several official and unofficial bundles.

Machine learning is defined as the ability of a machine to vary the outcome of a situation or behavior based on knowledge or observation which is essential for IoT solutions. It also has a larger role as the source of directed learning to modify behavior at the edge. The philosophy behind machine learning is to automate the creation of analytical models in order to enable algorithms to learn continuously with the help of available data. Artificial intelligence and machine learning are not novelty innovations. As early as 1959, Arthur Samuel defined the concept of machine learning as the ability of computers to learn to function in ways that they were not specifically programmed to do. Of course, the timeline from definition to implementation in everyday life can be a long one. Today, many factors have come together to make machine learning a reality, including large data sources that are great for learning, increased computational power for processing information in split seconds, and algorithms that have become more and more reliable.

Machine learning can be applied in cases where the desired outcome is known i.e., supervised Learning, or the data is not known beforehand (unsupervised learning), or the learning is the result of interaction between a model and the environment (reinforcement learning). Machine learning has experienced a boost in popularity among industrial companies through hype surrounding the Internet of Things (IoT). Many companies are already designating IoT as a strategically significant area, while others have kicked off pilot projects to map the potential of IoT in business operations. There are at least two main reasons why machine learning is the appropriate solution for the IoT universe. The first has to do with the volume of data and the automation opportunities. The second is related to predictive analysis.

There are two algorithms in Machine Learning for prediction of labeled datasets which come under supervised learning. They are:

- Classification
- Regression

Classification predictive modeling is the task of approximating a mapping function from input variables to discrete output variables. The output variables are often called labels or categories. The mapping function predicts the class or category for a given observation. A classification problem requires that examples be classified into one of two or more classes.

- A classification can have real-valued or discrete input variables.
- A problem with two classes is often called a two-class or binary classification problem.
- A problem with more than two classes is often called a multi-class classification problem.
- A problem where an example is assigned multiple classes is called a multi-label classification problem.

It is common for classification models to predict a continuous value as the probability of a given example belonging to each output class. The probabilities can be interpreted as the likelihood or confidence of a given example belonging to each class. A predicted probability can be converted into a class value by selecting the class label that has the highest probability.

Regression predictive modeling is the task of approximating a mapping function from input variables to a continuous output variable.

A continuous output variable is a real-value, such as an integer or floating-point value. These are often quantities, such as amounts and sizes.

- A regression problem requires the prediction of a quantity.
- A regression can have real-valued or discrete input variables.
- A problem with multiple input variables is often called a multivariate regression problem.
- A regression problem where input variables are ordered by time is called a time series forecasting problem.

An algorithm that is capable of learning a regression predictive model is called a regression algorithm. Some algorithms have the word "regression" in their name, such as linear regression and logistic regression, which can make things confusing because linear regression is a regression algorithm whereas logistic regression is a classification algorithm.

Classification predictive modeling problems are different from regression predictive modeling problems.

- Classification is the task of predicting a discrete class label.
- Regression is the task of predicting a continuous quantity.

There is some overlap between the algorithms for classification and regression. For example:

A classification algorithm may predict a continuous value, but the continuous value is in the form of a probability for a class label.

A regression algorithm may predict a discrete value, but the discrete value in the form of an integer quantity.

Some algorithms can be used for both classification and regression with small modifications, such as decision trees and artificial neural networks. Some algorithms cannot, or cannot easily be used for both problem types, such as linear regression for regression predictive modeling and logistic regression for classification predictive modeling.

Importantly, the way that we evaluate classification and regression predictions varies and does not overlap, for example:

- Classification predictions can be evaluated using accuracy, whereas regression predictions cannot.
- Regression predictions can be evaluated using root mean squared error, whereas classification predictions cannot.

Cloud computing is the delivery of different services through the Internet. These resources include tools and applications like data storage, servers, databases, networking, and software. Rather than keeping files on a proprietary hard drive or local storage device, cloud-based storage makes it possible to save them to a remote database. As long as an electronic device has access to the web, it has access to the data and the software programs to run it. Cloud computing is a popular option for people and businesses for a number of reasons including cost savings, increased productivity, speed and efficiency, performance, and security.

Cloud computing is named as such because the information being accessed is found remotely in the cloud or a virtual space. Companies that provide cloud services enable users to store files and applications on remote servers and then access all the data via the Internet. This means the user is not required to be in a specific place to gain access to it, allowing the user to work remotely. It takes all the heavy lifting involved in crunching and processing data away from the device you carry around or sit and work at. It also moves all of that work to huge computer clusters far away in cyberspace. The Internet becomes the cloud, and voilà—your data, work, and applications are available from any device with which you can connect to the Internet, anywhere in the world.

Clouds can be both public and private. Public cloud services provide their services over the Internet for a fee. Private cloud services, on the other hand, only provide services to a certain number of people. These services are a system of networks that supply hosted services. There is also a hybrid option, which combines elements of both the public and private services.

- Cloud computing is the delivery of different services through the Internet, including data storage, servers, databases, networking, and software.

- Cloud-based storage makes it possible to save files to a remote database and retrieve them on demand.
- Services can be both public and private—public services are provided online for a fee while private services are hosted on a network to specific clients.

There are various types of clouds, each of which is different from the other. Public clouds provide their services on servers and storage on the Internet. These are operated by third-party companies, who handle and control all the hardware, software, and the general infrastructure. Clients access services through accounts that can be accessed by just about anyone.

- Private clouds are reserved for specific clientele, usually one business or organization. The firm's data service center may host the cloud computing service. Many private cloud computing services are provided on a private network.
- Hybrid clouds are, as the name implies, a combination of both public and private services. This type of model allows the user more flexibility and helps optimize the user's infrastructure and security.

Cloud computing is not a single piece of technology like a microchip or a cellphone. Rather, it's a system primarily comprised of three services: software-as-a-service (SaaS), infrastructure-as-a-service (IaaS), and platform-as-a-service (PaaS).

- **Software-as-a-service (SaaS)** involves the licensure of a software application to customers. Licenses are typically provided through a pay-as-you-go model or on-demand. This type of system can be found in Microsoft Office's 365.
- **Infrastructure-as-a-service (IaaS)** involves a method for delivering everything from operating systems to servers and storage through IP-based connectivity as part of an on-demand service. Clients can avoid the need to purchase software or servers, and instead procure these resources in an outsourced, on-demand service. Popular examples of the IaaS system include IBM Cloud and Microsoft Azure.
- **Platform-as-a-service (PaaS)** is considered the most complex of the three layers of cloud-based computing. PaaS shares some similarities with SaaS, the primary difference being that instead of delivering software online, it is actually

a platform for creating software that is delivered via the Internet. This model includes platforms like Force.com and Heroku.

Cloud computing is far more than just accessing files on multiple devices. Thanks to cloud computing services, users can check their email on any computer and even store files using services such as Dropbox and Google Drive. Cloud computing services also make it possible for users to back up their music, files, and photos, ensuring those files are immediately available in the event of a hard drive crash.

2.2 Existing Systems

In the existing systems, humidity sensors and temperature sensors collected the data from the farm and the soil moisture sensor was deployed to collect the moisture levels of the soil. It is to be noted that the sensors collected data individually and were analyzed individually irrespective of the presence of other factors. But sometimes, it's the combination of these factors that define a farm's wellbeing. Hardware used in the existing system- microcontroller, soil moisture sensor, humidity sensor, GSM module. In the existing system farmers can access the information through SMS. There are many devices available in the market today that allow farmers to monitor the soil on a regular basis from the comfort of their home but of high cost.

Some of the articles related to the current point of interest are studied and discussed below.

[1] In this paper, QoS-aware cloud based autonomic information system (Agri-Info) for agriculture service has been proposed which manages various types of agriculture related data based on different domains, through different user preconfigured devices. Agri-Info uses autonomic resource manager for efficient resource allocation at infrastructure level after identification of QoS requirements for user's request. We have evaluated the performance of the system in cloud environment using Cloud Sim toolkit and experimental results show that this system performs better in terms of execution time, cost, network bandwidth and latency. The application model of Agri-Info is built on top of Cloud Sim in order to validate the system through real-time mobile and web application (in other words, data from the experiment is directly fed into the simulator

to provide edge-device operational behavior for the resource manager). The case study of Agri-Info is implemented in an Indian village to evaluate the customer satisfaction amongst farmers.

Experimental results showed that Agri-Info delivers a superior autonomic solution for farmers and approximate optimum solution for challenges of resource management, using the concept of autonomic computing.

Existing agriculture systems are not able to fulfill the needs of today's generation due to missing of important requirements like processing speed, lesser data storage space or latency and even the resources used in computer-based agriculture systems are not utilized efficiently. To solve the problem of existing agriculture systems, there is a need to develop a cloud based autonomic information system which delivers Agriculture as a Service. In this section, we present an architecture of QoS-aware cloud based autonomic information system for agriculture service called Agri-Info, which manages various types of agriculture related data based on different domains.

The main objectives of this system are:

- to get information from various IoT devices
- to analyze the information using Fuzzy logic and create fuzzy rules
- to store the user data and fuzzy rules in cloud repository (Fuzzy RuleBase) for future decisions
- to respond the user queries automatically based on the information stored in Fuzzy Rule Base
- to allocate the resources automatically based on QoS requirements of different requests.

An Android based application for handheld devices was developed for easy access. Main functions of this mobile application are: finding productivity, posting queries, weather forecasting, getting information about latest equipment and getting alerts about agriculture news and finding crop information. It shows the process of finding productivity based on the values of other parameters selected by the user. The value of the productivity is 'high' or 'low' based on the factors determining productivity. It also shows the process of posting query and posts query with an image while showing the latest information about equipment and latest news alerts. The output of weather forecasting and crop information is also given.

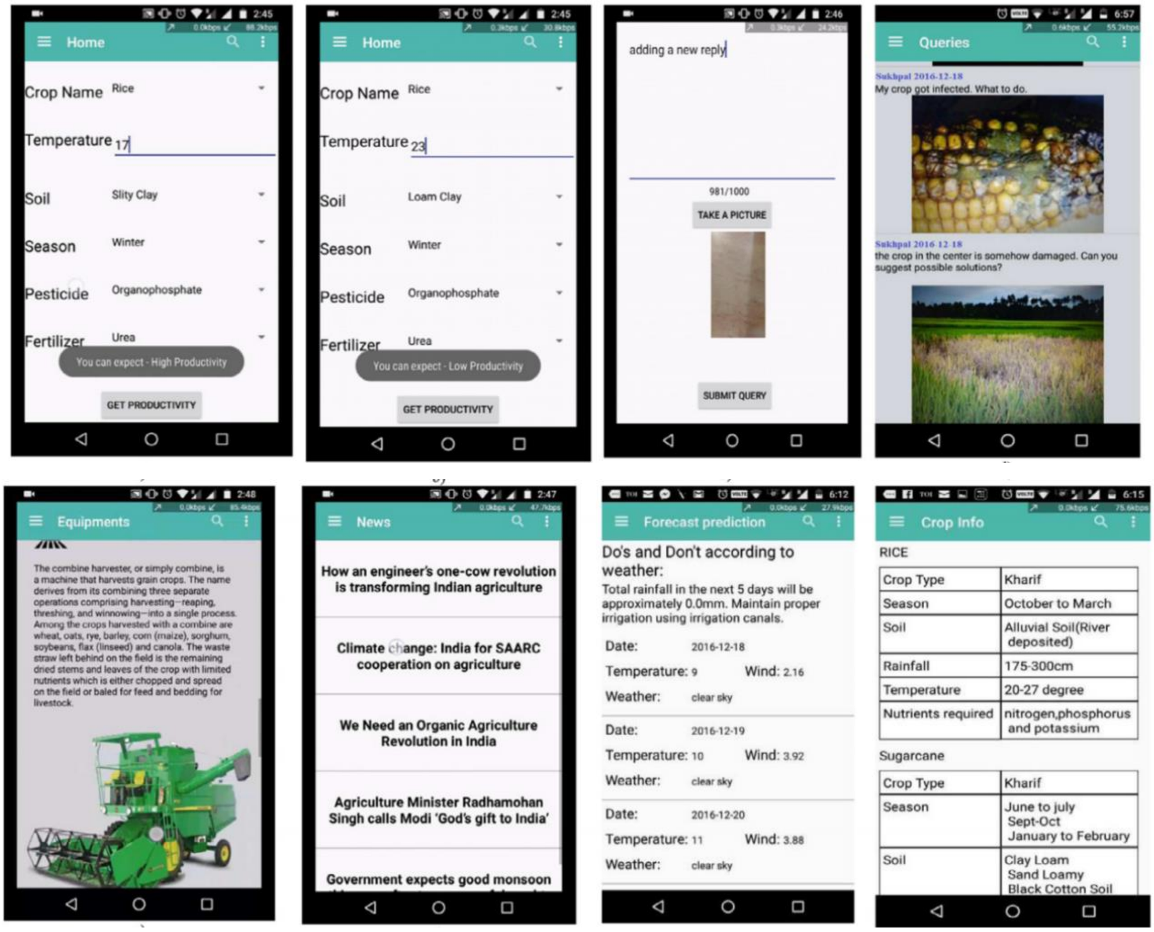

Fig 2.2.1: Main functions performed by the system

[2] This paper focuses on indulgence of cloud computing in agriculture along with IoT. Agriculture is becoming an important growing sector throughout the world due to increasing population. Major challenge in agriculture sector is to improve farm productivity and quality of farming without continuous manual monitoring to meet the rapidly growing demand for food. Apart from increasing population, the climate change is also a big concern in agricultural sector. Moreover, failures in current farming techniques bring in the need for more efficient and economical techniques to reach out farmers to a large extent. The purpose of this system is to propose a farming method, based Internet of Things (IoT) to deal with the adverse situations. The smart farming can be adopted which offer high precision crop control, collection of useful data and automated farming technique. This work presents an intelligent agriculture field monitoring system which monitors soil humidity and temperature. After processing the sensed data, it takes necessary action based on these values without human intervention.

Here, temperature and moisture of the soil are measured, and these sensed values are stored in ThingSpeak cloud for data analysis. In the most basic layer called sensor layer the sensors are connected to raspberry pi and are deployed in a farm. Then the middleware layer manages the hardware that is being deployed as a part of the system. The presence of Wi-Fi is a necessity as the live data has to be sent to the cloud continuously for the purpose of remote monitoring. In the top-most or the application layer which interacts with the user, the cloud retains all the data required by the farmer at a higher level and also certain graphical representations of the data are present.

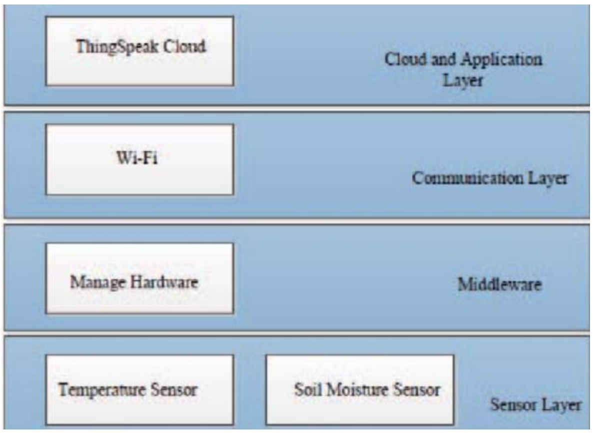

Fig 2.2.2: Layered architecture of the Intelligent IoT based system

[3] In this paper the three nutrients responsible for the pH of the soil are measured using fiber optic sensors to determine the value of NPK.A sensor can be made by integrating together the separate sensors available for the 3 nutrients, so that it can make better accurate results. The only factor that can change the yield of agriculture is soil nutrients. The most important factor that stipulates the condition of the soil is the pH. pH is the measure of the acid and base content of the soil. The methods to maintain soil pH by Nitrification and Liming. The main factor that changes the pH of the soil is the leaching happening to the soil. Due to the heavy rainfall, the leaching of the soil increases which results in the decrease of pH soil. The main nutrients that can increase the pH of soil is Nitrogen(N). Nitrogen is one of the main nutrients that has to be tested frequently to monitor the soil condition.

Plants cannot directly use nitrogen. The chemical nitrogen that we add convert the atmospheric nitrogen into the form that is usable by the plants. The overdose of nitrogen decreases the chlorophyll content and eventually the plant dries up. Next nutrient to be checked is phosphorus (P). It affects the yield of the plant. Similar to Nitrogen, excess amount of the Phosphorus dries up the plant. Potassium (K) in the soil is responsible for the overall development of the plant. Potassium deficiency affects the overall development of the plant its fruit, leaf, color, taste etc. Potassium helps in the intake of the CO2 from the soil. Moisture content of the soil is also an important factor that is needed to be monitored. This content changes with status of rain. The proposed model will:

1. Measure the above-mentioned nutrients' contents
2. Send the measured nutrient values to the nearest Crop Research Institute
3. The Institute would send back the recommendation which would fill the gap lacking in the soil.

Based on the measure of the nutrients individually they were divided into various classes of pH, N, P, K following the standard ranges shown in the below tables.

Table 2.2.3.1: pH dataset

Class No.	pH
0	Below 4
1	4.1-4.5
2	4.6-5.0
3	5.1-5.5
4	5.6-6.0
5	6.1-6.5
6	6.6-7.5
7	7.6-8.5
8	8.6-9.0
9	9.1-10.0

Table 2.2.3.2: Phosphorus dataset

Class No.	Available P(kg/ha)	Rating	
0	0.0-3.0	Low	128
1	3.1-6.5	Low	117
2	6.6-10.0	Low	106
3	10.1-13.5	Medium	94
4	13.6-17.0	Medium	83
5	17.1-20.5	Medium	71
6	20.6-24.0	Medium	60
7	24.1-27.4	High	48
8	27.6-31.0	High	37
9	31.1-34.0	High	25

Table 2.2.3.3: Nitrogen dataset

Available N(kg/ha)	Rating
280 kg/ha	Low
280-560 kg/ha	Medium
580 kg/ha	High

Table 2.2.3.4: Potassium dataset

Class No.	Available K(kg/ha)	Rating	
0	0.0-35.0	Low	128
1	36.0-75.0	Low	117
2	76.0-115.0	Low	106
3	116.0-155.0	Medium	94
4	156.0-195.0	Medium	83
5	196.0-235.0	Medium	71
6	236.0-275.0	Medium	60
7	276.0-315.0	High	48
8	316.0-355	High	37
9	356.0-395.0	High	25

Given these datasets, the values of N, P, K, pH from the soil are measured and is assigned the rating considering the obtained values and said what rate of sufficiency or deficiency is present and suggested required nutrients to be provided.

[4] In this paper, a system to test the soil fertility by using the principal of colorimetry is represented. Colorimetry is a technique in which we measure the amount of light absorbed by the color developed in the sample. An aqueous solution of the soil sample is prepared using extracting agents and is subjected to the photodiodes of the color sensor. The solution develops a color due to reaction of nutrients in the soil with chemicals. The output by the color sensor is calibrated with standard values present in the database. To verify the results obtained by the color sensor we use the Naive Bayes classification algorithm. This algorithm classifies the intensity values of the soil solutions into three class labels namely low, medium, high. After applying the Naive Bayes classifier, the accuracy of the system is predicted. The system is thus beneficial

to reduce the time required for testing the soil fertility and determining the accuracy of our results.

The steps can be broken into four major components namely soil sampling, soil pretreatment, actual measurement of NPK using color sensors and verification using Naïve Bayes classification algorithm.

- Soil sampling
- Soil pre-treatment
- Verification by Naive Bayes classification algorithm

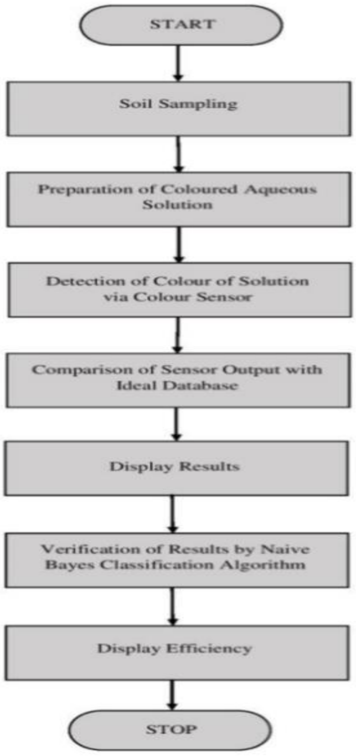

Fig 2.2.4: Flow diagram for determining soil fertility

The texture, color and structure of the soil samples are restricted to few categories. Also, the dataset used for the analysis of the proposed methodology was limited to the spatial soil characteristics of the region. For large soil datasets comprising of entire country, efficient methods could be created that utilize Data Mining/Big Data techniques to enhance the exactness of classification. Also, the system does not provide exact quantification of N, P, K contents in the soil.

[5] This paper presents an Internet of Things (IoT) based system by designing a novel Nitrogen-Phosphorus-Potassium (NPK) sensor with Light Dependent Resistor (LDR) and Light Emitting Diodes (LED). The principle of colorimetric is used to monitor and analyze the nutrients present in the soil. The data sensed by the designed NPK sensor from the selected agricultural fields are sent to Google cloud database to support fast retrieval of data. The concept of fuzzy logic is applied to detect the deficiency of nutrients from the sensed data. The crisp value of each sensed data is discriminated into five fuzzy values namely very low, low, medium, high and very high during fuzzification. A set of If-then rules are framed based on individual chemical solutions of Nitrogen (N), Phosphorous (P) and Potassium (K), Mamdani inference procedure is used to derive the conclusion about the deficiency of N.P and K available in soil chosen for testing and accordingly an alert message is sent to the farmer about the quantity of fertilizer to be used at regular intervals. The proposed hardware prototype and the software embedded in the microcontroller are developed in Raspberry pi 3 using Python. The developed model is tested in three different soil samples like red soil, mountain soil and desert soil. It is observed that the developed system results in linear variation with respect to the concentration of the soil solution. A sensor network scenario is created using Qualnet simulator to analyze the performance of designed NPK sensor in terms of throughput, end to end delay and jitter From the different variety of experiments conducted, it is noticed that the developed loT system is found to be helpful to the farmers for high yielding of crops.

The key features of this paper are:

- Design of novel NPK sensor by incorporating the colorimetric principle using Light Dependent Resistor (LDR) and Light Emitting Diodes (LEDs)
- Integration of NPK sensor with the software embedded in microcontroller unit for analyzing the sensed value
- Development of a fuzzy rule-based system to analyze the sensed value to decide upon the proportions of N, P and K in the soil at edge level.
- Internet-based user-friendly and automated intimation system at regular intervals of time to the farmers' electronics gadgets.

The data is sensed from the NPK sensor, then it is analyzed to find the level of nitrogen, phosphorous and potassium content in the soil. In general, there are two ways of analyzing the sensor data, (i) Fog Computing and (ii) Edge Computing. The former analyzes the sensed data using a processor that makes contact with the physical world, while the latter analyzes the sensor data over the internet at a centralized smaller edge server called as cloudlet. Since the data sensed by the proposed NPK sensor is the voltage based on the chemicals present in the soil solution, the analysis is performed at the edge level to determine the level of N, P and K in the soil.

Fuzzification partitions the input variable and assigning the linguistic label for each partition. In general, partitioning is carried out based on the range of the sensed data value with little overlap in the range of each linguistic label. The range of linguistic values assigned for 50gm of red soil dissolved in 100ml of water are as shown below. In this work, the crisp value of the NPK sensor data is divided into five linguistic labels namely very low, low, medium, high and very high.

Table 2.2.5: Fuzzification

NPK Sensor Value (Crisp Value)	**Fuzzy Range**	**Linguistics**
610-620	0-0.1	Very Low
620-625	0.1-0.3	Low
625-630	0.3-0.5	Medium
630-635	0.5-0.8	High
635-640	0.8-1	Very High

The SMS gateway is linked to receive the generated Message and forward to the farmers. Another optional way introduced in the application is to use HTTP to load the sensor data into Cloud server for future analytics of the soil nutrients. The sensed data is transferred to the Google cloud database to store the data for future analysis and to provide a monitoring of soil nutrients by the National or State level agriculture officials.

[6] This paper showed the design of a smart wireless sensor network (WSN) for an agricultural environment. The WSN is designed for supervising and controlling the variegated factors such as humidity, water level, temperature and human machination.

The WSN using ZigBee as the transmission medium consists of end devices with sensors, routers that propagate the network over larger distances, and a coordinator that communicates with the computer, which in turns illustrates the data and controls the entire

system. Sensors gather the various agricultural factors in real-time and transmit it using Internet of things, which collaborates with one another to perform action on behalf of people to reduce or eliminate the need of human labor. The agricultural monitoring system with wireless sensor networks is deployed at agricultural site and developed for agriculture monitoring. The base station will analyze the received data.

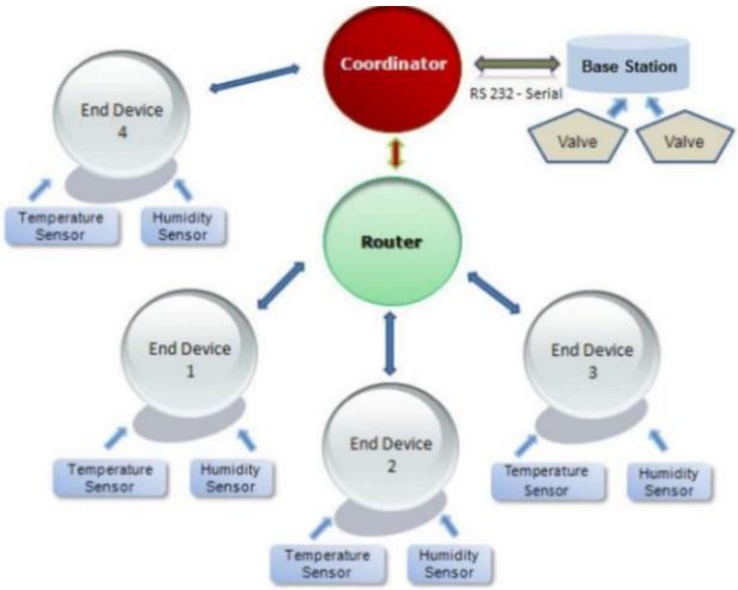

Fig 2.2.6: Toplogy of monitoring system

Linear network structure is a simply network can be used in rectangular greenhouses' digital agricultural data acquisition. The entire network, only a single path, determines the number of nodes equal to network layers or hop. Planar network structure such as large-scale farmland, pasture is complex, due to the multi-path network, scanning time analysis is also more complicated, this paper takes square instead of round for a brief analysis. Space network structure is more complex. It is conceivable that in the N-storey orchard, each storey with n-blocks of terrace set one node.

The automated agricultural system is used to proliferate the growth and flourishing of the cultivation crops. ZigBee based agriculture monitoring system as a reliable and efficient system for efficiently monitor the environmental conditions. In this paper, a wireless sensor network based on ZigBee that includes confirmation of valid data transmission from the base station is presented. At the same time, ZigBee network nodes deployment strategy is very important in the process of build digital agricultural data acquisition network.

The proposed system takes the features and ideas from the existing system while trying to overcome the defects and exemptions posed by them and minimizing cost constraints in all aspects and also increasing the efficiency in the current system. The above papers were analyzed and taken as references to build a new system having its data sent to the cloud enabling remote monitoring. In contrast to all these papers where required nutrients are provided based on the crop, here we select the crop based on the available resources and prevailing conditions. This adds value to the system providing it certain predictive power and a web application that effectively shows the collected data in various formats as required.

2.3 Tools Used

Hardware Requirements

- Raspberry PI 3
- DHT11 sensor
- Soil moisture sensor
- pH sensor

Software Requirements

- Raspbian OS
- ThingSpeak Cloud
- HTML/CSS

Languages Used

- Python

3.DESIGN OF THE PROPOSED SYSTEM

3.1 Block Diagram

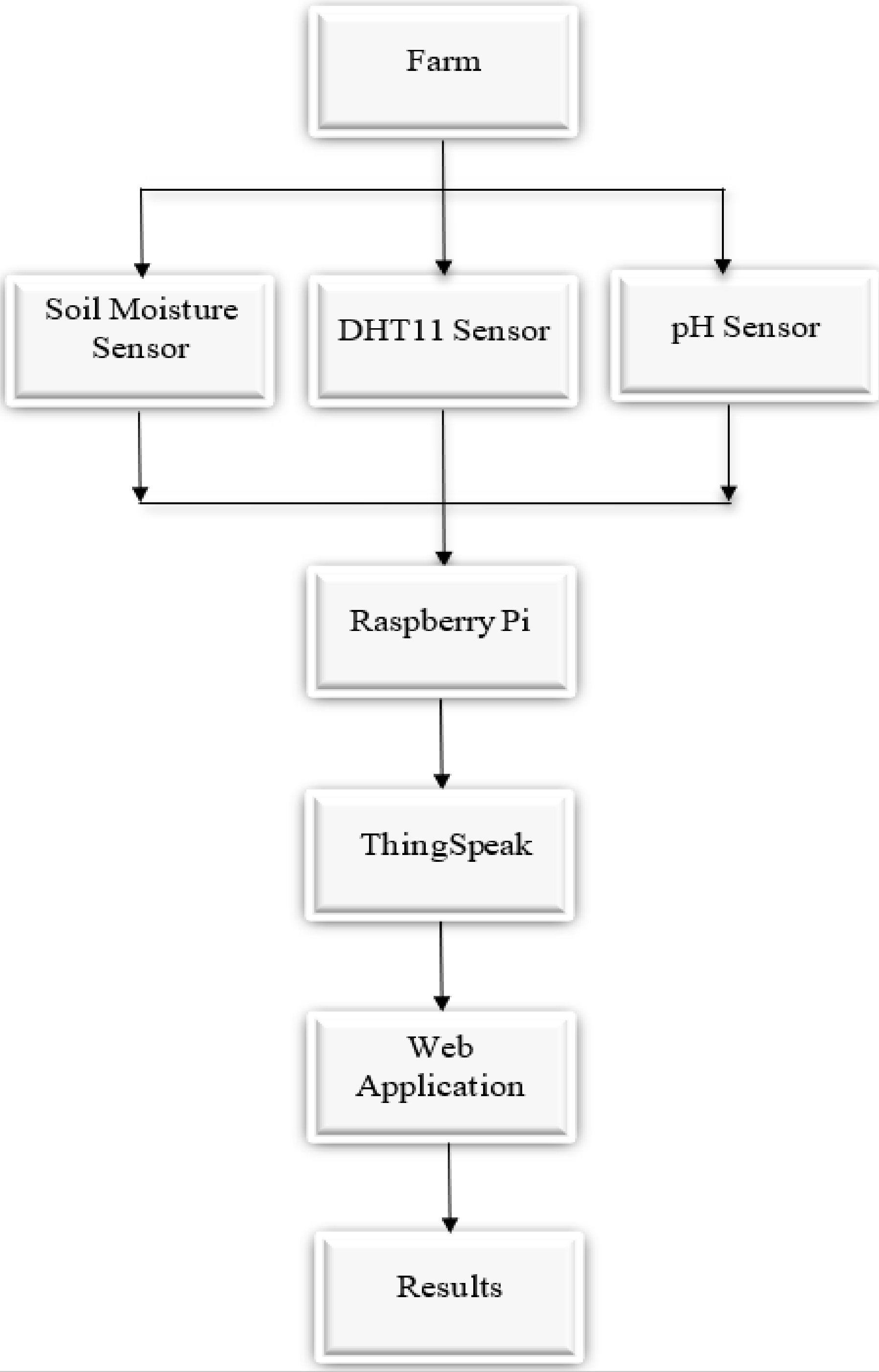

Fig 3.1 Block Diagram

The above block diagram gives description about the project. In the initial steps the required sensors are integrated to Raspberry Pi and are installed in the farm. The sensors used as a part of this project are DHT11 sensor (temperature and humidity), soil moisture sensor, pH sensor. Raspberry PI retrieves data from sensors and sends that data to cloud (ThingSpeak).

The data from sensors is measured against the standards set up for the farm and the farmer is given the privilege to monitor the farm periodically such that they get to know the changes in conditions and required measures can be taken.

Later live values will be retrieved from the cloud through web application and KNN algorithm is applied on them to predict the type of crop suitable for the farm.

3.2 Module Description

A modular approach towards any problem makes the solution simplified, organized and highly understandable such that it can be easily modified and adopted if needed. This project is divided into four significant modules.

- The first module deals with the connection of sensors with Raspberry Pi and make interactions with its components to work on it. The DHT11, soil moisture, pH sensors are connected to the Raspberry Pi and this in turn is connected to a monitor with a HDMI cable, this works as a means to display the actions performed on the microprocessor.
- In the second module, we embed our cloud credentials and requirements in the code by creating appropriate cloud channels and fields as required so that the generated data is accommodated in the respective fields in the cloud and this can be accessed through the cloud platform provided by thing speak.
- The third module is about building a model that has the capability to predict the crop type by applying KNN algorithm on the farm data values stored in the cloud against a predefined dataset. Once the analysis is performed the crop suitable for the soil is predicted and the model is first trained with the data we have and then the data we obtain through the farm is used for testing. The more data we get to collect, the more accurate the model will be.
- In the fourth module, we send the crop prediction result as a notification to the farmer through the mail.
- The visible module or the application is where we deploy a webpage and provide a privilege to the farmer to view his farm parameters at any time.

3.3 Theoretical Approach to the Prediction Module

For this system we propose to use the K- Nearest Neighbors algorithm for prediction of the crop type. K-Nearest Neighbor is a simple algorithm that stores all available cases and classifies new cases based on a similarity measure. The entire training dataset is stored. When a prediction is required, the k-most similar records to a new record from the training dataset are then located. From these neighbors, a summarized prediction is made. Similarity between records can be measured many different ways. A problem or data-specific method can be used. Generally, with tabular data, a good starting point is the Euclidean distance. KNN has been used in statistical estimation and pattern recognition already in the beginning of 1970's as a non-parametric technique.

Algorithm

***The k*-nearest neighbors algorithm** is a non-parametric method used for classification and regression. In both cases, the input consists of the *k* closest training examples in the feature space . The output depends on whether *k*-NN is used for classification or regression:

- In k-NN classification, the output is a class membership. An object is classified by a plurality vote of its neighbors, with the object being assigned to the class most common among its k nearest neighbors. If k = 1, then the object is simply assigned to the class of that single nearest neighbor.

- In k-NN regression, the output is the property value for the object. This value is the average of the values of k nearest neighbors.
- k-NN is a type of instance-based learning, or lazy learning, where the function is only approximated locally and all computation is deferred until function evaluation.
- Both for classification and regression, a useful technique can be to assign weights to the contributions of the neighbors, so that the nearer neighbors contribute more to the average than the more distant ones. For example, a common weighting scheme consists in giving each neighbor a weight of 1/d, where d is the distance to the neighbor.

- The neighbors are taken from a set of objects for which the class or the object property value is known. This can be thought of as the training set for the algorithm, though no explicit training step is required.

The training examples are vectors in a multidimensional feature space, each with a class label. The training phase of the algorithm consists only of storing the feature vectors and class labels of the training samples.

In the classification phase, k is a user-defined constant, and an unlabeled vector is classified by assigning the label which is most frequent among the k training samples nearest to that query point. A commonly used distance metric for continuous variables is Euclidean distance. For discrete variables, such as for text classification, another metric can be used, such as the overlap metric. In the context of gene expression microarray data, for example, k-NN has been employed with correlation coefficients, such as Pearson and Spearman, as a metric. Often, the classification accuracy of k-NN can be improved significantly if the distance metric is learned with specialized algorithms such as Large Margin Nearest Neighbor or Neighborhood components analysis.

The steps performed in this process are as followed:

1. The k-nearest neighbor algorithm is imported from the scikit-learn package.
2. Create feature and target variables.
3. Split data into training and test data.
4. Generate a k-NN model using neighbors' value.
5. Train or fit the data into the model.
6. Predict the future.

4.IMPLEMENTATION OF PROPOSED SYSTEM

4.1 Data Flow diagrams

A data flow diagram is a graphical representation of the "flow" of data through an information system, modelling its process aspects. A DFD is often used as a preliminary step to create an overview of the system without going into great detail, which can later be elaborated. DFD's can be used for the visualization of data processing. Data flow diagrams are also known as bubble charts. DFD is a designing tool used in the top-down approach to system design. This context-level DFD is next "exploded", to produce a level-1 DFD that shows some of the detail of the system being modelled. The Level-1 DFD shows how the system is divided into subsystems, each of which deals with one or more of the data flows to from an external agent, and which together provide all of the functionality of the system as a whole. It also identifies internal data stores that must be present in order for the system to do its job and shows the flow of data between the various parts of the system.

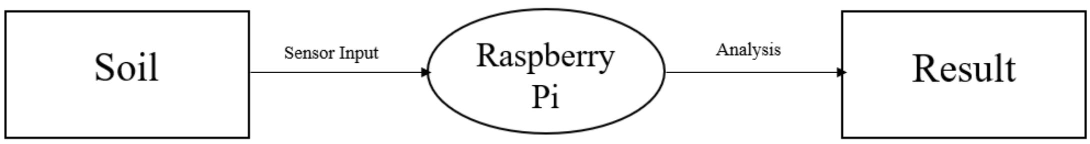

Fig 4.1.1: DFD Level-0

The DFD level-0 illustrates the flow of data from soil through sensors onto a Node MCU which are further analyzed to visualize output.

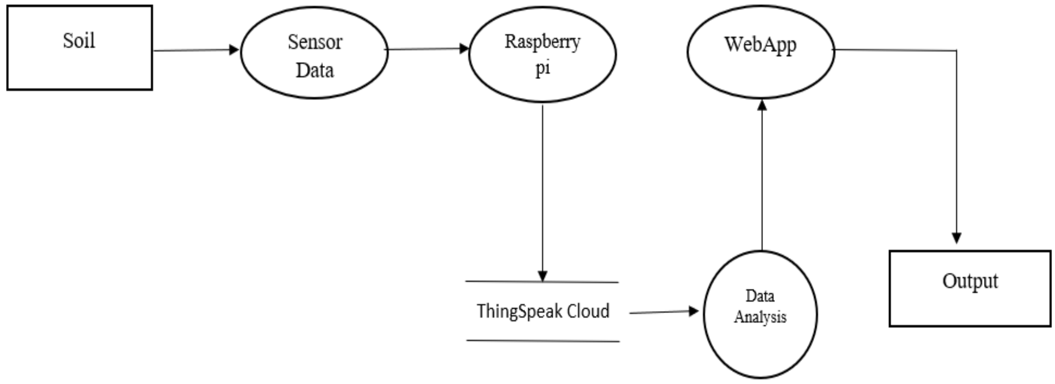

Fig 4.1.2: DFD Level-1

The DFD level-1 is the elaboration of DFD level-0 where the preprocessed data is uploaded on to a database which is analyzed using comparison algorithm and output is shown on a web-app.

4.2 UML Diagrams

UML stands for Unified Modelling Language. UML is a standardized general-purpose modelling language in the field of object-oriented software engineering. The standard is managed, and was created by, the Object Management Group.

The UML is a very important part of developing object-oriented software and the software development process. The UML uses mostly graphical notations to express the design of software projects.

Building blocks of the UML. The vocabulary of the UML encompasses three kinds of building blocks.

a. Things
b. Relations
c. Diagrams

Things in the UML:

Things are the abstractions that are first-class citizen in a model. There are four such things.

1. Structural Things
2. Grouping Things
3. Behavioral Things

These things are the basic object-oriented building blocks of the UML. You use them to write well-formed models.

Relationships in the UML

Things can be connected to logically be physically with the help of relationship in object-oriented modelling. These are four kinds of relationships in the UML.

1. Dependency
2. Association
3. Generalization
4. Realization

Diagrams in UML

A diagram is a graphical representation of a set of elements. These are nine kinds of diagrams in the UML.

a. Class diagram
b. ER diagram
c. Use Case diagram
d. Sequence diagram
e. Collaboration diagram

f. Activity diagram
g. Component diagram
h. State chart diagram

The diagrams of each of the above UML diagrams are depicted below.

4.2.1 Class Diagram

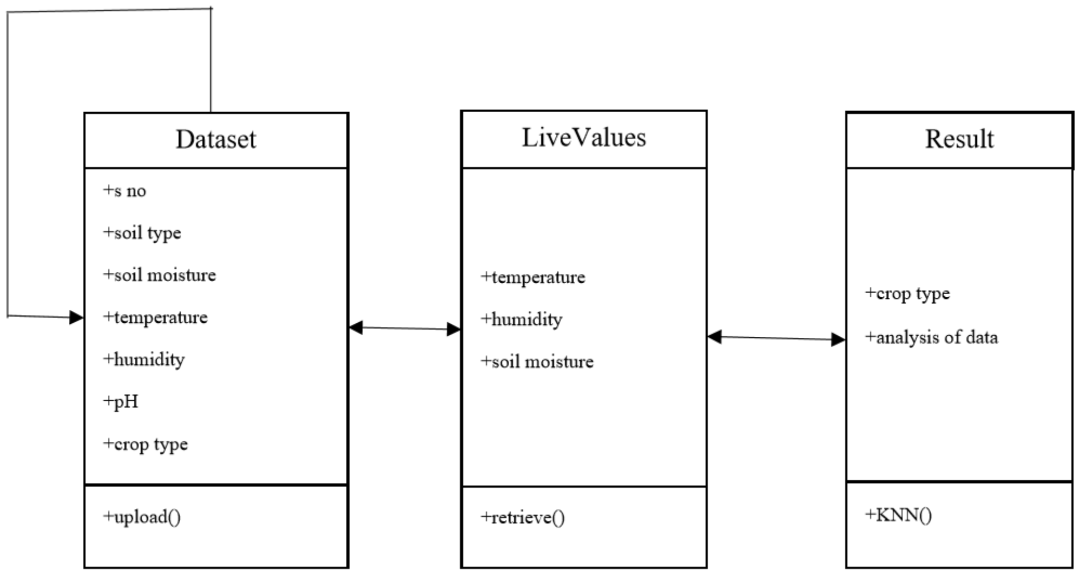

Fig 4.2.1: Class Diagram

The above class diagram shows how live values from sensors are collected and sent to cloud for analysis and sent to logistic regression algorithm for rain fall prediction.

4.2.2 Use Case Diagram

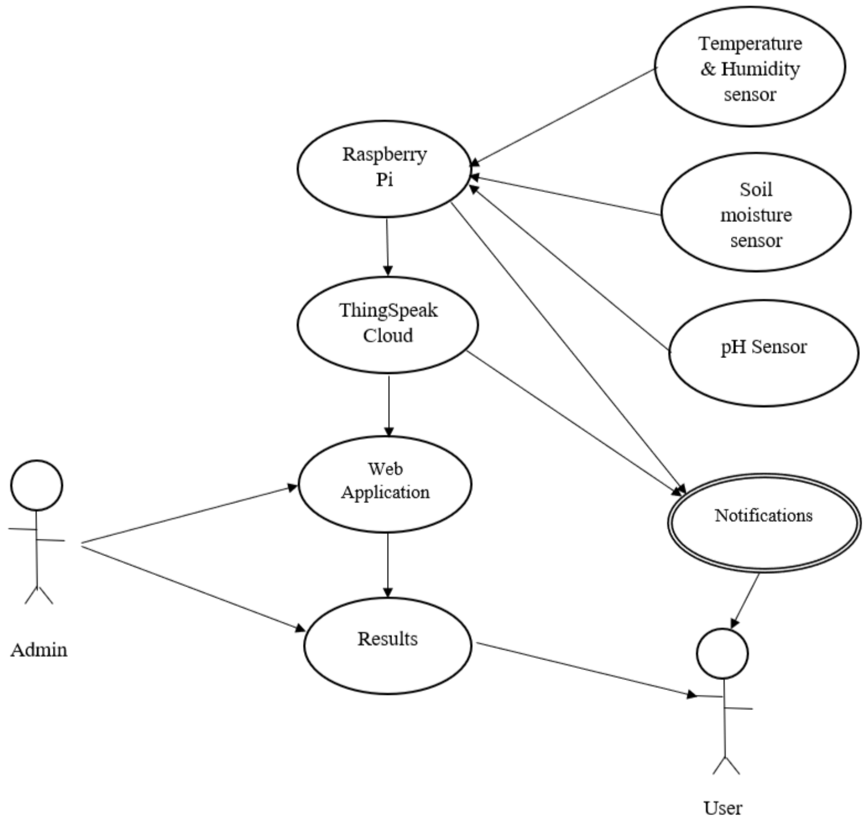

Fig 4.2.2: Use Case Diagram

The above Use Case diagram shows an idea of how the sensors take data and send back to Raspberry pi and how the admin has control over raspberry pi and cloud platform. After the analysis, the admin will send the result to the farmer through mail id.

4.2.3 State Chart Diagram

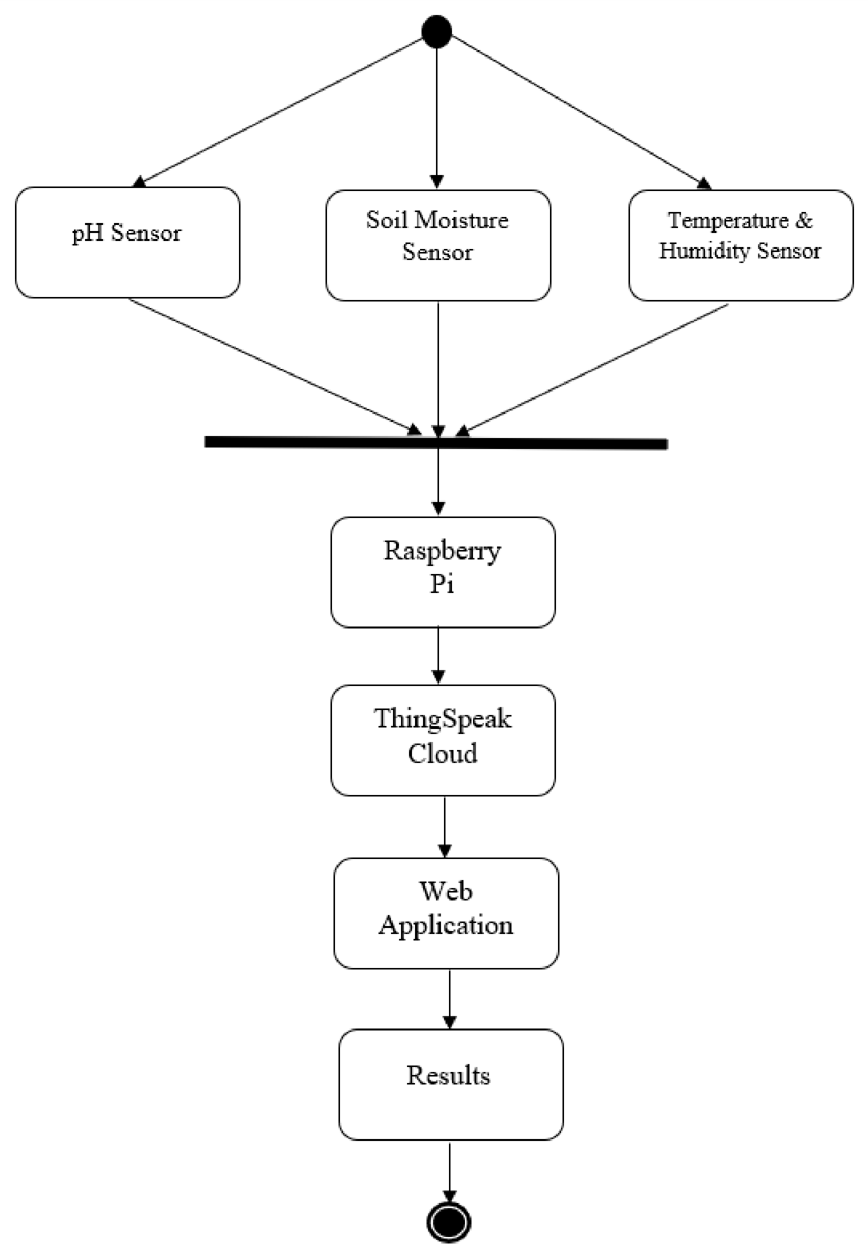

Fig 4.2.3: State Chart Diagram

4.3 Tools/Technologies used

4.3.1 Python Libraries

Python is one of the top four programming languages for building IoT solutions. Python is the official language to be used with Raspberry Pi. It is a wonderful and powerful programming language that's easy to use (easy to read and write) and, with Raspberry Pi, lets you connect your project to the real world. Python syntax is very clean, with an emphasis on readability, and uses standard English keywords. It is a lab-on-a-chip where it's more for educational use than anything. Here, we used certain python libraries to get our job done with the data processing, analysis etc., By default, Raspbian uses Python 2. However, versions 2 and 3 come installed by default. To open python in Raspberry Pi, we follow the following steps:

Step-1: Click the Raspbian logo.

Step-2: Navigate to Programming > Python 3 (IDLE).

Step-3: With the IDE loaded, click File > Open

Step-4: Navigate to the current Python program.

Step-5: When the file opens, we can run the program by clicking Run > Run Module or by pressing F5

on the keyboard.

4.3.2 Raspberry Pi

Raspberry Pi is a single board microcomputer which is a pretty useful device for embedded and IoT applications. The RaspberryPi3 has an on board wi-fi with which you can directly connect it to the internet for sending data (ex: sensor data) and receiving data (ex: your command via app or website) through the internet. It is a low cost, credit-card sized computer that plugs into a computer monitor or TV and uses a standard keyboard and mouse. It is a capable little device that enables people of all ages to explore computing, and to learn how to program in languages like Scratch and Python. It’s ideal and best suitable for projects where there is a computer requirement, but you don’t require much processing power, you want to keep the costs low and want to save on space.

Raspberry Pi has the ability to interact with the outside world and has been used in a wide array of digital maker projects, from music machines and parent detectors to weather stations and tweeting birdhouses with infra-red cameras.

The Raspberry Pi Foundation is a registered educational charity based in the UK. Our Foundation’s goal is to advance the education of adults and children, particularly in the field of computers, computer science and related subjects.

Models of Raspberry Pi

In the model B, you get an HDMI out, RCA video out, 2 USB ports, an SD card slot, a headphone jack, and an Ethernet port. The board itself has half a gigabyte of RAM and an onboard ARM processor. The model A has all of the same features of the model B minus one of the USB plugs, the Ethernet port, and half of the RAM.

Working of Raspberry Pi

- An SD card inserted into the slot on the board acts as the hard drive for the Raspberry Pi. It is powered by USB and the video output can be hooked up to a traditional RCA TV set, a more modern monitor, or even a TV using the HDMI port. This gives you all of the basic abilities of a normal computer.
- It also has an extremely low power consumption of about 3 watts. To put this power consumption in perspective, you could run over 30 Raspberry Pi’s in place of a standard light bulb.
- HDMI is the newest of the connectors out there and is great because it
- includes audio in the connection, unlike VGA and DVI.
- If you have an older TV that uses the standard yellow, red, and white RCA connections, then you will probably need that converter box. And to touch on audio, all you really need is a standard speaker extender that can go in the back of your computer or the headphone jack of a laptop.

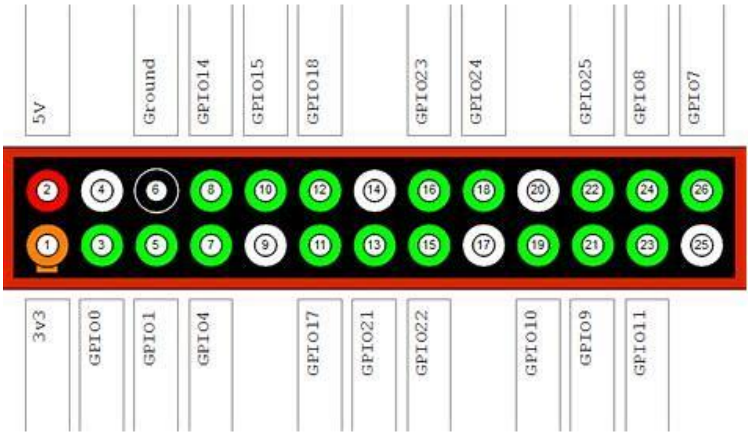

Fig 4.3.2.1: Raspberry Pi GPIO Pinout

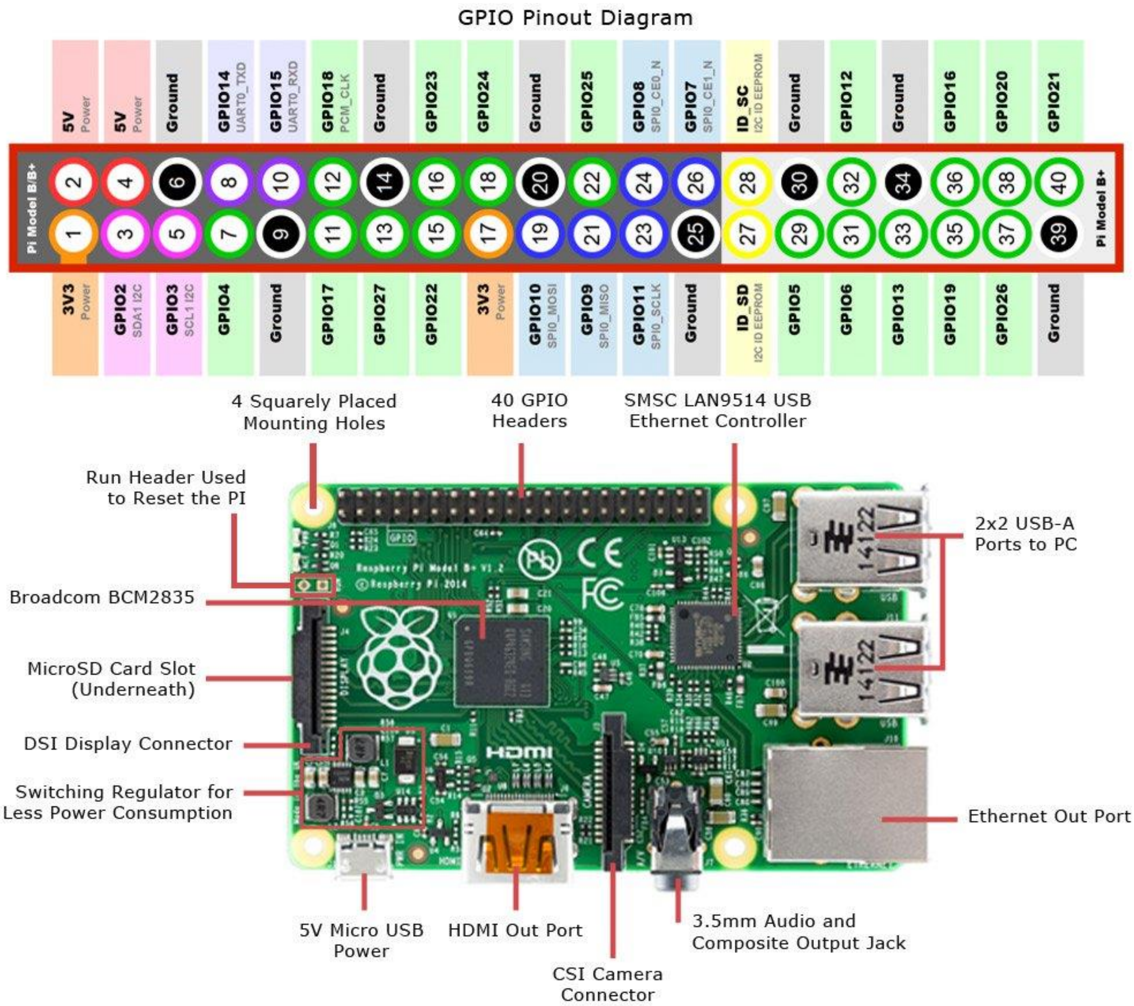

Fig 4.3.2.2: Raspberry Pi

- **Processor & RAM**: Raspberry Pi is based on an ARM processor. The latest version of Raspberry Pi (Mode B+) comes with boasting a 64-bit quad core processor running at 1.4GHz, dual-band 2.4GHz and 5GHz wireless LAN, Bluetooth 4.2/BLE, faster Ethernet, and PoE capability via a separate PoE HAT.
- **USB ports:** Raspberry Pi comes with two USB 2.0 ports. The USB on Raspberry Pi can provide a current up to 100µA. For connecting devices that draw current more than 100µA, an external USB powered hub is required.
- **Ethernet Ports:** Raspberry Pi comes with a standard RJ45 Ethernet port. You can connect Ethernet cable or a USB Wi-Fi adapter to provide Internet connectivity.
- **HDMI Output:** The HDMI port on Raspberry Pi provides both video and audio output. You can connect the Raspberry Pi to a monitor using an HDMI cable. For monitors that have a DVI port but no HDMI port, you can use an HDMI to DVI adapter/cable.
- **Composite Video Output**: Raspberry Pi comes with a composite video output with an RCA jack that supports both PAL and NTSC video output. The RCA jack can be used to connect old televisions that have an RCA input only.
- **Audio Output**: Raspberry Pi has a 3.5mm audio jack. This audio jack is used for providing audio output to old television along with the RCA jack for video. The audio quality from this jack is inferior to the HDMI output.
- **GPIO Pins:** Raspberry Pi comes with a number of general-purpose input/output pins. There are four types of pins on Raspberry Pi – true GPIO pins, I2C pins, SPI pins and serial RX and Tx pins.
- **Display Serial Interface (DSI):** The DSI interface can be used to connect an LCD panel to Raspberry Pi.
- **Camera Serial Interface (CSI):** The CSI interface can be used to connect a camera module to Raspberry Pi.
- **Status LED:** Raspberry Pi has five status LEDs.

Status LED	Function
ACT	SD card access
PWR	3.3V Power is present
FDX	Full duplex LAN connected
LNK	Link/Network activity
100	100 Mbit LAN connected

- **SD Card Slot**: Raspberry Pi does not have a built-in operating system and storage. You can plug-in an SD card loaded with a linux image to the SD card slot.
- **Power Input:** Raspberry Pi has a micro-USB connector for power input.
- **GPIO:** A powerful feature of the Raspberry Pi is the row of GPIO (general-purpose input/output) pins along the top edge of the board. A 40-pin GPIO header is found on all current Raspberry Pi boards (unpopulated on Pi Zero and Pi Zero W). Prior to the Pi 1 Model B+ (2014), boards comprised a shorter 26-pin header.
- **Voltages:** Two 5V pins and two 3V3 pins are present on the board, as well as a number of ground pins (0V), which are unconfigurable. The remaining pins are all general purpose 3V3 pins, meaning outputs are set to 3V3 and inputs are 3V3-tolerant.
- **Outputs**: A GPIO pin designated as an output pin can be set to high (3V3) or low (0V).
- **Inputs:** A GPIO pin designated as an input pin can be read as high (3V3) or low (0V). This is made easier with the use of internal pull-up or pull-down resistors. Pins GPIO2 and GPIO3 have fixed pull-up resistors, but for other pins this can be configured in software.
- As well as simple input and output devices, the GPIO pins can be used with a variety of alternative functions, some are available on all pins, others on specific pins.
- PWM (pulse-width modulation)

  Software PWM available on all pins.

Hardware PWM available on GPIO12, GPIO13, GPIO18, GPIO19.

- SPI

SPI0: MOSI (GPIO10); MISO (GPIO9); SCLK (GPIO11); CE0 (GPIO8), CE1 (GPIO7)

SPI1: MOSI (GPIO20); MISO (GPIO19); SCLK (GPIO21); CE0 (GPIO18);

CE1 (GPIO17); CE2 (GPIO16)

- I2C

Data: (GPIO2); Clock (GPIO3)

- EEPROM

EEPROM Data (GPIO0); EEPROM Clock (GPIO1)

- Serial

TX (GPIO14); RX (GPIO15)

Raspberry Pi Interfaces

- **Serial:** The serial interface on Raspberry Pi has receive (Rx) and transmit (Tx) pins for communication with serial peripherals.
- **SPI:** Serial Peripheral Interface (SPI) is a synchronous serial data protocol used for communicating with one or more peripheral devices. In an SPI connection, there is one master device and one or more peripheral devices.

There are five pins on Raspberry Pi for SPI interface:

i) MISO (Master in Slave Out): Master line for sending data to the peripherals.

ii) MOSI (Master Out Slave In): Slave line for sending data to the master.

iii) SCK (Serial Clock): Clock generated by master to synchronize data transmission

iv) CE0(Chip Enable 0): To enable or disable devices.

v) CE1(Chip Enable 1): To Enable or disable devices.

- **I2C:** The I2C interface pins on Raspberry Pi allow you to connect hardware modules. The I2C interface allows synchronous data transfer with just two pins –DA (data line) and SCL (clock line).

4.3.3 Sensors

DHT11 Temperature and Humidity sensor

The DHT11 is a basic, ultra, low-cost digital temperature and humidity sensor. It uses a capacitive humidity sensor and a thermistor to measure the surrounding air and spits out a digital signal on the data pin (no analog input pins needed). calculates relative humidity by measuring the electrical resistance between two electrodes. The humidity sensing component of the DHT11 is a moisture holding substrate with the electrodes applied to the surface. The change in resistance between the two electrodes is proportional to the relative humidity.

In the DHT11 sensor, work and storage conditions outside the sensor the proposed scope of work may lead to temporary drift of the signal up to 300%RH. return to normal working conditions, sensor calibration status will slowly toward recovery. To speed up the recovery process may refer to "resume processing". Avoid placing the components on the long-term condensation and dry environment, as well as the following environment. salt spray , acidic or oxidizing gases such as sulfur dioxide, hydrochloric acid Recommended storage environment Temperature: 10 ~ 40 degree centigrade Humidity: 60% RH or less .Therefore, in the measurement of humidity, should be to ensure that the work of the humidity sensor at the same temperature.

The temperature influences Relative humidity of the gas to a large extent dependent on temperature. Therefore, in the measurement of humidity, it should be to ensure that the work of the humidity sensor is at the same temperature. With the release of heat from electronic components that share a printed circuit board, the installation should be as far as possible from the sensor away from the electronic components and mounted below the heat source, while maintaining good ventilation of the enclosure.

To reduce the thermal conductivity sensor and printed circuit board copper plating should be the smallest possible, leaving a gap between the two. Light impact Prolonged exposure to sunlight or strong ultraviolet radiation, degrade performance. Resume processing Placed under extreme working conditions or chemical vapor sensor, which allows it to return to the status of calibration by the following handler. Maintain two hours in the humidity conditions of 45 degree centigrade and 70% RH humidity conditions to maintain more than five hours.

The impact of exposure to chemicals on the capacitive humidity sensor has a layer by chemical vapor interference, the proliferation of chemicals in the sensing layer may lead to drift and decreased sensitivity of the measured values.

In a pure environment, contaminants will slowly be released. Resume processing as described will accelerate this process. The high concentration of chemical pollution (such as ethanol) will lead to the complete damage of the sensitive layer of the sensor.

Soil Moisture sensor

The Soil Moisture Sensor uses capacitance to measure the water content of soil (by measuring the dielectric permittivity of the soil) which is a function of the water content. The sensor uses capacitance to measure dielectric permittivity of the surrounding medium. The sensor creates a voltage proportional to the dielectric permittivity, and therefore the water content of the soil. Simply insert this rugged sensor into the soil to be tested, and the volumetric water content of the soil is reported in percent.

Fig 4.3.3.1: Soil moisture sensor

pH sensor

A pH meter is a scientific instrument that measures the hydrogen-ion activity in water-based solutions, indicating its acidity or alkalinity expressed as pH. pH probes measure pH by measuring the voltage or potential difference of the solution in which it is dipped. Hence, a pH probe measures the potential difference generated by the solution by measuring the difference in hydrogen ion concentration using the Nernst equation and displays the pH as output. pH is a very important property to indicate if the medium is acidic, neutral or basic. It is proposed for the remote monitoring of the nutrient solution composition for the agricultural soil. pH is a key parameter for crop productivity. Every plant requires slightly acidic or neutral soil for efficient growth. A pH meter is a scientific instrument that measures the hydrogen-ion activity in water-based solutions, indicating its acidity or alkalinity expressed as pH. The pH meter measures the difference in electrical potential between a pH electrode and a reference electrode, and so the pH meter is sometimes referred to as a "potentiometric pH meter". The difference in electrical potential relates to the acidity or pH of the solution.The pH meter measures the difference in electrical potential between a pH electrode and a reference electrode, and so the pH meter is sometimes referred to as a "potentiometric pH meter".

Potentiometric pH meters measure the voltage between two electrodes and display the result converted into the corresponding pH value. They comprise a simple electronic amplifier and a pair of electrodes, or alternatively a combination electrode, and some form of display calibrated in pH units.

It usually has a glass electrode and a reference electrode, or a combination electrode. The electrodes, or probes, are inserted into the solution to be tested.

The design of the electrodes is the key part: These are rod-like structures usually made of glass, with a bulb containing the sensor at the bottom. The glass electrode for measuring the pH has a glass bulb specifically designed to be selective to hydrogen-ion concentration. On immersion in the solution to be tested, hydrogen ions in the test solution exchange for other positively charged ions on the glass bulb, creating an electrochemical potential across the bulb. The electronic amplifier detects the difference in electrical potential between the two electrodes generated in the measurement and converts the potential difference to pH units.

Fig 4.3.3.2: pH sensor

The reference electrode is insensitive to the pH of the solution, being composed of a metallic conductor, which connects to the display. This conductor is immersed in an electrolyte solution, typically potassium chloride, which comes into contact with the test solution through a porous ceramic membrane. The display consists of a voltmeter, which displays voltage in units of pH.

On immersion of the glass electrode and the reference electrode in the test solution, an electrical circuit is completed, in which there is a potential difference created and detected by the voltmeter. The circuit can be thought of as going from the conductive element of the reference electrode to the surrounding potassium-chloride solution, through the ceramic membrane to the test solution, the hydrogen-ion-selective glass of the glass electrode, to the solution inside the glass electrode, to the silver of the glass electrode, and finally the voltmeter of the display device.

The voltage varies from test solution to test solution depending on the potential difference created by the difference in hydrogen-ion concentrations on each side of the glass membrane between the test solution and the solution inside the glass electrode. All other potential differences in the circuit do not vary with pH and are corrected for by means of the calibration.

4.3.4 ThingSpeak Cloud

ThingSpeak is an open data platform for the Internet of Things that allows us to aggregate, visualize and analyze live data streams in the cloud. It enables sensors, instruments, and websites to send data to the cloud where it is stored in either a private or a public channel. It is an API to store and retrieve data from things using the HTTP and MQTT protocol over the Internet or via a Local Area Network. Our device or application can communicate with ThingSpeak using a RESTful API, and we can either keep our data private or make it public. In addition, we use the cloud to analyze and act on our data in order to perform online analysis and processing of the data as it comes.

Some of the key capabilities of ThingSpeak include the ability to:

- Easily configure devices to send data to ThingSpeak using popular IoT protocols.
- Visualize your sensor data in real-time.
- Aggregate data on-demand from third-party sources.
- Use the power of MATLAB to make sense of your IoT data.
- Run your IoT analytics automatically based on schedules or events.
- Prototype and build IoT systems without setting up servers or developing web software.
- Automatically act on your data and communicate using third-party services like Twilio® or Twitter®

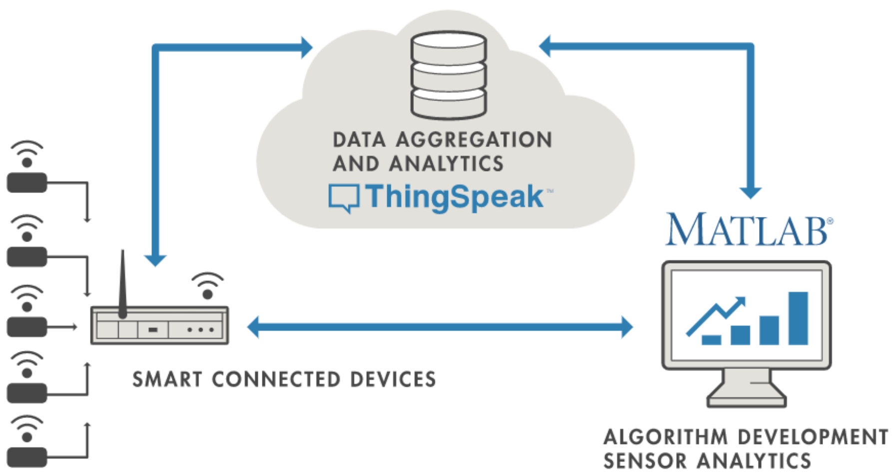

Fig 4.3.4: Thing Speak architecture

On the left, we have the smart devices (the "things" in IoT) that live at the edge of the network. These devices collect data and include things like wearable devices, wireless temperature sensors, heart rate monitors, and hydraulic pressure sensors, and machines on the factory floor.

In the middle, we have the cloud where data from many sources is aggregated and analyzed in real time, often by an IoT analytics platform designed for this purpose.

The right side of the diagram depicts the algorithm development associated with the IoT application. Here an engineer or data scientist tries to gain insight into the collected data by performing historical analysis on the data. In this case, the data is pulled from the IoT platform into a desktop software environment to enable the engineer or scientist to prototype algorithms that may eventually execute in the cloud or on the smart device itself. An IoT system includes all these elements. ThingSpeak fits in the cloud part of the diagram and provides a platform to quickly collect and analyze data from internet connected sensors.

4.4 System Hardware setup

Connection of Raspberry Pi with sensors

- The pH sensor, soil moisture sensor and DHT11 sensor are connected to Raspberry Pi through the GPIO pins. Male- male, female- female, male- female jumper wires are used to make the connections.
- This microprocessor-sensor setup in turn is connected to the monitor with a HDMI cable through the HDMI out port.
- The keyboard and mouse are connected to Raspberry Pi through the two USB-A ports. The SD Card for the Raspbian OS is inserted in the micro SD card slot.
- The internet or Wi-Fi connection is supplied to the microprocessor through Ethernet Out port. Finally, the power supply is provided with the USB charging cable through 5V Micro USB Power port.

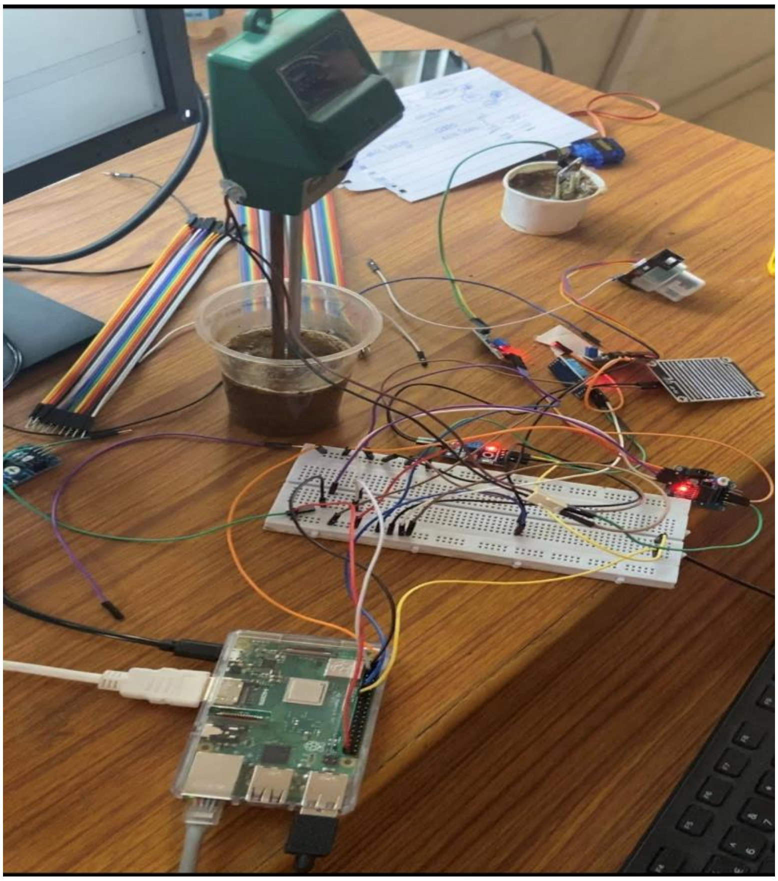

Fig 4.4: Hardware connection setup

4.5 Dataset Description

The dataset represents different values of six parameters.

- Soil type
- Soil moisture
- Temperature
- Humidity
- pH
- Crop type

Based on the above parameters we are set to determine the suitable crop type for a given farmland.

The soil moisture is the moisture content that is present inside the ground and determines how wet or dry the soil is. This is considered as a factor for prediction because some crops require much water content at the roots while the others may require less moisture or dry conditions. If the moisture content of a soil is optimum for plant growth, plants can readily absorb soil water. Much of water remains in the soil as a thin film. Soil water dissolves salts and makes up the soil solution, which is important as medium for supply of nutrients to growing plants. So, based on the moisture levels of the soil a suitable crop is chosen.

Humidity greatly affects your plant's growth in such a way that it prevents plants from experiencing water evaporation. Keeping the humidity in the environment high allows the stomata (the pores in a plant that absorb water vapor in the air and enable gas exchange) to be open, hence keeping the photosynthesis working by absorbing CO2. In the case of Anthurium, good humidity around the plant is even more important than for most other crops, because the plant can only absorb a reduced amount of humidity and hence has less water evaporation than most plants. Most plants need humid air in order to thrive. That's because the pores through which they breathe lose most of their moisture when the surrounding air is dry, a loss that the plant can't always replace through the water its roots absorb.

Temperature is another factor that determines plant growth and development. Along with the levels of light, carbon dioxide, air humidity, water and nutrients, temperature influences plant growth and ultimately crop yields. Water temperature for your plants. When watering your plants, it is essential to use water at the right temperature. This is because the roots of your plants are very sensitive to extremes of temperature. Using water that is too hot or too cold can put your plant under stress and cause damage.

Soils are often described as acidic or alkaline. Unfortunately, the meanings of these terms and their relationship to plant growth are not clearly understood by some gardeners. The relative acidity or alkalinity of soil is indicated by its pH. The pH scale runs from 0 to 14. Any pH reading below 7 is acidic and any pH above 7 is alkaline. A pH of 7 indicates a neutral soil. The pH is important because it influences the availability of essential nutrients. Most horticultural crops will grow satisfactorily in soils having a pH between 6 and 7.5. Since most garden soils are in this range, most gardeners experience few problems with soil pH.

Parameter	Highly Suitable (S1)	Moderately suitable (S2)	Marginally suitable (S3)	Not suitable (N)
Climate				
Temperature (°C) growing season	25-32	33-35	36-38	>38
Total rainfall (mm)	750-900	900-1200	500-600	
Land				
Length growing Period	>150	120-150	90-120	<90
Soil Drainage(class)	well drained	Moderate	poor	verypoor
Soil				
Texture (class)	l,scl,cl,sil.	Sl,sc,sic.	C(ss),ls,s.	
pH	6-7	7.1-8.0	8.1-9.0	>9
OC	>0.75	0.5-0.75	<0.5	
Effective soil depth	>75	50-75	25-50	<25
Coarse fragments	<15	>15-35	>35	
Salinity EC(dS m^{-1})	Non saline	1-2	2-4	<4
Sodicity ESP (%)	Non sodic	5-10	10-15	
slope	<3	3-5	5-10	

Parameter	Highly Suitable (S1)	Moderately suitable (S2)	Marginally suitable (S3)	Not suitable (N)
Climate				
Temperature (°C) growing season	30-34	35-38 21-29	39-40 15-20	<40 >15
Total rainfall (mm)	1110-1250	900-1110	750-900	<750
Land				
Slope (%)	0-1	1-3	3-5	>5
Drainage (class)	Imperfectly drained	Moderately drained	well drained	excessively drained
Free from flooding (months)	>4	3-4	2-3	-
Depth of water table (cm)	<10	10-20	>20-40	>40
Soil				
Texture (class)	c, sic, cl, sicl, sc	scl, sil,l.	sl, ls	s
Depth (cm)	>75	51-75	25-50	<25
$CaCO_3$ (%)	<10	10-25	25-40	>40
Soil fertility				
pH	5.5-6.5	6.4-7.5 4.5-5.4	7.6-8.5	>8.5 <4.5
Salinity/Sodicity				
EC (dS m^{-1})	<3	3-6	6-10	>10
ESP (%)	<15	15-40	40-50	>50

Fig 4.5.1: Suitability criteria for chilli   Fig 4.5.2: Suitability criteria for paddy

The suitable conditions for the growth and produce of the crops chilli and paddy are shown in the above image. It gives the matching criteria for the soil and crop based on the conditions prevailing.

The suitable conditions for the growth of cotton are mentioned in the below image.

Soil-site characteristics			Rating			
		Unit	Highly suitable S1	Moderately suitable S2	Marginally suitable S3	Not suitable N
Climatic regime	Mean temperature in growing season	°C	16-25	26-30 13-15	31-32 10-12	>32 <10
Land quality	Land characteristics					
Oxygen availability to roots	Soil drainage	Class	Well drained	Moderately/ Imperfectly	Poorly drained	Very poorly drained
Nutrient availability	Texture- surface	Class	sl, l, ls	s, scl	sil, cl	Heavy c
	Sub-surface texture	class	scl, sil	s, sil	s	Heavy c
	pH	1:2.5	5.5-6.5	6.6-8.2 5.0-5.4	>8.2 <5.0	
	CEC	C mol (p+)/kg	>16	<16	<5	
	OC	%	High	Medium	Low	
Rooting conditions	Effective soil depth	cm	75-100	50-75	25-50	<25
	Stonnines	%	0-10	10-15	15-35	>35
Soil toxicity	Salinity (EC saturation extract)	dS/m	>16	<16		
	Sodicity (ESP)	%	Non Sodic	10-15	>15	
Erosion hazard	Slope	Hills %	<5	5-10	10-15	>15
		Plains %	<3	3-5	5-8	>8

Fig 4.5.3: Suitability criteria for cotton

Soil-site characteristics			Rating			
		Unit	Highly suitable S1	Moderately suitable S2	Marginally suitable S3	Not suitable N
Climatic regime	Mean temperature in growing season	°C	24-33	22-23 34-35	20-21 36-40	<20 >40
	Total rainfall	mm	750-1000	500-750	250-500	<250
Land quality	Land characteristics					
Moisture availability	Length of growing period	Days	>110	90-110	60-90	<60
Oxygen availability to roots	Soil drainage	Class	Well drained/ mod. Well drained	Imperfectly drained	Poorly drained	V.Poorly
Nutrient availability	Texture	Class	sl, l, scl, cl, sc	sic, sicl c, ls	Heavy clays (ss), s	-
	pH	1:2.5	6.0-8.0	8.1-8.5 5.5-5.9	8.6-9.5 4.0-5.4	>9.5 <4.0
	$CaCO_3$ in root zone	%	<15	15-25	25-30	>30
Rooting conditions	Effective soil depth	cm	>75	50-75	25-50	<25
Soil toxicity	Salinity (EC saturation extract)	dS/m	<1.0	1.0-2.0	2.0-4.0	>4
	Sodicity (ESP)	%	<10	10-15	15-20	>20
Erosion hazard	Slope	%	<5	5-10	10-15	>15

Fig 4.5.4: Suitability Criteria for wheat

The conditions required by the crop wheat is mentioned in the above image.

The characteristics of crop growth for Corn is detailed in the image below.

Soil-site characteristics			Rating			
		Unit	Highly suitable S1	Moderately suitable S2	Marginally suitable S3	Not suitable N
Climatic regime	Mean temperature in growing season	°C	20-35	18-19 36-40	15-17 41-45	<15 >45
	Total rainfall	mm	750-900	600-750	500-600	<500
Land quality	Land characteristics					
Moisture availability	Length of growing period	Days	>120	90-120	70-90	<70
Oxygen availability to roots	Soil drainage	Class	Well drained/ Moderately well drained	Imperfectly drained	Poorly drained	Very poorly drained
Nutrient availability	Texture	Class	l, sl, scl, cl, sc	ls, sic, sicl c,ls	Heavy clays (>60%)	-
	pH	1:2.5	6.0-8.5	8.5-9.0 5.5-5.9	9.1-9.5 5.0-5.4	>9.5 <5.0
	$CaCO_3$ in root zone	%	<15	15-25	25-30	>30
Rooting conditions	Effective soil depth	cm	>75	50-75	25-50	<24
Soil toxicity	Salinity (EC saturation extract)	dS/m	<1.0	1.0-2.0	2.0-4.0	>4
	Sodicity (ESP)	%	<10	10-15	15-20	>20
Erosion hazard	Slope	%	<5	5-10	10-15	>15

Fig 4.5.5: Suitability criteria for Corn

Soil site characteristics	Degree of limitation				
	0(None)	1(Slight)	2(Moderate)	3(Severe)	4(Very severe)
Rainfall (mm)	>1000	750-1000	500-750	<500	
Temp. (°C) during growing period	10 to 18	18-22	22-26	26-29	>29
Slope (%) -Plain irrigated	<1	1 to 3	3 to 5	5 to 10	>10
-Hilly unirrigated	<3	3 to 8	8 to 15	15-25	
Drainage	Well	Moderately well	Imperfect	Poor	Very poor
Flood hazards	Nil	Slight	Moderate	Severe	Very severe
Erosion	None	Slight	Moderate	Severe	Very severe
Soil depth (cm)	>80	50-80	30-50	15-30	<15
Soil texture	sil,l,sicl,fsl,cl	scl,sc,sl	sic,c,ls	fs	s
Surface stoniness (%)	<3	3 to 15	15-40	>40	
Relief	Normal	Flat	Concave	Concave	
NPK rating	HHH	MMM	MLL	LLL	
pH	5.8-6.5	6.5-7.0	7-7.5	>7.5	
Base saturation (%)	>70	50-70	35-50	<35	

Fig 4.5.6: Suitability criteria for Sunflower

The growth parameters required by a sunflower crop are mentioned above.

The parameters as defined by the crop institute for the growth of the crop coconut are shown below.

Soil site characteristics	Degree of limitation				
	0(None)	1(Slight)	2(Moderate)	3(Severe)	4(Very severe)
Rainfall (mm)	>1000	750-1000	500-750	<500	
Temp. (°C) during growing period	18.5-26	26-30	30-33	33-40	>40
Slope (%) -Plain irrigated	<1	1 to 3	3 to 5	5 to 10	>10
-Hilly unirrigated	<3	3 to 8	8 to 15	15-25	
Drainage	well	Moderately well	Imperfect	Poor	Very poor
Flood hazards	Nil	Slight	Moderate	Severe	Very severe
Erosion	None	Slight	Moderate	Severe	Very severe
Soil depth (cm)	>100	80-100	50-80	30-50	<30
Soil texture	sil,l,sicl,fsl,cl	scl,sc,sl	sic,c,ls	fs	s
Surface stoniness (%)	<3	3 to 15	15-40	>40	
Relief	Normal	Flat	Concave	Concave	
NPK rating	HHH	MMM	MLL	LLL	
pH	5.5-6.5	6.5-7.5	7.5-8.0	>8	
Base saturation (%)	>70	50-70	35-50	<35	

Fig 4.5.7: Suitability criteria for Coconut

Table 4.5.1: Dataset

Sno	soil type	soil moistu	temperatu	humidity	ph	croptype
1	clay	60	25	80	6	garlic
2	sandy	50	5	70	7	sunflower
3	clay	70	15	75	8.5	rice
4	alluvial	60	26	60	7	garlic
5	sandyloam	70	28	75	7	rice
6	clayloam	50	12	65	6	sunflower
7	sandyloam	55	7	78	6	cotton
8	sandyloam	70	0	68	6.5	rice
9	sandy	60	15	75	7.2	garlic
10	clay	50	7	60	5.5	tamarind
11	compost	50	27	50	6	tamarind
12	rocksandy	60	12	45	9	garlic
13	clayloam	65	20	55	7.6	chilli
14	acidic	40	24	45	6	corn
15	siltyloamsc	55	10	60	7	cotton
16	potting	50	21	65	5.5	tamarind
17	siltyloamsc	55	30	50	7	cotton
18	welldraine	60	25	65	6	garlic
19	loamysoil	70	17	75	7	rice
20	sandyloam	75	15	60	6	custard apple
21	sandyloam	75	-13	50	5	custard apple
22	clayloam	60	28	60	7	garlic
23	sandy	75	20	50	8	custard apple
24	clayloam	70	18	65	6	rice
25	welldraine	50	12	45	4	tamarind
26	sandy	80	30	65	6	coconut
27	clayloam	70	47	55	5	rice
28	welldraine	50	10	60	5.5	tamarind
29	lateritesoil	75	8	15	5	groundnut
30	sandyloam	65	3	70	5	chilli
31	welldraine	55	10	35	6	cotton
32	gardensoil	50	24	60	6.5	tamarind
33	compost	60	23	70	6.5	ginger
34	sandy	50	15	8	7	tamarind
35	deepsandy	70	40	70	5.5	rice
36	sandyloam	80	10	65	5	coconut
37	alluvial	70	15	80	4.5	rice
38	clayloam	60	30	60	7	ginger
39	saline	75	35	15	5.5	groundnut
40	sandy	65	18	35	6.5	chilli
41	sandyloam	60	21	55	7	ginger
42	welldraine	40	10	60	5.5	sunflower
43	loamysoil	65	5	60	6	chilli
44	moistrichs	80	10	70	6	coconut
45	sandy	80	20	70	6	coconut
46	welldraine	40	35	50	8	sunflower

47	sandy	75	22	60	7	groundnut
48	compost	65	13	60	6	chilli
49	sandyloam	60	13	60	5	ginger
50	clay	55	-20	50	6.5	cotton
51	loamysoil	65	20	60	5	rice
52	redlaterite	70	38	70	6	rice
53	lateritesoil	70	13	80	5.2	rice
54	loamysoil	55	9	50	6	garlic
55	loamysoil	20	15	70	6	corn
56	blacksoil	60	21	75	5.5	ginger
57	sandyloam	60	4.4	70	6.5	ginger
58	sandyloam	50	24	60	6	mustard
59	sandyloam	70	25	50	6.2	rice
60	loamysoil	80	13	50	6.5	orange
61	sandy	50	10	50	6	mustard
62	welldraine	60	18	75	6.5	wheat
63	sandyloam	70	25	70	6	banana
64	clayloam	60	20	65	5.5	wheat
65	sandyloam	80	15	50	5.8	orange
66	lightsandyt	70	10	60	6.5	banana
67	saline	60	15	70	6.5	wheat
68	heavyclay	50	25	50	5.5	mustard
69	loamysoil	12	20	250	6.3	corn
70	sandyloam	50	36	500	5.5	mustard
71	clayloam	60	35	60	5	wheat
72	sandy	70	25	70	6	banana
73	sandy	50	21	50	6.3	mustard
74	loamysoil	60	45	70	4.8	wheat
75	sandyloam	75	18	80	6	groundnut
76	clayloam	60	18	80	6.2	wheat
77	sandy	50	19	50	5.5	mustard
78	sandy	70	20	80	6	banana
79	sandy	50	25	70	6.5	mustard
80	sandyloam	75	28	39	6	groundnut
81	deppresse	70	20	50	8	banana
82	blacksoil	50	6	30	7	cotton
83	sandyloam	60	20	50	7	wheat
84	alluvial	40	25	40	6	sunflower
85	acidic	10	8	60	6	corn
86	clayloam	60	20	50	4.5	wheat
87	heavyclay	8	25	70	6.5	corn
88	alluvial	15	10	50	5.5	corn
89	loamysoil	60	20	70	6	chilli
90	sandy	60	27	80	6.7	chilli
91	slightlyacid	50	18	50	6.5	cotton
92	redsoil	60	30	60	7	chilli
93	blacksoil	45	21	45	21	sunflower

94	blacksoil	75	24	70	6 coconut
95	redsoil	50	20	50	4 cotton

In the above mentioned dataset, as one can notice temperatues ranging from as high as 47C to as low as -20C i.e., almost freezing temperature, difeerent crops that can be grown under these temperatures are collected from the nearest crop research institutes in Hyderabad based on which the above dataset was generated. The data contains the crop types that are staple to the states of Telangana and Andhra Pradesh and a few places in India. To justify the extremely low temperatures, the conditions were created in a green house.

Fig 4.5.8: Greenhouse internal view

Fig 4.5.9: Greenhouse external view

A greenhouse is a structure with walls and roof made chiefly of transparent material, such as glass, in which plants requiring regulated climatic conditions are grown. These structures range in size from small sheds to industrial-sized buildings. A miniature greenhouse is known as a cold frame. The interior of a greenhouse exposed to sunlight becomes significantly warmer than the external temperature, protecting its contents in cold weather. Many commercial glass greenhouses or hothouses are high tech production facilities for vegetables, flowers or fruits. The glass greenhouses are filled with equipment including screening installations, heating, cooling, lighting, and may be controlled by a computer to optimize conditions for plant growth. Different techniques are then used to evaluate optimality-degrees and comfort ratio of greenhouseate i.e., air temperature, relative humidity and vapor pressure deficit, in order to reduce production risk prior to cultivation of a specific crop.

4.6 Cloud Setup

ThingSpeak is an IoT analytics platform service that allows us to aggregate, visualize, and analyze live data streams in the cloud. We can send data to ThingSpeak from devices, create instant visualization of live data, and send alerts. To induce Thing speak into our system, we create required channels separately for storing data related to temperature, humidity, soil moisture. Once we decide upon the required channels, thing speak generates a unique API key and this is embedded in our code in order to access the cloud from Raspberry Pi.

Channel Setup

- A new channel is created for monitoring temperature and humidity and another channel for soil moisture.
- For the creation of channel certain details like name of the channel and description of the channel must be given appropriately.
- A channel has at most eight fields which are the attributes we wish to have in that channel. Here we take two fields, temperature and humidity.
- The other details can be given if required, else can be ignored as we do here.
- Then the channel is saved. All the channels we create this way are reflected in my channels section.

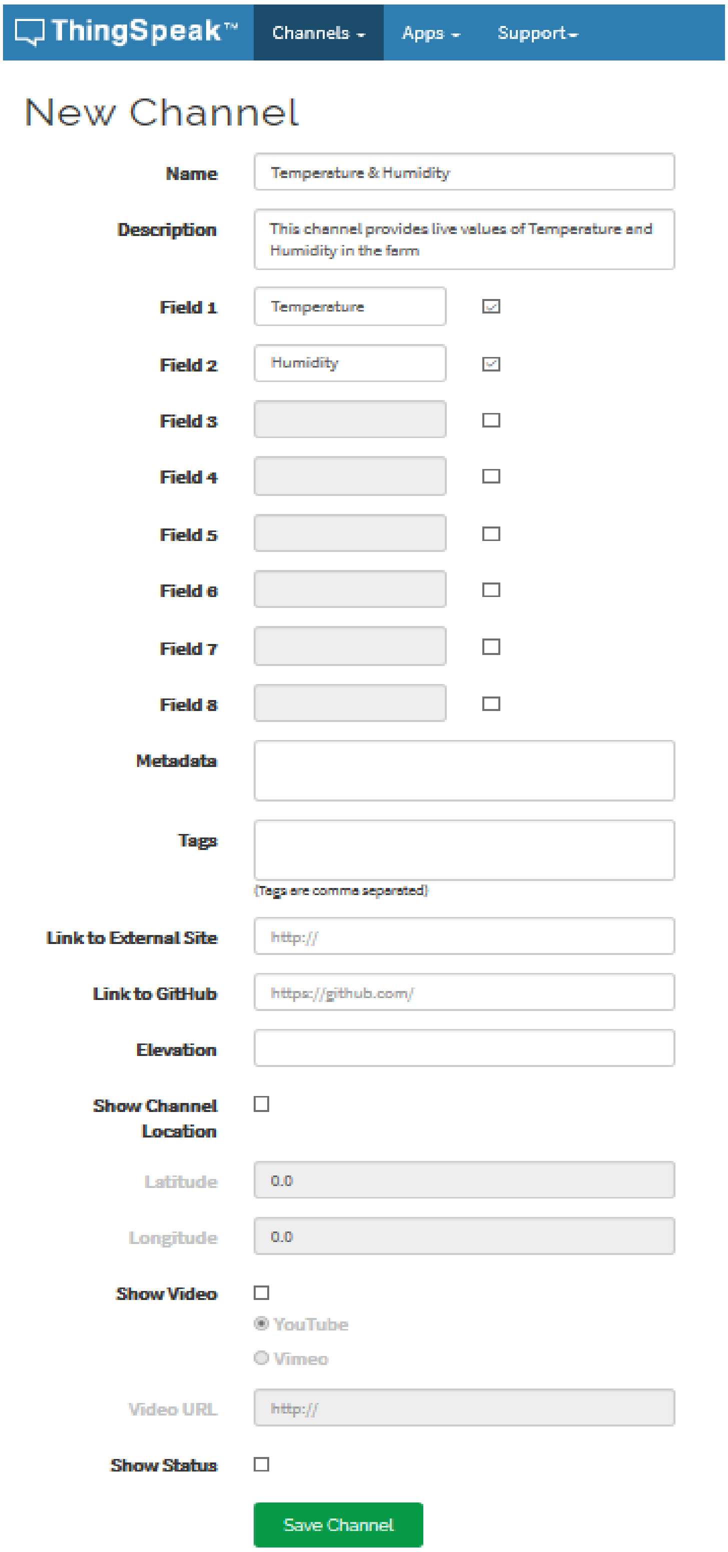

Fig 4.6.1: Creating a new channel on Thing Speak

Here, in the first phase we create a new channel with the appropriate name and decription that suits and best defines our requirements.

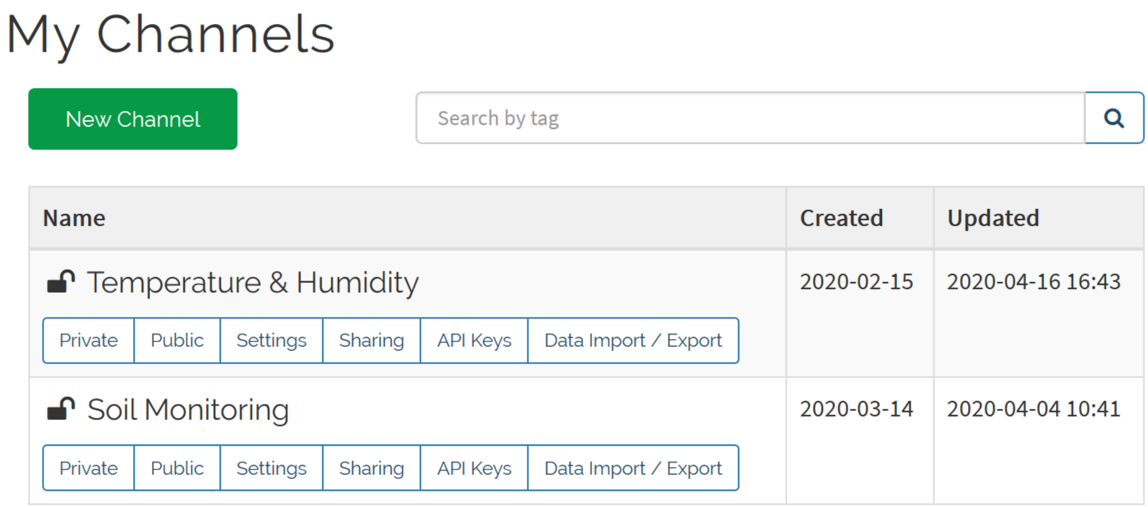

Fig 4.6.2: Channels in Thingspeak used in this system

Once we create the channels, we can view them under My Channels on the Thingspeak platform.

API Key Generation

For each channel we create API keys are generated and we use these API keys in our code to connect the Raspberry Pi to cloud.

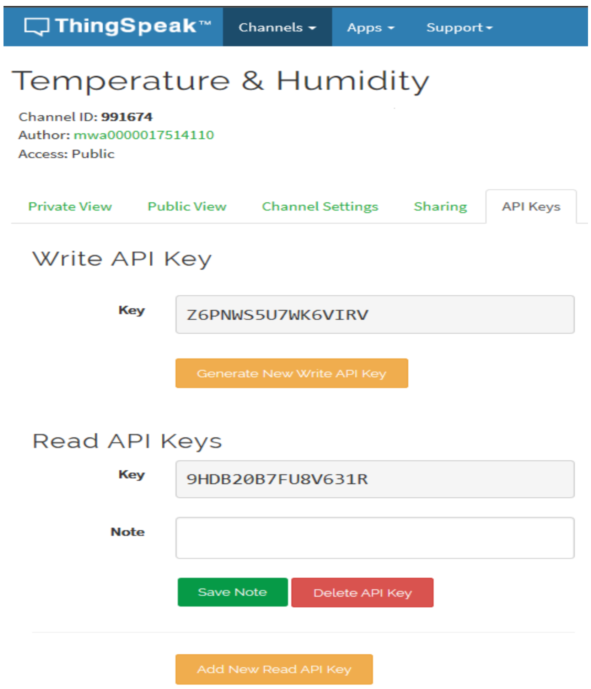

Fig 4.6.3: API keys of the Cloud channel

For each channel we create, we get the API keys and add them in our code to access from our setup.

4.7 Pseudo Code

4.7.1 Temperature and Humidity from DHT11 Sensor

```
#!/usr/bin/python

import sys

import Adafruit_DHT

while True:

   humidity, temperature = Adafruit_DHT.read_retry(11, 4)

   print 'Temp: {0:0.1f} C  Humidity: {1:0.1f} %'.format(temperature, humidity)
```

4.7.2 Soil Moisture from Soil Moisture sensor

```
#! /usr/bin/python

import RPi. GPIO as GPIO

import time

#GPIO SETUP

channel = 21

GPIO. setmode (GPIO.BCM)

GPIO. setup(channel, GPIO.IN)

def callback(channel):

      if GPIO. input(channel):

            print "Water Detected!"

      else:

            print "Water Detected!"

GPIO. add_ event_ detect (channel, GPIO.BOTH, bouncetime=300)
```

```
GPIO. add_ event_ callback (channel, callback)
while True:
        time. sleep (1)
```

4.7.3 pH values from pH sensor

```
import serial
import time
port= serial. Serial("/dev", baudrate= 9600, timeout=1.0)
def pH():
        x=10
        for i in range(x):
                rcv=port.readline()
                if len(rcv) > 4:
                        a= True
                        break
                else:
                        x= x+ 1
        while a:
                value= rcv. split(',')
                ph= value[0]
                return ph
```

4.7.4 Sending data from sensors to Thing speak

```
import RPi.GPIO as GPIO
import Adafruit_ DHT
import urllib2
import time
GPIO. setwarnings (False)
DHT_SENSOR = Adafruit_DHT.DHT11
myAPI = "Z6PNWS5U7WK6VIRV"

DHT_PIN = 4
baseURL = 'https://api.thingspeak.com/update?api_key=%s' % myAPI

while True:
   humidity, temperature = Adafruit_DHT.read(DHT_SENSOR, DHT_PIN)

urllib2.urlopen(baseURL+"&field1=%s&field2=%s&"%(temperature,humidit
y))
if humidity is not None and temperature is not None:

print("Temp={0:0.1f}CHumidity={1:0.1f}%".format(temperature,humidity))
      time.sleep (3);
```

4.7.5 Crop prediction

```
#import all the required modules
import numpy as np
import pandas as pd
from sklearn.preprocessing import StandardScaler
from sklearn.model_selection import train_test_split
from sklearn.neighbors import KNeighborsClassifier
#load dataset into a dataframe
df=pd.read_csv(r"C:\Users\l\Desktop\smaratdataset.csv")
X=df.iloc[:95, 2:6].values
y=df.loc[:,'croptype'].values
#perform scaling
scaler=StandardScaler()
scaler.fit(X)
X=scaler.transform(X)
#split into training and test data
X_train,X_test,y_train,y_test=train_test_split(X,y,test_size=0.20)
#define k- nearest neighbors classifier and fit data
classifier=KNeighborsClassifier(n_neighbors=5)
classifier.fit(X,y)
#predict the test data using the classifier
y_pred=classifier.predict(X_test)
```

4.7.6 Sending notifications to e-mail

The notification regarding the croptype is sent to the user's mail by the following steps

- Install SMTP related packages

  sudo apt-get install ssmtp

  sudo apt-get install mailutils

- Edit the SSMTP configuration file

  sudo nano /etc/ssmtp/ssmtp.conf

- Include the following lines in ssmtp.conf

  root=postmaster
  mailhub=smtp.gmail.com:587
  hostname=raspberrypi
  AuthUser=AGmailUserName@gmail.com
  AuthPass=TheGmailPassword
  FromLineOverride=YES
  UseSTARTLS=YES

- Save and Exit
- Then use the following command to send notifications

  – echo "hello" | mail –s "test" xyz@gmail.com

4.8 Testing Process

Testing is an activity to check whether the actual results match the expected results and to ensure that the system is Defect free. It involves execution of a software component or system component to evaluate one or more properties of interest. Testing also helps to identify errors, gaps or missing requirements in contrary to the actual requirements. The testing phase in the project involves identifying the areas which must be tested, the criteria for a positive test and the method of testing which must be taken up. There are various aspects that needs to be tested as a part of this system.

The first thing that needs to be tested is the values that we obtain from the sensors. For this purpose, we collect and data from one sensor and we cross checked the values with results from another sensor. The results from two sensors are checked for proximity so that they can be correct and valid. And the time to time variation of data should be valid. For example, during a sunny day the temperature of the day should be greater than that in the night. For a soil moisture sensor, the moisture levels should be high at the time of watering and should fall when the soil dries up.

The next aspect that is tested is the flow of data continuously to the cloud, to ensure that live values can be monitored properly. It so happens sometimes that due to certain issues the data is incapable of reaching the cloud, be it internet connectivity or any other. This results in data loss and such conditions can be eliminated by maintaining a queue and send that data when connection is re-established.

The most important criterion, to test whether the crop predicted is correct or not. All other test criteria supplement this one in some way or other. The system is run on all the test cases. All the test cases use the same k value i.e. the number of nearest neighbours. Since all the class labels are unique in the dataset, the tested data is compared against the trained data to check whether predicted crop is correct or not.

In this manner, each part of the entire system was tested in various phases and the system is giving adequate outputs with an accuracy of 82%.

5.RESULTS

5.1 Data collected from sensors

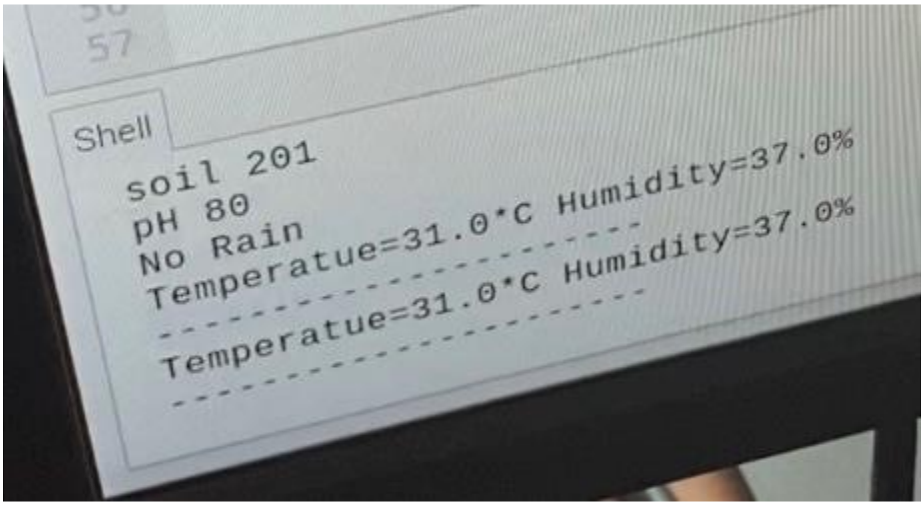

Fig 5.1: Sensor data

The above image shows the data generated continuously from the IoT sensors installed in the field.

5.2 Prediction results

```
y_pred=classifier.predict(X_test)
print(y_pred)
```

```
['groundnut' 'sunflower' 'rice' 'rice' 'chilli' 'coconut' 'mustard'
 'wheat' 'garlic' 'coconut' 'mustard' 'tamarind' 'corn' 'orange' 'ginger'
 'banana' 'cotton' 'cotton' 'banana']
```

```
xpre=[[55,77,80,6]]
ypre=classifier.predict(xpre)
print(ypre)
```

```
['rice']
```

```
print(accuracy)
```

```
0.823790463729103
```

Fig 5.2: Crop prediction results

The crop is predicted in the above image based on the prevailing conditions in the farm. The results were given with an accuracy of 82% by the model.

5.3 Web Page

5.3.1 Front Page

Fig 5.3.1: Front page

A webpage is designed to monitor and manage the activities enabled by the system. The above image is the main oage for the webpage we created.

5.3.2 Generic Page

Fig 5.3.2: Generic page

In this page, there is a link that directs to the dataset used for the prediction of crops.

5.3.3 Farm Monitoring Page

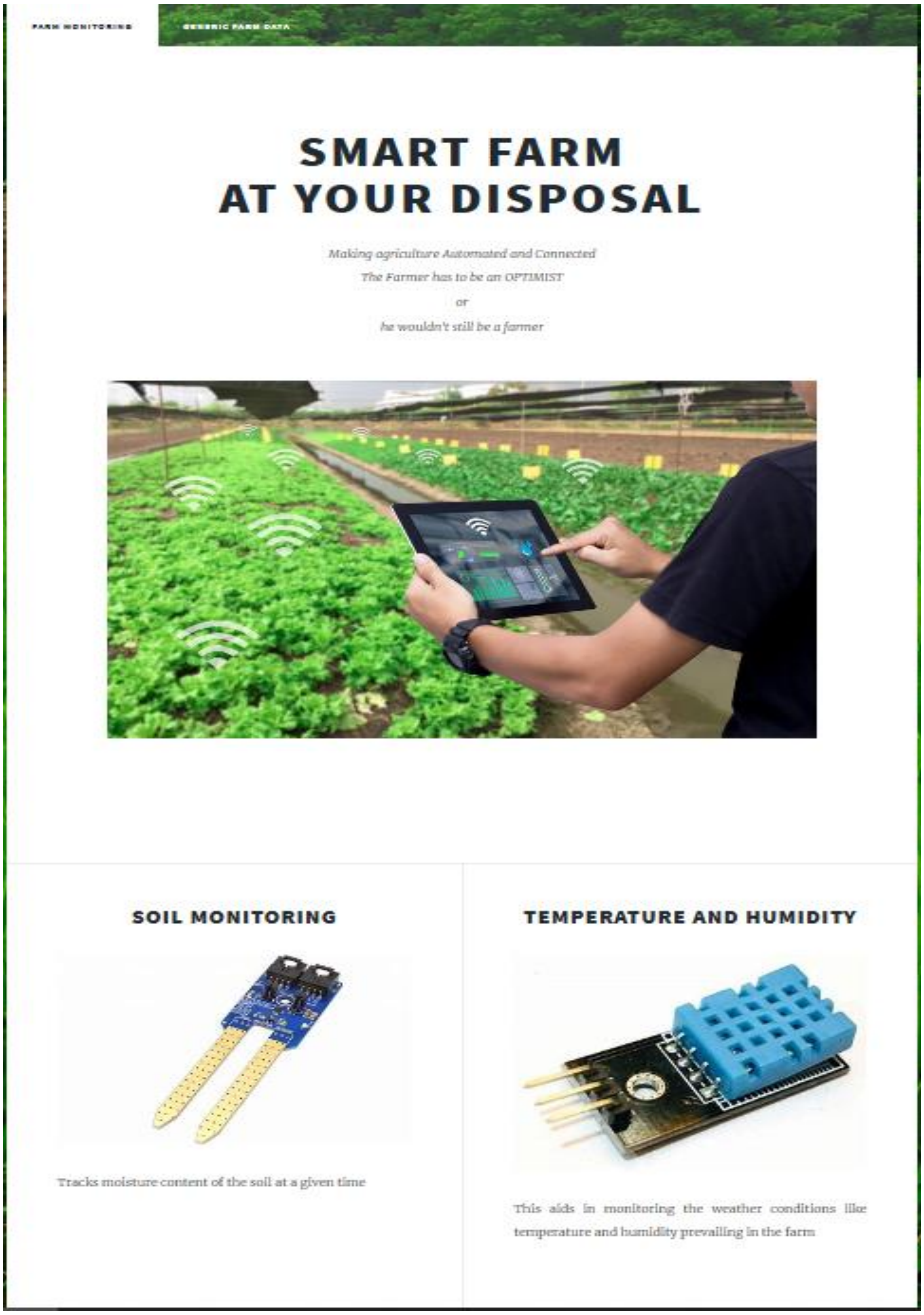

Fig 5.3.3: Farm Monitoring page

In the farm monitoring page, the live values and graphical representations of sensor data in the farm can be viewed through the links there.

5.4 Farm Monitoring Graphs

5.4.1 Temperature Analysis

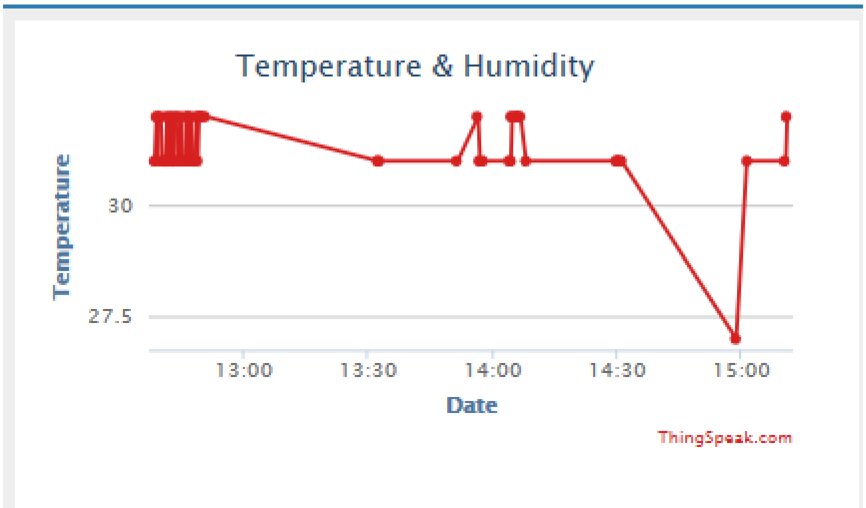

Fig 5.4.1: Temperature Analysis

The temperature data that is collected all the time is monitored through these graphs. This shows the variation in temperature graphically.

5.4.2 Humidity Analysis

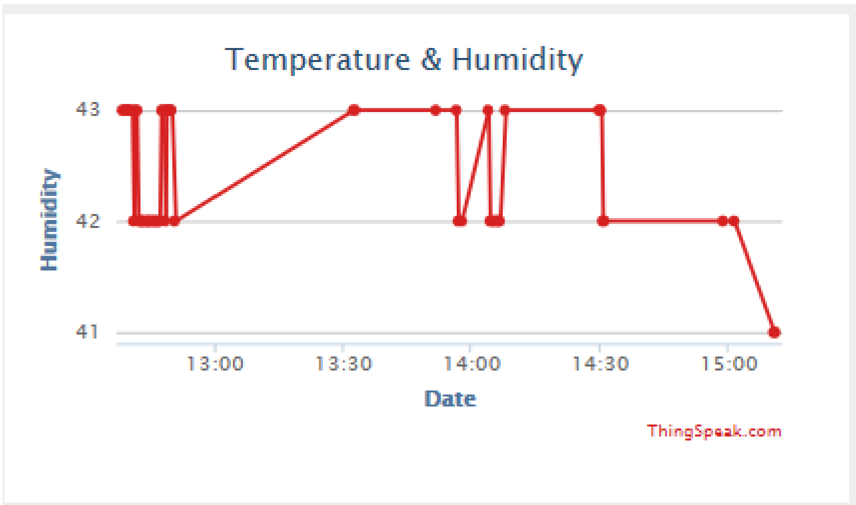

Fig 5.4.2: Humidity Analysis

The humidity prevailing the farm is observed through the graph generated above considering the humidity data collected periodically.

5.4.3 Soil Moisture Analysis

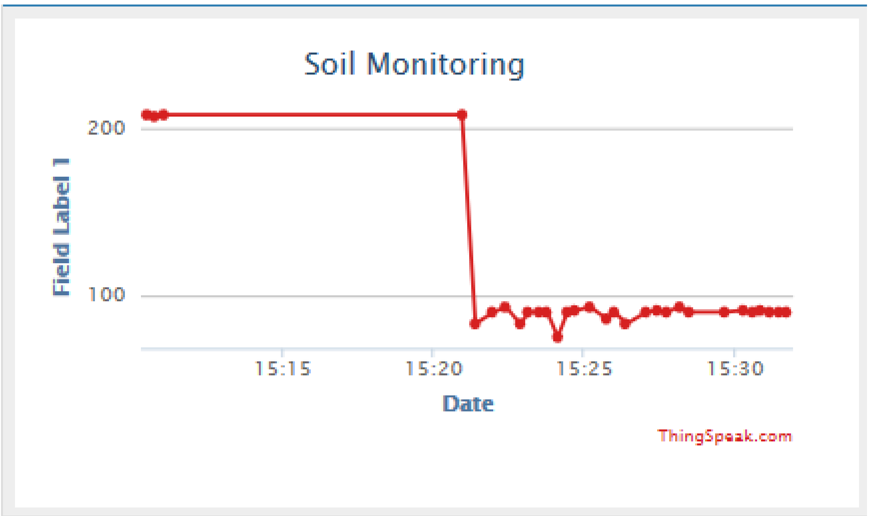

Fig 5.4.3: Soil Moisture Analysis

The above image shows the moisture levels of the soil being monitored continuously such that when the moisture levels begin to drop it indicates the need for watering.

6. CONCLUSION

This project is developed using Raspberry Pi and various sensors whose values are necessary to estimate the conditions of the farm. K-Nearest Neighbor(KNN) algorithm provided us easy measure to compare each set of data with the trained data using any kind of distance function and further classify in to one cluster.

As a part of our project, we studied a set of related papers and research articles that demonstrated various techniques of better agriculture practices and maintenance measures. The losses that farmer had to face due to wrong decisions on selection of crop and estimation of conditions went adverse. This project has the potential to help the farmers to minimize such losses. Instead of making conditions favourable to crop we select, we can rather go in a reverse idealogy making the selection of crops based on the resources we have. This pre- estimation of crop is economical, also preserves the fertility of the land.

6.1 Limitations

This project has few exceptions and limitations.

- One of the limitations is that, it is a single entity user project since it can only send result to single user at a time.
- The number of sensors used in this project are limited. So, cannot monitor other related factors and only major parameters are considered.
- Single controller for USB and Ethernet. This limits the system by dividing the speed between Ethernet and USB.
- Here, the hardware setup is installed at one point in the farm, in this way the other parts of the farm which may have relatively different values are not being considered. So, as a whole to the farm it may not judge uniformly.

6.2 Future Scope

- This project can be extended by adding other types of sensors like soil quality sensor, NPK sensor, turbidity sensor, Air quality sensor. This can help in managing the farm as a whole considering all aspects.
- In the current system, only one Raspberry Pi is installed at a point in the farm. Instead, if we can increase the number of nodes like this by placing them at various points on the farm, even the unevenness and other changes can be noticed and the aggregate of values given by all the nodes can be taken as the final condition.

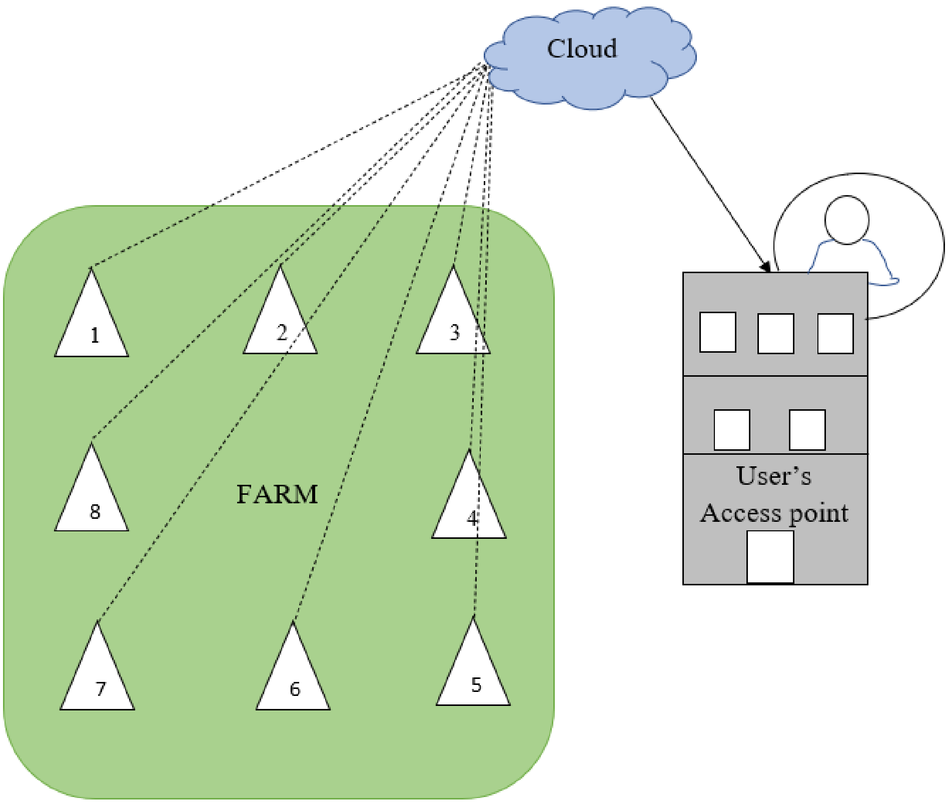

Fig 6.2.1: Installation of multiple Nodes

In the above figure, the farm is installed with nodes at various points on the farm. The data from all the nodes is sent to the cloud and is aggregated there. From the cloud, the user can access the results anywhere and anytime.

- A camera and fiber optic sensor can be connected with the nodes. In this scenario, the camera can capture images periodically and image processing is done on these images to determine the growth of the crop.

Fig 6.2.2: Images of various phases of growth

Another use of camera module would be to identify pests and crop diseases. The images that are captured by the camera, upon processing can reveal a few interesting results. We can notice the part of the farm that has been affected and can take measures accordingly. If the entire farm is uniform, it can be compared with the previous images to judge if it has been infected or attacked by pests.

Fig 6.2.3: Identifying Crop diseases

- Along with other sensors, if a fiber optic sensor is also installed it can be used to determine the fertility of the soil by measuring the micronutrients of the soil. This can be very helpful in deciding the fertilizers that are to be used for the proper growth and productivity of the soil.

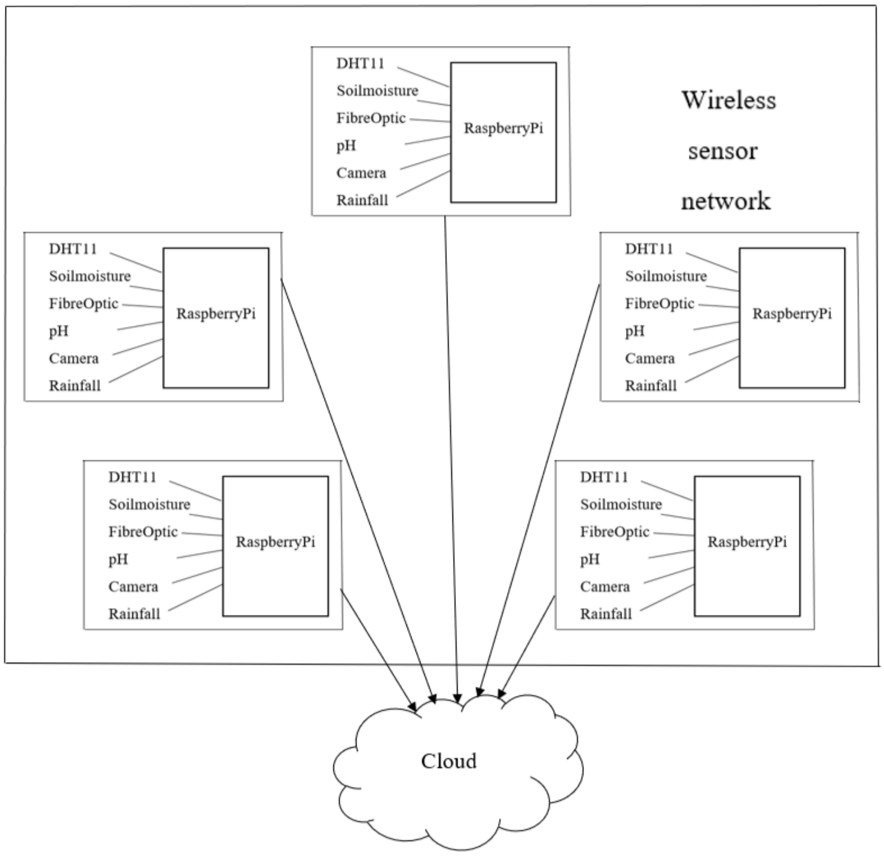

Fig 6.2.4 System with fiber optic sensors and camera

OUTCOME OF THE PROJECT

Title : Current research on Internet of Things (IoT) in Agriculture

Authors : K Spandana, Suresh Pabboju, Likhitha Chinthakuntla, G Sai Chandana

The paper is accepted for appearance in Volume-9 Issue-8 June 2020, IJITEE, Scopus indexed with impact factor of 6.03.

The paper is a research on the current works and technologies emerging in agriculture, which gave us the motivation to work towards the current project. It comprises a colloid of solutions being properly blend as to work towards the betterment of the future of agriculture.

From: **IJITEE Journal** <submit2@ijitee.org>
Date: Wed, May 20, 2020 at 8:48 AM
Subject: Re: Change of contents in the paper
To: Smt K.Spandana Assistant Professor <kspandana_cse@cbit.ac.in>

Dear Author (s),
We have received your final paper G5904059720 along with all supporting documents.
Congratulations...
This paper is considered for the publication of Volume-9 Issue-8 June 2020. It will be appeared in the journal website between 10 June to 15 June 2020. The URL of the Volume/issue is given below:
https://www.ijitee.org/download/volume-9-issue-8/
You can download your paper from above URL after publication.

Fig 7.0 Confirmation for research paper acceptance

REFERENCES

[1] Sukhpal Singh, Inderveer Chana, Rajkumar Buyya,” Agri-Info: Cloud Based Autonomic System for Delivering Agriculture as a Service”, Elsevier Research article, Internet of Things 9, 100131, 2019.

[2] Md. Ashifuddin Mondal, Zeenat Rehena,” IoT based Intelligent Agriculture field Monitoring System”, 8th International conference on Cloud Computing, Data Science and Engineering, IEEE 2018.

[3] Akhil R, Gokul M S, Sruthi Menon and Lekshmi S Nair, “Automated Soil Nutrient Monitoring for Improved Agriculture”, IEEE-0688, International Conference on Communication and Signal Processing, April 3-5, 2018, India.

[4] Surili Agarwal, Neha Bhangale, Kameya Dhanure, Shreeya Gavhane, V.A.Chakkarwar, Dr. M.B.Nagori, “Application of Colorimetry to determine Soil Fertility through Naive Bayes Classification Algorithm”, IEEE=43488, 9th ICCCNT 2018 July 10-12, 2018, IISC, Bengaluru.

[5] Lavanya G, Rani C, Ganeshkumar P,” An Automated Low Cost IoT based Fertilizer Intimation System for Smart Agriculture”, Sustainable Computing: Informatics and Systems (2019).

[6] Ziang Zhou, Kun Xu, Dingyun Wu, “Design of Agricultural Internet of Things Monitoring System Based on ZigBee”, CHEMICAL ENGINEERING TRANSACTIONS, VOL. 51, 2016.

[7] Mohanraj I, KirthikaAshokumar, NarenJ,” Field Monitoring and Automation using IOT in Agriculture Domain”, 6th International Conference on Advances In Computing & Communications, ICACC 2016, September 2016, pg:931-939.

[8] Deepa V. Ramane, Supriya S.Patil,A. D. Shaligram, “Detection of NPK nutrients of soil using Fiber Optic Sensor”, International Journal of Research in Advent Technology (E-ISSN: 2321-9637) Special Issue National Conference “ACGT 2015”, 13-14 February 2015

[9] Rashmi Ranjan, Mehajabeen Fatima," Survey Paper on IOT and Image Processing Based Crop Disease Identification System", IJCSE, VOL-6, Issue-8, Aug 2018, E-ISSN:2347-2693.

[10] Jirpond Muangprathub, Nathaphon Boonam, Siriwan Kajornkasirat, Narongsak Lekbangpong, Apirat Wanichsombat, Pichetwut Nillaor, "IoT and Agricuture Data Analysis for Smart Farm", Computer and Electronics in Agriculture, ELSVIER Journal, 0618-1699.

[11] Deepak Vasisht, Zerina Kapetanovic, Jong-ho Won, Xinxin Jin, Ranveer Chandra, Ashish Kapoor, Sudipta N. Sinha, Madhusudhan Sudarshan, Sean Stratman,," FarmBeats : An IoT Platform for Data-Driven Agriculture", 14th UNISEX Symposium on Networked System Design and Implementation,pg:515-529.

[12] Zhao Liqiang, Yin Shouyi, Liu Leibo, Zhang Zhen, Wei Shaojun, "A Crop Monitoring System Based on Wireless Sensor Network", Procedia Environmental Sciences 11 (2011) 558 – 565.