(IJIASE) 2019, Vol. No. 5, Jan-Dec **e-ISSN: 2454-9258, p-ISSN: 2454-809X**

"DEVELOPING AN INTEGRATED MODEL BASED ON MACHINE LEARNING AND INTERNET OF THINGS (IOT) TO FORECAST CRITICAL FACTORS TO OPTIMIZE AGRICULTURAL YIELDS"

Unnati Gupta

KR Mangalam World School, Vaishali, Ghaziabad

ABSTRACT

As the worldwide population is increasing rapidly, so it is obvious that the demand for food also increase—likewise, nations like India whose 1/third capital originated from cultivating needs to improve the techniques for cultivating. We propose an IoT based framework, which gathers different physical elements continuously, for example, temperature, humidity, moisture and other things. This learning can be concentrated to get diagnostic bits of knowledge. We attempt to apply different cutting-edge AI methods on this clustered information to forecast the yield appropriate for given elements.

I. INTRODUCTION

Farming is the primary root of nourishment for human species. As the population is expanding the sources and conventional ways to deal with cultivating should be updated. Development in the farming division is important for the advancement of the monetary state of the nation. In India, a large portion of the ranchers produce customary harvests (those which are delivered by their predecessors). In any case, for a comparable state of being, there might be different potential harvests that can be delivered on a similar land, which may bring about more noteworthy benefit.

IoT and ML have risen with superior processing which makes it possible to gather information and investigations it. It very well may be applied in different areas, for example, horticulture, clinical and so on to assemble brilliant frameworks. ML procedures can be utilized to build up a prescient model.

ANN organically motivated programming worldview which empowers a PC to gain from observational information; ANNs are excellent in taking in designs from nonlinear information.

Numerous investigates are done in the field of horticulture, which implies the utilization of sensors to gather information. The gathered information can be dissected to separate some valuable insights. Additionally, ML can be utilized to build up a model to anticipate different Factors, for example, reasonable yield, water necessities, assessed cost and so on which can improve the manner in which the cultivating is finished.

In this paper, we'll just think about the issue of foreseeing appropriate yield for given states of being. Counterfeit Neural Network is utilized for order of the yield.

A. Factors Influencing Production of Crop

1. Genetic factor also called Internal Factor.

2. Environmental factor that is External Factor.

In this paper, we discard all the internal factors. We just think about climatic elements and edaphic factors (soil factors) which go under outer variables. Almost half of the yield is credited to the impact of climatic components that influence the efficiency of the harvest.

Characterized

1. Moisture

2. Soil Condensation

3. Soil Variety

4. Month

5. Sunlight

6. Weather situation

7. Heat

8. Elevation

9. Atmospheric pressure

II. RELATED WORK

A. Agriculture Field Monitoring

Rather than watching the farms constantly, this paper proposes the structure to screen similar traits utilizing a remote sensor design. The farmer may furthermore have recognized by means of SMS by a specialist.

B. Environment Monitoring

The paper proposed a framework that gathers different natural details like temperature, mugginess, enlightenment, voltage and so on from nursery and from that point it sends the information to the closest worker by means of GPRS. The framework incorporates a web application which can show the nursery status using maps.

C. Farm Automation

The number of tasks of the farm can be robotised like water system framework; a temperature-controlled framework for animals and ranch item. This paper presents programmed house temperature control, lighting framework control, programmed sprinkler framework and security in farmhouses.

D. Crop yield Prediction

This paper present determining strategies for assessing crop yield measures utilising time series models to anticipate crop yields.

This paper expands the IoT based monitoring frameworks with AI calculations to make forecasts on what yields are reasonable for the given condition.

III. PROPOSED WORK

The IoT gadget gathers information with a sensor cluster and sends it to the worker; a cooperative model uses this information to appraise the models and "future" information for expectation. The framework comprises of two significant parts:

1. Managing environmental data,

2. Implementing machine learning methods on gathered information.

The first module consists of following:

- Sensors to gather information, for example, soil dampness sensor, mugginess sensor, temperature sensor, barometric sensor, Altitude and light sensor.
- For connectivity Wi-Fi module.
- For processor Microcontroller.
- Solar Panel|Power Supply.

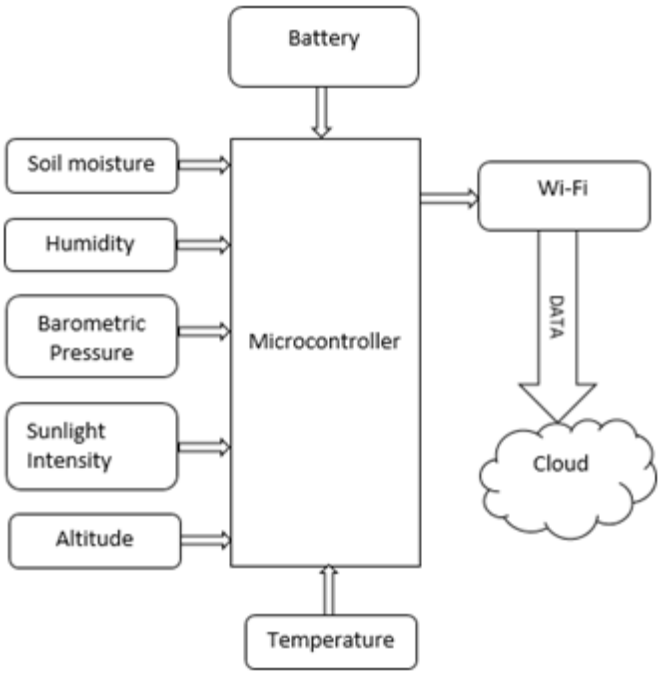

Figure 1: Architecture of the device.

The engineering of the gadget can be perceived with the assistance of the above outline.

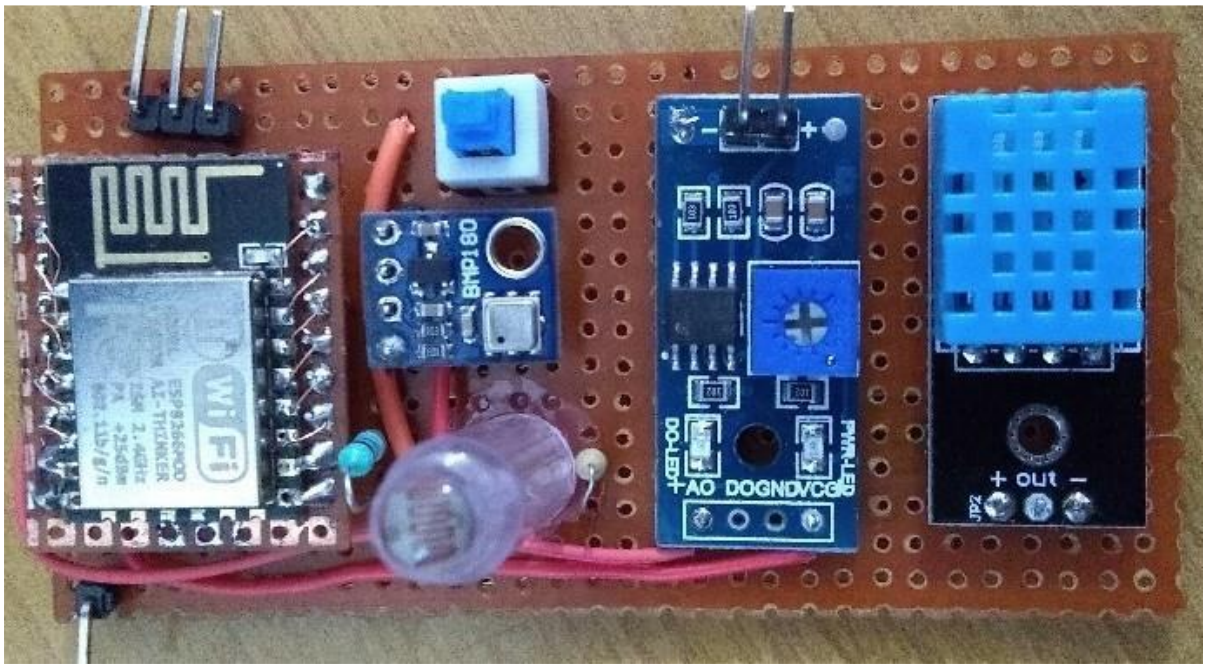

Figure 2: Embedded Circuit with Sensors.

The second module consists of:

- Pre-processor Data: Pre-Processing is an initial step in ML as the nature of information and the helpful data that can be gotten from it legitimately influences the capacity of our model to learn.
- ML model to make expectations The model is prepared by thinking about 9 highlights, the element vector is given to model to forecast the best crops.

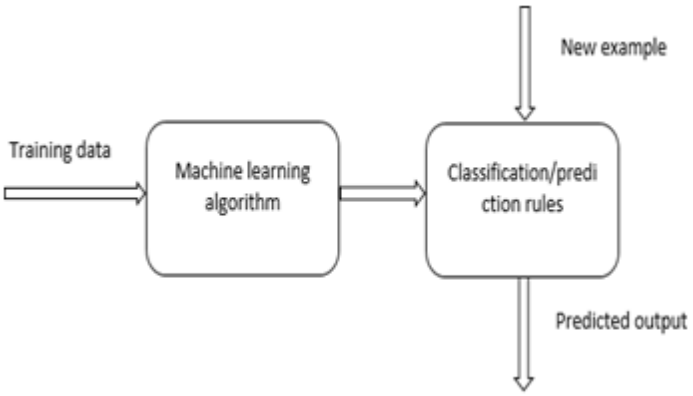

Figure 3: Model based learning.

IV. METHOD OVERVIEW

The goal of the proposed work is to forecast the yield name utilizing ML methods; this incorporates the accompanying stages:

A. Dataset

The information is gathered from 50 unique fields of chosen crops in Dehradun area.

Info information comprises of 10 boundaries, in particular: season, soil type, stickiness, elevation, climate condition, daylight, environmental weight and the yield mark (ground truth). For effortlessness, we have just thought to be 3 yields: paddy, wheat and sugar-stick. The dataset contains 1/third information point for each yield. The dataset is isolated into two gatherings: 75% for preparing the model, which incorporates the approval set also and 25% for testing the model. The primary motivation behind utilising the testing dataset is to test the speculation capacity of a prepared model. Each harvest class contains information focuses on both training and test dataset to dodge irregularity. Kindly don't change them. You may note quirks. For instance, the head edge in this layout gauges proportionately more than is standard. This estimation and others are intentional, utilising details that foresee your paper as one aspect of the whole procedures, and not as a free record. Kindly don't amend any of the current assignments.

1. Sample: The information pre-handling is the method of introducing the information into the necessary arrangement for better understanding and the better outcomes. All things considered, it might vary from circumstance to circumstance and as far as information to information. Distinctive information requires various types of approach for the pre-preparing, for example, for mathematical information; the pre-handling methods can be including scaling. Interestingly, on account of the absolute information, the pre-preparing methods differ, for example, mark encoding, one-hot encoding and so forth.

Feature	Type	Range
Temperature	Numeric	-100-200
Weather	Categorical	clear, clouds, scattered, clouds, broken clouds, shower, rain,
Month	Categorical	Jan, Feb, Mar, Apr, May, Jun, Jul, Aug, Sep, Oct Nov, Dec
Soil Type	Categorical	Loam, Clay, Sand, Silt, Peat, Chalk
Humidity	Numeric	0-100
Soil Moisture	Numeric	0-100
LUX	Numeric	> 0
Pressure	Numeric	>0
Altitude	Numeric	c

humidity	Soil moisture	Soil type	sunlight	Weather condition	Temp (°C)	altitude	pressure	month	Crop label
0.74	0.08	alluvial	365.81	sunny	24	655	1007	November	wheat
0.41	0.41	loam	612.56	sunny	35	667	993	June	paddy
0.91	0.78	clay-loam	599.98	passing clouds	26	667	996	June	paddy
0.67	0.46	alluvium	654.89	sunny	11	657	1015	February	sugar-cane
0.49	0.69	loam	587.90	Passing clouds	34	671	996	June	paddy
0.62	0.45	alluvium	600.54	sunny	21	657	1015	march	sugar-cane
0.32	0.21	alluvium	488.34	partly sunny	17	657	1015	February	sugar-cane

In this dataset, three kinds of pre-preparing strategies have been utilised, for example,

1. Highlight Scaling since the scope of estimations of crude information fluctuates broadly; the information is standardized for better precision of the classifier. Numeric highlights are mean standardized,

$$x' = \frac{x - \text{average}(x)}{\max(x) - \min(x)}$$

Mean normalization

2. Name Encoding: Label Encoding is utilized to change over all-out information into mathematical information. Every class is encoded to some remarkable number worth. For example, male or female can be encoded to 0 or 1 individually.

3. One Hot Encoding: One hot encoding is utilized to doubles downright information by making a spurious variable. As this type of portrayal is better reasonable by the ML calculations.

3. Artificial Neural Network.

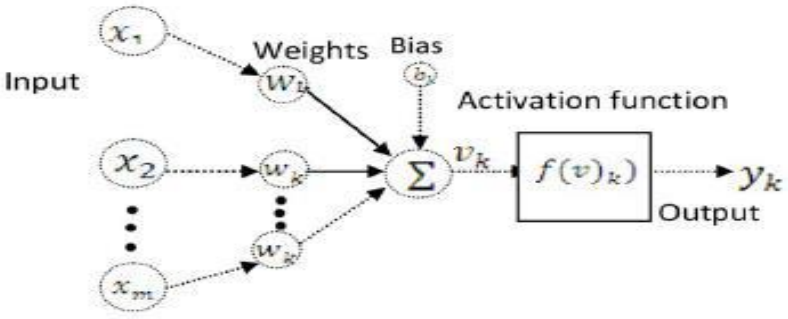

Figure 5

ANN comprises of the accompanying significant parts:

1. Information Layer: It contains hubs which compare to various highlights, the info includes vector R9.

2. Concealed Layer: Only 1 shrouded layer is utilized. It has 10 neurons.

3. Yield Layer: R3 autonomous probabilities for comparing crops for thought about harvests.

Initiation Function: In shrouded layers, ReLu (Rectified Linear Units) is utilised. The greatest bit of leeway of ReLu is surely non-immersion of its inclination, which incredibly quickens the intermingling of stochastic slope plunge contrasted with the sigmoid.

$$f(x) = max(0, x) = \begin{cases} x_i, & if\ x_i \geq 0 \\ 0, & if\ x_i < 0 \end{cases}$$

$$f(x) = \left(\frac{1}{(1 + exp^{-x})} \right).$$

In the yield layer, sigmoid is utilised in light of the fact that its yield is between 0 to 1, which is deciphered as free probabilities of ground truth classes.

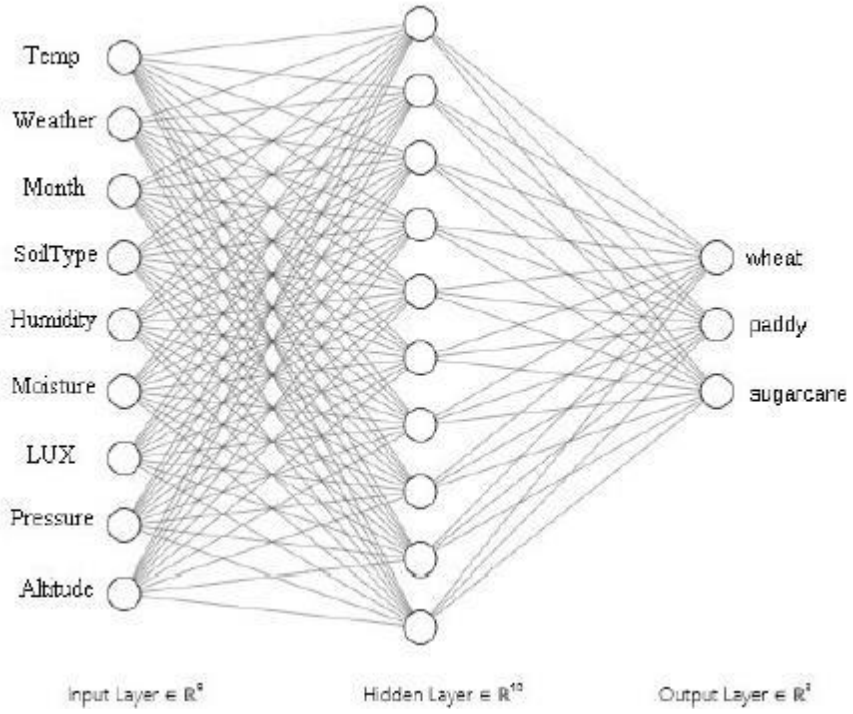

Figure 6

We prepared our model utilising stochastic slope plummet with a clump size of 128 models. As the preparation dataset is little, we start with a huge learning pace of 0.1 at that point progressively decline the rate to 0.01. The age size is set to 10.

The hyper parameters of the model are tuned utilising k overlap approval with k = 3. In the wake of tuning, the model is then prepared all in all preparation dataset which gets around 13% misfortune on the test dataset.

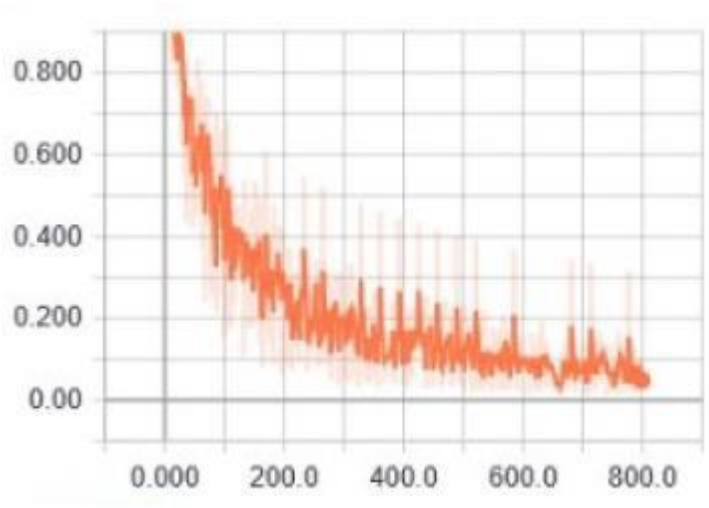

Figure 7: Learning Curve.

The x-axis of this bend tells to steps, and the y-pivot tells to blunder rate.

V. CONCLUSION

In India, the normal salary of a farmer is assessed at Rs. 77,976 every year, as per the Dalwai board of trustees' report [Jan 30, 2018]. Present-day methods can be utilised to improve the cultivating business. IoT and ML-based arrangements could help inappropriate use of assets.

The determination of yield is a repetitive and tedious assignment for a rancher. The rancher, for the most part, chooses crop as per their conventional information and past encounters. The proposed framework predicts all the appropriate harvests. Along these lines, it disposes of the human predispositions.

This examination shows that IoT gadgets can be utilised to gather information, and ML calculations can be applied to get scientific bits of knowledge. By and by, there can be a ton of other down to earth factors which we did exclude, accessibility of assets, trimming framework alternatives, market request and accessibility, Govt. Approaches, and so on.

VI. FUTURE WORK

A wide range of variations, tests, and trials have been left for the future because of the absence of time (for example the examinations with genuine information are normally very tedious, requiring even days to complete a solitary run).

(IJIASE) 2019, Vol. No. 5, Jan-Dec **e-ISSN: 2454-9258, p-ISSN: 2454-809X**

Gathering the right information is a tedious assignment. In this way, we just apply these techniques on the trial dataset. The nature of the outcomes can be additionally improved by gathering more information. Additionally, another sort of sensors could be utilised; for example, electrochemical sensors can be utilised to give constant observing of soil supplement information.

The accompanying thoughts could be tried:

• Market investigation should be possible on anticipated harvests, to do hazard examination, which is likewise a significant factor during the choice of yields.

• Effect of climatic condition on specific yields could be dissected.