JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

FARMING FOR FUTURE: ENHANCED CROP YIELD PREDICTION

[1]Laxmi M C, [2]Amulya Y R, [3]Aishwarya, [4]M R Adhithi, [5]Jyothi JR Sahani

[1]Assistant professor, [2,3,4,5]Student
[1]Computer Science and Engineering,
[1]City Engineering College, Bangalore, India

Abstract: Agriculture is the most important sector of Indian Economy. Indian agriculture sector accounts for 18 percent of India's GDP and provides employment to 50% of the country’s workforce. Agricultural researchers insist on the need for an efficient mechanism to predict and improve the crop growth and Majority of research work in agriculture focus on biological mechanisms to identify crop growth and enhance in yield. The outcome of crop yield primarily depends on parameters such as variety of crop, seed type and environmental parameters such as sunlight (Temperature), soil (ph.), water (ph.), rainfall and humidity. By analyzing the soil and atmosphere at particular region best crop in order to have more crop yield and the net crop yield can be predicted. With the price in the market at periodic time of modal price of the crop in certain season and year based with the name of crop. This prediction will help the farmers. To choose appropriate crops for their farm according to the soil type, temperature, humidity, water level, spacing depth, soil PH, season, fertilizer and months. And based on their price prediction the farmers can change the crop based on profit or loss before cultivation to get more profits.

Index Terms— Arduino; Humidity sensor; N, P, K sensor; Moisture sensor.

I.INTRODUCTION

Agriculture is a developing topic of study. Sustainable agriculture helps in feasible and maintains soil quality, reduces soil degradation, saves water resources, improves land biodiversity, and ensures a natural and healthy environment. Specifically, agricultural crop prediction is crucial and primarily dependent on soil and environmental factors like temperature, humidity, and rainfall. Agriculture stands out as a primary sector of societal importance, given its crucial role in supplying a substantial amount of food. Currently, a growing global population faces hunger in numerous nations, primarily attributable to shortages or deficiencies in food supply. Sustainable agriculture involves the practice of preserving the environment without sacrificing the fundamental needs of future generations, all while enhancing farming efficiency. Traditionally, farmers have chosen what to grow based on knowledge of the local climate, seasonal cycles, and folklore. Conventional approaches, however, have drawbacks. Crop yield predictions grow more difficult, particularly when dealing with pests, diseases, and shifting climatic trends. In a fast changing agricultural world, traditional agriculture frequently lacks the accuracy and efficiency needed to maximize yields and optimize resource usage. The agricultural industry has seen a radical transformation with the introduction of contemporary technology, especially sensor technologies. High-tech sensors that measure temperature, humidity, crop health, and soil moisture give real-time data that greatly increases crop forecast accuracy. Today's crop prediction is enhanced by the combination of machine learning and sensor data. To build prediction models, machine learning algorithms can examine past data, meteorological trends, and real-time sensor inputs. Crop production is determined by various factors, including crop genotype, environmental conditions, and management practices. Fluctuations in climatic conditions, both spatially and temporally, can lead to diverse agricultural yields from one year to the next. In such scenarios, precise yield forecasting plays a crucial role in global food production. Reliable predictions enable informed import and export decisions. As a result, farmers can leverage the projected yield for improved management and financial decision-making. The performance of hybrids can be anticipated to be effective in novel and unexplored environments. The future of agriculture faces adverse effects due to shifting environmental conditions, especially with the impact of global warming and climate variability. Precisely predicting crop production through statistical models is a laborious and demanding process. It requires considerable time and effort to forecast crop production using statistical models. Based on the challenges presented in the study and the research questions at hand, a machine learning model can serve either a descriptive or predictive purpose.

1.1 Data and Sources of Data

Stock the information associated with the yield and manure in the csv which contains of the state, district, crop year, season, crop, area, production, and another training data include phosphorous level, level of potassium, level of nitrogen in the soil, how much quantity of phosphorous, potassium, nitrogen should be utilized to maximize soil richness.

1.2 Literature Survey

1) Revolutionizing Agriculture: Harnessing AI and IoT for Automation and Digitization (2023)

Agriculture automation, Artificial intelligence, Deep learning Internet of things, Smart farm machinery. The DL is a subsidiary of ML, was recently utilized for the issue of crop yield projections and is thought to be very convincing. The major difference between ML and DL is the low performance of the DL network with a small training sample. In addition, the DL approach can extract key functions from input data automatically; however, in other studies, the features are manually extracted, and effective DL.

2) Crop Prediction Based on Characteristics of the Agricultural Environment Using Various Feature Selection Techniques and Classifiers (2022)

The predictive methodology, influenced by Myers et al. and Muriithi, utilizes statistical and mathematical tools are integral in the improvement of both new and existing products. NN has been used in SF for different purposes including soil or atmospheric parameter forecasting optimal Environmental Conditions prediction and artificial neural networks (ANN) based controller to optimize irrigation. The desire of achieving tasks is to produce accurate results using NN, while minimizing the training time made the literature use tailored architectures.

II RESEARCH METHODOLOGY

The methodology section outline the plan and method that how the study is conducted. This includes Universe of the study, sample of the study, Data and Sources of Data, study's variables and analytical framework. The details are as follows;

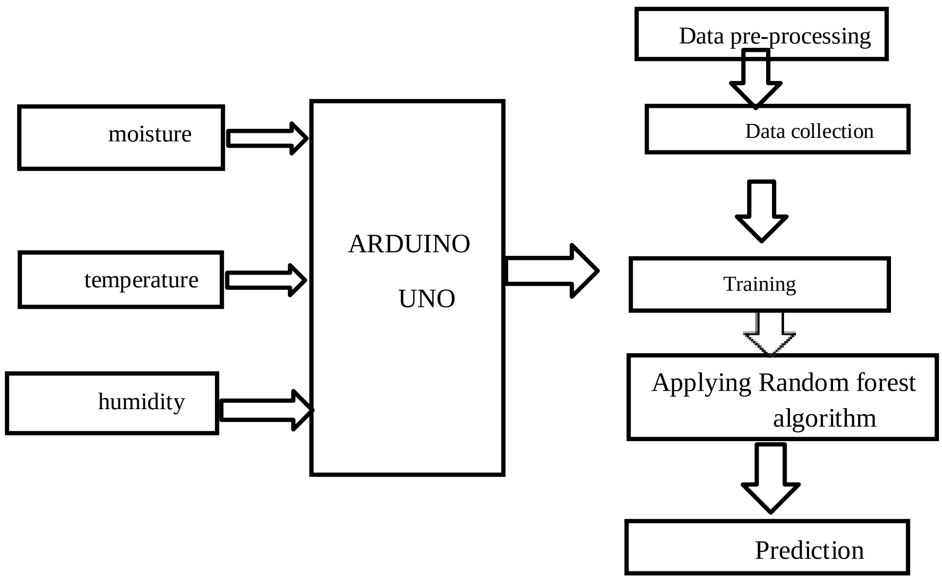
2.1 Data Pre-processing: Here the natural information in the harvest information is polished and the metadata is attached to it by clearing the things which are modified tothe whole number. So, the data can be effortlessly trained. Right now, we first deliver the metadata into this and after that this metadata will be joined to the information and substitute the transformed information with metadata. Then this information will be moved further and clear the unfavourable information in the list and it will partition the information into the train and the test data.

2.2 Fertilizer utilization using Back- propagation: The Fertilizer information set which is accessible in the CSV arrangement is pre-operated and made ready to develop the model with that dataset. Firstly, the data collection is partitioned into 80% for composing information and 20% for the test information. The calculation used to produce the dataset into a model is the Backpropagation calculation backpropagation algorithm is the scheme from the several layer perceptron in the artificial neural system. The backpropagation calculation is utilized for gigantic datasets which have no right connections between the credits of the dataset to shape a system model via developing the dataset and determining the yield. This algorithm contains three layers in the system model, they are the information layer, concealed layer, and yield layer.

2.3 Random Forest

Random Forest algorithm is a supervised classification algorithm. We can see it from its name, which is to create a forest by some way and make it random. There is a direct relationship between the number of trees in the forest and the results it can get: the larger the number of trees, the more accurate the result. But one thing to note is that creating the forest is not the same as constructing the decision with information gain or gain index approach.

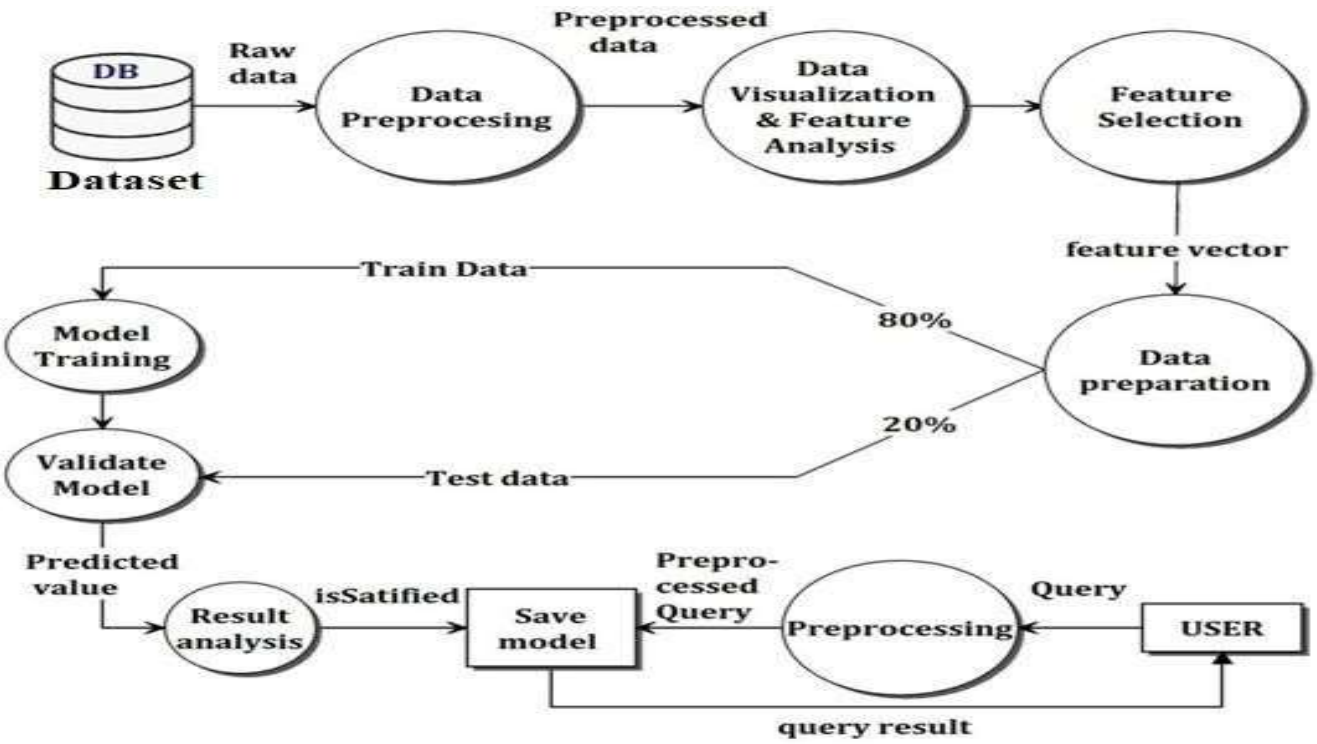
2.4 System Architecture



Analysis is the process of breaking a complex topic or substance into smaller parts to gain a better understanding of it. Analysts in the field of engineering look at requirements, structures, mechanisms, and systems dimensions. Analysis is an exploratory activity. The Analysis Phase is where the project lifecycle begins. The Analysis Phase is where you break down the deliverables in the high-level Project Charter into the more detailed business requirements. The Analysis Phase is also the part of the project where you identify the overall direction that the project will take through the creation of the project strategy documents. Gathering requirements is the main attraction of the Analysis Phase. The process of gathering requirements is usually more than simply asking the users what they need and writing their answers down. Depending on the complexity of the application, the process for gathering requirements has a clearly defined process of its own. This process consists of a group of repeatable processes that utilize certain techniques to capture, document, communicate, and manage requirements. Systems design is the process of defining the architecture, components, modules, interfaces, and data for a system to satisfy specified requirements. Systems design could see it as the application of systems theory to product development. There is some overlap with the disciplines of systems analysis, systems architecture and systems engineering. If the broader topic of product development "blends the perspective of marketing, design, and manufacturing into a single approach to product development," then design is the act of taking the marketing information and creating the design of the product to be manufactured. Systems design is therefore the process of defining and developing systems to satisfy specified requirements of the user.

III ANALYSIS AND PREDICTION

In this module, patterns in data is recognized, percentage correlation between various factors affecting crop yield and price are determined. Various data visualization techniques are used to study the patterns in data and factors causing change. Algorithms like Multiple Linear Regression and Random Forest are used to predict crop yield and price. The accuracy of these algorithms are compared using mean absolute percentage error thus helping us determine the most suitable approach for prediction.



During the detailed phase, the view of the application developed during the high-level design is broken down into modules and programs. Logic design is done for every program and then documented as program specifications. For every program, a unit test plan is created. The entry criteria for this will be the HLD document. And the exit criteria will the program specification and unit test plan (LLD). A data flow diagram is the graphical representation of the flow of data through an information system. DFD is very useful in understanding a system and can be efficiently used during analysis. A DFD shows the flow of data through a system. It views a system as a function that transforms the inputs into desired outputs. Any complex systems will not perform this transformation in a single step and a data will typically undergo a series of transformations before it becomes the output. With a data flow diagram, users are able to visualize how the system will operate that the system will accomplish and how the system will be implemented, old system data flow diagrams can be drawn up and compared with a new systems data flow diagram to draw comparisons to implement a more efficient system.

IV. IMPLEMENTATION

Implementation is the carrying out, execution or practice of a plan, a method, or any design, idea, model, specification, standard or policy for doing something. As such, implementation is the action that must follow any preliminary thinking in the order for something to actually happen. Implementations allow the users to take over its operation for use and evaluation. It involves training the users to handles the system and plan for a smooth conversion.

Implementation is a process of ensuring that the information system is:

1)Constructing a new system from scratch.

2)Constructing a new system from the existing system.

- **HARDWARE IMPLEMENTATION**

1. ARDUINO

The Arduino Uno is a microcontroller board based on the ATmega328. It has 14 digital input/output pins (of which 6 can be used as PWM outputs), 6 analog inputs, a 16 MHz ceramic resonator, a USB connection, a power jack, an ICSP header, and a reset button. It contains everything needed to support the microcontroller; simply connect it to a computer with a USB cable or power it with a AC-to-DC adapter or battery to get started.

2. Moisture Sensor SA SM01

In this sensor we are using 2 Probes to be dipped into the Soil As per Moisture We will get Analogue Output variations from 0.60volts - 12volts. Soil moisture sensors measure the water content in soil. A soil moisture probe is made up of multiple soil moisture sensors. One common type of soil moisture sensors in commercial use is a Frequency domain sensor such as a capacitance sensor. Another sensor, the neutron moisture gauge, utilizes the moderator properties of water for neutrons.

3. Humidity sensor

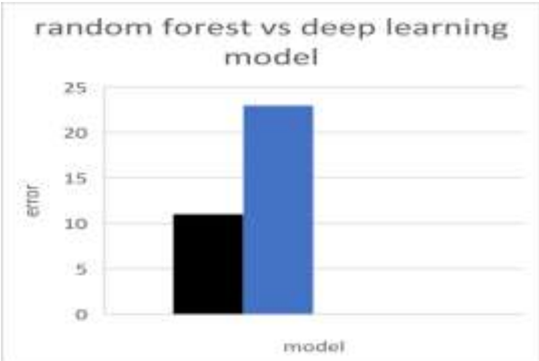
A humidity sensor (or hygrometer) senses, measures and reports both moisture and air temperature. The ratio of moisture in the air to the highest amount of moisture at a particular air temperature is called relative humidity. Relative humidity becomes an important factor when looking for comfort.

- **SOFTWARE IMPLEMENTATION**

The Arduino Uno has a number of facilities for communicating with a computer, another Arduino, or other microcontrollers. The ATmega328 provides UART TTL (5V) serial communication, which is available on digital pins 0 (RX) and 1 (TX). An ATmega16U2 on the board channels this serial communication over USB and appears as a virtual com port to software on the computer. The '16U2 firmware uses the standard USB COM drivers, and no external driver is needed. However, on Windows, a .inf file is required. The Arduino software includes a serial monitor which allows simple textual data to be sent to and from the Arduino board. The RX and TX LEDs on the board will flash when data is being transmitted via the USB-to-serial chip and USB connection to the computer (but not for serial communication on pins 0 and 1). A Software Serial library allows for serial communication on any of the Uno's digital pins. The ATmega328 also supports I2C (TWI) and SPI communication. The Arduino software includes a Wire library to simplify use of the I2C bus; see the documentation for details. For SPI communication, use the SPI library. The Arduino Uno can be programmed with the Arduino software. The ATmega328 on the Arduino Uno comes pre burned with a bootloader that allows you to upload new code to it without the use of an external hardware programmer. It communicates using the original STK500 protocol (reference, C header files).

V. RESULTS AND DISCUSSION

In this paper, attempt is made orderly to know the yield manufacturing inspection and is operated by applying both the Random Forest calculation and Backpropagation calculation. These models were investigated with various kind of crops in different regions across India to anticipate the harvest. Indeed, even manure information was developed utilizing the back-propagation calculation and estimated to acquire the output of how much nitrogen, phosphorus is necessary for the sector of land. Both the models for the crop creation were examined in expecting the yield and by different parameter concerning the fault rate. We inspected the mistake rate gained while focusing at the random forest calculation and backpropagation where we obtained the misconception rate lesser to the random forest than back propagation while determining the yield for both of the models and the solicitude is plotted in the chart Fig. 3. For prediction of the yield, the client will enter the data as indicated in Fig 4. The client must have to enter the particulars consecutively. The result of harvest anticipation is indicated in Fig 5. The input of the compost information is given as shown in fig.6 and output of the fertilizer information is shown in Fig. 7



Enter nitrogen:

Enter phosphorus Enter potassium

Amount of nitrogen fertilizer:45.02 Amount of

phosphorus:28.09 Amount of potassium:35.10

REFERENCES

[1] Niketa Gandhi et al," Rice Crop Yield Forecasting of Tropical Wet and Dry Climatic Zone of India Using Data Mining Techniques", IEEE International Conference on Advances in Computer Applications (ICACA), 2021.

[2] K.E. Eswari. L.Vinitha. "Crop Yield Prediction in Tamil Nadu Using Bayesian Network ", International Journal of Intellectual Advancements and Research in Engineering Computations, Volume-6, Issue-2, ISSN: 2348- 2079.

[3] Shruti Mishra, Priyanka Paygude, Snehal Chaudhary, Sonali Idate "Use of Data Mining in Crop Yield Prediction" IEEE Xplore Compliant - Part Number:CFP18J06-ART, ISBN:978-1-5386-0807-4; DVD Part Number: CFP18J06DVD , ISBN:978-1-5386-0806-7.

[4] Anna Chlingaryana, Salah Sukkarieha, Brett Whelanb — Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review, Computers and Electronics in Agriculture 151 (2022) 61–69, Elisver, 2022.

[5] Adegbeye M J, Reddy P R, Obaisi A I, Elghandour M M, Oyebamiji K J, Salem A Z, Morakinyo-Fasipe O T, Cipriano-Salazar M & Camacho-Díaz L M, Sustainable agriculture options for production, greenhouse gasses and pollution alleviation, and nutrient recycling in emerging and transitional nations-An overview, *J Cleaner Produc*, **242** (2020) 18319.