

# Novel Methods for Crop Prediction based on Influencing Parameters in Indian Regions using Data Mining



Albin Joseph, Manohar

**Abstract:** Agriculture plays a crucial role for the production of food in Indian regions. Indian regions mainly produces crops like rice, wheat, maize and many other types of crop. There are several factors required for the productivity of any harvest, but we know that soil, climate, pesticides, Fertilizers and ground water is most influencing essential factor for the productivity of any harvest. Let us consider soil which is the key element as it provides nutrients for proper development and growth of crops. Secondly, climate is also having major role in agriculture as crop growth depends on rainfall, humidity, temperature etc. Thirdly, Pesticides is widely used to control pest and prevents the damage of crops. Fourthly, Fertilizers can improve the quality of crops. Finally, ground water which will enrich nutrients in soil. The current preparation centers around different information mining procedures utilized in various regions of India and anticipate future harvest along with reasonable information mining procedure saw during the period(1920-2019).The parameters considered for the examination were soil, atmosphere, water thickness, pesticides and composts and Crop informational collection. The Classification calculations utilized in preparation were Adaptive boosting classification, Excess tree classification, neural based classification, Multiple Process classification, Decision making classification, K-closest neighbors, Bayesian theory classification, decision Forest classification, support group machine, and Randomized Gradient Classification. By using the techniques mentioned above we can improve the harvest prediction using information mining techniques which in turn help the farmers to take better decisions in future and it can be used in other technologies like image analyzing etc. The Experimental results show predicted crop, suitable algorithm and algorithm accuracy in that particular state of India respectively.

**Keywords:** Agriculture, crop, climate, fertilizers, groundwater, Machine Learning, Classifiers, pesticides, soil.

## I. INTRODUCTION

Information mining is used for separating valuable and important data from huge information. Information mining are used to tell future harvest creation which will assist ranchers with taking most suitable choice for their yields.

Information mining and AI systems are used to consider the impact of different data and make expectations of the harvest creation, India has the most elevated creation of numerous yields and significant yields developed in India are Paddy, grain and corn. Based on various development seasons in India, harvests can be partitioned into summer, winter and seasonal crops. In India, Paddy is a summer and winter crop, which is developed between (June to December) for summer season and for winter season. Corn is a summer and winter crop, which is developed between for summer Season and for winter season. Grain is a winter crop, which is a spring harvest or winter crop and developed between [1]. Soil living beings are a crucial segment of living systems, yet their exercises get little information in rural area procedures. Here we organize the ability of soil living creatures to improve natural framework organization transport and display that earth biodiversity propels various condition limits simultaneously (i.e., organic framework multi convenience). The possibility of common acceleration to earth components and strategies for centred maltreatment of earth natural qualities. The promising methods were to manage improve agrarian supportability through the headway of soil biodiversity and concentrated on the leading group of soil Network association [2]. In controlled condition considers, warm temperatures extended the pace of phonological headway; regardless, there was no effect on leaf zone or vegetative biomass diverged from ordinary temperatures. The huge impact of more sweltering sunlight was during the ongoing phase of headway and in all cases grain yield was basically diminished from an ordinary temperature framework. Sunlight impacts are extended by liquid inadequacies and earth liquid indicating that analysing the coordinated effort of sunlight and liquid will be required to develop progressively amazing modification methods to offset the impacts of increasingly noticeable temperature preposterous events related with a changing air [3]. Groundwater is a significant water asset for rural water system and modern and household utilizes the world over, particularly in parched locales. China's Yellow River Basin (YRB) is an average bone-dry and semiarid region where farming water system relies significantly upon groundwater capacity [4]. African countries, for instance, were among the least made countries on the planet. Provincial activities are the major wellspring of their economy and enormous proportions of pest killing products have been imported for an impressive period of time to shield crops from hurt by vermin and sicknesses. A part of these pesticides ended in limit before they were used and their evacuation transformed into a huge issue.

Revised Manuscript Received on February 05, 2020.

\* Correspondence Author

**Albin Joseph**, PG Student, Dept. of Computer Science and Engineering, CHRIST( Deemed to be University), Bangalore, India Email: albin.joseph@mtech.christuniversity.in

**Dr Manohar M**, Associate Professor, Dept. of Computer Science and Engineering, CHRIST( Deemed to be University), Bangalore, India Email: manohar.m@christuniversity.in

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](https://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC BY-NC-ND license ([http://creativecommons.org/licenses/by-nc-nd/4.0/](https://creativecommons.org/licenses/by-nc-nd/4.0/))



Further, these countries report different systems related to pesticide the board [5].

The joined usage of manufactured and normal fertilizers is seen as an OK procedure to proceed with high collect yield and redesign soil common carbon (SCC), yet it is up 'til now indistinguishable when and to what degree creation manures could be superseded by regular composts [6]. The present preparation focuses on how crop prediction and data mining technique used in various states of India by taking some of the influencing parameters of soil, climate, pesticides, Fertilizers and groundwater.

## II. LITERATURE FOUNDATION

The author describes the Controlled traffic cultivating (CTF) confines traffic to perpetual paths, consequently making traffic free beds for crop generation [7]. The author describes about the soil measure tool that is planned for use in business soil analyzing research offices and usages fast, sharp methods [8]. The author describes Parasitic people group assume a basic job in directing soil supplement cycling and plant development. Be that as it may, the impacts of long haul mineral and natural treatment on parasitic networks in Ultisols remain to a great extent obscure [9]. The author describes the development of corn cultivators in agricultural zones of West Africa, a met outskirts architecture is utilized to assess the normal specialized efficiencies and corn profitability holes over the agro environmental zones [10]. The author describes Data mining is a rising field of research in Information Technology just as in horticulture [11]. The author describes information mining counts were used to make architecture to measure the consequence of leaf roller bug watching decisions on horticulture crops in Auckland [12]. The author describes Soil disintegration by water has impact on crop profitability [13]. The author describes Egypt is confronting an issue of nourishment frailty joined with destitution particularly in provincial Upper Egypt [14]. The author describes three SOM (self-Organizing based models), in particular, managed Kohonen systems (SKN), and counter-spread counterfeit systems (CP-ANN) and XY-combination (XY-F), which utilize Supervised Learning to relate high goals information on soil and harvest with is frequency classes of wheat yield efficiency were use [15]. The author describes Plant development advancing microscopic organisms give various advantages in cultivating by reproducing crop health and more, vitamins and smothering help in improving earth quality [16]. The author describes the importance of the Indian economy in horticulture, which is molded by the poor soil quality [17]. The author describes the examination included data collection from 100 paddy-developing nuclear families during the 2010 tempest season across five town tracts in Myanmar [18]. The author describes Programmed water framework in the earth liquid process device is enacted by using plant liquid weight and earth liquid lack liquid framework booking [19]. The author describes Impractical overexploitation of groundwater for rural water system has prompted quick groundwater exhaustion and serious natural harm in the semi-dry Hebei Plain of China [20]. The author describes the paper gives an intensive preparation of composing related to the assessment of ecological disaster on harvest change using air, water and gather yield architecture [21]. The author describes low-level farmers in east Africa face a wide extent of agrarian age threats. Natural change and

variability present new risks and vulnerabilities [22]. The author describes Environmental change is set to be especially problematic in poor agrarian networks. We evaluate the variables affecting ranchers' decision of environmental change adjustment rehearses and related effects on family unit nourishment security and neediness in Pakistan utilizing thorough information from 950 ranchers from its significant territories [23]. The author describes as indicated by the ebb and flow water emergency and spend in excess of 94 percent of water in farming the motorized water system frameworks, and changed the genuine plant water estimation are required it is encourage to foresee precipitation in the developing season [24]. The author describes viable precipitation assumes a colossal job in the development of harvests in farming during all phases of yield development [25]. The author describes Irregularity in atmosphere systems of precipitation is a wellspring of living components and non-living component stresses in horticultural frameworks around the world [26]. Climate significantly affects crop development and advancement. This paper exhibits an outline of yield demonstrating and uses of harvest models, and the noteworthiness of climate identified with these applications [27]. The author describes universe is experiencing a high weather extravagant and its disaster on every single normal structure, spaces and zones, yet the nature and level of insufficiency to organic change isolates from adaptable generally outrageous of individual, society likewise, zone [28]. The author describes Checking water quality with traditional techniques isn't a simple assignment. Remote detecting with wide inclusion and various fleeting checking is good choice for ground-water process [29]. The author describes Groundwater for agrarian reasons for existing is of most extreme significance in the Indian province of Gujarat [30]. The author describes Farming water protection is significant in parched and semi-bone-dry zones where water assets are inadequate and horticulture water system represents a high extent of water use [31]. The author describes this examination assesses the impact of common solid waste on its incorporating surface water Nature of town Varanasi, India [32]. The author describes to restrict and counter decline of surface water-levels and improve the availability of water for crop age, Water recharge mediations are commonly gotten across India [33]. The author describes the improvement of water system water accessible to producers has a great impact on harvest just as other harvest production [34]. The author describes Data mining is the way toward extricating valuable or concealed data from an enormous database [35]. The author describes Water shortage has consistently been a genuine test for horticulture and nourishment security. Consequently, various techniques are considered to spare more water and build water efficiency [36]. The author describes Vietnam is confronting genuine difficulties concerning the sum and poisonous quality of the pesticides utilized [37]. The author describes need to assess pesticide use in country populaces, especially in creating nations, is pressing. Pest killing and related dangerous observations were considered among large arbitrarily ranchers from few areas in cotton export of Chandigarh, Islamabad [38].

The author describes Evaluating dermal definition to pesticides in cultivating frameworks in creating nations is of unique intrigue for the estimation of potential wellbeing dangers, particularly when there is an absence of word related cleanliness guidelines. [39].

The author describes this article looks at both the executives rehearses prescribed to smallholder ranchers by augmentation operators, and the factors influencing the equivalent [40]. The author describes the absolute convergence of large pesticides estimated in needled leaves from large examining destinations in Spain [41]. The author describes pesticides are applied for securing the amount and nature of farming harvests. The national specialists ought to give cutting-edge, exact, and simple to comprehend data in the preparation of ranchers in the utilization of PPE (Personal Protective Equipment). This paper features the significance of PPE in pesticide applications in horticulture [42]. The author describes Chinese ranchers are frequently blamed for abusing pesticides that assume a significant job in improving harvest yield by decreasing misfortunes to trim nuisances [43]. The author describes washing agent got from increasing and decreasing of vegetable oil from genuinely refined vegetable oil is normally arranged off at dump disposal at a significant expense [44]. The author describes the impact of natural compost and synthetic manure on development and advancement of Sunflower plant had been examined in exploratory plots in Agribusiness University [45]. The author describes expanded nourishment request from the quickly developing human populace which escalated field progress from arid regions to harvest land in dry districts of Asian countries. The outcomes propose that in recently developed farmland, long haul natural excrement and incorporated preparation can produce advantage for soil chemical compound aggregation, what's more, convey extra influence on physical properties [46]. The author describes red plant is a annual plant of more financial worth.[47].The author describes an old field experiment was directed in Asian country regions to explore plant yield reaction to preparation exercise [48]. The author describes to improve soil accumulation through appropriate preparation is significant for upgrading soil quality and yield Profitability [49]. The author describes the execution of compelling administration practices, for example, joined utilization of compost and natural revisions prompted upgraded earth quality and related increments in earth chemical compounds sequestration in harvest editing framework in the Himalayan fields of India[50]. The author describes huge dry land territory of the Loess Plateau (China) is subject of creating techniques for a maintainable yield generation [51]. The author describes the use of nitrogen (N) composts in agribusiness has been expanding significantly since 1970s [52]. Based on the literature review discussed above from we proposed a model which is used to predict the crop based on the influential parameters of soil, climate, underground-water, pesticides and fertilizers etc.

### III. INFORMATION COLLECTION

This incorporates subtleties of the informational indexes, information pre-handling, information mining methods and building forecast model utilizing Python library.

#### A. Material collected

All the materials were collected from the available database of the government website of 25 states for range of 100 years

from 1920 to 2020, which includes yearly data of various parameters such as soil, climate, Pesticides, Fertilizers, groundwater and crop data etc. (Table I) shows us the all parameters used for crop prediction.

- **Soil (kg):** The acidity-basicity meter, carbon element, Electric sensitivity and chemical compounds of several regions of India was included for the examination.
- **Climate (degree):** Rainfall (mm), Humidity, Wind, Maximum temperature, Minimum temperature and Air Pressure of several regions of India was included for the examination.
- **Pesticides (kg):** The Malathion, Dimethoate, DDVP, Copper Oxy-chloride, Carbaryl, Carbosulphan of several regions of India was included for the examination.
- **Fertilizers (kg):** NPK, MAP, MOP, TSP, Urea and DAP several regions of India was included for the examination.
- **Groundwater (mgbl):** Rise (depth of water), Fall (depth of water), Rise (Fluctuation), Fall (Fluctuation), pH, Electrical Conductivity (EC) of several regions of India was included for the examination.
- **Crop(harvest/ton):** 25 crop data which includes Rice, Wheat, Jowar, Bajra, Maize, Ragi, Small Millets, Barley, Tur, Gram, Other Pulses, Groundnuts, Sessasum, Sugarcane, Cotton, Jute, Mesta, Tobacco, Tea, Banana, Potato, Coconut, Pepper, Chilies, Turmeric and Ginger of several regions of India was included for the examination.

#### B. Material Pre-processing

Labeled training data is one of the machine learning technique which allows us to collect the data and predict new data from the previous data. Classification algorithm is one of the machine Learning/Data mining that takes the including data's so that the excluding data separates one phase into good response and the other into bad response. Considering the dataset the input variable is Soil, Climate, Pesticides, Fertilizers and Groundwater and output (target variable) is Crop dataset. Classification algorithms were used on Soil, Climate, Pesticides, Fertilizers and Groundwater parameters to predict the crop from the Crop data set of duration (1920-2019).Classification algorithms used in this Research were Adaptive boosting classification, Excess tree classification, Multiple Process classification, neurological classification, Decision making classification, K-closest neighbors, Bayesian theory classification, decision Forest classification, support group machine, and Randomized Gradient Classification. The Section discusses about the prediction model or architecture model for Classification algorithms using Scikit-learn. Python Library is an open resource database for Data digging and Data exploration via continuous change in python .It is licensed and certified and includes labeled and unlabeled algorithm.



**Table- I: Parameters for Crop Prediction**

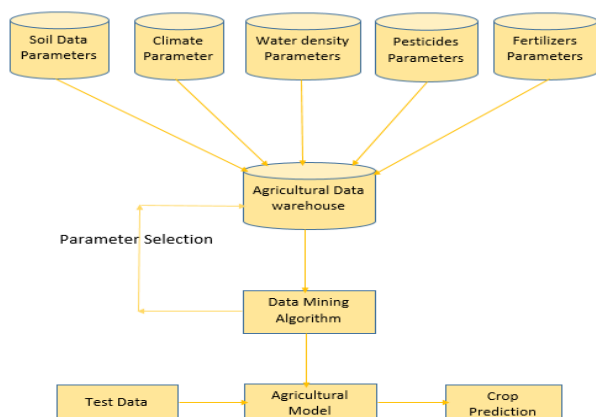
| Parameter description | Climate                                      | Soil | Groundwater          | Fertilizers | Pesticides          | Crops |
|-----------------------|--|------|----------------------|-------------|---------------------|-------|
| Parameter one         | Rainfall                                     | pH   | Rise(depth of water) | NPK         | Malathion           | all   |
| Parameter two         | Humidity                                     | OC   | Fall(depth of water) | MAP         | Dimethoate          | all   |
| Parameter three       | Wind   | EC   | Rise(Fluctuation)    | MOP         | DDVP                | all   |
| Parameter four        | Max temp                                     | N    | Fall(Fluctuation)    | TSP         | Copper Oxylchloride | all   |
| Parameter five        | Min temp                                     | P    | pH                   | Urea        | Carbaryl            | all   |
| Parameter six         | Air Pressure                                 | K    | EC                   | DAP         | Carbosulphan        | all   |
|                       |  |      |                      |             |                     |       |
| Parameter description |  |      |                      |             |                     |       |
| all                   | Supports all the above parameter description |      |                      |             |                     |       |
| Crops                 |  |      |                      |             |                     |       |
| all                   | Includes all the Crop data                   |      |                      |             |                     |       |

## IV. METHODOLOGY

### A. Problem Statement

The objective of Project is to come with an idea of predicting a crop in several regions of Indian states using data mining algorithms. Influential data's which alter growth in crops for a particular regions, we will prepare those factors and select the influencing factors with the help of data mining algorithms. We finally will predict the crops for all regions, factors influencing, data mining algorithms and accuracy value.

### B. Architectural Model



**Diag.1 Architecture of crop Prediction**

(Diagram 1) shows the architecture of crop prediction in which we have soil data, climate, water density, Pesticides and Fertilizers parameters. We take different input parameters of soil ,climate, water density ,pesticides and Fertilizers and store it in the agricultural data warehouse ,then we apply data mining classification and prediction algorithms to check whether the whether the input parameters are supporting or not ,if not supporting then we will go for the reselection of the parameters ,if it is supporting then we take it as the test data and build agricultural model and predict the crop .we will do this process with several data mining algorithms and select the best data mining algorithm and crop predicted from all in different states of India which will more yield in the future .

### C. Algorithm Justification

In our proposed model we have used ten data mining classification algorithms which are as follows Adaptive boosting classification, Excess tree classification, Multiple

Process classification, neurological based classification, Decision making classification, K-closest neighbors, Bayesian theory classification, decision Forest classification, support group machine, and Randomized Gradient Classification respectively. We will be justifying the importance of each algorithm. The author describes the Yield forecast methods is used to see the suitable data by matching out different data of soil and more data that is identified with environment. Database which includes earth component, acidity-basicity meter, chemical compounds, precipitation and moistness. For that cause we copy neurological based system [54].The author describes Adaptive Boosting, or Adaptive, is a calculation planned for improving the exhibition of gatherings of frail students by gauging the information itself just as the students. Two forms of Adaptive boosting classification were applied, and other calculation joining were also included in this process, for anticipating harvest and mineral content on the side accuracy horticulture. Every one of the three calculations use input Neurological based system prepared for back propagation as the frail architecture. Information from different areas were assembled because of on-ranch tests of various nitrogen rate applications utilizing haphazardly stratified preliminaries dependent on earlier annual harvest and minerals. The three Adaptive boosting classification calculations are combined with a basic neurological classification with a solitary shrouded base. Outcomes affirm past discoveries in various fields, where troupe strategies beat single models. Outcomes are improved by medium units for harvest expectation, through little rate for protein forecast [55].The author describes a certified pondering of experiments and their presentation in harvest expectation in similar farming. Experiments are actualized in a database entry obtained after large duration, while in harvest expectation for one crop. The Experiments utilized for harvest forecast supported Support group Machine, decision Forest, Neurological based classification and ensemble methods. The final data obtained is the best experiment for harvest [56].The author describes Agribusiness is brimming with vulnerability because of environmental change, ground water insufficiency, precipitation and advancement of new irritations. Harvest yield expectation in agribusiness is an exceptionally enormous difficulty and there is colossal dataset where ranchers discover hard to anticipate the yield what's more, seed choice. In the present circumstance because of increment in the populace the generation of grains and horticultural items needs to be expanded at the same time to satisfy the needs of the individuals. Past cultivating procedures should be joined with the present day innovation to improve the outcomes. A few Environment and monetary variables like atmosphere, downpour and pesticide are capricious that influences the yield development which thus influences the efficiency. Measurable and scientific apparatuses can be utilized to measure the yield utilizing past information. The outcome can help ranchers in crop decision and give a knowledge of the efficiency that outcome in expanded benefit. Informational index has been gathered from Tamil Nadu and factual sites for searching and Machine learning calculations and their job in farming were searched.

This paper center around the Machine learning calculations like Naïve Bayesian, Support vector machine, neural systems, Decision tree, K Nearest Neighbor in crop yield forecast [57]. The author describes the importance of SGD(Stochastic Gradient Descent) classifier, Stochastic Gradient Descent, the whole info set isn't considered. Any arbitrary quantities of info factors are considered, which the real number of info factors are not exactly [58]. Rural uses of the Gaussian process classifier (GPC) based strategies is considered. A strategy of ordering crops from multi-worldly Landsat 8 satellite symbolism is proposed. The strategy utilizes the model of unearthy highlights dependent on GP relapse with steady desire and square exponential covariance capacities. Principle steps of the classification strategy and instances of acknowledgment of culture species are spoken to. The author describes ground based information are utilized for quantitative approval of the proposed classification technique. The most elevated generally speaking classification precision in three classes of yields is 77.78% [59]. An extra-trees classifier. This class actualizes a Meta estimator that fits various randomized choice trees (a.k.a. extra-trees) on different sub-tests of the dataset and utilizations averaging to improve the prescient exactness and authority over-fitting.

## V. INFORMATION PROCESSING

Obtained data represents the data which we got for Crop prediction and Test data prediction which includes various states of India, predicted crop, best algorithm, best parameter and accuracy of algorithm highlighted in (Table II) and (Table III). (Table IV) highlights the best crop predicted (2020), test data Crop predicted, best algorithm and accuracy of both predicted Crop and Test Data Crop used respectively. (Diagram 2) shows the graphical process of Crop prediction and Test data Prediction of various states of India.

**Table- II: Processing of Crop Prediction**

| CROP PREDICTION(2020) |            |                |           |          |
|-----------------------|------------|----------------|-----------|----------|
| STATE                 | CROP       | BEST ALGORITHM | PARAMETER | ACCURACY |
| Kerala                | Rubber     | Naive Bayes    | 1,2,3,5,6 | 99.00%   |
| Andhra Pradesh        | Ragi       | ANN            | 1,2,5,6   | 98.00%   |
| Arunachal Pradesh     | Banana     | J48            | all       | 99.00%   |
| Assam                 | Tumeric    | SVM            | all       | 99.00%   |
| Bihar                 | Maize      | ETC            | all       | 99.00%   |
| Chattisgarh           | Rice       | Adaboost       | all       | 96.00%   |
| Goa                   | Coconut    | Random Forest  | all       | 99.00%   |
| Gujarat               | Cotton     | Random Forest  | all       | 99.00%   |
| Haryana               | Pulses     | GPC            | all       | 57.00%   |
| Himachal Pradesh      | Barley     | Adaboost       | all       | 99.00%   |
| Jammu and Kashmir     | Millet     | Naive Bayes    | all       | 99.00%   |
| Karnataka             | Ragi       | Adaboost       | all       | 99.00%   |
| Jharkhand             | Rice       | ANN            | all       | 99.00%   |
| Madhya Pradesh        | Sugarcane  | J48            | all       | 99.00%   |
| Maharashtra           | Jowar      | ANN            | all       | 99.00%   |
| Meghalaya             | Pepper     | Adaboost       | all       | 96.00%   |
| Orissa                | Tumeric    | KNN            | all       | 88.00%   |
| Punjab                | Chillies   | SGD            | all       | 70.00%   |
| Rajasthan             | Chillies   | GPC            | all       | 88.00%   |
| Sikkim                | Potato     | KNN            | all       | 88.00%   |
| Tamil Nadu            | Ginger     | Naive Bayes    | all       | 99.00%   |
| Tripura               | Jute       | KNN            | all       | 99.00%   |
| Uttar Pradesh         | Sugarcane  | Adaboost       | all       | 99.00%   |
| Uttarakhand           | Groundnuts | KNN            | 2,3,4,5,6 | 99.00%   |
| West Bengal           | Tea        | J48            | all       | 88.00%   |

**Table- III: Processing Test Data Prediction**

| TEST DATA PREDICTION(1950-1980) |            |                |           |          |
|---------------------------------|------------|----------------|-----------|----------|
| STATE                           | CROP       | BEST ALGORITHM | PARAMETER | ACCURACY |
| Kerala                          | Rubber     | Naive Bayes    | 1,2,3,4   | 99.00%   |
| Andhra Pradesh                  | Ragi       | ANN            | 2,4       | 90.00%   |
| Arunachal Pradesh               | Banana     | J48            | all       | 99.00%   |
| Assam                           | Tumeric    | SVM            | all       | 99.00%   |
| Bihar                           | Maize      | ETC            | all       | 99.00%   |
| Chattisgarh                     | Rice       | Adaboost       | all       | 91.00%   |
| Goa                             | Coconut    | Random Forest  | all       | 99.00%   |
| Gujarat                         | Cotton     | Random Forest  | all       | 99.00%   |
| Haryana                         | Pulses     | GPC            | all       | 54.00%   |
| Himachal Pradesh                | Barley     | Adaboost       | all       | 99.00%   |
| Jammu and Kashmir               | Millet     | Naive Bayes    | all       | 99.00%   |
| Karnataka                       | Ragi       | Adaboost       | all       | 90.00%   |
| Jharkhand                       | Rice       | ANN            | all       | 99.00%   |
| Madhya Pradesh                  | Sugarcane  | J48            | all       | 99.00%   |
| Maharashtra                     | Jowar      | ANN            | all       | 99.00%   |
| Meghalaya                       | Pepper     | Adaboost       | all       | 99.00%   |
| Orissa                          | Tumeric    | KNN            | all       | 99.00%   |
| Punjab                          | Chillies   | SGD            | all       | 90.00%   |
| Rajasthan                       | Chillies   | GPC            | all       | 99.00%   |
| Sikkim                          | Potato     | KNN            | all       | 99.00%   |
| Tamil Nadu                      | Ginger     | Naive Bayes    | all       | 99.00%   |
| Tripura                         | Jute       | KNN            | all       | 99.00%   |
| Uttar Pradesh                   | Sugarcane  | Adaboost       | all       | 90.00%   |
| Uttarakhand                     | Groundnuts | KNN            | 2,3,4,5,6 | 99.00%   |
| West Bengal                     | Tea        | J48            | all       | 99.00%   |

**Table- IV: Processing of Crop Prediction and Test Data Prediction**

| STATE             | PREDICTED CRO | TEST CROP  | BEST ALGORITHM | ACCURACY PREDICTED | ACCURACY TEST |
|-------------------|---------------|------------|----------------|--------------------|---------------|
| Kerala            | Rubber        | Rubber     | Naive Bayes    | 99.00%             | 99.00%        |
| Andhra Pradesh    | Ragi          | Ragi       | ANN            | 98.00%             | 90.00%        |
| Arunachal Pradesh | Banana        | Banana     | J48            | 99.00%             | 99.00%        |
| Assam             | Tumeric       | Tumeric    | SVM            | 99.00%             | 99.00%        |
| Bihar             | Maize         | Maize      | ETC            | 99.00%             | 99.00%        |
| Chattisgarh       | Rice          | Rice       | Adaboost       | 96.00%             | 91.00%        |
| Goa               | Coconut       | Coconut    | Random Forest  | 99.00%             | 99.00%        |
| Gujarat           | Cotton        | Cotton     | Random Forest  | 99.00%             | 99.00%        |
| Haryana           | Pulses        | Pulses     | GPC            | 57.00%             | 54.00%        |
| Himachal Pradesh  | Barley        | Barley     | Adaboost       | 99.00%             | 99.00%        |
| Jammu and Kashmir | Millet        | Millet     | Naive Bayes    | 99.00%             | 99.00%        |
| Karnataka         | Ragi          | Ragi       | Adaboost       | 99.00%             | 90.00%        |
| Jharkhand         | Rice          | Rice       | ANN            | 99.00%             | 99.00%        |
| Madhya Pradesh    | Sugarcane     | Sugarcane  | J48            | 99.00%             | 99.00%        |
| Maharashtra       | Jowar         | Jowar      | ANN            | 99.00%             | 99.00%        |
| Meghalaya         | Pepper        | Pepper     | Adaboost       | 96.00%             | 99.00%        |
| Orissa            | Tumeric       | Tumeric    | KNN            | 88.00%             | 99.00%        |
| Punjab            | Chillies      | Chillies   | SGD            | 70.00%             | 90.00%        |
| Rajasthan         | Chillies      | Chillies   | GPC            | 88.00%             | 99.00%        |
| Sikkim            | Potato        | Potato     | KNN            | 88.00%             | 99.00%        |
| Tamil Nadu        | Ginger        | Ginger     | Naive Bayes    | 99.00%             | 99.00%        |
| Tripura           | Jute          | Jute       | KNN            | 99.00%             | 99.00%        |
| Uttar Pradesh     | Sugarcane     | Sugarcane  | Adaboost       | 99.00%             | 90.00%        |
| Uttarakhand       | Groundnuts    | Groundnuts | KNN            | 99.00%             | 99.00%        |
| West Bengal       | Tea           | Tea        | J48            | 88.00%             | 99.00%        |



**Diag.2 Graphical Processing of Crop Prediction and Test Data Prediction of various states of India**

## VI. CONCLUSION

Classification algorithm were taken as future modelling for future of crop production. Ten classification algorithms namely were Adaptive boosting classification, Excess tree classification, Multiple Process classification, neurological classification, Decision making classification, K-closest neighbors, Bayesian theory classification, decision Forest classification, support group machine, and Randomized Gradient Classification have been used. From the above experimental results we obtained the results of Crop Predicted and best algorithm for various states of India, and we also did checking with test data crop, after obtaining the Graphical Processing for predicted crop and test data crop we observed that out of 25 states of India 14 states showed very good accuracy with 99 percent, 10 states showed average accuracy between 70 and 99 percent and only one state with poor accuracy's of 57 and 54 percent. The states with very good accuracies were Kerala, Arunachal Pradesh, Assam, Bihar, Goa, Gujarat, Himachal Pradesh, Jammu and Kashmir, Jharkhand, Madhya Pradesh, Maharashtra, Tamil Nadu, Tripura and Uttarakhand, average accuracies were Andhra Pradesh, Chhattisgarh, Karnataka, Meghalaya, Orissa, Punjab, Rajasthan, Sikkim, Uttar Pradesh and West Bengal, and state which had poor accuracies was Haryana. It was also observed that out of the ten classification algorithms, nine classification algorithms showed good accuracies which are Naïve Bayes, J48, SVM (Support Vector Machine), and Extra-Tree Classifier (ETC), Random Forest, Ada-Boost, ANN (Artificial Neural Network), K-nearest neighbor (KNN) for different states and one algorithm with poor accuracy which is Gaussian Process Classifier for State Haryana respectively. While comparing with other research methods for crop prediction, researchers couldn't able to do prediction for all crops as they only predicted for certain crop which were Rice, Wheat and Maize as described by the author [1].



In our proposed model for the prediction of crop, we were able to predict crops for each states of India along with best algorithm used, one of challenges faced was obtaining the crop data for all states, and drawback of our model was we couldn't able to find crop data and other parameter data for states like Mizoram, Nagaland, Telangana, Manipur respectively. In future we will be using this proposed model to predict crop in other Countries.

## ACKNOWLEDGMENT

This research work is carried out in School of Engineering and Technology, Christ (Deemed to be University). I am thankful to Dr Manohar M (Guide, Associate Professor) for his valuable efforts that helped me in the Completion of Research Journal, and I am also thankful to Mithun BN (Assistant Professor) for assisting in collection of Data Sets in Research work.

## REFERENCES

1. Suvidha Jambekar, Shikha Nema, Zia Saquib, "Prediction of Crop Production in India Using Data Mining Techniques", 2018 Fourth International Conference on Computing Communication Control and Automation (ICCCUBEA), India.
2. S. Franz Bender, Cameron Wagg Marcel G.A. van der Heijden, "An Underground Revolution: Biodiversity and Soil Ecological Engineering for Agricultural Sustainability", 2016, Cell Press, Trends in Ecology Evolution.
3. Jerry L. Hatfield, John H. Prueger, "Temperature extremes: Effect on plant growth and development", 2015, Elsevier, Weather and Climate Extremes, 2110 University Blvd, Ames, IA, USA.
4. Kang Zhang, Xianhong Xie, Bowen Zhu, Shanshan Meng, Yi Yao "Unexpected groundwater recovery with decreasing agricultural irrigation in the Yellow River Basin", 2018, Elsevier, Agricultural Water Management, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China.
5. Kumelachew Mulu Loha, Marja Lamoree, Jana M. Weiss, Jacob de Boer, "Import, disposal, and health impacts of pesticides in the East Africa Rift (EAR) zone: A review on management and policy analysis", 2018, Elsevier, Crop Protection 112, Dept. Environment Health, Vrije Universiteit, De Boelelaan 1085, 1081 HV Amsterdam, The Netherlands, Dept. Aquatic Sciences and Assessment, Swedish University of Agricultural Sciences (SLU), Box 7050, 750 07 Uppsala, Sweden.
6. [6] LI Hui, FENG Wen-ting, HE Xin-hua, ZHU Ping, GAO Hong-jun, SUN Nan, XU Ming-gang, "Chemical fertilizers could be completely replaced by manure to maintain high maize yield and soil organic carbon (SOC) when SOC reaches a threshold in the Northeast China Plain", 2017, Elsevier, Journal of Integrative Agriculture, China.
7. LI Hui, FENG Wen-ting, HE Xin-hua, ZHU Ping, GAO Hong-jun, SUN Nan, XU Ming-gang, "Chemical fertilizers could be completely replaced by manure to maintain high maize yield and soil organic carbon (SOC) when SOC reaches a threshold in the Northeast China Plain", 2017, Elsevier, Journal of Integrative Agriculture, China.
8. Richard L. Haney, Elizabeth B. Haney, Douglas R. Smith, R. Daren Harmel, Mike J. White "The soil health tool—Theory and initial broad-scale application", 2017, Elsevier, Applied Soil Ecology, USA.
9. Guiping Ye, Yongxin Lin, Jiafa Luo, Hong J. Di, Stuart Lindsey, Deyan Liu, Jianbo, Fan Weixin Ding, "Responses of soil fungal diversity and community composition to long-term fertilization: Field experiment in an acidic Ultisol and literature synthesis", 2019, Elsevier, Applied Soil Ecology, China.
10. Bright O. Asante, Omphile Temoso, Kwabena N. Addai, Renato A. Villano – "Evaluating productivity gaps in maize production across different agroecological zones in Ghana", 2019, Elsevier, Agricultural Systems, USA.
11. Rupanjali D. Baruah, R.M. Bhagat, Sudipta Roy, L.N. Sethi, "Use of Data Mining Technique for Prediction of Tea Yield in the Face of Climate Change of Assam, India", 2016, IEEE, International Conference on Information Technology, Assam, India.
12. M.G. Hill, P.G. Connolly, P. Reutemann, D. Fletcher "The use of data mining to assist crop protection decisions on kiwifruit in New Zealand", 2014, Elsevier, Hamilton 3240, New Zealand.
13. V.N. Sharda, P. Dogra, C. Prakash, "Assessment of Production losses due to water erosion in rainfed areas of India", 2010, Journal of Soil and Water Conservation.
14. Ayman E. Khedr, Mona Kadry, Ghada Walid "Proposed framework for

- implementing data mining techniques to enhance decisions in agriculture sector", 2015, Elsevier, Ministry of Agriculture, Egypt.
15. X.E. Pantazi, D. Moshou, T. Alexandridis, R.L. Whetton, A.M. Mouazen "Wheat yield prediction using machine learning and advanced sensing techniques", 2015, Elsevier, Computers and Electronics in Agriculture, UK.
16. Wusirika Ramakrishna, Radheshyam Yadav, Kefeng Li "Plant growth promoting bacteria in agriculture: Two sides of a coin", 2019, Elsevier, Applied Soil Ecology, India.
17. M.S. Sirsat, E. Cernadas, M. Fernández-Delgado, R. Khan "Classification of agricultural soil parameters in India", 2017, Elsevier, Computers and Electronics in Agriculture, India.
18. Hla Myo Thwe, Paul Kristiansen, David F. Herridge "Benchmarks for improved productivity and profitability of monsoon rice in lower Myanmar", 2019, Elsevier, Field Crops Research, NSW, 2351 Australia.
19. Bhumiika Uniyal, Jörg Dietrich "Modifying Automatic Irrigation in SWAT for Plant Water Stress scheduling", 2019, Elsevier, Agricultural Water Management, Institute of Hydrology and Water Resources Management, Universität Hannover, Germany.
20. Honglin Zhong, Laixiang Sun, Günther Fischer, Zhan Tian, Zhuoran Liang "Optimizing regional cropping systems with a dynamic adaptation strategy for water sustainable agriculture in the Hebei Plain", 2019, Elsevier, Agricultural Systems, USA.
21. Yinhong Kang, Shahbaz Khan, Xiaoyi Ma "Climate change impacts on crop yield, crop water productivity and food security – A review", 2009, Elsevier, Progress in Natural Science 19, China.
22. Drake N. Mubiru, Maren Radeny, Florence B. Kyazze, Ahamada Zziwa, James Lwasa, James Kinyangi, Catherine Mungai "Climate trends, risks and coping strategies in smallholder farming systems in Uganda", 2018, Elsevier, Climate Risk Management, Uganda.
23. Akhter Ali, Olaf Erenstein "Assessing farmer use of climate change adaptation practices and impacts on food security and poverty in Pakistan", 2016, Elsevier, Climate Risk Management, Pakistan.
24. Kaveh Ostad-Ali-Askari, Mohammad Shayannejad and Hossein Ghorbanizade Kharazi "ASSESSMENT OF ARTIFICIAL NEURAL NETWORK PERFORMANCE AND EXPONENTIAL REGRESSION IN PREDICTION OF EFFECTIVE RAINFALL", 2015, International Journal of Development Research, Iran.
25. Abishek B, R. Priyatharshini, Akash Eswar M, P. Deepika "Prediction of Effective Rainfall and Crop Water Needs using Data Mining Techniques", 2017, IEEE, International Conference on Technological Innovations in ICT For Agriculture and Rural Development, Chennai, India.
26. Md Abiar Rahman, SuChul Kang, Nidhi Nagabhatla and Robert Macnee "Impacts of temperature and rainfall variation on rice productivity in major ecosystems of Bangladesh", 2017, Agriculture Food Security.
27. Gerrit Hoogenboom "Contribution of agro meteorology to the simulation of crop production and its applications", 2000, Elsevier, Department of Biological and Agricultural Engineering, the University of Georgia, Griffin, GA, USA.
28. Ram Babu Singh and Ajay Kumar "Climate variability and water resource scarcity in drylands of Rajasthan, India", 2015, Springer.
29. Ersan Batur and Derya Maktav "Assessment of Surface Water Quality by Using Satellite Images Fusion Based on PCA Method in the Lake Gala, Turkey", 2018, IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING.
30. Pennan Chinnasamy, Gourav Misra, Tushaar Shah, Basant Maheshwari, Sanmugam Prathapar "Evaluating the effectiveness of water infrastructures for increasing groundwater recharge and agricultural production – A case study of Gujarat, India", 2015, Elsevier.
31. Xiaoyu Gao, Yining Bai, Zailin Huo, Xu Xu, Guanhua Huang, Yuhong Xia, Tammo S. Steenhuis "Deficit irrigation enhances contribution of shallow groundwater to crop water consumption in arid area", 2017, Elsevier, China.
32. Sachin Mishra, Dhanesh Tiwary, Anurag Ohri, Ashwani Kumar Agnihotri "Impact of Municipal Solid Waste Landfill leachate on groundwater quality in Varanasi, India", 2019, Elsevier.
33. S. Prathapar, S. Dhar, G. Tamma Rao, B. Maheshwari "Performance and impacts of managed aquifer recharge interventions for agricultural water security: A framework for evaluation", 2015, Elsevier, India.

34. Kishan S. Rawat , Sudhir Kumar Singh , Sandeep Kumar Gautam "Assessment of groundwater quality for irrigation use: a peninsular case study" ,2018, Springer,India.
35. Faisal Aburub, Wael Hadi "Predicting Groundwater Areas Using Data Mining Techniques: Groundwater in Jordan as Case Study" ,2016, International Journal of Computer and Information Engineering.
36. Leila Goli Raeisi, Saeed Morid, Majid Delavar, Raghavan Srinivasan, "Effect and side-effect assessment of different agricultural water saving measures in an integrated framework",2019, Elsevier,India.
37. Pham Van Hoi , Arthur Mol, Peter Oosterveer "State governance of pesticide use and trade in Vietnam" ,2013, Elsevier, Hanoi University of Agriculture, Gialam Hanoi, Vietnam.
38. Muhammad Khan , Hafiz Zahid Mahmood , Christos A. Damalas , "Pesticide use and risk perceptions among farmers in the cotton belt of Punjab, Pakistan" ,2014, Elsevier, Department of Management Sciences, COMSATS Institute of Information Technology, Lahore, Pakistan.
39. Camilo Lesmes-Fabian , Glenda Garc'ia-Santos , Fanny Leuenberger , David Nuytens , Claudia R. Binder, "Dermal exposure assessment of pesticide use: The case of sprayers in potato farms in the Colombian highlands" ,2012, Elsevier, Switzerland.
40. Willis N. Ochilo, Miriam Otupa, MaryLucy Oronje, Florence Chege, Eunice K. Lingeera, Evelyn Lusenaka, and Edward O. Okonjo, "Pest Management Practices Prescribed by FrontlineExtension Workers in the Smallholder Agricultural Subsector of Kenya" ,2018, Journal of Integrated Pest Management.
41. Nuno Ratola , Vera Homem , José Avelino Silva , Rita Araújo , José Manuel Amigo , Lúcia Santos , Arminda Alves , "Bio monitoring of pesticides by pine needles — Chemical scoring, risk of exposure, levels and trends" ,2014, Elsevier, Spain.
42. Nigar Yarpuz-Bozdogan, "The importance of personal protective equipment in pesticide applications in agriculture" ,2018, Elsevier, Turkey.
43. ZHANG Chao, Guanming Shi , SHEN Jian, HU Rui-fa, " Productivity effect and overuse of pesticide in crop production in China" ,2015, Elsevier.
44. Soh Kheang Loh , Stephen James , Muzzamil Ngatiman, Kah Yein Cheong , Yuen May Choo , Weng Soon Lim, " Enhancement of palm oil refinery waste – Spent bleaching earth (SBE) into bio organic fertilizer and their effects on crop biomass growth" ,2013, Elsevier, Malaysia.
45. Xiangyang Liu, Guangxi Ren, Yan Shi "The effect of organic manure and chemical fertilizer on growth and development of Stevia rebaudiana Bertonii" ,2011, Elsevier, China.
46. YANG Rong, SU Yong-zhong, WANG Tao, YANG Qin "Effect of chemical and organic fertilization on soil carbon and nitrogen accumulation in a newly cultivated farmland" ,2016, Elsevier, China.
47. Sigit Soeparjono "The Effect of Media Composition and Organic Fertilizer Concentration on The Growth and Yield of Red Ginger Rhizome",2015, Elsevier, Faculty of Agriculture University of Jember, Indonesia..
48. Zhenwei Song, Hongjun Gao, Ping Zhu, Chang Peng, Aixing Deng, Chengyan Zheng, Md Abdul Mannaf, Md Nurul Islam, Weijian Zhang, "Organic amendments increase corn yield by enhancing soil resilience to climate change" ,2015, Elsevier,China.
49. Z.C. Guo, Z.B. Zhang, H. Zhou, M.T. Rahman, D.Z. Wang, X.S. Guo, L.J. Li, X.H. Peng, "Long-term animal manure application promoted biological binding agents but not soil aggregation in a Vertisol" ,2018, Elsevier,China.
50. Subhadip Ghosh, Brian Wilson, Subrata Ghoshal, Nimai Senapati , Biswapati Mandal , "Organic amendments influence soil quality and carbon sequestration in the Indo-Gangetic plains of India" ,2012, Elsevier, India.
51. Xueyun Yang, Pingru Li, Shulan Zhang, Benhua Sun, and Chen Xin-ping, "Long-term-fertilization effects on soil organic carbon, physical properties, and wheat yield of a loess soil" ,2010.
52. Cui Xin, Yan Qing-wei, Sun Jia-lin, Xiao Shuang, Xie Fu-chun, and Chen Ya-jun, "Research Progress on Nitrogen Use and Plant Growth",2013, Journal of Northeast Agricultural University ,China.
53. Vuong M. Ngo, Nhien-An Le-Khac, M-Tahar Kechadi, "An Efficient Data Warehouse for Crop Yield Prediction" ,2018, 14th International Conference on Precision Agriculture Montreal, Quebec, Canada.
54. Miss.Snehal S.Dahikar, Dr.Sandeep V.Rode, "Agricultural Crop Yield Prediction Using Artificial Neural Network Approach", 2014, international journal of innovative research in electrical, electronics, instrumentation and control engineering,India.
55. Amy Peerlinck, John Sheppard, Jacob Senecal, "AdaBoost with Neural Networks for Yield and Protein Prediction in Precision Agriculture", International Joint Conference on Neural Networks, 2019, Hungary.
56. Rohit Kumar Rajak, Ankit Pawar, Mitalee Pendke , Pooja Shinde, Suresh Rathod, Avinash Devare, "Crop Recommendation System to

Maximize Crop Yield using Machine Learning Technique",2017, International Research Journal of Engineering and Technology, India.

57. k. porchilambi, dr. p. sumitra, "machine learning algorithms for crop yield prediction: a survey", 2019, jetir, Elayampalayam.
58. tanmay banavlikar, aqsa mahir, mayuresh budukh, soham dhodapkar, "crop recommendation system using neural networks",2018, international research journal of engineering and technology.
59. Anastasiia N. Safonovaa and Yegor V. Dmitrievb, " Classification of Agricultural Crops from Middle-Resolution Satellite Images Using Gaussian Processes Based Method",2018, Journal of Siberian Federal University.

## AUTHORS PROFILE



**Albin Joseph**, was brought in Kerala and currently lives in Bangalore, Karnataka. He did his Schooling in Delhi. He was qualified Bachelor in CSE in the year 2016 at Anna University, Chennai. He is currently pursuing M.tech in Computer Science at Christ (Deemed to be University), Bangalore. He worked as a Navision trainee in Corporate Serve Solutions Private Limited for 6 months. He has completed Project on A mechanism to Reduce Packet Delay in Wireless Transmission in Anna University, Chennai in the year 2016. He has also completed Project on Color Detection in images using OpenCv and Python in Oscki Labs, Bangalore in the Year 2019.



**Manohar M**, was brought up in Udupi and currently he lives in Bangalore, Karnataka. He was qualified Bachelor in CSE in the year 1999 at Dr AIT, Bangalore.. He has his Master in CSE in the year 2004 at Dr AIT, Bangalore. He completed his Ph.D. in CSE in the area of data mining and big data in the year 2016 at Jain University, Bangalore. He is an Associate Professor in the Computer Science and Engineering Department at Faculty of engineering of CHRIST (Deemed to be University), Bangalore. He is an educator by choice and vocation, with an experience of 19 years in Teaching. He has many publications in both conference and Journals. He is a member of Professional body IACSIT, IET.