iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** VII **Month of publication:** July 2022

DOI: https://doi.org/10.22214/ijraset.2022.45564

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue VII July 2022- Available at www.ijraset.com

Recommendation of Crops Based on Season and Productivity using Decision Tree

Ashwini M N[1], Dr. Shankaragowda B. B[2]
[1, 2]*Department of Master of Computer Application, BIET, Davangere*

Abstract: ***Crop recommendation system or prediction system is the art of predicting crop yields to improve the production and production before the harvest actually takes place, it takes typically a couple of months in advance. Crop prediction depends on the computer programs that describe the plant-environment and the soil features interactions in quantitative terms. The soil testing will start with the collections of a soil sample from the field. The first basic principles of the soil testing is that a field can be sampled in such a way that by getting a chemical analysis of the soil sample and also majorly depend on temperature and rainfall will accurately reflect the field's true nutrient status on a particular area to help out farmers to improve the production.***
Keywords: Crop Recommendation, Rainfall, Chemical analysis and Soil Testing.

I. INTRODUCTION

India is one among the oldest countries which is still practicing agriculture. But in recent times the trends in agriculture has drastically evolved due to globalization. Various factors have affected the health of agriculture in India. Many new technologies have been evolved to regain the health. One such technique is precision agriculture. Precision agriculture is budding in India. Precision agriculture is the technology of "site-specific" farming. It has provided us with the advantage of efficient input, output and better decisions regarding farming. Although precision agriculture has delivered better improvements it is still facing certain issues. There exist many systems which propose the inputs for a particular farming land. Systems propose crops, fertilizers and even farming techniques. Recommendation of crops is one major domain in precision agriculture. Recommendation of crops is dependent on various parameters.

Agriculture makes a dramatic impact in the economy of a country. Due to the change of natural factors, Agriculture farming is degrading now-a-days. Agriculture directly depends on the environmental factors such as sunlight, humidity, soil type, rainfall, Maximum and Minimum Temperature, climate, fertilizers, pesticides etc. Knowledge of proper harvesting of crops is in need to bloom in Agriculture. India has seasons of

1) Winter which occurs from December to March
2) Summer season from April to June
3) Monsoon or rainy season lasting from July to September and
4) Post-monsoon or autumn season occurring from October to November.

Due to the diversity of season and rainfall, assessment of suitable crops to cultivate is necessary. Farmers face major problems such as crop management, expected crop yield and productive yield from the crops. Farmers or cultivators need proper assistant regarding crop cultivation as now-a-days many fresh youngsters are interested in agriculture.

Impact of IT sector in assessing real world problem is moving at a faster rate. Data is increasing day by day in field of agriculture. With the advancement in Internet of Things, there are ways to grasp huge data in field of Agriculture. There is a need of a system to have obvious analyses of data of agriculture and extract or use useful information from the spreading data.

Impact of IT sector in assessing real world problem is moving at a faster rate. Data is increasing day by day in field of agriculture. With the advancement in Internet of Things, there are ways to grasp huge data in field of Agriculture. There is a need of a system to have obvious analyses of data of agriculture and extract or use useful information from the spreading data.

Extracting knowledge from the data set is the process of mining. It aims to give accurate results to farmers. It finds hidden patterns. It discovers useful knowledge from the tremendous data set. It is one of the processes in Knowledge Discovery in Databases(KDD). Apart from the KDD process, in recent days with the development in IT world, Machine Learning has emerged to handle big volume of data and involves high performance computing too. Application of Machine Learning in Agriculture peaks up day by day.

©IJRASET: All Rights are Reserved

Machine Learning techniques are used in crop management, livestock management, water management and soil management. One kind of machine leaning technique is using a recommendation algorithm. They provide personalized products in E-Commerce. These recommendation concepts are used in agriculture in this paper to provide crops to sow. Simple Data Analytics is used on crop dataset and personalization of agricultural crops are suggested to farmers.

A. Problem Statement

Agriculture in India plays a predominant role in economy and employment. The common problem existing among the Indian farmers are they don't choose the right crop based on their soil requirements. Due to this they face a serious setback in productivity. This problem of the farmers has been addressed through machine learning based crop recommendation system.

B. Objectives

The aim of this work is to develop a crop recommendation system based of productivity and season. Therefore following are the objectives of our proposed system:

1) To collect the agriculture data from last decades of year from different regions and crops respectively.
2) Pre-process the data and extract the attributes which affects the crop cultivation and production.
3) Train the Machine learning model with these features and recommend the crop based on season and productivity.
4) To evaluate performance analysis of the proposed recommendation system.

C. Proposed System

Recommender systems have lent its hands to users to choose items they like. Recommendation system is the approach to provide the suggestions to the users of their interest. This can be practiced for agricultural use too. Based upon the factors of agriculture, farmers are given with ideas for their cultivation process. New techniques to increase crop cultivation can also be recommended. Farmers will be given recommendation by considering the season of crop production.

II. LITERATURE SURVEY

In 2016, K. He, X. Zhang, S. Ren, and J. Sun present a residual learning framework to ease the training of networks that are substantially deeper than those used previously. They explicitly reformulate the layers as learning residual functions with reference to the layer inputs, instead of learning unreferenced functions. They provide comprehensive empirical evidence showing that these residual networks are easier to optimize, and can gain accuracy from considerably increased depth. They also present analysis on CIFAR-10 with 100 and 1000 layers. The depth of representations is of central importance for many visual recognition tasks. Solely due to our extremely deep representations, they obtain a 28% relative improvement on the COCO object detection dataset.

In 2015, H. A. Nugroho, A. Nugroho, J. Grafika, and N. Bulaksumur provides various segmentation methods applied on the same image like Watershed, Active contour, SVM (Support vector machine), Morphological operation etc., and give a comparison of their output. This helps the doctor to focus only on their area of interest. Key Words: Segmentation, Watershed algorithm, Active contour, Morphological Operations.

In 2016, Hanung Adi Nugroho, Made Rahmawaty, Yuli Triyani, Igi Ardiyanto research proposes a classification of thyroid ultrasound images by using some texture features into two classes. The dataset consists of 39 ultrasound images which grouped into 25 cystic cases and 14 solid cases. An initial step of image pre-processing is conducted to enhance the detection capability. Afterwards, followed by some methods of morphological operation, that is active contours without Edges (ACWE) and histogram equalization. The feature extraction is developed based on texture analysis by using Gray Level Co-occurrence Matrix (GLCM), Histogram and Gray Level Run Length Matrix (GLRLM). Finally, Multilayer Perceptron (MLP) is used to classify cystic nodule from solid nodule. The result shows that the proposed method achieves the accuracy of 89.74%, sensitivity of 88.89%, specificity of 91.67%, positive predictive value (PPV) of 96.00% and negative predictive value (NPV) of 78.57%. This indicates that the proposed method is excellent in classifying thyroid ultrasound images.

In 2018, Ankita Tyagi, Ritika Mehra and Aditya Saxena provides the analysis and classification models that are being used in the thyroid disease based on the information gathered from the dataset taken from UCI machine learning repository. It is important to ensure a decent knowledge base that can be entrenched and used as a hybrid model in solving complex learning task, such as in medical diagnosis and prognostic tasks. They also proposed different machine learning techniques and diagnosis for the prevention of thyroid. Machine Learning Algorithms, support vector machine (SVM), K-NN, Decision Trees were used to predict the estimated risk on a patient's chance of obtaining thyroid disease.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

In 2019, Naghmeh Mahmoodial, Prabal Poudel, Alfredo Illanes and Michael Friebe a novel approach for thyroid texture characterization based on extracting features utilizing higher order spectral analysis (HOSA) was used. A Support Vector Machine (SVM) was applied on the extracted features to classify the thyroid texture. Since HOSA is a well suited technique for processing non-Gaussian data involving non-linear dynamics, good classification of thyroid texture can be obtained in US images as they also contain non-Gaussian Speckle noise and nonlinear characteristics. A final accuracy of 93.27%, sensitivity of 0.92 and specificity of 0.62 were obtained using the proposed approach.

In 2020, Q. Zhang, J. X. Hu1 and S. Zhou proposed method for the detection of hyperthyroidism based on the modified LeNet-5. Three test groups were designed by randomized controlled experiment design for the detection of hyperthyroidism, in which the control group was tested manually by experienced doctors, the LeNet-5 network learning group adopted the classical LeNet-5 network learning algorithm and the experimental group used the modified LeNet-5 network learning algorithm. Evaluation indices included the accuracy, detection efficiency, sensitivity, and specificity and F1 score of the detection. At the same time, differences between the two algorithms in the detection of thyroid nodules were compared. There was no significant difference between the 3 groups in the accuracy and specificity of detection of hyperthyroidism. In terms of the detection efficiency and sensitivity, the performance of the network learning algorithm group and learning group was better than that of the control group. Both the network learning algorithm group and experimental group could detect the thyroid nodules accurately, but there was no significant difference in the accuracy of detecting the type of thyroid nodules

III.SYSTEM DESIGN

A. System Architecture

Crop production depends on many agricultural parameters. Proposed work is based on the production of crops in previous years, crops can be recommended to the farmers. This kind of suggestions will make farmer to know that whether that particular is yielding a good production in recent years. Production of crops may become less due to any crop disease, water problem and many other factors. While considering about the production, farmers may get knowledge about which crop is in high volume in the market in that year. Based on this farmer can take decision of trend on crops in recent years. Farmers will be given recommendation by considering the season of crop production.

Prediction of crops was done according to farmer's experience in the past years. Although farmer's knowledge sustains, agricultural factors has been changed to astonishing level. There comes a need to indulge engineering effect in crop prediction. Data mining plays a novel role in agriculture research. This field uses historical data to predict; such techniques are neural networks, K-nearest Neighbour. K-means algorithm does not use historical data but predicts based on-computing centres of the samples and forming clusters. Computational cost of algorithm acts as a major issue. Use of Decision tree machine learning classifier is a boon to agriculture field which computes accurately even with more input. The block diagram for proposed system is shown in below figure1.1..

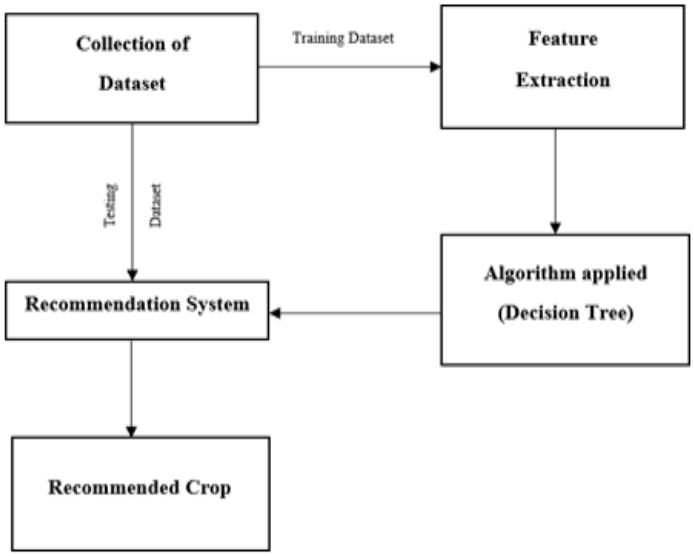


Fig. 1 Architecture of proposed crop recommendation System

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue VII July 2022- Available at www.ijraset.com

B. Dataflow Diagram

Figure 2 depicts DFD is also called as bubble chart. It is a simple graphical formalism that can be used to represent a system in terms of input data to the system, various processing carried out on this data, and the output data is generated by this system.

1) *Training Phase*: A set of past 10 years data is pre-processed for feature extraction. The training phase can be summarized as follows:

a) Extract features like rainfall, temperature, soil type and season.

b) Train a Decision tree classifier using this feature set.

The output of the training phase is a trained classifier capable of recommending crop based on season and productivity

The performance of the trained classifier can be evaluated using measures like accuracy, sensitivity and specificity.

2) *Classification:* This phase can be summarized as follows:

a) Take as input, feature extracted parameters.

b) Use the trained classifier to recommend the crop based on the input parameter.

The output of this phase is a crop name with duration of yield.

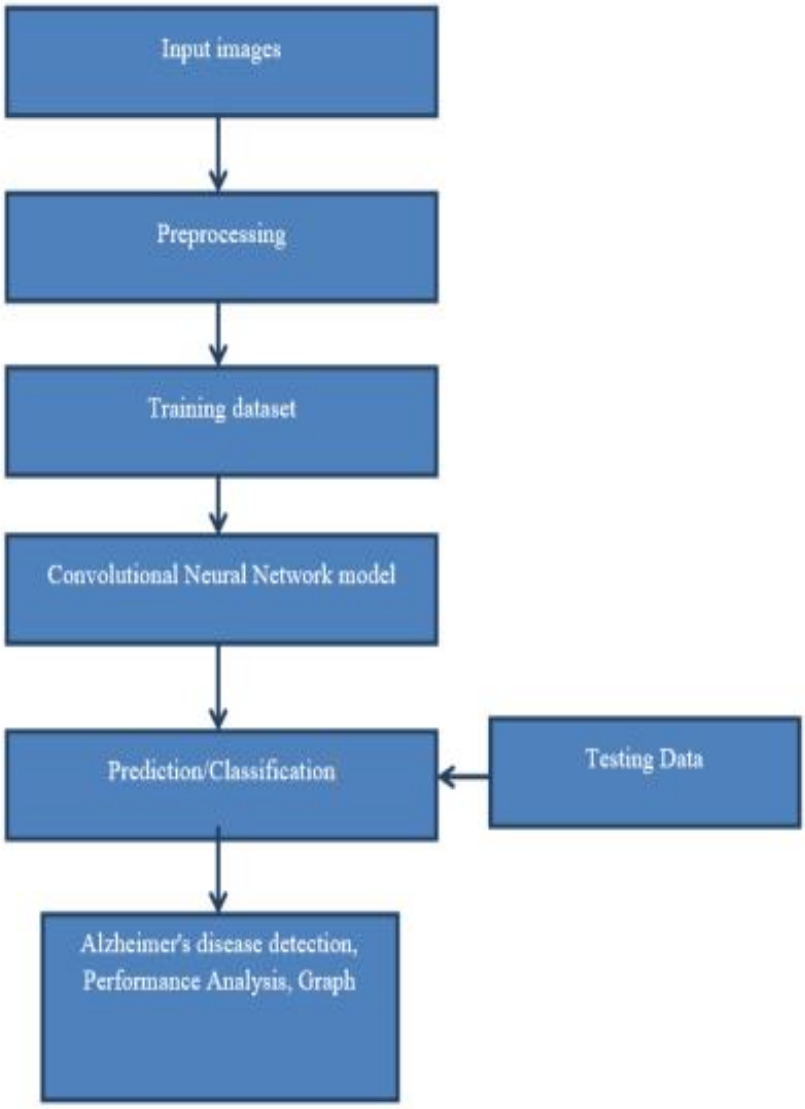


Fig. 2 Dataflow Diagram

©IJRASET: All Rights are Reserved

C. Activity Diagram

Figure 3 represents Activity diagrams are graphical representations of workflows of stepwise activities and actions with support for choice, iteration and concurrency. In the Unified Modelling Language, activity diagrams can be used to describe the business and operational step-by-step workflows of components in a system. An activity diagram shows the overall flow of control.

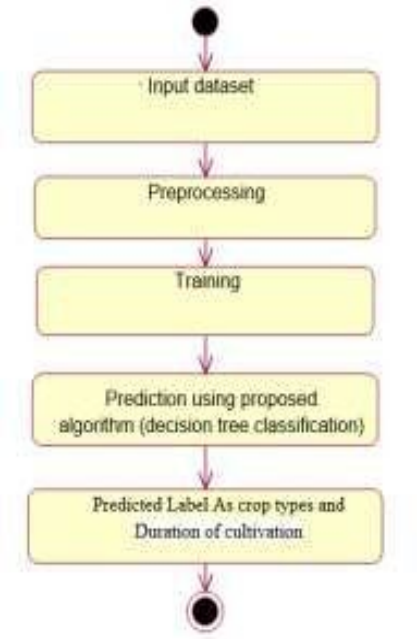


Fig. 3 Activity Diagram

D. Sequence Diagram

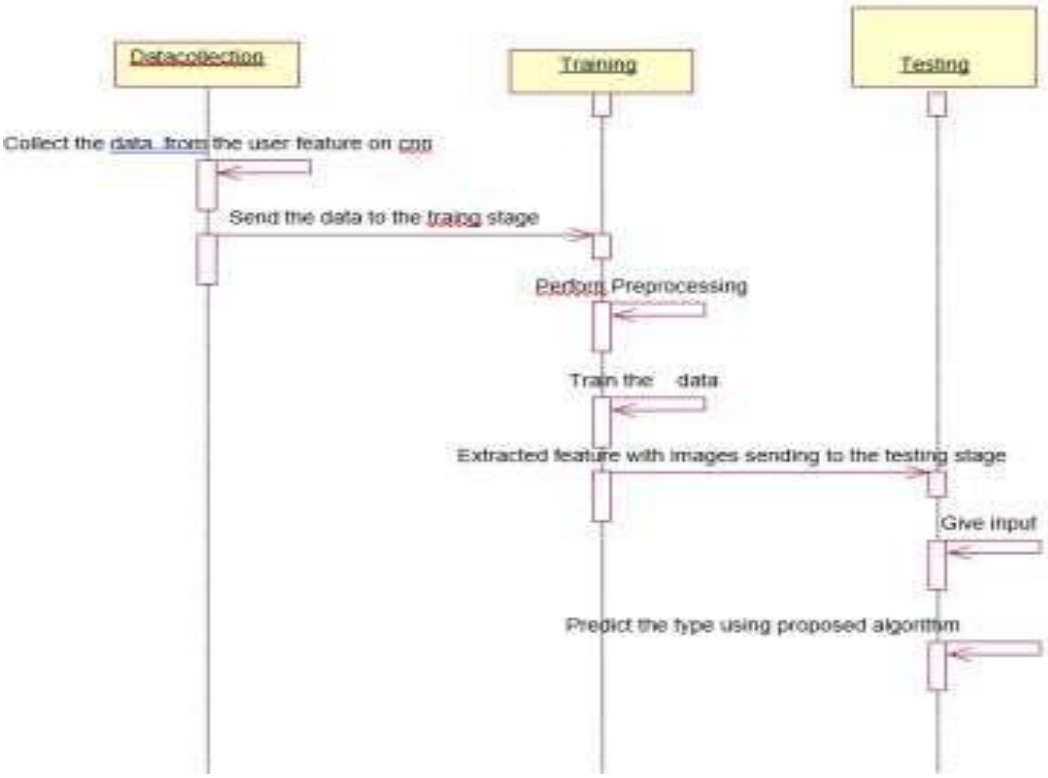


Fig. 4 Sequence Diagram of Proposed System

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue VII July 2022- Available at www.ijraset.com

IV. IMPLEMENTATION

A. Data Collection

This is the first real step towards the real development of a machine learning model, collecting data. This is a critical step that will cascade in how good the model will be, the more and better data that we get, the better our model will perform. There are several techniques to collect the data, like web scraping, manual interventions and etc. The dataset used in this crop recommendation in India taken from some other source

Dataset: The dataset consists of 821 individual data. There are 14 columns in thedataset, which are described below.

1) *States:* total number of states in India
2) *Rainfall:* rainfall in mm
3) *Ground Water:* Total ground water level
4) *Temperature:* temperature in degree Celsius
5) *Soil type:* Number of soil types
6) *Season:* Which season is suitable for crops
7) *Crops:* Types of crops
8) *Fertilisers Required:* Types of Fertilisers required
9) *Cost of Cultivation:* Total cost for cultivation
10) *Expected Revenues:* Total expected revenues
11) *Quantity of Seeds per Hectare:* seeds for quantity per hectare
12) *Duration of Cultivation:* number of day for duration of cultivation.
13) *Demand of Crop:* demand of crop (High, low)
14) *Crops for mixed Cropping:* which crop can mixed for cropping.

B. Data Preparation

Wrangle data and prepare it for training. Clean that which may require it (remove duplicates, correct errors, deal with missing values, normalization, data type conversions, etc.)

Randomize data, which erases the effects of the particular order in which we collected and/or otherwise prepared our data Visualize data to help detect relevant relationships between variables or class imbalances (bias alert!), or perform other exploratory analysis Split into training and evaluation sets

C. Model Selection

A decision tree is a flowchart-like tree structure where an internal node represents feature(or attribute), the branch represents a decision rule, and each leaf node represents the outcome. The topmost node in a decision tree is known as the root node. It learns to partition on the basis of the attribute value. It partitions the tree in recursively manner call recursive partitioning. This flowchart-like structure helps you in decision making. It's visualization like a flowchart diagram which easily mimics the human level thinking. That is why decision trees are easy to understand and interpret. Decision Tree is a white box type of ML algorithm. It shares internal decision-making logic, which is not available in the black box type of algorithms such as Neural Network. Its training time is faster compared to the neural network algorithm. The time complexity of decision trees is a function of the number of records and number of attributes in the given data. The decision tree is a distribution-free or non-parametric method, which does not depend upon probability distribution assumptions. Decision trees can handle high dimensional data with good accuracy. The decision rules are generally in form of if-then-else statements. The deeper the tree, the morecomplex the rules and fitter the model.

Before we dive deep, let's get familiar with some of the terminologies:

- Instances: Refer to the vector of features or attributes that define the input space
- Attribute: A quantity describing an instance
- Concept: The function that maps input to output
- Target Concept: The function that we are trying to find, i.e., the actual answer
- Hypothesis Class: Set of all the possible functions
- Sample: A set of inputs paired with a label, which is the correct output (also known as theTraining Set)
- Candidate Concept: A concept which we think is the target concept
- Testing Set: Similar to the training set and is used to test the candidate concept and determineits performance

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

D. Analyze and Prediction

In the actual dataset, we chose only 7 features:

1) *States:* total number of states in India.
- Rainfall: rainfall in mm.
- Ground Water: Total ground water level
- Temperature: temperature in degree Celsius
- Soil type: Number of soil types
- Season: Which season is suitable for crops
- Crops: Types of crops

2) *Accuracy on Test Set:* We got a accuracy of 90.7% on test set.
3) *Saving the Trained Model:* Once you're confident enough to take your trained and tested model into the production-ready environment, the first step is to save it into a .h5 or. pkl file using a library like pickle. Make sure you have pickle installed in your environment. Next, let's import the module and dump the model into.pkl file.

V.RESULTS AND SCREENSHOTS

The figure 5 depicts the output screen of the proposed system. It is the home page in which we get login button. We use django framework for fron-end. Once we start the Flask server we get this page in our local host port.

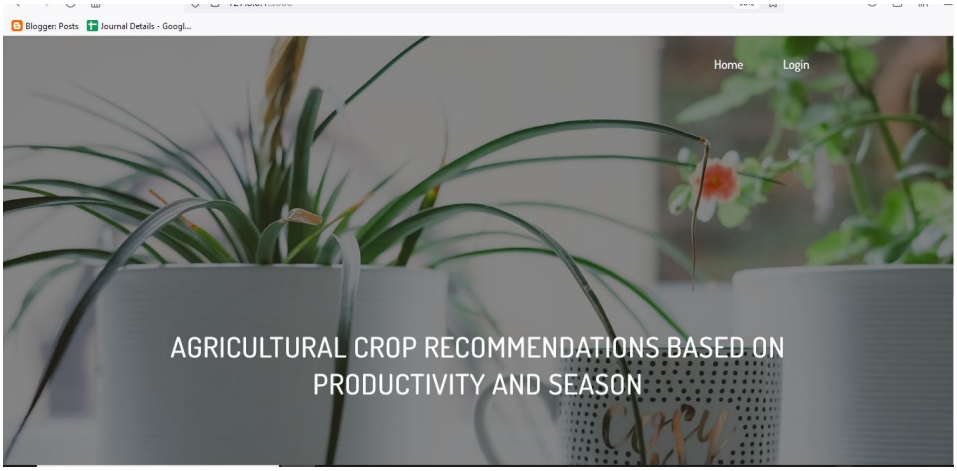


Fig. 5: Home Page

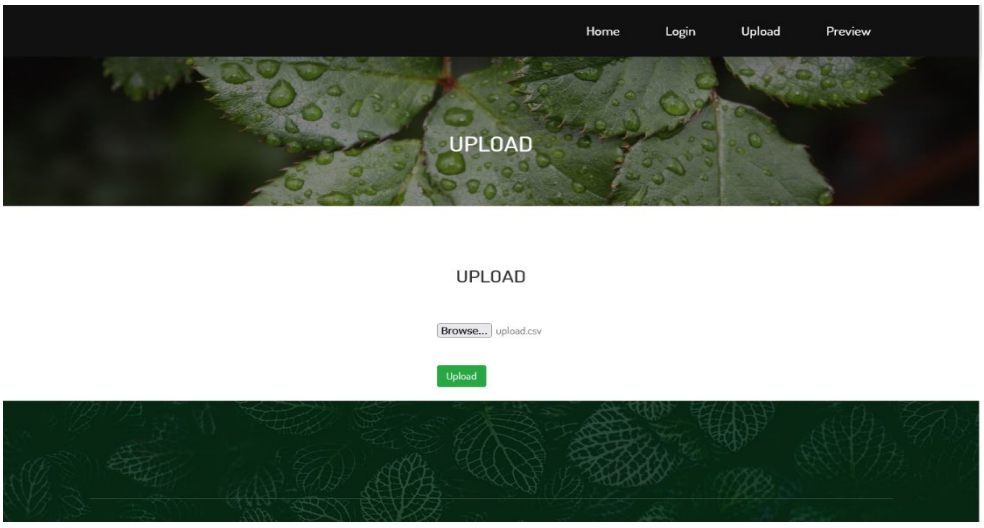


Fig. 6: Dataset uploading for Training

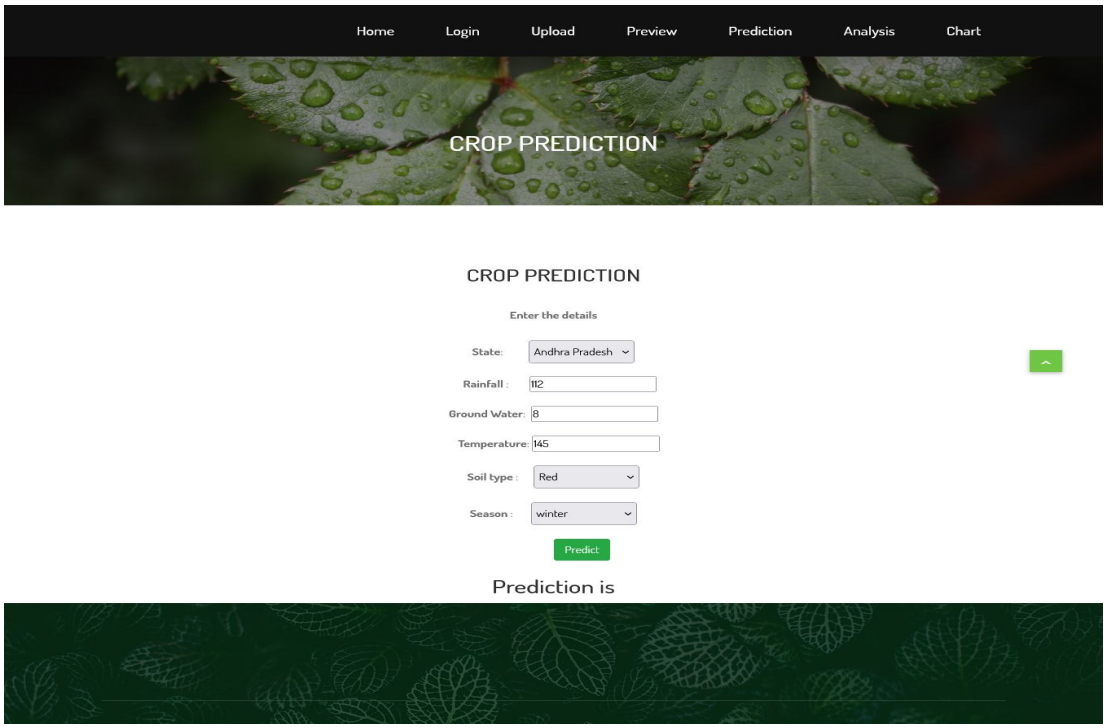


Fig. 7: Giving Input Parameters for crop recommendation

Figure 7 depicts the input parameters which are consider for the recommendation of the crops and Figure 6 depicts the uploading of the past years dataset for extracting the features and training the model

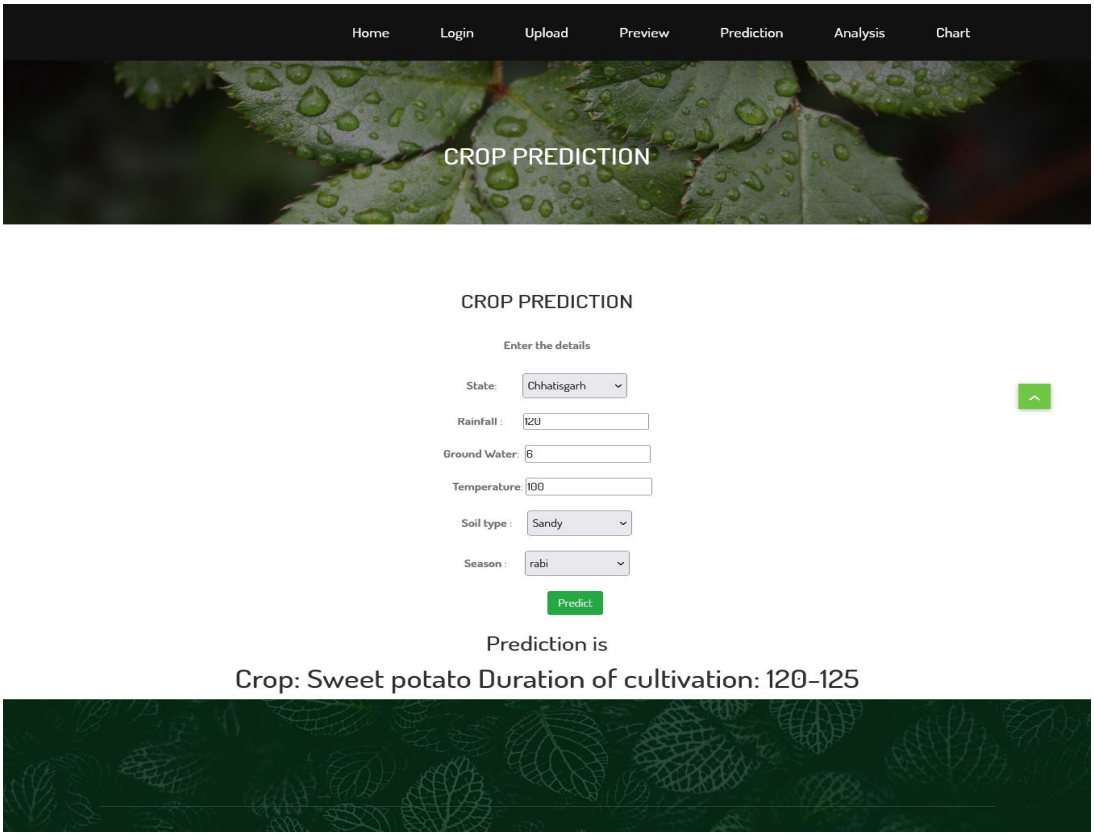


Fig. 8: Crop Prediction with Cultivation Duration

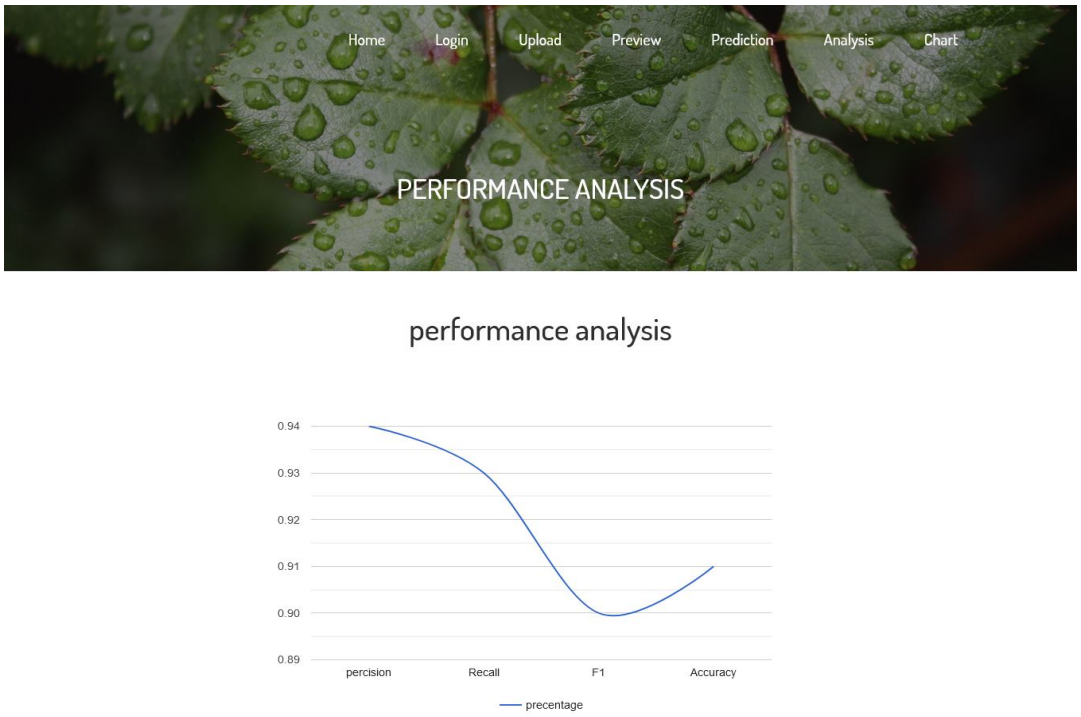


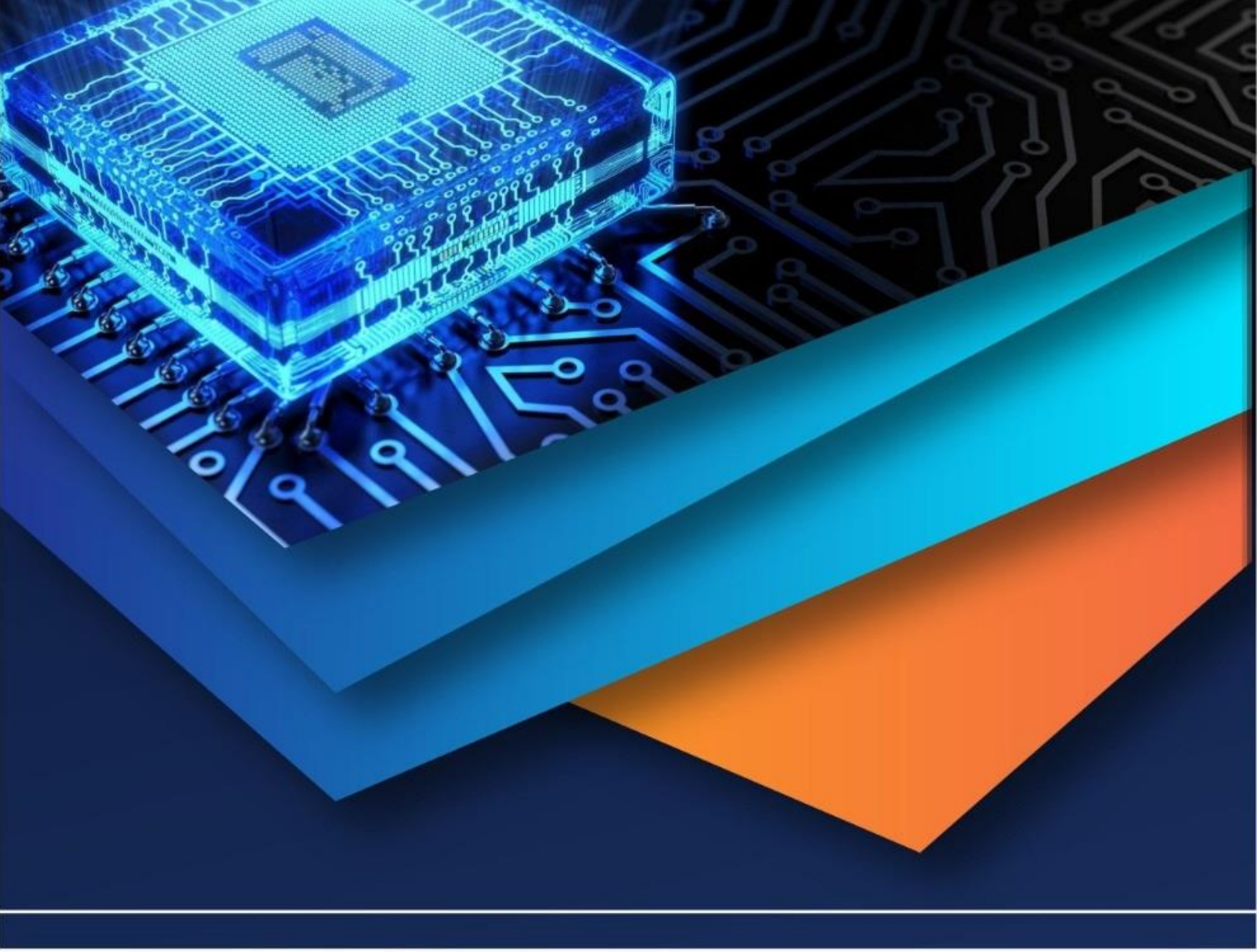
Fig. 9: Performance analysis of the proposed model

VI. CONCLUSION

In this Work, significance of management of crops was studied vastly. Farmers need assistance with recent technology to grow their crops. Proper prediction of crops can be informed to agriculturists in time basis. Many Machine Learning techniques have been used to analyse the agriculture parameters. Some of the techniques in different aspects of agriculture are studied by a literature study. Blooming Neural networks, Soft computing techniques plays significant part in providing recommendations. Considering the parameter like production and season, more personalized and relevant recommendations can be given to farmers which makes them to yield good volume of production.

REFERENCES

[1] Shreya S. Bhanose, Kalyani A. Bogawar (2016) “Crop And Yield Prediction Model”, International Journal of Advance Scientific Research and Engineering Trends, Volume 1,Issue 1, April 2016
[2] A.Swarupa Rani (2017), “The Impact of Data Analytics in Crop Management based on Weather Conditions”, International Journal of Engineering Technology Science and Research, Volume 4,Issue 5,May.
[3] Pritam Bose, Nikola K. Kasabov (2016), “Spiking Neural Networks for Crop Yield Estimation Based on Spatiotemporal Analysis of Image Time Series”, IEEE Transactions On Geoscience And Remote Sensing.
[4] Priyanka P.Chandak (2017),” Smart Farming System Using Data Mining”, International Journal of Applied Engineering Research, Volume 12, Number 11.
[5] Vikas Kumar, Vishal Dave (2013), “KrishiMantra: Agricultural Recommendation System”, Proceedings of the 3rd ACM Symposium on Computing for Development, January.
[6] Ramesh A.Medar (2014), ”A Survey on Data Mining Techniques for Crop Yield Prediction”, International Journal of Advance Research in Computer Science and Management Studies, Volume 2, Issue 9, September.
[7] Shreya S.Bhanose (2016),”Crop and Yield Prediction Model”, International Journal of Advence Scientific Research and Engineering Trends, Volume 1,Isssue 1,ISSN(online) 2456-0774,April.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)