Volume 9, Issue 3, March – 2024 International Journal of Innovative Science and Research Technology
ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24MAR402

Prediction of Crop Compliant System Using Machine Learning Techniques

Dr. A. Sudhir Babu
Dean of Academics
Dhanekula Institute of Engineering and Technology
Vijayawada, India

Ch. Yoditha
Student (CSE)
Dhanekula Institute of Engineering and Technology
Vijayawada, India

G. Navya
Student (CSE)
Dhanekula Institute of Engineering and Technology
Vijayawada, India

Abstract:- Machine learning techniques are used to analyze agriculture production using the Agriculture Production dataset. Crop compliant prediction, essential for informed agricultural decisions, relies on factors such as weather conditions and crop management practices. Through analysis of the dataset, various machine learning models are applied to provide insights crucial for farmers and agricultural stakeholders. These insights aid in crop selection and farming practices, optimizing decisions based on climatic conditions and soil characteristics. The trained model facilitates crop prediction and recommendation, mitigating financial risks for farmers and promoting optimal crop yields.

Keywords:- *Machine Learning, Crop Yield Prediction, Weather Conditions, Climatic Conditions, Soil Characteristics.*

I. INTRODUCTION

Agriculture, as the backbone of economies around the world, plays a vital role in providing sustenance and livelihoods for millions of people. As the global population continues to grow, the demand for food production intensifies, necessitating the optimization of agricultural practices. Crop yield prediction [2] stands as a fundamental aspect of this optimization process, offering insights that can revolutionize the agricultural landscape. In the highly technologically developed world of today, data-driven strategies have become effective methods for revealing complex patterns buried in agricultural datasets. By utilizing these strategies to their full capacity, we may allocate resources more effectively while minimizing our negative effects on the environment and improving food security. The fusion of data science and agriculture [11] has the potential to transform current farming methods into ones that are more intelligent, effective, and sustainable. Accurate crop yield prediction is the specific issue that project aims to solve. Farmers can relate to this issue since they need accurate data to plan their planting times, efficiently use their resources, and maximize their crops. The initiative aims to close the gap between the complexity of agricultural systems and the demand for useful insights by utilizing the capabilities of machine learning.

A. Existing System

In earlier days, the crops will be grown by farmers without any data as how much water, pesticides, fertilizers are used to grow a crop. The farmers depend on the climatic rainfall which is not possible for them to know when it is about to rain, after that climatic condition prediction is used to know whether the rainfall is possible for the required time to grow a crop it is done using collection of data for years. After that there are multiple prediction systems are formed in which the prediction of crop is possible. Many of those prediction systems are formed using collection of data for years. In that way different advanced prediction systems [12] are formed to make the farmers receive profits in the crop they selected or by preferring the crop which can give them profit.

B. Proposed System

The machine learning algorithms are good for predicting the crop which will give higher production and profit for farmers in agriculture. In this the random forest algorithm is used as it is the best algorithm compared to other algorithms used.

II. LITERATURE SURVEY

In suggested system, it employed a vast dataset that included all of India's states, whereas in the old system, just a single state was considered. These suggestions may be extracted and used to educate the farmers. The farmer can have a better understanding of the crops to cultivate by using a pictorial depiction. Machine Learning Techniques develops a well-defined model with the data and helps us to attain predictions. Agricultural issues like crop prediction, rotation, water requirement, fertilizer requirement and protection can be solved. Data Analytics paves a way to evolve useful extraction from agricultural database. Crop Dataset has been analyzed and recommendation of crops is done based on productivity and season [1].

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24MAR402

Applying data mining techniques to predict annual yield of major crops and recommend planting different crops in different districts. The data mining techniques are used to predict the annual yield of crops and it will recommend according to the output after applying data mining technique which crop is used in that area [3].

Judge Yield forecasting of spring maize using remote sensing and crop modeling. The sensors are used to sense the forecast of spring maize and how to plant it according to the crop modeling [4].

Grassland biomass estimation by using multitemporal remote sensing data machine learning approach. The sensing of data is done using machine learning approaches and in crop yield prediction using machine learning algorithms the machine learning algorithms are used [5].

Yield Prediction Using Image Analysis of Apple Fruit and Tree Canopy Features with Neural Networks, the apple fruit yield is predicted using neural networks for image analysis [6].

Strawberry Yield Prediction Based on a Deep Neural Network Using High-Resolution Aerial Orthoimages, the strawberry yield prediction is done using deep neural network using orthoimages [7].

Neural Networks and Image Visualization for Early Forecast of Apple Yield, the apple yield should be predicted using neural networks and image visualization for early forecast [8].

Deep Learning on Tomato Yield Prediction and Factor Interpretation, the tomato yield prediction is done using deep learning and also factor interpretation is done [9].

Crop Yield Prediction Using Deep Reinforcement Learning Model for Sustainable Agrarian Applications, the crop yield prediction is done using reinforcement learning model for sustainable agrarian applications [10].

III. METHODOLOGY

After opening the input page, the input should be given according to the shown values in it. The shown values in it are nitrogen, phosphorus, potassium, temperature, humidity, ph, rainfall.

The values are given according to the survey done by government for farmers to know the values in soil such as nitrogen, phosphorus, potassium, ph. The remaining values should be known by temperature news etc.

By giving the values in the input and clicking on predict the predicted crop will be generated on the output screen.

IV. MACHINE LEARNING ALGORITHMS

There are four types of algorithms in machine learning [5]. They are;

A. Supervised Algorithms

In supervised algorithms the data is present in organized form, in organized data the missing value and other are removed by using data pre-processing.

- *Logistic Regression*

Logistic regression is a supervised learning classification algorithm used to predict the probability of a target variable. In this algorithm the S shaped curve is used in this the data is divided in the top and bottom of the shape if the test data is in the middle of the S shaped curve, then it is not predicted correctly. It is used for solving regression problems. There are two types of regressions. They are linear regression, logistic regression. In linear regression the data is stored in a line and in logistic regression the data is stored in s shaped curve.

- *Random Forest*

A Random Forest Algorithm is a supervised machine learning algorithm that is extremely popular and is used for Classification and Regression problems in Machine Learning. In this algorithm the data is divided as small as possible. Random forest is a commonly used machine learning algorithm, that combines the output of multiple decision trees to reach a single result. Its ease of use and flexibility have fueled its adoption as it handles both classification and regression problems. So, prediction can be as accurate as possible.

- *Decision Tree*

Decision Tree algorithm is a supervised machine learning algorithm. It is used in both classification and regression models, but mostly preferred for classification problems. In this algorithm, the decision tree is used. In decision tree, there are two nodes, which are decision node and leaf node. Decision nodes are used to make any decision and the leaf nodes which show the output of the decisions.

- *K-Nearest Neighbor*

K-Nearest Neighbor is a supervised machine learning algorithm. It is also called as K-NN algorithm. This algorithm assumes the similarity between new data and available data and put the new data into the category that is most similar to the available categories. It is a non-parametric algorithm, which means it does not make any assumption on underlying data.

B. Unsupervised Algorithms

In unsupervised algorithms the data is in unorganized form. So, first the data should be organized, this is done by clustering the similar data.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24MAR402

➢ *K-Means*

In K-Means, k represents number of cluster present for the unorganized data provided. In this the clusters are divided using the similarities of the objects present in it. In K-Means the elbow method is used to show the number of clusters present in it.

C. Semi-Supervised Algorithms

In this algorithm the data is present in two forms; it is in both organized and unorganized form. So, it is a hybrid technique between supervised and unsupervised algorithms.

D. Reinforcement Algorithm

It also has a short form of RL algorithm. It is a machine learning technique that trains software to make decisions to achieve the most optimal results [10].

In these supervised algorithms are used. The accuracy of the algorithms present in supervised machine learning algorithms are checked and the best algorithm is used for crop yield prediction.

V. WORKING PROCESS

Random forest will be used the following steps;

A. Importing Libraries and Reading Dataset

Importing of libraries are done for usage of those libraries in the crop yield prediction using machine learning algorithms. The crop recommendation dataset is in the form of csv which is coma separated values is read as follows in the below line,

data= pd.read_csv('Crop_recommendation.csv').

where pd is pandas imported.

B. Data Preprocessing

In data preprocessing there are 4 steps they are Data Cleaning, Data Integration, Data Transformation, Data Reduction. By using these four steps the dataset will be perfect without any missing values and the dataset is useful for code.

C. Splitting Data Set Into Training and Testing Set Handling

The dataset is divided into training and testing sets in the ratio of 7:3 which is 70% for training and 30% for testing X_train,X_test,y_train,y_test=train_test_split(X,y,test_size=0 .3, random_state=42). This is the code to training and test set.

D. Implementing a Random Forest Classifier

This is how the implementing of random forest classifier is done log= RandomForestClassifier().

E. Predicting Test Cases Using Random Forest

The prediction of test cases in random forest is done as log.fit(X_train,y_train), pred=log.predict(X_test) by this the prediction is done successfully.

F. Checking the Accuracy Score

The accuracy of the random forest algorithm is shown by using this code accuracy_score(y_test,pred).

The output of this in crop yield prediction is 99%.

So, the random forest algorithm is used for crop yield prediction using machine learning algorithms.

VI. MODELING AND ANALYSIS

A. Upload Values

The dataset which is used in crop yield prediction is crop recommendation.csv in which the data is stored in the form of comma separated values. The data is stored by separating each of it with commas. So, it is stored as .csv. The preprocessing and other steps which deals with the data are present in those steps. The dataset is divided into two parts of 7:3 ratio. The 70% of the divided dataset is used for training and remaining 30% is used for testing purpose. The values present in the dataset are nitrogen, phosphorus, potassium, temperature, humidity, ph, rainfall. The crops present in the dataset are rice, maize, banana, mothbeans, mungbean, blackgram, pomegranate, mango, grapes, watermelon, apple, orange, papaya, coconut, cotton, jute, coffee, muskmelon, lentil, pigeonpeas, kidneybeans, chickpea, this are the crops on which the prediction is done .

The values should be uploaded according to the measurements given to the farmer by Indian government. The values will be entered without any doubt as the values are given by government.

B. Predict Crop

For crop prediction, the supervised algorithms are used. In prediction, the supervised algorithms accuracy is checked and the accuracy for which algorithm is more that algorithm is used to write code for crop yield prediction the algorithm which has more accuracy is random forest when comparing some of the supervised machine learning algorithms.

In random forest algorithm the dataset is divided into two parts of 7:3 ratio the training should be done using 70% of the dataset and the remaining dataset is used for testing purpose. In such a way random forest algorithm is used for predicting the crop.

After the values are entered, the user will press predict and the predicted crop will be shown on the output screen.

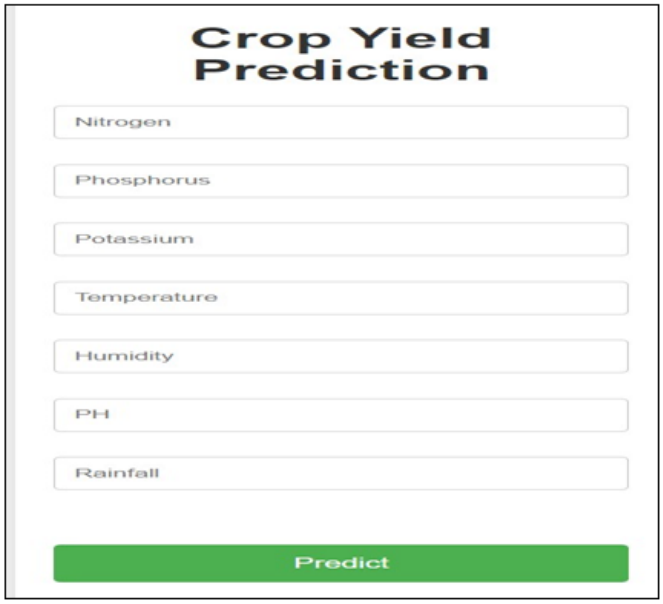
VII. OUTPUT

The crop compliant prediction using machine learning algorithm is done using the best fit algorithm is random forest algorithm which gives an accuracy of 99%.

So, by using the best fit algorithm the code is written and the flask code is used to prepare the web page, in the flask the code is written in such a way that the web input page to show the values to be entered the needed data will be shown in the block in which the data is to be entered such as nitrogen in the block where the nitrogen values should be entered.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24MAR402

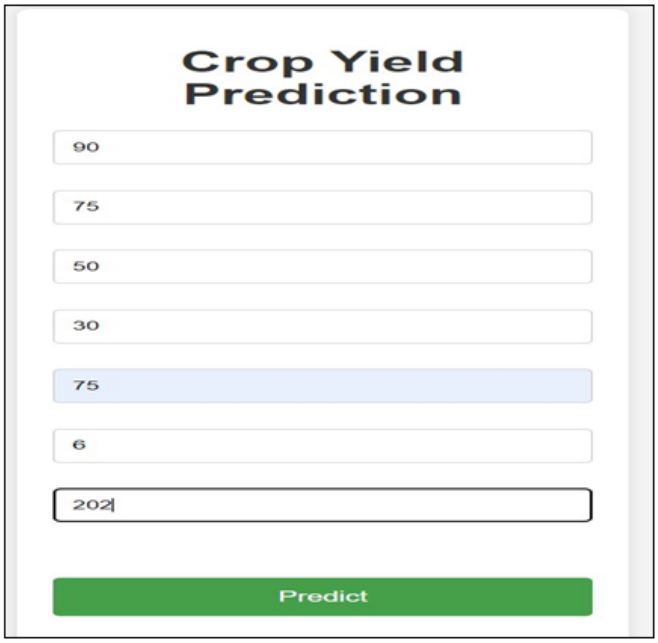
It is the output after using random forest algorithm which is best fit and flask is used to create this web page. The rainfall value must be given in millimeters or inches. The humidity must be given in grams. The nitrogen, phosphorus, potassium will be entered according to the measurements given by government. The input screen is shown as follows;



Screenshot 1: Input Screen for Crop Yield Prediction

In the above screenshot 1 user can enter the values for nitrogen, phosphorus, potassium, temperature, humidity, ph, rainfall. The values the user enter should be in the form of the data given by government.

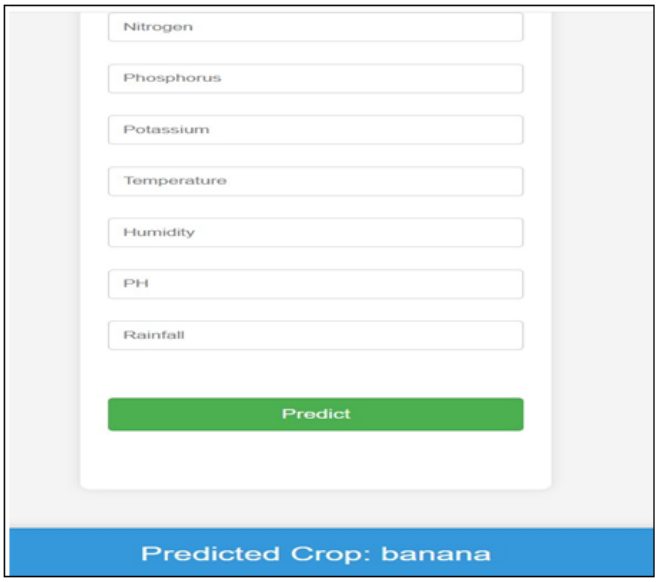
After entering the values in screenshot 1 it is shown as follows in screenshot 2.



Screenshot 2: Given Input Values in Web Page

In the above screenshot 2 the values are given after giving the values the predict is to be clicked. The data entered in the screenshot 2 must be in the form of data given by the government to the farmer.

The below screenshot 3 is the output for values given in screenshot 2.



Screenshot 3: Output Screen for Crop Yield Prediction

In the above screenshot 3 the crop recommended is estimated to give profit according to the values. The screenshot 3 shows the predicted crop as banana for the values given in screenshot 2. There are 22 crops in the dataset. So, the user can literally know which crop can bring profit to them according to the values entered. The 22 crops will be grown mostly in India in different seasons and different areas. So, the prediction will be done in any part of the country.

VIII. CONCLUSION

In this crop compliant prediction, it has analyzed the Agriculture Production dataset using various machine learning techniques. The insights gained from the analysis can help farmers and agricultural stakeholders make informed decisions about crop selection and farming practices based on different climatic conditions and soil characteristics. The trained model can be utilized for crop prediction and recommendation purposes. The code and the report are available on GitHub for reference and further exploration. It can further be extended by using conventional neural network in this method they can use images of the crop to predict the growth of the crop. In this algorithm usage it can also use Internet of things to measure the data like percentage of nitrogen, potassium, phosphorus. Which can be used to measure the values in soil. So, it can also construct an IOT to measure and test the data by comparing with training data present with us. It is all about crop yield prediction and how the user predicts the crop for next agriculture time.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24MAR402

REFERENCES

[1]. "Agriculture Crop Yield Prediction Using Machine Learning" published by Firdous Hina1, Dr. Mohd. Tahseenul Hasan Anjuman College of Engineering and Technology, Sadar, Nagpur-17(2022).

[2]. 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, April2016.

[3]. A.T.M.S. Ahamed, N.T. Mahmood, N. Hossain, M.T. Kabir, K. Das, F. Rahman, R.M. Rahman "Applying data mining techniques to predict annual yield of major crops and recommend planting different crops in different districts in Bangladesh" 2015 IEEE/ACIS.

[4]. I. Ahmad, U. Saeed, M. Fahad, A. Ullah, A. Ahmad, J. Judge "Yield forecasting of spring maize using remote sensing and crop modeling" (10) (2018).

[5]. I. Ali, F. Cawkwell, E. Dwyer, S. Green "Modeling managed grassland biomass estimation by using multitemporal remote sensing data machine learning approach" IEEE J. 10 (7) (2017).

[6]. Cheng, H.; Damerow, L.; Sun, Y.; Blanke, M. Early "Yield Prediction Using Image Analysis of Apple Fruit and Tree Canopy Features with Neural Networks". J. Imaging 2017.

[7]. Chen, Y.; Lee, W.S.; Gan, H.; Peres, N.; Fraisse, C.; Zhang, Y.; He, Y. "Strawberry Yield Prediction Based on a Deep Neural Network Using High-Resolution Aerial Orthoimages". Remote Sens. 2019.

[8]. Crtomir, R., Urka, C., Stanislav, T. et al. "Application of Neural Networks and Image Visualization for Early Forecast of Apple Yield". Erwerbs-Obstbau 54, 6976 (2012).

[9]. Duo Attention with" Deep Learning on Tomato Yield Prediction and Factor Interpretation" PRICAI 2019: Trends in Artificial Intelligence, 2019.

[10]. D. Elavarasan and P. M. D. Vincent, "Crop Yield Prediction Using Deep Reinforcement Learning Model for Sustainable Agrarian Applications," in IEEE Access, vol. 8, pp. 86886-86901, 2020.

[11]. Tripathy, A. K., et al. "Data mining and wireless sensor network for agriculture pest/disease predictions." Information and Communication Technologies (WICT), 2011 World Congress on IEEE.

[12]. Priyanka P. Chandak (2017), "Smart Farming System Using Data Mining", International Journal of Applied Engineering Research, Volume 12.