International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 07 | July 2021 www.irjet.net p-ISSN: 2395-0072

CROP YIELD PREDICTION USING MACHINE LEARNING

Mr. Mahesh B.L[1], Ms. Aditi[2], Ms. Aisha Reza GD[3], Mr. Akhil Roy[4], Mr. Nikhil M[5]

[1]*Assistant Professor, Dept. of CSE, Yenepoya Institute of Technology, Moodbidri, India-574225*
[2,3,4,5]*Students, Dept. of CSE, Yenepoya Institute of Technology, Moodbidri, India-574225*

Abstract-*India being an agricultural country, its economy mainly depends on agriculture yield growth and allied agro-industry products. In India agriculture is largely influenced by rain water which is highly unpredictable. Agriculture growth depends on diverse soil parameters like nitrogen, phosphorous, potassium, crop rotation, soil moisture, surface temperature. It also depends on weather aspects which include temperature, rainfall etc. Agriculture is one of the major fields in our country and also plays a major role in our country's economy. India is the second –largest producer of agriculture crops and agriculture is one of the major and least paid occupation in India. Variability in seasonal climate conditions can have harmful effects, with incidents of drought reducing production. Developing better techniques to predict crop productivity in various climatic conditions can help farmer and other stakeholders in their decision making in terms of agronomy and crop choice*

***Key Words*: Indian Agriculture, Machine Learning Techniques, Crop selection method, KNN, SVM, RF**

1. INTRODUCTION

Agriculture is demographically the broadest economic sector and it plays a major role in the overall socio-economic fabric of India. It also contributes a large portion of employment. As the time passes the need for production has been increasing exponentially. With the advent of new technologies and overuse of non-renewable energy resources patterns of rainfall and temperature are disrupted. The inconsistent trends developed from the side effects of global warming make it difficult for the farmers to clearly predict the temperature and rainfall patterns thus affecting their crop yield productivity.

Crop yield prediction is nothing but forecasting the yield of the crop from past historical data which includes factors such as temperature, humidity, PH, rainfall, crop name. Machine learning can bring a revolution in agricultural field by changing the income scenario through growing an optimum crop. Along with all advances in the machines and technologies used in farming, useful and accurate information about different matters also plays a important role in it. Our project will focus on predicting the yield of the crop by applying various machine learning techniques so that it can solve many agriculture and farmers problem. The prediction made by machine learning algorithms will help the farmers to decide which crop to grow to get the maximum yield and hence it will improve Indian economy. Machine learning is found to be a very pleasing field that can contribute to the agriculture field. The various models built using Machine learning can take various inputs to give some concrete output.

In order to perform accurate prediction and handle inconsistent trends in temperature and rainfall various machine learning algorithms can be applied. It will complement the agricultural growth and all together augment the ease of living for farmers. The main goal of agricultural planning is to achieve maximum yield rate of crops by using limited number of land resources. Whenever there is loss in unfavourable conditions we can apply crop selecting method and reduce the losses, thus it will help to increase the crop yield rate and this inturn helps in improving countries economy.

LITERATURE SURVEY

In [1] the authors NiketaGndhi& Amiya Kumar Tripathi&OwaizPetkar&Liesa J Amstronghas concluded that Support Vector Machines (SVM's) a supervised machine learning technique. There are number of examples of where it has been used in the agriculture domain. Tripathi reported on how SVMwas applied for reduction of precipitation for climate change scenarios to minimize the generalization error bound and to achieve generalized performance. SVM was used to forecast to demand and supply of pulpwood. SVM was also

applied to provide insights into crop response patterns related to climate conditions by providing the features contribution analysis for agricultural yield prediction for classification of agricultural datasets the use of discretization base support vector machines was used. Hung reported the use of SVM to model urban land use conversion. The study reported a relationship between various factors and rural urban land use. SVM has also been applied for the estimation of crop biophysical parameters with the use of aerial hyper spectral observations.

In [2] the authors**Rakesh Kumar, MP Singh, Prabhat Kumar, JP Singh**has concluded that Indian agriculture is highly dependent on summer rainfall. The correlation between summer rainfall and agriculture product production. This paper presents an analysis of crop climate relationship using past crops data. Correlation analysis tells that the monsoon rainfall, Pacific and Indian Ocean sea-surface temperatures and drawing sea-level pressure directly influences the crop prediction in India. Results show that the state-level crop production statistics and sub divisional monsoon rainfall are consistent with all-India result, except few cases. Crop sequencing technique is used to improve net yield rate of crop over season. It uses a method called Crop Selection Method(CSM).

In [3] the authors **Prof. D. S. Zingade, OmkarBuchade, NileshMehra, ShubhamGhodekar, Chandan Mehta** has concluded that there is no existing system which recommends crops based on multiple factors such as Nitrogen, Phosphorus and Potassium nutrients in soil and weather components which include temperature and rainfall. The proposed system suggests an android based application, which can precisely predict the most profitable crop to the farmer. The user location is identified with the help of GPS. According to user location, the feasible crops in the respective location is identified from the soil and weather database. These soils are compared with past year production to identify the most profitable crop in the current location. After this processing is done at server side, the result it sent to the user's android application Location is the only input for the extrapolation system. Depending on the numerous scenarios and additional filters according to the user requirement the most producible crop is suggested.

In [4] the authors **Ramesh Medar& Vijay S. Rajpurohit& Shweta**has concluded that it will review that various applications of machine learning in the former sector and also provides insights into the troubles faced by the farmers and method help in increasing the farmer sector in the countries and apply the more machine learning application. The algorithm used are Artificial Neural Networks, Bayesian belief networks, Decision tree algorithm, clustering regression analysis.

2. METHODOLOGY

The proposed methodology have two phases, training phase and test phase. In the training phase data will be collected and preprocessed. The preprocessed data will be clustered using K-Means algorithm. The training phase stops with number of generated rules. In the testing phase yield value is predicted based on the generated rules. The work starts with preprocessing step. In this the collected data was preprocessed. In the preprocessing, some data was removed from the dataset. Some of the area will not be suitable for crop production so it will be removed. The proposed methodology will be conducted using **Python matplotlib**and **Seaborn** which is used for data visualization. Data preprocessing are performed by using the **Pandas** library of python.Basically broad five steps are used:

- Data collection
- Data wrangling
- Data preprocessing
- Data visualization
- Exploratory data analysis (EDA)

1. KNN (K- Nearest Neighbor)

K-nearest neighbor method can be used for both regression and classification predictive problems. This method helps in interpret output, calculate time and predictive power. The Machine learning techniques are used in various fields. KNN is also one of the machine learning method. This is also called as method of sample based learning. This will contain the data of past datasets and can be used while predicting the new datasets. This will apply function called as distance function like

Manhattan or Euclidean distance. This can be used to compute distance from samples to all other training samples. It calculates the target value for new samples. The target vale will be the weighted sum of target values of the k nearest neighbors. The valve of K can be directly proportional to the prediction. Whenever the valve of K is small this indicates there is high variance and there is low bias. If the valve of the K is larger than this indicates that there is low variance and high bias. The main advantage of this KNN is it does not require any training or the optimization. This KNN uses data samples when predicting the new datasets. Hence it is having higher complexity and also more time consumption.

2. Support Vector Machine

Support vector machine used in crop yield prediction is called support vector regression. The aim of the support vector technique is to obtain non-linear function using kernel function (a linear function or polynomial function). The radial basis function and the polynomial function are the widely used kernel function. The merit of support vector regression is to avoid difficulties of using linear function in large input samples space and optimization of a complex problems transformed into simple linear function optimization.

3. Random Forest Algorithm

Random Forest is a ML algorithm. At training situation multitude decision trees are built and the output will be divided based on number of classes i.e., classification, prediction of class i.e., regression. The number of trees is proportional to accuracy in prediction. The dataset comprises factors like rainfall, perception, temperature and production. These factors in dataset are used for training. Only two-third of the dataset is considered. Remaining dataset is used for experimental basis. The algorithm random forest consists of 3 parameters like: n-tree which describes the n number of trees which need to grow, m try - mentions how many variables need to be taken at a node split. Node size - In terminal nodes it advise us the number of observation need to take.

3. SYSTEM DESIGN

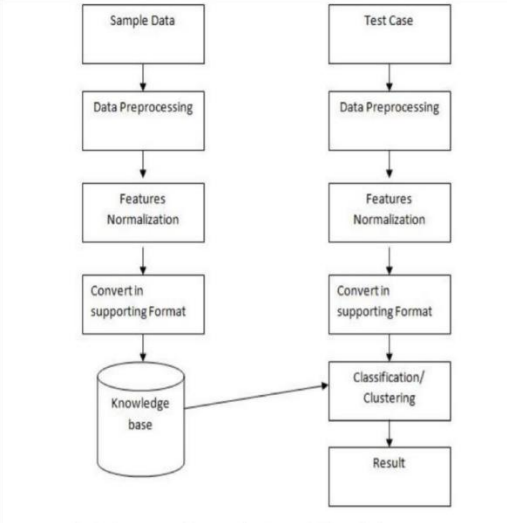

Fig-1: System design for Crop Yield Prediction.

4. RESULTS

To predict the crop yield rate anywhere in the Karnataka state, a web page is created. The web page is given with three attributes namely location, soil type and area. Whenever a wrong soil type is given for a specific location, it will prompt to give the correct soil type for the location. The web page consists of 3 methods namely KNN, SVM and Random Forest and theses three methods gives the name of the crop that is suitable for the given area along with the estimated price of the crops and the yield rate of the crop. The web application also compare these three methods to show which method gives better accuracy. The accuracy may vary with each crops and location.

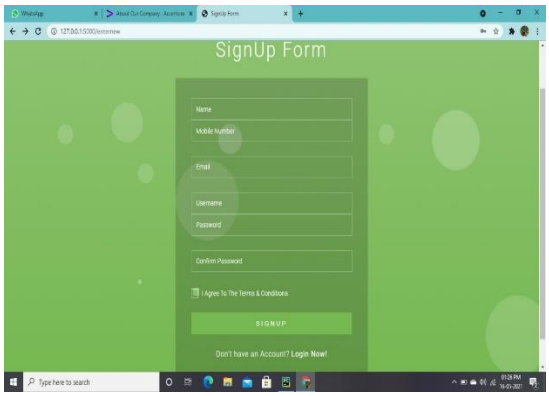

Fig.-2: Sign-up page of the web page

This is the signup page. Here we can create an account for the webpage.

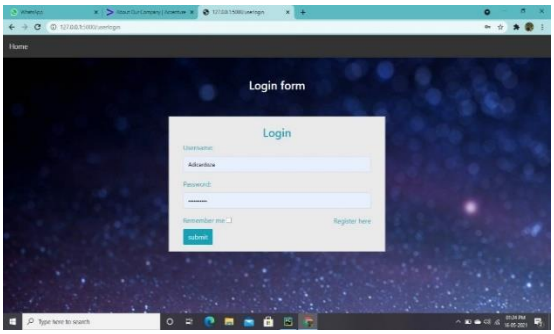

Fig.-3: Login page of the web application

After creating the account, we can login with the details in the login page.

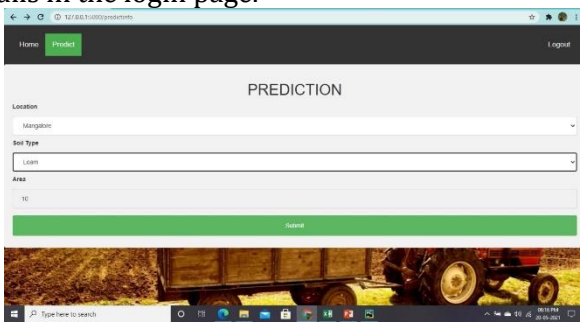

Fig.-4: Input page of the web application

The above picture shows the prediction page. Here we can give the input (location, soil type, area) to predict the crop.

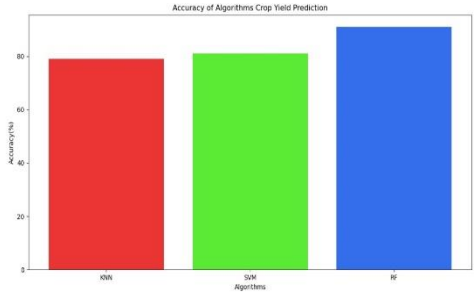

Fig.-5: Result of checking accuracy

This picture shows the accuracy comparison of the algorithms. The accuracies of the algorithms may change in each prediction.

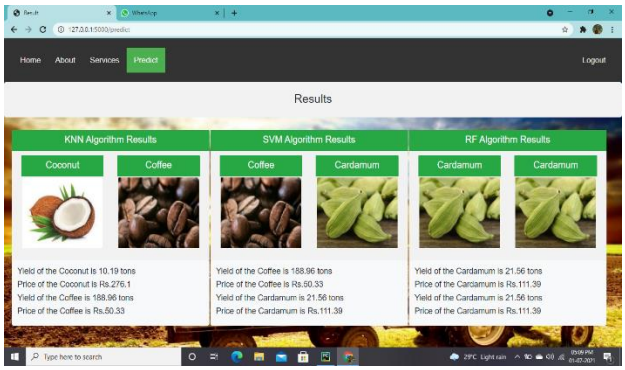

Fig-6: Output page of the web application

The Output page shows the suitable crops along with the crop yield rate.

5. CONCLUSION

Agriculture is the field which aids in economic growth of our country. But this is lacking behind in applying new technologies of machine learning. Thus our farmers should know all the new technologies of machine learning and other new techniques. These techniques help in getting maximum yield of crops. Many techniques of machine learning are applied on agriculture to improve yield rate of crops. These techniques also help in solving problems of agriculture. We can also get the accuracy of yield by checking for various methods. Hence we can improve the performance by checking the accuracy between various crops. Sensor technologies are implemented in countless farming sectors. This paper aids in getting maximum yield rate of the crops. Also assists in selecting proper crop for their selected land and

selected season. These techniques will resolve the problems of farmers in agriculture field. This will aid in improving the economic growth of our country.

REFERENCES

[1] Niketa Gandhi, OwaizPetkar, Leisa J Armstrong "Rice crop yield prediction using Support Vector Machines" 2016 IEEE Technological Innovations in ICT for Agriculture and Rural Development

[2] J.P. Singh, M.P. Singh, Rakesh Kumar and Prabhat Kumar "Crop Selection Method to Maximize Crop Yield Rate using Machine Learning Technique", International Journal on Engineering Technology, May 2015.

[3] Prof. D. S. Zingade, OmkarBuchade, NileshMehra, ShubhamGhodekar, ChandanMehta "Crop Prediction system using machine Learning".

[4] Ramesh Medar, Vijay S.Rajpurohit, Shweta "Crop yield Prediction using Machine Learning Techniques" 2019 IEEE 5th International Conference for Convergence in Technology (I2CT)

[5] S. S. Kale and P. S. Patil, "A Machine Learning Approach to Predict Crop Yield and Success Rate," 2019 IEEE Pune Section International Conference (PuneCon), Pune, India, 2019.

[6] Gour Hari Santra, Debahuti Mishra and Subhadra Mishra, Applications of Machine Learning Techniques in Agricultural Crop Production, Indian Journal of Science and Technology, October 2016.

[7] Karan deep Kauri, Machine Learning: Applications in Indian Agriculture, International Journal of Advanced Research in Computer and Communication Engineering, April 2016.

[8] S. Djodiltachoumy, A Model for Prediction of Crop Yield, International Journal of Computational Intelligence and Informatics, March 2017.

[9] Nishit Jain, Amit Kumar, SahilGarud, Vishal Pradhan, Prajakta Kulkarni, Crop Selection Method Based on Various Environmental Factors Using Machine Learning, Feb -2017.

[10] D.Sindhura, B.Navya Krishna, K.SaiPrasanna Lakshmi, B.Mallikarjun Rao, Dr. J Rajendra Prasad, Effects of Climate Changes on Agriculture International Journal of Advanced Research in Computer Science and Software Engineering, March 2016.

[11] T.GiriBabu, Dr.G.AnjanBabu, Big Data Analytics to Produce Big Results in the Agricultural Sector, March 2016.

[12] D Ramesh , B Vishnu Vardhan, Analysis Of Crop Yield Prediction Using Data Mining Techniques, International Journal of Research in Engineering and Technology, Jan-2015,

[13] AnshalSavla, HimtanayaBhadada, ParulDhawan, Vatsa Joshi, Application of Machine Learning Techniques for Yield Prediction on Delineated Zones in Precision Agriculture, May 2015.

[14] SitiKhairunniza-Bejo, SamihahMustaffha and Wan Ishak Wan Ismail , Application of Artificial Neural Network in Predicting, Journal of Food Science and Engineering, January 20, 2014