ZKG INTERNATIONAL

ISSN: 2366-1313

SMART FARMING USING MACHINE LEARNING BASED DECISION TREE REGRESSION

[1]V.B.N. AJAY, [2]K. SWATHI

[1]Assistant professor, Dept. of CSE, NVR College of Engineering & Technology, Tenali-Guntur (Dt)

vbnajaycse@gmail.com

[2]Assistant professor, Dept. of CSE, NVR College of Engineering & Technology, Tenali-Guntur (Dt)

swathimca54@gmail.com

Abstract: *Agriculture plays an important role in Indian economy. But now-a-days, agriculture in India is undergoing a structural change leading to a crisis situation. The only remedy to the crisis is to do all that is possible to make agriculture a profitable enterprise and attract the farmers to continue the crop production activities. As an effort towards this direction, this research paper would help the farmers in making appropriate decisions regarding the cultivations with the help of machine learning. This framework utilizes Machine Learning's Decision Tree Regression Algorithm to predict crop price. The attributes considered for prediction are rainfall, wholesale price index, month, and year. Consequently, the system gives an advance forecast to the farmers' which grows the speed of profit to them and consequently the country's economy.*

Keywords: *Agriculture, Machine learning, Crop Recommendation, crop price prediction, Decision Tree Regression model.*

I. INTRODUCTION

Agriculture is the backbone of the Indian economic system. But agriculture in India is experiencing a structural trade which is leading to a crisis situation. The relative contribution of agriculture to the GDP has been steadily declining over the years. It is disturbing that India has gone from being a food self-sufficient country to an online food importer. All these trends indicate that the agricultural sector in India is facing crisis at present. It is said that the impact of an agricultural disaster in India can be quite massive and is likely to affect all the different sectors and the country's economy in different ways. The most practical remedy for disaster is to do everything possible to make agriculture a worthwhile agency and to attract farmers to maintain crop production activities [1].

ISSN: 2366-1313

In the afterlife, the cultivators were projecting their performance from performance reviews for the past year. So, for this form of data analysis in crop forecasting, there are unusual techniques or algorithms, and with the help of those algorithms, we forecast crop yields. Today, privileged humans do not have any awareness of planting plants at the right time and the right place. Considering all the issues and issues like weather, temperature, and many more factors, there is no suitable solution and technology to overcome the situation you are facing. Accurate crop yield history data is a vital component in making determinations about controlling agricultural threats. Therefore, this paper proposes a concept for predicting harvest and crop yield based on weather conditions and historical records associated with harvesting. The farmer will monitor the crop production by acre before planting the field [2].

Being a rural country, India's economy is largely based on the development of agricultural crops and standardized agro-industrial things. He is currently rapidly progressing towards a juvenile specialist turnaround. India is now making unexpected progress towards technological development. Smart farming is changing the face of agriculture in India. Technology can provide a way to address most of the challenges farmers face. It can help them predict the weather more accurately, reduce waste, improve production, and increase their profit margins. With the current popularity, growers and consumers in the real world find it difficult to determine the correct costs of plants without having to do so.[3]

Advance knowledge of volatile trend rates or weather conditions. Thus, innovation will grow to be beneficial to agriculture. The objectives of the paper are to forecast farming expenditures from before. This work relies on locating appropriate regional datasets that help us achieve high accuracy and higher overall performance. The Agro-Genius device uses machine learning to build a price prediction model. In recent years, little fluctuation in crop prices has been observed. This has led to an increase in the rate of damage to the crops produced each year. The overriding purpose of this forecasting machine is to ensure that farmers get a better understanding of their yield and that they face price risk.

The proposed system will also forecast the weather helping the farmer make correct decisions regarding field ploughing, field harvesting etc. Similarly, fertilizers play an important role. Fertilizers load the soil

ISSN: 2366-1313

with the required nutrients that the crops eliminate from the soil. Crop yields and production will be fundamentally decreased if fertilizers are not used. That is the reason fertilizers are utilized to enhance the soil's supplement stocks with minerals that can be immediately assimilated and utilized by crops. Our system will provide fertilizer consumption based on different crops and provide a portal to buy the fertilizers and seeds from the user's location. They can even get the exact location along with the address of the fertilizer and seed shop. The provided fertilizers will get more profit to the farmers on the growing system suggested crop. It will also show the best suited crop based on cultivation date and month and location details, thereby maximizing the yield.

II. LITERATURE SURVEY

The Project work on Smart Farming System based on Machine learning been carried out based on following literatures reviewed. We will be now be discussing in brief the literatures pertaining to effect of temperature rise, yield variation and predictive analysis based on Machine Learning algorithm.2

The following articles focused on predicting crop load using machine learning and consequences. As of April 2019, exploration targets have identified both the cost and benefit of the crop before planting. The received setup datasets provide sufficient insights to correctly predict cost and demand within business enterprise sectors. The authors predicted maximum value crops and their expected rate during harvest time according to location, by predicting other ancient raw data sets using exceptional machine learning algorithms.

In [4] Predicting the yield of the crop using a machine learning algorithm. International Journal of Engineering Science Research Technology. This paper focuses on predicting the yield of the crop based on the existing data by using the Random Forest algorithm. Real data of Tamil Nadu were used for building the model

In [5]. Machine learning approach for forecasting crop yield based on parameters of climate. The paper was provided at the International Conference on Computer Communication and Informatics (ICCCI). In the current research, a software tool named Crop Advisor has been developed as a user-friendly web page for predicting the influence of climatic parameters on crop yields.C4.5 algorithm is used to produce the most influencing climatic

ISSN: 2366-1313

parameter on the crop yields of selected crops in selected districts of Madhya Pradesh.

In [6]. Analysis of Crop Yield Prediction by making Use of Data Mining Methods. IJRET: The paper provided in the International Journal of Research in Engineering and Technology. In this paper, the main aim is to create a user-friendly interface for farmers, which gives the analysis of rice production based on the available data. For maximizing the crop productivity various Data mining techniques were used to predict the crop yield.

In [7]. Random Forests for Global and Regional Crop Yield Predictions. institute on the Environment, University of Minnesota, St. Paul, MN 55108, United States of America. The generated outputs show that RF is an effective and different machine-learning method for crop yield predictions at regional and global scales for its high accuracy.

In [8] Crop Prediction using Machine Learning This research work helps the beginner farmer in such a way to guide them for sowing the reasonable crops by deploying machine learning. Naive Bayes, a supervised learning algorithm puts forth in the way to achieve it. The proposed supervised machine learning using naive Bayes Gaussian classifier with boosting algorithm is developed to predict the crop at high accuracy.

The Naïve bayes works on the basis of Bayes theorem

$$P(C/X) = P(X/C)\ P(C)\ /P(X)$$

$P(C|X)$ is the posterior probability of class (c, target) given predictor (x, attributes).

P(C) is the prior probability of class.

$P(X|C)$ is the likelihood which is the probability of predictor given class.

$P(X)$ is the prior probability of predictor. Based on the posterior probability the future of data can be predicted. Naïve bayes work well for large data set

In [9] Crop Prediction on the Region Belts of India: A Naïve Bayes MapReduce Precision Agricultural Model The planned work introduces an efficient degree economical crop recommendation system. From the yield graphs, the simplest time of sowing, plant growth, and gathering of plants may be known. Conjointly the best and worst condition may also be incurred. The model focuses on all styles of farms, and smaller farmers may also be benefitted. This model may be increased to seek out the yield of each crop, and for a chemical recommendation.

ISSN: 2366-1313

The work shown by Nishiba [10] is the expected utilization of data mining procedures in foreseeing the harvest yield dependent on the input parameters average rainfall and area of the field. The easy-to-use website page created for anticipating crop yield can be utilized by any client by giving the normal precipitation and region of that place. Different Data Mining techniques are applied to different datasets.

III. PROPOSED SYSTEM

We have used Python for basic programming in all modules. Flask is used for hosting. Socket Programming is used for a chat application. Chart.js is used for visualizing the maps. JavaScript is used for validation purposes. For Weather Forecast and fertilizer shop location, we have used APIs. Using the self-made dataset and concept of linear regression in machine learning we have implemented a Crop recommendation model so that a farmer can learn about the best suited crop for a particular region. In Fertilizer Recommendation we have used a dataset for predicting which fertilizer should be used for the disease present on crops. Socket programming is used for farmers interaction using provided chat application. Google API is used for providing a multilingual website for ease to read.

Decision Tree Regression Algorithm

The Decision Tree algorithm is a member of the family of supervised learning algorithms. Unlike other supervised learning algorithms, the clustering algorithm technique may also be used to address regression and classification tasks.

The decision tree regression machine-learning technique watches features of an item and trains a model in the structure of a tree to anticipate data later on to make significant nonstop output. Continuous output means that the output is not discrete, a known set of numbers or values.

The input to the algorithm is:

1. Input parameter

2. Training dataset

Formulas used for prediction

$$SSE = \sum_{ies1}(y_i - y1) + \sum_{ies2}(y_i - y2)$$

Where y1and y2 are the values of the dependent variable in group s1and s2 that is wholesale price index parameter in the dataset.

SYSTEM ARCHITECTURE

ISSN: 2366-1313

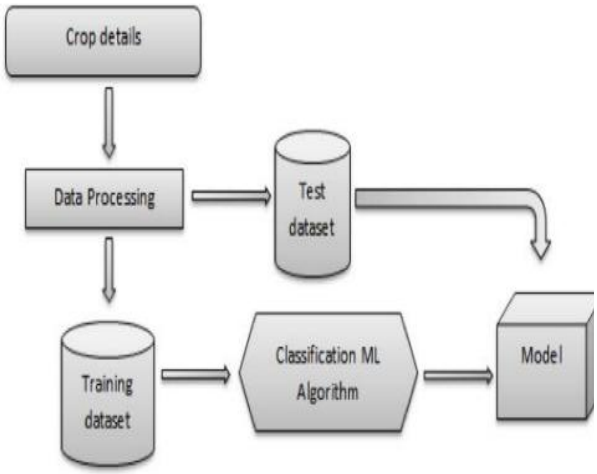


Fig.1 System architecture

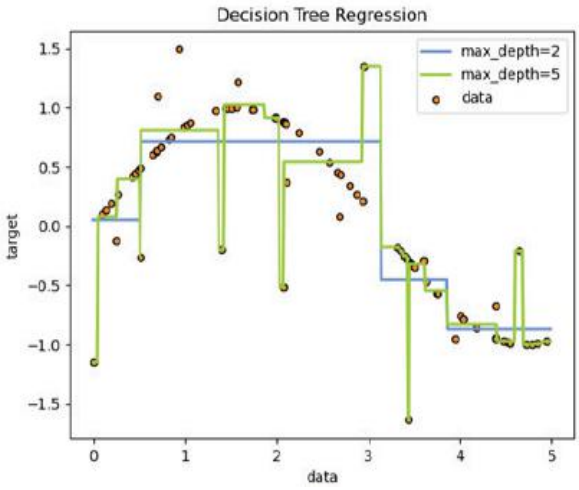
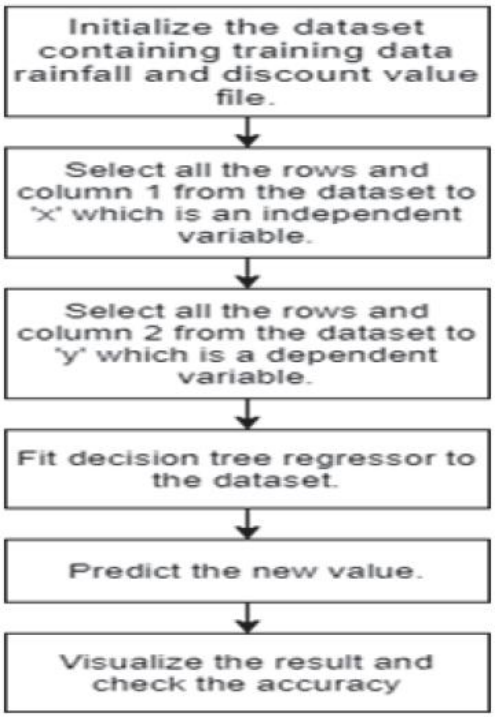


Fig.2 Decision Tree regression

The crop details are stored in a dataset. These have different attributes which are used for comparing and building relationships in between the attributes. Once the behaviour is compared, Data Transformation takes place. As the dataset is huge, we build a training data set using for example only the first thousand values of the dataset. So, we have a Model Training data set. Using a particular algorithm, we predict the results with maximum accuracy.

This system basically helps a farmer predict as to what crop to sow and also predicts the cost so that the farmer knows how much profit he would make approximately.

Steps to implement algorithm



IV. IMPLEMENTATION

The dataset which is used is obtained from the official site of the government of India. We trained the model on KNN, Naive Bayers, Decision Tree Regression algorithms and found that the Decision

ISSN: 2366-1313

Tree Regression algorithm reduced the overfitting problem. It also improved accuracy significantly. We performed testing and training on our dataset. The model was trained and thus the results obtained were noted. Then, compared the expected result with the initial dataset. Later, we used the test samples to estimate the accuracy of the model. We predicted the accuracy of the model with different algorithms. Out of the 3 algorithms used, we concluded that the Decision Tree Regression gave the best results and hence used the same to train our model. In Weather Forecast, using the open weather map API, we are forecasting the weather by taking the city name as an input from the user. For Crop Recommendation, the user has to give input parameters such as Date of Cultivation and Location. Using the self-made dataset of Maharashtra state and applying linear regression, crop recommendation model which will display most suitable and least suitable crops. For Fertilizer Recommendation, we are taking crop name and disease present on the crop as an input and by using the dataset we are displaying the appropriate recommended fertilizer. To see nearby Agri-Shop, we initialized google API and using location and recommended fertilizer, location of nearby shops is displayed. For Chat, we initialized socket and established a connection between client and server.

V. RESULTS

Utilizing this framework, we ought to get the same accuracy indeed when an information autonomous framework is utilized. For testing purposes, we have calculated the mean absolute error, Coefficient of determination R2 and Variance score of both training and testing dataset. Along with that we have calculated accuracy for the test dataset.

For Paddy:

R2 of the Test Set: ~0.9999

R2 of the Train Set: 1.0

Mean absolute error test set: 1.64

Test Variance score: 0.98

Mean absolute error Training Set: 4.72

Training Variance score: 0.70

Test Set Accuracy ~ 0.9773

ISSN: 2366-1313

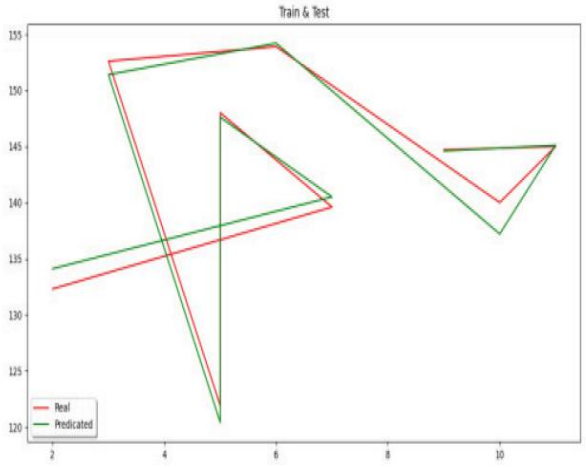


Fig.3 Real Values versus Predicted Values

It basically tells you how much better your classifier is performing over the performance of a classifier that simply guesses at random according to the frequency of each class.

In this experimental we built the following models and evaluated their performance

- KNN,
- Naive Bayes
- Decision Tree algorithms

Table.1 shows the performance evaluation of the various machine learning algorithms.

Author with year	Classification model	Performan Accuracy
Karthikeya et al. [2020]	KNN	63.63
Setiadi et al.[2020]	Naïve Bayes	85.71
Proposed method	Decision Tree	91.1

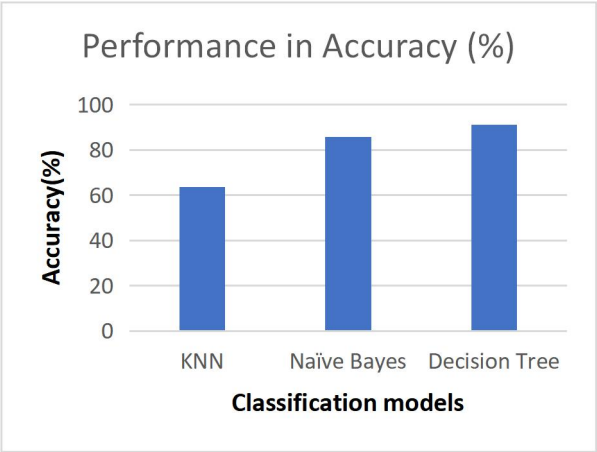


Fig.4 Accuracy of various classification models.

VI. CONCLUSION

This project is undertaken using machine learning and evaluates the performance by using KNN, Naive Bayes, and Decision Tree algorithms. In our proposed model among all the three algorithm Decision Tree gives the better yield prediction with 91.1% accuracy as compared to other algorithms as most extreme sorts of harvests will be secured under this system, farmers may become more acquainted with the yield which may never have been developed. The work exhibited the expected utilization of machine learning methods in foreseeing the harvest cost dependent on the given attributes. The created web application is easy to understand and the testing accuracy is over 91%.

ISSN: 2366-1313

REFERENCES

[1] Rachana, Rashmi, Shravani, Shruthi, Seema Kousar, Crop PriceMForecasting System Using Supervised Machine Learning Algorithms, International Research Journal of Engineering and TechnologyM(IRJET), Apr 2019

[2] Nishiba Kabeer, Dr.Loganathan.D, Cowsalya.T, Prediction of Crop Yield Using Data Mining, International Journal of Computer Science and Network, June 2019

[3] J. Vijayalakshmi, K. PandiMeena, Agriculture TalkBot Using AI, International Journal of Recent Technology and Engineering (IJRTE), July 2019

[4] Gamage, A., & Kasthurirathna, D. Agro-Genius: Crop Prediction Using Machine Learning, International Journal of Innovative Science and Research Technology, Volume 4, Issue 10, October – 2019

[5] Vohra Aman, Nitin Pandey, and S. K. Khatri. "Decision Making Support System for Prediction of Prices in Agricultural Commodity." 2019 Amity International Conference on Artificial Intelligence (AICAI). IEEE, 2019.

[6].. Rashmi Priya, Dharavath Ramesh.2018." Crop Prediction on the Region Belts of India: A Naïve Bayes MapReduce Precision Agricultural Model".

[7]. Igor Oliveira, Renato L. F. Cunha, Bruno Silva, Marco A. S. Netto.2018." A Scalable Machine Learning System for PreSeason Agriculture Yield Forecast."

[8] Renato L. F. Cunha, Bruno Silva, Marco A. S. Nett. A Scalable Machine Learning System for Pre-Season Agriculture Yield Forecast. 2018 IEEE 14th International Conference on e-Science (e-Science)

[9] Pandit Samuel, B.Sahithi , T.Saheli , D.Ramanika , N.Anil Kumar, Crop Price Prediction System using Machine learning Algorithms, Quest Journals Journal of Software Engineering and Simulation, 2020.

[10] Rubhi gupta, Review on weather prediction using machine learning, International Journal of Engineering Development and Research, 2020.

[11] H. K. Karthikeya, K. Sudarshan, 2020, "Prediction of Agricultural Crops using KNN Algorithm", pp.1422-1424.

[12] Tedy Setiadi, Fiftin Noviyanto, 2020, "Implementation of Naïve Bayes Method

ISSN: 2366-1313

in Food Crops Planting Recommendation",
pp. 4750-4755.