International Journal of Computer Engineering in Research Trends
Multidisciplinary, Open Access, Peer-Reviewed and fully refereed
Research Paper Volume-7, Issue-6, 2020, Regular Edition E-ISSN: 2349-7084

Dynamic Crop-yield and Price Forecasting using Machine Learning

Chaitra Kotnikall[1*], Jayateerth Vadavi[2]

[1*]*Student, Dept. Of CSE, Shri Dharmasthala Manjunatheshwara College of Engineering, Dharwad, India.*
[2]*Associate Professor, Dept. Of CSE, Shri Dharmasthala Manjunatheshwara College of Engineering, Dharwad, India.*

E-mail: chaitrak446@gmaill.com, jvvadavi@gmail.com

Available online at: http://www.ijcert.org

Received: 27/06/2020 Revised: 28/06/2020 Accepted: 04/07/2020 Published: 14/07/2020

Abstract:- Background: Guaranteeing food profitability is a significant issue for creating nations like India, where more than 33% of the individuals is live in neediness. To estimate cost, there is no system in place to advise farmers what crops to grow. Hence, this paper explains the attempt to predict crop price that a farmer can obtain from his land by analyzing patterns in past data.

Methods/Statistical analysis: This method makes use of several data such as rainfall, temperature, market prices, and past yield of a crop. The supervised machine learning algorithm, namely, the Decision tree algorithm and analyse the data and predict for the new set of data, is implemented. It also predicts the price and the gain for the next twelve months over the past twelve months and gives the time series analysis of the same.

Findings: The proposed model is developed to help farmers make better decisions concerning which crop is most suitable during his desired time of sowing and the location. This System predicts the yield and price of the crop of choice, giving the farmer useful information well before starting the process of cultivation.

Improvements: The System can introduce and make available climate-aware cognitive farming techniques and identifying systems of crop monitoring, early warning on pest/disease outbreak based on advanced AI innovation.

Keywords: Crop-yield, Supervised Machine learning in Agriculture, Decision tree, Forecasting, Analysis.

1. Introduction

Agriculture is the most vital sector of the Indian Economy. The Indian agriculture sector accounts for 18% of India's gross domestic product and employs to five-hundredths of the country's hands. However, the latest studies have shown a gentle decline within the contribution created by agriculture to the Asian nation economy though it is demographically the broadest economic sector and plays a significant role within the overall socio-economic cloth of India.

Contribution of the agricultural sector to the gross domestic product of the Asian nation is also the dearth of adequate crop designing by farmers further as by the govt.

The gross domestic product is one in all the most indicators that want to notice the health of a country's economy. Fast

fluctuations in crop costs area unit common within the market. This fluctuation in costs is primarily owing to the lack of previous design. It will lead the crops to be extremely priced, being a drawback for the patron once the value hikes and farmers to suffer from a loss in investment once the value drops. In such a state of affairs, it's tough for a farmer to create an informed alternative of the crop to grow in his land or to estimate the yield and worth to expect from it. The intention of this project is to assist the farmer to build higher

selections by analysing historical yield and worth knowledge victimization machine learning.

Machine Learning is an application of Artificial Intelligence that has proven to produce good prediction models in various aspects such as stock market, weather, the outcome of decisions, crop, and in our case, crop yield and price. Machine learning algorithms are divided into three main groups, as shown in figure 1.1, based on their purpose:

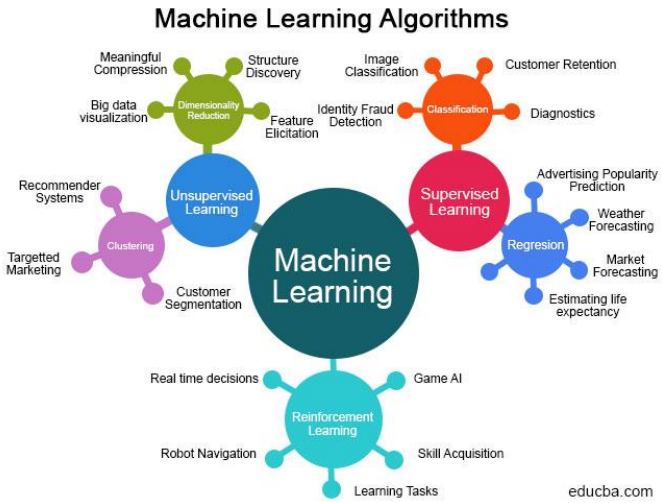

Figure 1.1 Type of Machine Learning

Today the farmers cultivate crops supported the expertise gained from the previous generation. Since the usual technique of farming is practiced, there exists an Associate in Nursing excess or deficiency of crops while not meeting the particular demand. The farmers aren't aware of the demand that takes place within the current agricultural economy. This ends up in the loss to the farmers. The expressed reasons so as of importance behind farmer suicides were – atmosphere, low turnout costs, stress and family responsibilities, poor irrigation, and an increase within the price of cultivation. The most reason is that the low costs of the merchandise and therefore the accumulated price of cultivation. The value of crops is decided by economic demand and, therefore, the limits of the assembly.

Yield prediction is a vital agricultural drawback. Each farmer is curious about knowing what proportion yield he is concerning expect. Within the past, yield prediction was performed by considering farmer's previous expertise on a specific crop.

In this paper, the main aim is to create a user-friendly interface for farmers, which gives the analysis of several crop production based on available data. Different Machine learning techniques were used to predict the crop yield for maximizing crop productivity.

2. Background

Nowadays, many researchers have implemented several models to predict crop yield predictions. However, these models have several drawbacks due to the incorrect use of the algorithms.

In [1], a predictive model is proposed to identify the crop yield using a Hybrid neural network considering the soil parameters and the external climate conditions. However, the model fails to predict the real-time fluctuations in the climate and soil parameters. Hence the model does suit real-time analysis.

In [2], A Prediction system is developed using KNN and Apriori algorithm to analyse and recommend the crops to the farmers. The System has a well-developed Interface to input crop, which allows the user to input the crop names and which it outputs the crop yield. However, the System does not have the provision to predict the price at the same time.

In [3], an automated farming crop prediction system is developed using the KNN algorithm and Multi-linear regression for Bangladesh countries. But the model is considered to be an initial step in the advancement since it doesn't identify any new research gaps.

In [4], the System is designed to predict crop yield and the fertilizer recommendation, which could be used for analysis soil and the current growth of crops, based on which the fertilizer is recommended, which is not suitable for the real-time analysis.

3. Objectives

The central focus of this research work is to predict crop prices based on the previous trends in weather, yield and price. In an Agri-based country like India, predicting crop yield and price can aid many people whose sole survival is dependent on the crop that they plan to grow.

The primary objectives of the proposed System include

- To provide the farmer with the yield of a crop, based on rainfall and WPI using machine learning.
- To predict the future market price of crops by taking previous crop prices and sowing month into consideration.
- To predict the crop price for all around the country for the next year.

 https://doi.org/10.22362/ijcert/2020/v7/i06/v7i0607

4. Proposed Work

This work focuses on investigating the prediction of Crop yield and cost estimation. The proposed methodology uses the tree algorithm to predict the results efficiently and proves to best suitable for the research work. The data collected is analyzed and worked on predicting the yield and the cost of the crops at any given time.

The research work majorly involves the following implementation modules.

1. Data Acquisition
2. Data Exploration
3. Machine learning prediction
4. Web application

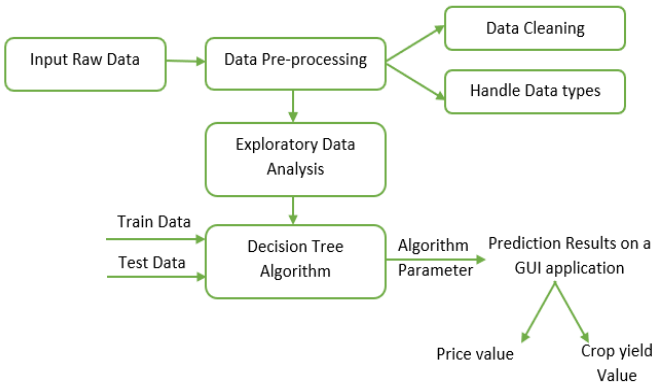

Figure 4.1 Proposed Systems

Figure 4.1 shows the architecture diagram of the proposed work; functionalities are as follows:

Module 1: Data Acquisition

Dataset is prepared by collecting the crop data obtained from the public repository, data.gov.in. There are a handful of datasets that contain data. We obtained the data which contains the details of the rainfall of the individual crops.

A sample of an acquired set of data and their attribute are shown in figure 4.2 below, where WPI is represented as Wholesale Price Index.

Month	Year	Rainfall	WPI
4	2012	47.5	104.2
5	2012	31.7	104.2
6	2012	117.8	107
7	2012	250.2	111
8	2012	262.4	114
9	2012	193.5	116.1
10	2012	58.7	116.6
11	2012	30.7	116.3
12	2012	11.7	115.9
1	2013	11.3	116.7
2	2013	40.1	117.9
3	2013	15.7	118.7
4	2013	30.4	120.4
5	2013	57.8	121.9
6	2013	219.8	124.9
7	2013	310	129.4
8	2013	254.7	132
9	2013	152.7	131.3
10	2013	129.4	131
11	2013	14	131.3

Figure 4.2 Dataset

Module 2: Data Exploration

Exploratory Data Analysis (EDA) is an essential advance that happens when components coming up with and getting info, and it needs to be done before any demonstrating. This can be because it's essential for Associate in nursing info research workers to virtually comprehend the thought of the data while not creating suspicions. The impact of knowledge investigation is often terribly useful in obtaining a handle on the structure of the data, the appropriation of the qualities, and interrelationships within the informational index.

The purpose of EDA is:

- To utilize summary measurements and representations to all ora lot of seemingly comprehend data, discover items of data regarding the inclinations of the data, its quality, and to detail suppositions and, therefore, the speculation of our hypothesis.
- For knowledge preprocessing to be effective, it is required to own a general image of your data, and Basic accurate portrayals van utilized to differentiate properties of data and the knowledge and have that information esteems got to be treated as commotion or exceptions.
- Visualization is the method toward anticipating the data, or components of it, into a mathematician area or into dynamic photos. Within the data mining method, data investigation is used during a big selection of steps together with preprocessing, modeling, and interpretation of results.

During this process of analysis, Univariate and Bivariate analysis is done. Figure 4.3 shows the distribution of rainfall across all the states of India.

 https://doi.org/10.22362/ijcert/2020/v7/i06/v7i0607

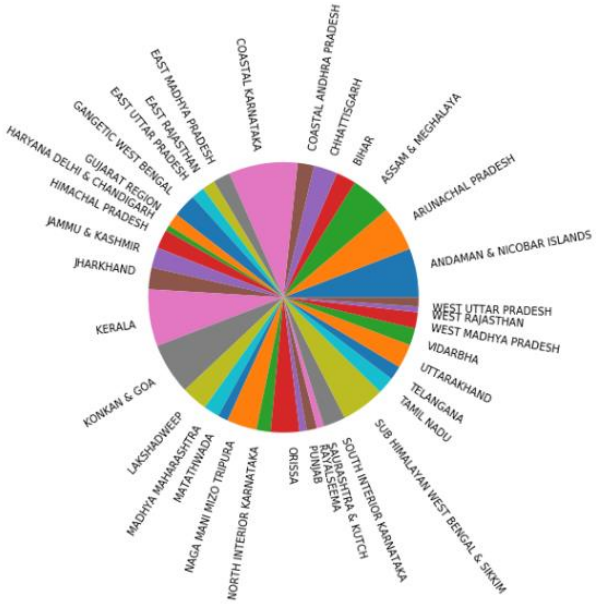

Figure 4.3 Distribution of Rainfall across India

Figure 4.4 represents the Top 3 crops of the Kharif season crops produced in the highest crop-producing states.

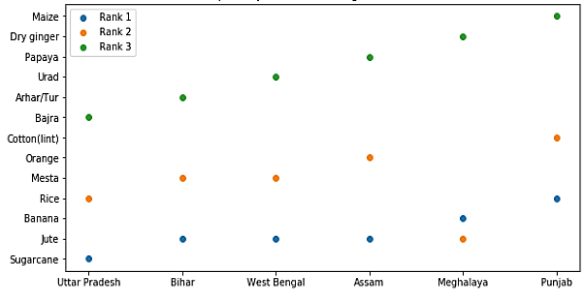

Figure 4.4 Top 3 crops of Kharif season Crop Distribution

Figure 4.5 represents the Top 3 crops of the Rabi season crops produced in the highest crop-producing states.

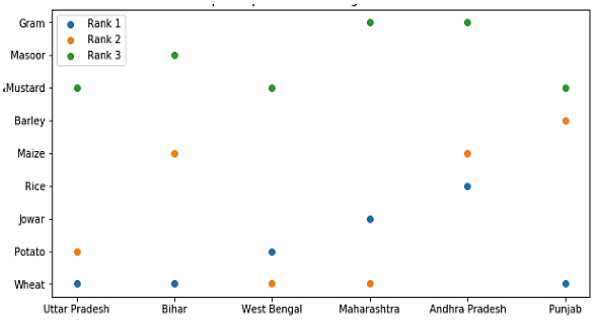

Figure 4.5 Top 3 crops of Rabi season Crop Distribution

Module 3: Machine learning Prediction

Machine learning prediction has these following steps:

1. Divide data into two parts: training and testing data.
2. We are Defining the algorithm namely Decision tree algorithm.
3. Training and testing data against the algorithms.
4. We are updating the User Interface with the calculated values.

Decision tree Algorithm

Decision tree is associate degree formula that uses a tree-like graph or model, their potential outcomes to predict the ultimate decision, this formula uses conditional management statement. A call tree is associate degree formula for approaching discrete-valued target functions, during which call tree is denoted by a learned perform.

For inductive learning these styles of algorithm's area unit terribly notable and are with success applied abroad vary of tasks. We tend to offer a label to a brand new dealing that's whether or not it's legit or fraud that category label is unknown then dealing worth is tested against the choice tree, and at that time from the root node to output/class label for that dealing a path is copied. Call rule determines the result of the content of the leaf node. Normally rules have the shape of 'If condition one and condition two however not condition three then outcome'. Call tree helps to work out the worst, best, and expected values for various situations simplified to grasp and interpret and permit the addition of the latest potential situations.

Steps for creating a call tree area unit that foremost to Calculate the entropy of each attribute victimization the dataset in downside then the dataset is split into subsets victimization the attribute that gain is most, or entropy is minimum at that time to form a call tree node containing that attribute, and in conclusion, the rule is performed on subsets victimization remaining attributes to make a call tree.

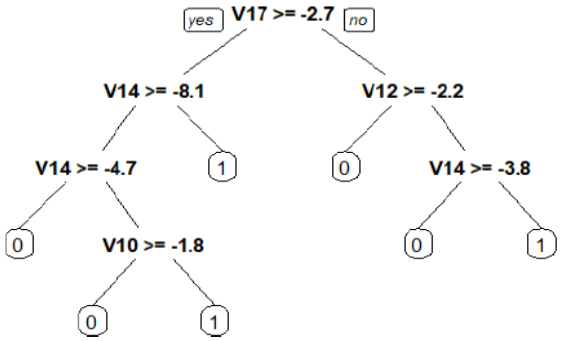

Figure 4.6 Decision tree sample

5. Results and Evaluation

In the evaluation, we want to understand, for a number of metrics, whether our method works well for the problem statement we are trying tackle. We calculate the crop yield, its increase or decrease and also its price. Figure

5.1 shows the top gaining crops obtained using Decision tree algorithm.

Top Gaining Crops

Item Name	Price (per Qtl.)	Change
Maize	₹1427.62	3.32% ▲
Sunflower	₹3748.1	2.43% ▲
Safflower	₹3445.0	2.38% ▲
Groundnut	₹4058.9	2.05% ▲
Cotton	₹4917.6	2.02% ▲

Figure 5.1 Top 5 gaining crops

Figure 5.2 shows the top losing crops obtained using the Decision tree algorithm.

Top Losing Crops

Item Name	Price (per Qtl.)	Change
Niger	₹4959.5	-7.81% ▼
Moong	₹3934.0	-6.41% ▼
Masoor	₹3298.4	-2.4% ▼
Arhar	₹3622.4	-2.33% ▼
Urad	₹4553.7	-2.31% ▼

Figure 5.2 Top 5 losing crops

Figure 5.3 shows the different crops explored in our study.

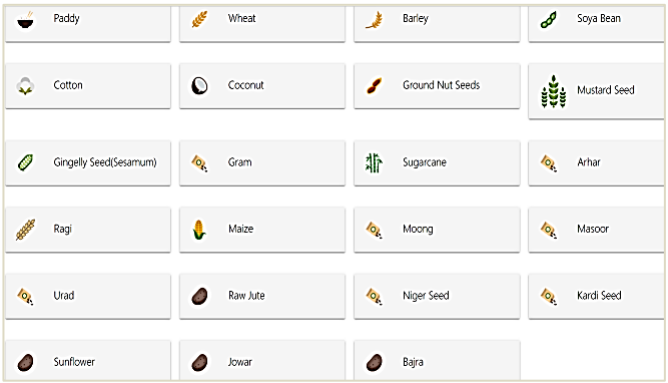

Figure 5.3 Kharif/Rabi crops

Figure 5.4 shows the 12month forecast for Rice crop, where we do the same for all the crops.

Forecast Trends

Month	Price (per Qtl.)	Change
Jul 20	₹1376.9	-2.34% ▼
Aug 20	₹1376.9	-2.34% ▼
Sep 20	₹1376.9	-2.34% ▼
Oct 20	₹1376.9	-2.34% ▼
Nov 20	₹1419.87	0.71% ▲
Dec 20	₹1543.17	9.45% ▲
Jan 21	₹1431.08	1.5% ▲
Feb 21	₹1451.01	2.92% ▲
Mar 21	₹1452.25	3.0% ▲
Apr 21	₹1436.06	1.86% ▲
May 21	₹1443.53	2.39% ▲
Jun 21	₹1409.91	0.0% ▲

Figure 5.4 Crop Yield/Price Forecast

Figure 5.5 shows the time-series graph of next and previous 12-month forecast for Rice crop, where we can show for all the crops.

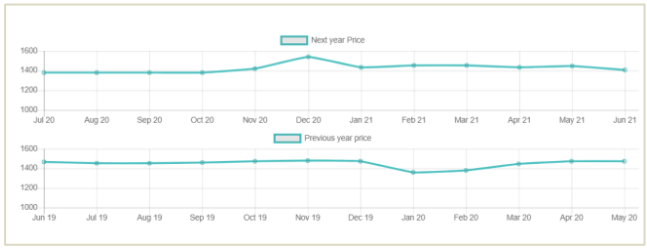

Figure 5.5 Time-series Yield/Price Forecast

6. Conclusion

The proposed model is developed to help farmers make better decisions concerning which crop is most suitable during his desired time of sowing and the location. This System predicts the yield and price of the crop of choice, giving the farmer useful information well before starting the process of cultivation.Numerous prediction algorithms can be used for crop yield and price prediction, such as decision trees, neural networks, SVM. Our model uses a Decision tree. It is trained on several Kharif and Rabi crops (like paddy, arhar, bhajra, barley, etc.), providing good accuracy.

7. Future Work

To be able to feed a growing population responsibly, farmers must increase food production on existing farmland to avoid deforestation. The longer-term work done ought to optimize farming practices to extend yields, crop quality and incomes in an exceedingly property manner.

An Advanced value prognostication system may be developed wherever a dashboard can predict the market

 https://doi.org/10.22362/ijcert/2020/v7/i06/v7i0607

value trends' statistical exploitation method for a minimum of a period and, therefore, the production pattern of various crops. A Platform for Agriculture ought to use massive information, AI, Machine Learning, satellite representational process, and weather information to assess the land area and monitor crop health on a true-time basis. So it will notice cuss and sickness infestations, estimate the crop output and yield, and conjointly forecast costs.

Other key inputs, such as the prices in markets of neighboring States, can also be factored into the price forecast. The System should be developed in a way that it does a real-time analysis to get an accurate solution to farmers' problems and optimize their farming practices. The System can introduce and make available climate-aware cognitive farming techniques and identifying systems of crop monitoring, early warning on pest/disease outbreak based on advanced AI innovation.

References

[1] S. Kulkarni, S. N. Mandal, G. S. Sharma, M. R. Mundada and Meeradevi, "Predictive Analysis to Improve Crop Yield Using a Neural Network Model," *2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, Bangalore, 2018, pp. 74-79.

[2] S. V. Bhosale, R. A. Thombare, P. G. Dhemey and A. N. Chaudhari, "Crop Yield Prediction Using Data Analytics and Hybrid Approach," *2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*, Pune, India, 2018, pp. 1-5.

[3] T. Siddique, D. Barua, Z. Ferdous and A. Chakrabarty, "Automated farming prediction," *2017 Intelligent Systems Conference (IntelliSys)*, London, 2017, pp. 757-763.

[4] S. Bhanumathi, M. Vineeth and N. Rohit, "Crop Yield Prediction and Efficient use of Fertilizers," *2019 International Conference on Communication and Signal Processing (ICCSP)*, Chennai, India, 2019, pp. 0769-0773.

[5] P. Bose, N. K. Kasabov, L. Bruzzone and R. N. Hartono, "Spiking Neural Networks for Crop Yield Estimation Based on Spatiotemporal Analysis of Image Time Series," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 11, pp. 6563-6573, Nov. 2016.

[6] N. Hemageetha and G. M. Nasira, "Radial Basis Function Model for Vegetable Price Prediction", International Conference on Pattern Recognition, Informatics and Mobile Engineering, pp. 424-428, 2013

[7] Yung-Hsing Peng, Chin-shun Hsu and Po-Chuang Huang, "An investigation of Spatial approaches for crop price forecasting in different Taiwan markets",2015.

[8] Prof A K Mariappah and J Austin Ben Das, "A Paradigm for rice yield prediction in Tamil Nadu", IEEE International Conference on Technological Innovations in ICT For Agriculture and Rural Development (TIAR), pp. 18-21, 2017.

[9] J. Hartigan, Clustering Algorithms, John Wiles & Sons, New York, 1975.

[10] Fagerlund S Bird species recognition using Support Vector Machines. EURASIP J Adv Signal Processing, Article ID 38637, p 8, 2007.

[11] Holmgren P, Thuresson T Satellite remote sensing for forestry planning: a review. Scand J for Res 13(1):90– 110, 1998.

[12] Das KC, Evans MD Detecting fertility of hatching eggs using machine vision II: Neural Network classifiers. Trans ASAE 35(6):2035–2041, 1992.

Authors Profile

Chaitra Kotnikall, Student, M.Tech, Dept. Of CSE, Shri Dharmasthala Manjunatheshwara College of Engineering, Dharwad, India.

Jayateerth Vadavi, Associate Professor, Dept. Of CSE, Shri Dharmasthala Manjunatheshwara College of Engineering, Dharwad, India.

