A SURVEY ON CROP ANALYSIS & AGRICULTURE COMMODITIES PRICE PREDICTION USING MACHINE LEARNING TECHNIQUES

Prashantha S[1], Shravan C Y[2], Bharath B[3], Bharghavachar B N[4], Prof. Shilpa B L[5]
Department of Computer Science and Engineering. Vidya Vikas Institution of Engineering and Technology, Mysore-570028.

Abstract: Agriculture is the backbone of our country. Agriculture plays an important role in economy of the country. The demand of agricultural products continuously increases with increase in population. Farmers need to think of increase in crop yield with the limited amount of land. The suicide rate is increasing with every passing year because the farmers aren't able to get the desired price for their crops and farmers need to predict the yield of the crop before cultivating into agricultural land. Farmers are not getting the proper price for which they have cultivated. Yield of the crops depends on soil parameters, rainfall, and soil moisture. Price prediction in agriculture commodity has been a major problem for the farmers. The main aim is to provide new framework and develop a system with more efficient price prediction. Using machine learning techniques the price prediction and crop Analysis can be done which reduces the farmer effect.
Keywords: crop Analysis, price prediction, machine learning technique.

I. INTRODUCTION

Agriculture is practiced in India from ancient period. [1]Around 60% agricultural land used by the farmers. Farmers yield different types of crops in various parts of our country. Crops can be cultivated based on different aspects such as soil parameters, rainfall, humidity etc. Dataset consists of 7418 rows and 10 columns of data. There are two phases first phase is crop analysis and next phase is agricultural price prediction

In crop analysis using machine learning technique we calculate which state having more demand for the crops like in Karnataka, Arecanut crop is cultivated more, in Jharkhand sugarcane crop cultivated more. Using previous years dataset the future crop price prediction is done .For price prediction the crop parameters such as timestamp, variety, minimum prize, maximum price of different crop is considered. We use machine learning methods for analyzing the crop yield prediction. In the second phase also predicting the profit of the particular crop. Arrival time of crop in the market plays prominent role in the crop price for the farmers.

[2]As there is no synchronization in the production of and demand for the agricultural commodities either farmer fail to get good market prices for their products or consumer suffers high prices due to less production. Hence, in order to reduce the mismatch in demand and supply of food crops effectively, the expected demand for various food commodities needs to be predicted to guide the farmers

accordingly. This could eliminate the gap between the consumer's demand and producer's supply and reducing the loss for both consumers and farmers. So there is a need to forecast the demand of food commodities that could help the farmers in selecting and growing the crops to satisfy the actual demand of the society.

1.1 Background

Agriculture places an important role of economy. [3]As the population of the country increases, demand for the food also increases. So we mainly focus on crop analysis and predicting the crop price using machine learning technique. Food is the most important aspect of our life, without it, people can't survive. Therefore, preserving food becomes an absolute aspect for countries all over the world, especially the developing countries. During the natural disaster preserving enough food for the people plays an important role for the government of such countries. Prices of crops, vegetables and fruits are very important necessity and people's lives are very closely related to it. Social and economic factors are responsible for the fluctuation of prices of agricultural commodities. Various Agricultural commodities are available in the market. There are many factors that influence the prices of agricultural commodities and even the same commodities can have different price in diverse markets.

II. RELATED WORK AND DETAILS

Machine learning (ML) is a function of artificial intelligence that contributes the system to learn automatically and improve from experience without being programmed. Machine learning focuses on
accessing data for the development of computer programs. Machine Learning deals with problems where the relation between input and output variables is not known or hard to obtain. Here the "learning" term denotes the automatic acquisition of structural descriptions from examples of which is been described. ML does not make assumptions about the correct structure of the data model, which describes the data. Machine learning characteristic is useful to model complex non-linear behaviors, such as a function for crop yield prediction. For Crop Yield Prediction the ML techniques are more useful.

2.1 Analysis and Prediction of Future Market Price for Agricultural Commodities

Farmers are the backbone of our India and agricultural

sector. For the agriculture sector agribusiness is one of the major segments. This paper mainly aims helps the farmer community to minimize agribusiness risk. Using soft computing technique we can predict agriculture commodity market price. With this knowledge farmer community can easily make the decision to achieve the goals of profit making.

The prices of agricultural commodity have a unstable nature which may rise or fall differently causing negative effect on the economy. The work completed here for predicting costs of horticultural wares is helpful for the farmers as they can sow crops depending upon its future cost. Farming items have regular rates; these rates are spread over the whole year. In the event that these rates are known to farmers ahead of time, it will be guarantying on Rate on Investments (ROI). Horticultural specialists can pursue these charts and anticipate advertise rates which can be informed to farmers. The outcomes will be given dependent on the area of the clients of this application.

2.2 Machine Learning in Agriculture on Crop Price Prediction

Commodity	Variety	Min_price	Max_price	Modal_price
Ginger(Green)	GreenGinger	1000	1000	1000
Arhar Dal	Arhar Dal	5500	6800	6150
Avare Dal	Avare	2341	2341	2341
Bengal Gram	BengalGram	4400	5800	5100
Bengal Gram	Average	4000	4600	4300
Corriander	Coriander	5500	14000	9750
Dry Chillies	Local	4833	4833	4833
Garlic	Other	1500	5000	3250
Green Gram	GreenDal	7200	7600	7400
Ground Nut	Ground Nut	7600	9000	8300
Ginger(Green)	GreenGinger	1000	1000	1000
Arhar Dal	ArharDal	5500	6800	6150

Table 1-Commodity and Variety with Prices

As shown in the Table (1), the average price has been calculated by taking minimum price and maximum price of different commodities. This model price can help the farmers to predict the price of crops in a better way.
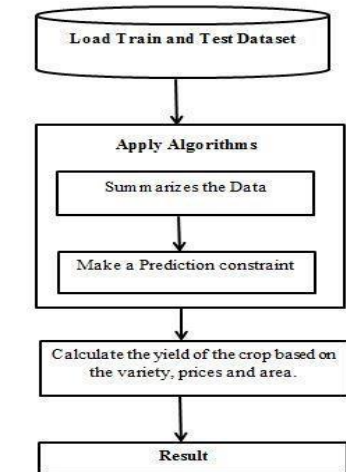

Fig.1. Predictive Models and Techniques

Predictive technique play an major role in decision making System. [4]Predictive models will build the abstract model to express the output variable as a function of descriptive variables to predict the future. In general, the predictive models can be categorized four major parts, namely, Load train & dataset, Apply algorithms, Calculate with predict and Result display techniques as shown in Fig. 1.

2.2.1. Linear Regression models
[3]Linear Regression is a machine learning algorithm which performs a regression task. Using regression models a target prediction value based on self-determining variables. The relationship between variables and forecasting is find out using Linear Regression models. Regression model fits a linear line with the data points, so that the distances of data points from the line is minimized. This model is represented by a linear equation (1). Different regression models differ for both dependent and independent variables. Using independent variable (x) the dependent variable value can be performed. The relationship between x (input) and y(output) find out using regression technique. so, the name is Linear Regression.

$$D= I +S*X +e \quad (1)$$
Where,
- D is the output variable or dependent variable
- X is the input variable or independent variable
- I is an Intercept of the line
- S is the slope of the line
- e is the random error
- N is number of observations

2.2.2. Logistic Regression models

[5]Logistic Regression is another machine learning algorithm which is used to predict the probability of a categorical dependent variable. The logistic regression has a binary variable that contains data coded as 1 or 0, this model also predicts Y=1 as a function of X. The liner Regression checks whether it is true or false using previous year data set.

$$Y=b_0+b_1x_1+\dots\dots\dots+b_ix_i.$$
Where,
Y= Dependent variable.
b_0=Constant
$b_1,...b_i$=coefficient of variables $x_0,x_1,....x_i$.
x_1=Independent variable.

2.2.3. Decision Tree Learning
[5]Decision tree learning is a method for approaching discrete-valued target functions. Decision trees categorize instances by arranging them down the tree from the root to some sub node and it provides the classification of the instance.

Decision Tree algorithm comes under the family of supervised learning algorithms. Rather than other supervised learning algorithms, the decision tree algorithm which is

used to solve regression and classification problems. The Decision Tree mainly helps in creating the training model that can use to predict the class or value of the target variable by learning simple decision rules contingent from training data. Decision Trees will predict the class label from main root node of the tree. We have compare the values of the root attribute with the record's attribute. They have the comparison, we follow the branch corresponding to that value and jump to the next node.

Decision trees will sort root to some leaf node, with the leaf node providing the classification of the example. Each node in the tree acts as a test case for some attribute, and each edge descending from the node corresponds to the possible answers to the test case. This process is recursive in nature and is repeated for every sub tree rooted at the new node.
Root Node: It is parent node and it comes first in the tree representation.
Splitting: It is a process of distributing a node into two or more sub-nodes.
Decision Node: When a sub-node splits into further sub-nodes, then it is called the decision node.
Leaf / Terminal Node: If the node doesn't split is called Leaf or Terminal node.
Pruning: In a tree, we remove sub-nodes of a decision node is called pruning.
Branch / Sub-Tree: Subsection of the entire tree is called branch or sub-tree.
Parent and Child Node: In a node, which is divided into sub-nodes, then we say parent node of sub-nodes whereas sub-nodes are the child of a parent node.

2.2.4. K-nearest neighbor's classifier
[5]The k-nearest neighbors (KNN) algorithm is a machine learning algorithm, which is used to solve both classification and regression problems. Rather than calculating distance between points in normal method, KNN algorithm finds out the distance between the points in a simplified method. KNN helps in solving problems that have solutions that depend on identifying similar objects. Though, the straight-line distance is a popular and convenient choice.

KNN algorithm works as stated below
1. KNN will first load the data
2. Later predict the total number training data points
3. Calculates the distance between the data points
4. KNN sorts distances in ascending order
5. Later sort the array

III. PROPOSED MODEL
Crop analysis and agriculture commodities price prediction is an important aspect for the farmers. Prediction is a statement about future events. Price prediction for agricultural commodity has become the need of the hour for farmers. Although future events are uncertain, so accurate prediction is not possible. This paper includes a decision making support model that can be helpful for farmers to predict prices. For the price prediction the system uses classification rules. System predicts crop prices based on the attributes

such as date, commodities, yield, maximum and minimum trades.
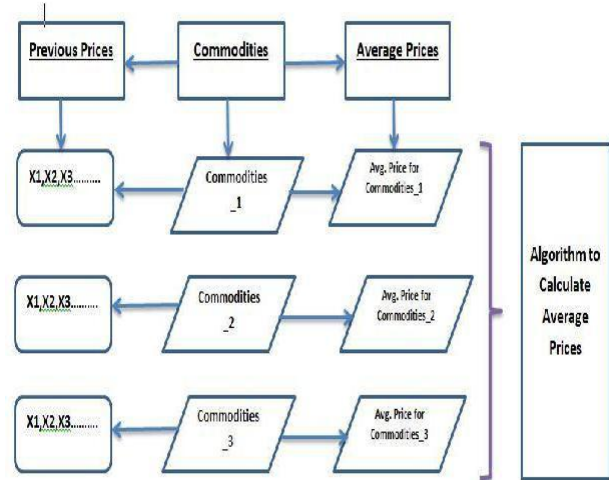

Figure-2: Price Prediction Support System for better price prediction in agricultural commodity.

As shown in the Fig(2), we are calculating average prices of different commodities using machine learning algorithms. Farmers have to get commodity name and previous selling price of the crop. Based on the previous prices, this model will be able to provide average prices for a particular crop which will be beneficial for farmers to make better decisions and predict prices.

IV. CONCLUSION
This method is to provide help to the farmers for expecting the better amount for their crops and for predicting the best price for the crops. This also helps the farmers to check previous prices of different commodities. This system will allow farmers to make better decisions for bidding the better prices for their crops in the market. Finding crops that have special changing patterns over the years like decrease in production can help us understand the reasons behind them in a more specific way. We can also use these data to recommend new crops that can be grown in places which has the suitable weather and economy conditions. The above results can be combined to help farmers decide their crops on more monetary gain and with little risk. Government can be better prepared for anomalies with better resource arrangements such as insurances, logistics and resources.

REFERENCES
[1] Zingade, D. S., Omkar Buchade, N. Mehta, Shubhm Ghodekar, and Chandan Mehta. "Crop prediction system using machine learning." Int. J. Adv. Eng. Res. Dev. Spec. Issue Recent Trends Data Eng 4, no. 5 (2017): 1-6.
[2] BV, Balaji Prabhu, and M. Dakshayini. "Performance analysis of the regression and time series predictive models using parallel implementation for agricultural data." Procedia computer science 132 (2018): 198-207.
[3] 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.

[4] Priya, P., U. Muthaiah, and M. Balamurugan. "Predicting yield of the crop using machine learning algorithm." International Journal of Engineering Sciences & Research Technology 7.1 (2018): 1-7.

[5] Tom M. Mitchell, Machine Learning, India Edition 2013, McGraw Hill Education.

[6] Madaan, Lovish, Ankur Sharma, Praneet Khandelwal, Shivank Goel, Parag Singla, and Aaditeshwar Seth. "Price forecasting & anomaly detection for agricultural commodities in India." In Proceedings of the 2nd ACM SIGCAS Conference on Computing and Sustainable Societies, pp. 52-64. 2019.

[7] Zhang, Yongli, and Sanggyun Na. "A novel agricultural commodity price forecasting model based on fuzzy information granulation and MEA-SVM model." Mathematical Problems in Engineering 2018 (2018).

[8] Dey, Umid Kumar, Abdullah Hasan Masud, and Mohammed Nazim Uddin. "Rice yield prediction model using data mining." 2017 International Conference on Electrical, Computer and Communication Engineering (ECCE). IEEE, 2017.

[9] Rajeswari, S., and K. Suthendran. "Developing an Agricultural Product Price Prediction Model using HADT Algorithm."

[10] Shalev-Shwartz, Shai, and Shai Ben-David. Understanding machine learning: From theory to algorithms. Cambridge university press, 2014.