Optimizing the Productivity and Increasing the Profitability of Oil Seed Corps using Data Analytics

K.Baskar, S.Arivalagan, P.Sudhakar

Abstract*: This study discuss the about the optimization of the yield of the oil crops by using Adaboost technique for predicting the desired results. Big data analyses are used to collect and process the required information. Agriculture is the backbone on the Indian economy. The yield of the various crops depends on the rainfall, weather, types of soil, fertilizer used, irrigation methods and cultivating procedure. Every year, the mismatch in the demand and supply results in major loss to the crop producers, the issues can be addressed by preprocessing the data scientifically. For this process a novel OSM method is used. Our proposed method gives suggestion about when they sow Oilseeds and Sell grain in Market (OSM) that they required more information. In this study data related to various crops has been collected from different sources. Adaboost technique is used to process the data for predicting the required information for the optimum benefit of the producers. This method of analysis gave a better solution to the producers to gain maximum yield and profit.

Index Term: Big data analytics, Adaboost, Yield, Preprocessing, oilseeds marketing price.

I. INTRODUCTION

This paper discuss about the problems associated with the increase in the population of the world; mankind must adopt modern policy framework in different methods of agricultural activity. The existing scenario faces 30 years sustainability around the world and 70 percent more food demand [1]. In future some of existing resources goes into starvation like reduced fertile land, climate change, and high cost seeds - all of which severely affect the production of many food products [18]. The concept of culinary oilseeds crop and smart agriculture are based on the use of innovative solutions by farmers, such as increasing yields and costs in such conditions. According to future market intelligence analysts, this type of home care is used to predict the future oilseeds market trends. Most of the market (about 53%) is now located in United Nations Organization (UNO) listed, which is of interest to Indian farmers' "smart" innovations [2].

Revised Manuscript Received on September2, 2019.

K.Baskar, Research Scholar, Dept. of Computer Science and Engineering, Annamalai University, Chidambaram, Tamilnadu State, India. (email: baskar_ka@yahoo.com)

Dr. S. Arivalagan, Asst. Professor, Dept. of Computer Science and Engineering, Annamalai University, Chidambaram, Tamilnadu State, India. (email: arivucseau@ gmail.com)

Dr. P. Sudhakar, Associate Professor, Dept. of Computer Science and Engineering,, Annamalai University, Chidambaram, Tamilnadu State, India. (email: kar.sudha@ gmail.com)

Intellectual technologies are introducing a small range of field farms, livestock breeding and fisheries industries. There is tremendous progress in spreading "smart" solutions in the field of grain cultivation. Consider the major innovative solutions of the concept of "smart" agriculture. This paper finds a solution in novel ways for existing farmer problems. So, this paper uses the OSM method of analyzing big data of oilseeds information, the result of analyzing big data information available from this will be a source of farming for farmers. The farmers have been advised to be more interested in management and production of oilseeds Groundnut, sesame, castor and sunflower are being the main oilseeds grown in India. The farmers are required to follow the procedures to increase the crop production rate likes spraying of fertilizers, micronutrients, nutrient mixture spraying, seed size, seed treatment, weed control and irrigation are advised by the Agriculture Department for better yield [3] [4]. Assistance schemes are offered by the Government of India for the farmers who need to cultivate oilseeds, Furthermore, the Agriculture Department said that additional information on this may be related to Agricultural Officers in the respective area. But, farmers cultivate sugarcane due to the low cost of sugarcane cultivation in India.

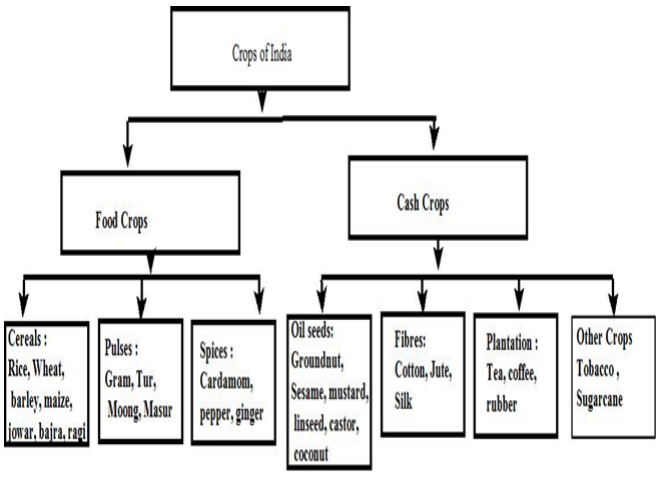

Fig. 1. List of Food Crops and Cash Crops in India

Sugarcane production was huge than the oilseeds production. Sugarcane and oilseed crops are both heat-bearing sowing plants. Sugarcane and oilseed crops are two ways to reduce the cost, which means that the cost of maintenance is low as it

includes only fertilizer. The lack of new sensor technological advancement in high yields of oilseeds is one of the main causes of growth retardation. Oilseeds span increases the cost of cultivation. The oilseed crops price is very high compared to other food crops like wheat and rice. In much of the 20th century and early periods, India was an important exporter of vegetable and oil trading that occupied an important place internationally [5]. In addition, this paper is trying to make a new way of thinking as a solution, such that the oilseeds production rate of the past and the use of new sensor technologies, from the point of time. In the farmer who lets us grow up will be able to gain more profit which is shown in Fig. 1. Thus, in first type of analyzing process is obtained big data information in the past history of oilseeds production, rainfall, and weather condition. Second type of analyzing process is obtained big data information from a particular oilseeds agricultural land Farmers face certain problem of cultivating oilseed crops with fixing new type of sensors on frame. So, the Indian government has provided a solution. According to the food security Act., the government of India has framed the Indian Government's policy to overcome natural calamity happening here. So, this will enable farmers to make a profit [13] [21].

What was the biggest challenge to seeing self-sufficiency in food production fifty years ago and what is the current challenge?: There are two big challenges now. One is ecology, and the other is dependent on the economy. The first challenge is protecting the basic assets of our agricultural fields, such as land, water, and biodiversity. The next challenge is to continue Frame is a farmer who continuously cultivates same kind of oilseeds in his land, say to increase the productivity without contaminating the environment. The groundwater level is vital for the green revolution but some states follow stringent practices like rain water harvesting while in other states it went to much low level. During the Green Revolution the total population of the country is between 40 crores and 50 crores. Now it is 130 crores. In 2030, this research focuses it to be, say 150 crores. Increasing the population is essential for increasing the potential of cultivation [6]. The growth rate of oilseeds crops has been increased due to the population growth in this category. This is possible only if the modern theory has had the knowledge of the growth of oilseeds [19].

There are different opinions about Genetic Change Technology (GCT). Can GCT crops help to resolve food shortages?: The production of Oilseeds follows older sowing format in the existing scenario and hence food production is very low due to their oldest practices. The action taken by the Indian government to create awareness on recycling and hybrid recycled oilseeds. There are many ways to increase the reproductive capacity of crops. Genetic transformation has advantages and disadvantages. One has to decide what is the most beneficial. Genetically modified crops require quality control system in our country. In order to consolidate the biodiversity conservation, the research program aims at integrating genetically engineered crops for development. Farmers need to find a way to achieve only the benefits of genetic crops in the absence of hazards. Now, the technology is developing a new problem when dealing with genetic technology. The complex given above is extracted from this paper [6].

Is organic farming sufficient for oilseeds in crop production?: There are three aspects of natural agriculture. Farmers should own cattle to produce natural manure, the second aspect: They have the ability to control the insects and third aspect: diseases that occur in the oil seeds crop. They need to handle crop rotation. Genetically, the ability to resist pests and diseases can help them [7].

Droughts and floods are caused by changes in the rainfall due to seasonal changes; Thus agriculture has become an unpredictable job; how do farmers face these challenges?: Low rainfall and an increase in average temperatures are affecting agriculture. We see the drought, the overwhelming rainfall, and the rising sea levels which is shown in Fig. 2 and Table 1. This guideline is provided by the Agriculture Department of the Government of India. From this paper has given our main points with important information about sowing oil seeds. This paper sets some guidelines for adjusting to this position and recovering from its impact in agricultural fields.

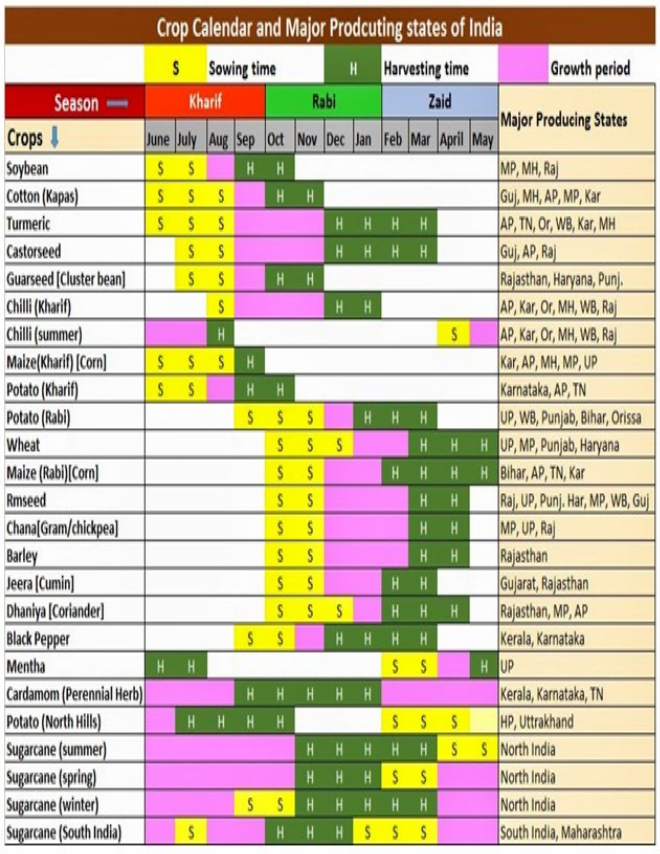

Fig. 2. Crop Calendar and Major Production States of India

A climate management committee with multiple officers are involved in each states affected by drought. The timely counseling of farm families should be carried out to increase its benefits during normal rainy season and reduce the drought impacts. There are three main seasons, 1. Kharif Season = June to September, 2. Rabi Season = October to Jan, 3. Zaid Season = February to May [8]. Maintaining animal care camps and saving them from destruction can also help the situation. People should be empowered to take care of the

management of food in a crisis situation. If the temperature increases, the wheat yield will decrease. More varieties that produce more than one yield are required to produce high growth daily. These will give a higher yield in short time. The top of the underlying issue of the paper is from taken from the [8].

India is in the worst place in the past to reduce starvation. What should this paper do to fix this?: India's record in yield production is backward. One side of the farmland has grown tremendously, and on the other side the poor are starving. The food safety law should be implemented properly. Smaller grains with regular yield capacity should be included. Development of agricultural productivity is very important. Growing agricultural profits and fast growing demands exist to meet the growing global needs. The future of agriculture relies on growing crops with the understanding of crop and many environmental aspects. Based on the data collected by Agricultural Research, crop efficiency, its eco-cultivarability, and the ability to protect their diseases in different environments (Eg. soil quality and environmental conditions) can be traced to crop performance [10]. The manually collected crop performance is very low in quality because its earlier conditions are not taken into consideration by individual operators. New technologies, devices (Eg. wireless sensor networks, network-connected weather stations, cameras, and smart telephones, etc.),

Big data and IOT-based technologies can collect a wide range of harvested and environmental performance data, apart from spam cameras, cellular smart-phone software and personal observations recorded by blue tooth. Such data can then filter out the right data and calculate crop recommendations for any particular farm [11]. In this report is study for the selection of environmental, soil, fertilization and irrigation information; data and filter out incorrect data from the viewpoint of crop measurement associated with the machine, crop projections and customized harvest recommendations are introduced to the specific farm. It is extracted from this paper [9].

List of India Major Crops	Temperature (Min. and Max.)	Rainfall	Soil
Rice	22-32 °C	150-300 cm	Deep clay and brown soil
Wheat	21-26 °C	75-100 cm	The well-arid fertile salt and clay loamy
Millets	27-32 °C	50-100 cm	They are less sensitive to soil deficiencies. They can be grown in inferior alluvial or loamy soil They are less sensitive to soil defects. They are grown in brown soil also.
Grams	20-25 °C	40-45 cm	Loamy Soil
Sugar Cane	21-27 °C	75-150 cm	Deep rich loamy soil
Cotton	21-30 °C	50-100 cm	Black soil of Deccan and Malwa Plateau. However, it grows well in the soil of the Sutlaj-Ganga plain and red and late soils.
Oilseeds	20-30 °C	50-75 cm	Well-drained light sand poles, red, yellow and black soils are suitable for its cultivation.
Tea	20-30 °C	150-300 cm	Well drained, deep gelatin mud
Coffee	15-28 °C	150-250 cm	Well drained, deep gelatin mud

Table 1. List of India Major Crops with Soil, Temperature and Rainfall The Table 2 represents different type of crops and its total area used in the Indian hectare, production and yield. The total oilseed production in India was expected as 32.10 million tonnes in 2016-17, which is higher than the 6.85 million tonnes of production produced throughout 2015-2016. It maintains agriculture as an occupation for the agricultural farmer. When a large population demands food and agricultural products are increasing their efficiency and wage increases, small farmers are unable to make adequate income for a decent livelihood.

II. RELATED WORK

The Government of India has provided high technology applications for Indian agriculture. Thus, Indian agriculture has been benefited greatly. That is, weather data and historical data of old agricultural production of oil seeds, indicate that the Indian peasants reap integral benefits. In the present time, there is a wide range of big data on farming for farmers. This

Crops	Area (Lakh Hectare)			Production (Million Tonnes)			Yield (Kg / Hectare)		
	2014-15	2015-16	2016-17*	2014-15	2015-16	2016-17*	2014-15	2015-16	2016-17*
Rice	441.10	434.99	431.94	105.48	104.41	110.15	2391	2400	2550
Wheat	314.65	304.18	305.97	86.53	92.29	98.38	2750	3034	3216
Coarse cereals	251.70	243.89	247.71	42.86	38.52	44.19	1703	1579	1784
Pulses	235.54	249.12	294.65	17.15	16.35	22.95	728	656	779
Food grains	1243.00	1232.18	1280.26	252.02	251.57	275.68	2028	2042	2153
Oilseeds	255.96	260.87	262.06	27.51	25.25	32.10	1075	968	1225
Sugarcane	50.66	49.27	43.89	362.33	348.45	306.72	71512	70720	69886
Cotton@	128.19	122.92	108.45	34.80	30.01	33.09	462	415	519
Jute & Mesta#	8.10	7.82	7.66	11.13	10.52	10.60	2473	2421	2490

*4th advance estimates @Production in million bales of 170 Kg each. # Production in million bales 180 Kg. each. Table 2: Annual report 2017-2018, Department of agriculture and cooperation, Area, production and yield of major crops can be achieved by linking with the knowledge of the farmers and with the addition of new sensor technologies with big data analytics tools. Even if the farmer is not expecting the same kind of information, it cannot be used to make the best results in real time for farming. It is a very important challenge to extracting knowledge from big data. Currently the use of sensor is controlled and convenient in agriculture. The sensors hand over their information to its control tool. Most of the information is stored in the Big data. This requires a high efficiency technology to compute it. 60-80% penetrate precise agricultural technologies in the other country most advanced agricultural areas, the most common being computer with high-speed internet access, soil samples analysis (98%); Yield Maps, Crop Monitoring, GPS Navigation Systems (~ 80%); virtual Reality (VR) techniques and recommendation cards (recommendations) are used by over 60% of respondents, synthetic disorders and plant analysis of vegetation are used by 30% of farmers. Although new advances in use of drones are skiing, the use of software for data collection and processing technologies is the most common practice (80%), followed by the use of VR technology for the use of nutrients and fertilizers, utilization and seeds sowing (50-60%).

Restrictions and Opportunities for Production of Oilseeds in

India: Restrictions over the growth of oilseeds in India can be broadly classic Improvement in technology [5]. The complex given above is extracted from this paper [20]

A. Threats

Low cultivation of oilseeds in field as a sub-crops or main crops, which may be affected by problems like poor watering and pests. So, farmers can get lack of benefit in various ways I) Since massive production, distribution of seeds and other inputs are not adequately arranged, II) cultivation-land for finance and corporate credit facilities, storage, marketing and processing facilities and technology , III) The worst economic condition for farmers and the risk of oil seeds and IV) Adoption of agricultural technology and sample agriculture, various group activities, poor performance of crops [15].

B. Future Opportunities

Growths of oilseeds in industrialized countries have been down in recent years. American manufacturing scenario and other high industrialized countries have significantly set back. This is important because of the unchanging land and basic economic activity. In addition, the need for vegetable oil is increasing in third world countries. In this manner, cultivation in countries like India is the best in the vast areas. Export of cooking and industrial oil in the world is likely to meet growing demands elsewhere. Insecticides, groundnut, flax and castor seed products have shown a decline since 1983-84 as per annual report of Indian agriculture reports. Production trends in sunflower and palm kernel case, production showed positive trend. Vegetable oil production is normal in the case of large developed countries. Our country's oilseed crop production is low when compared with other countries. To avoid this paper has to follow the customs. In the US and Canada, France helps to maintain production levels using 'progressive technologies'. However, countries such as India have higher demand. Natural resources and cultivated lands in developing countries should follow advanced technologies. Lands can be utilized with the latest technologies to ensure proper production. Continuous crops are practiced by creating harvesting and proper crop management techniques [15] [20].

C. Data Mining Techniques

Data execution reveals a trend to help in the processing of information. There are many data mining algorithms, modeling and processing to keep the existing information out of the big data storage. Intelligence analysis based data mining is one of the Big data analysis technologies that give them the opportunity to earn more profit from the existing farms and know how to increase crop production [22]. This is not an exaggeration: business analysis is a small part of this powerful tool. From this, the price of oilseed crops and the price of the crops in the past can be understood. The farmer's oilseed production will not be able to perceive the fall. Data mining policies can be long-known, but with larger data evolving, they are still widely used. Relatively simple and straightforward statistics are not enough when working with large data sets.

Sometimes there may be error, studied the data received from two different information database. The information this paper receive is accurate when this researchers store all the information in one place. The farmers can use this information to generate greater profit in the future. These requirements have created complex data mining processes and a great way of giving big data technology. To solve problems, analytical data is required, a description of which is designed to describe, and the final statement was created. The large data combines three dimensions: volume, speed and variety. Volume: the use of sensors on agriculture land will increase the information received from it. When companies fill all data types, it is growing and can easily reach terabytes and even petabytes. Speed: Within 2 minutes or less. The data is stored at various point of time during oilseed plantation. So the same event will have to be stored multiple times with different times. In time-sensitive processes such as fraud detection, large data should be used when entering the company in order to increase their value. Variety: can be any configured or unstructured type, which contains large data from the oilseed crops cultivation, sensors, sound, video recording, visited web pages lists, protocol files and other data text data. When these data are analyzed in different ways, new ideas are born. Intelligence analysis based data mining is made by simple tools using relatively simple database systems and one's own built or prepared software packages. Sophisticated data mining is based on defined past experiences and protocols using existing software and packages, and various specialized tools are related to different methods. A server with large data sets, as well as their processing allows one to create complex generalizations of data mining results with groups and data comparisons. New systems and tools, such as integrated storage and processing systems, are now available. Big databases are strictly regulating and strictly follow the specification, facilitate queries to them, and analyze data with known format. Consider some important data mining technologies. The association (or relationship) is the most popular method of data mining. Using this type of method can help the farmer to get the information needed quickly; this method is often the same type when compared to two or more elements. The classification method is used to describe a number of attributes to identify a particular class, as well as other methods of input. Decision is made to determine the classification used by trees. Each class has some characteristics that represent its meanings. This is a good example of a client's credit rating.

D. Proposed System

There are many problems in Indian agriculture; some of them are natural, some are human made. During the period 1970-71, 162.1 million hectares of land area got increased to 165.5 million hectares in 1990-91 and decreased to 159.4 million hectares in 2000-01 due to construction activity in cultivable lands for living [23].
Our proposed work solves in big data analytic on optimizing way which is helpful to farmer.

Farmers obtain their profit from predictive analysis in big data analytics.

In First Type Analyzing Process is obtained from agriculture land: Analyzing and receiving information of oilseeds is understandable and reasonable. Now learning techniques have been used. It is used to differentiate herbicide and crop varieties. The use of water from irrigation to the environment is used for neural network. This is not the only reason for the use of the sensor; it is also used to monitor the seeds of sunflowers to prevent the species from being destroyed. For seed crops, the number of seeds and amount of sand on the earth, Sensor is used to prevent the pig breeding of crops; The use of the sensors is in preventing the pain caused by the pig land and the nature of the air in the ground normally. The sensor is used to distinguish between soyabean growths. The artificial neural network is used with this sensor.

E. Dataset Collection

This paper has gathered oilseeds data set from Indian agriculture department. Information inoilseeds field is derived from the production center of oilseeds in various districts. Our data set has information of production rate from past 1992 to 2018. Oilseeds production is obtained from each district which has contributed information to the Indian agricultural sector. Weather and rainfall information is available from the Indian Weather Department which is from 2006 to 2018 in which data field is from January to December. This paper found the level of water using the sensor in the oil seeded farming for agriculture and find the value of pH of the soil.

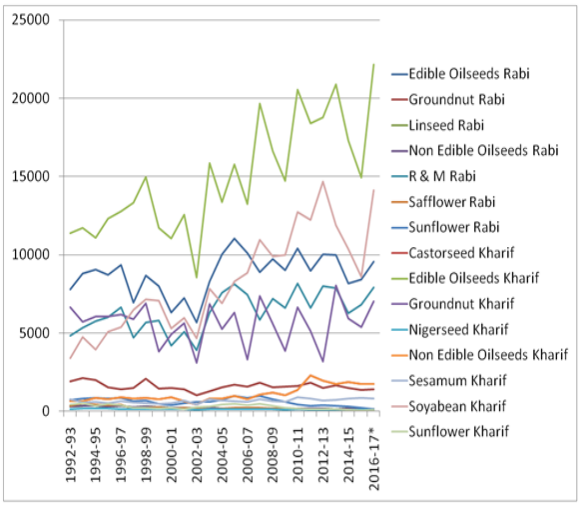

Fig. 3.B Various Type of Oilseeds

In Second type of Analyzing Process: Oil seeds are an important place in Indian agriculture. Seed crops are an important source of high yield and sustainable growth in agricultural production. Production of quality oil seeds is important. Unfortunately, good quality oil crops have reached the majority of the peasants, especially small and marginal farmers, mainly the price of seeds and crops. In 1964-67, the high yielding oil crops was started as the main thrust of the crop growth in the country.

Figure 3 shows different types of oilseeds sowed in India, cultivated in different climates. When new technologies are used in agriculture, the consequence results are varied since 2010as shown in Fig. 3. The edible oilseeds and soya bean produce good yields in two different periods.

F. Quality of Fertilizer

52 Fertilizer Control Labs have been set up in various parts of the country to maintain the quality of fertilizers. Additionally, there is a Federal Fertilizer Control and Training Institute in Faridabad with three regional centers in Mumbai, Kolkata and Chennai. Pests, germs and weeds cause one third of the total loss of production. Crops (pesticides, herbicides and chemotherapy) are used to protect oilseed crops and avoid losses. The increased use of these inputs is stored in unnecessary waste from crops, especially food crops. But the indiscriminate use of organisms results in widespread environmental pollution, which takes its own count. Figure 4 shows the average yield of oilseeds from 1992 -2018, edible oilseeds rules the majority of the growth rate from 1992-2018, R& M rabi and soya bean kharif oil seeds occupy the second and third place of oilseed growth rate in India.

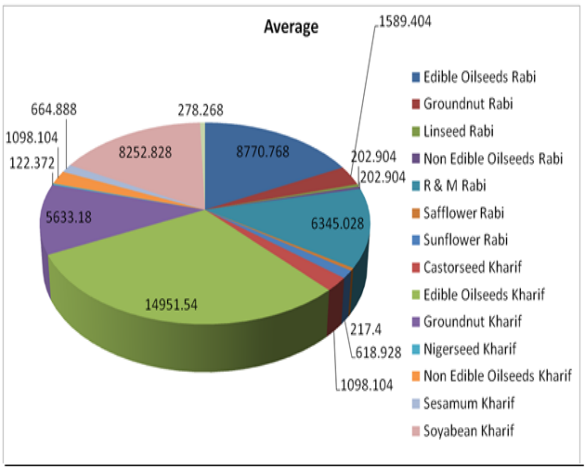

Fig. 4.B Average Yield of Oilseeds from 1992 -2018

Figure 5 shows 33 states' oilseeds production from 1992 to 2018. Most importantly Madhya Pradesh and Maharashtra are the two states playing a very important role in the production of oilseeds. Additionally, Rajasthan and Punjab states play a very important role in production of oilseeds in India.

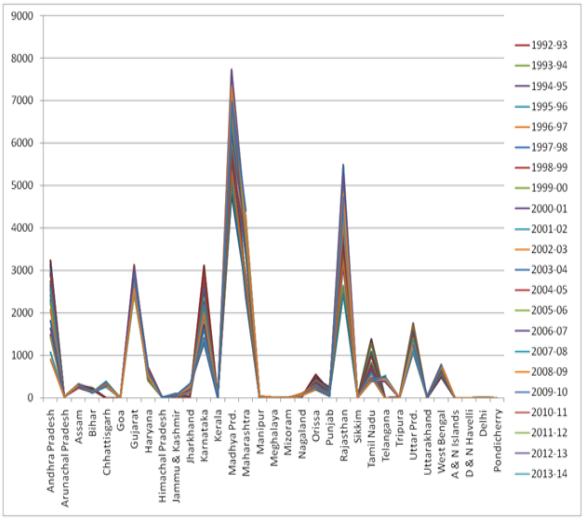

Fig. 5. B States' Oilseeds Production from 1992 to 2018

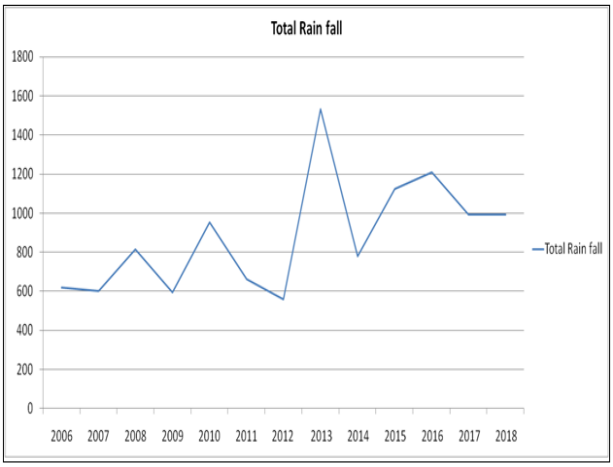

Fig. 6. The Average Rainfall from 2006 to 2018 in India

G. Irrigation and Rainfall

Although India has world's second largest irrigation land after China, only one third of the irrigation area is being deployed. Rainfall is uncertain, unbelievable and embarrassing In India, tropical rainfall produces the most important agricultural products in India. In India, the traditional irrigation methods are not followed. However, irrigable, areas must be protected against adverse effects. Irrigation is very important for crop and oilseeds. If the sunflower plant needs more water and it develops in dry
weather, the growth and production rate will decrease. Groundnut does not need heavy rains which will reduce. Soyabean crop requires moderate temperature. Therefore, irrigation is important to balance the temperature. To understand this, we use this new technology sensor to determine the level of water in the agriculture land and make it the best way to do it. Larger channels in Punjab and Haryana are ineffective due to improper irrigation (due to salt, alkalanite and hydrofluoric areas).

Figure 6 shows the average rainfall from 2006 to 2018, in the period from 2012 to 2014, the rain fall is high and the indicated rain fall in the following periods is average. When farmers see the growth of oilseeds in various districts it depends on the rainy season.

H. Technologies in Agriculture

New technology is very important for oilseeds to grow and increase its growth rate. Agricultural operations are necessary for mechanization, thus avoiding the reduction of the labor force, and agriculture is convenient and efficient. Agricultural Tools and Machines are a major input for efficient and timely agricultural activities, simplification of many cultivation and productivity. Unless new technologies are used, such a situation arises because poor farmers are waiting for a long time to harvest their crops.

The big farmer uses new technology and they reap the benefits, but the small farmers do not use it because they do not have enough money to use new technology. Our new knowledge technique is the best way for them. Small farmers are aware of crop yielding rate, sales rates , climate description and rain fall in advance. He is forced to sell his obligations and pay debt, and sells the poor farmer and precious things at his price.

The rural credit survey report, generally producers, sells their products in unnecessary places, buying unwanted terms, usually unwanted terms. In the absence of an organized marketing system, private traders and arbitrators dominate the marketing and trade of agricultural products. Although the producer does not receive similar benefits, the services paid to the middlemen by the burden on consumers. Table 3 shows the average production of oilseeds ('000 tonnes), weather(°C), and rainfall(mm) from 2006 to 2018 in India. It is prepared from Annual Report of 2017-2018, Department of agriculture and cooperation When the oilseed production is too high, the rainfall is moderate, and also weather condition is average. When rainfall is too high, the average of oilseeds production is 4125.3('000 tonnes) and weather condition is 19.83°C in year 2013. When weather condition is high, the average oilseeds production and rainfall were above average.

Year	Oilseeds production ('000 tonnes)	Rainfall (mm)	Weather (°C)
2006	3730.387	618.7	20.07
2007	3238.587	601.6	19.69
2008	3967.373	815.0	19.60
2009	3695.867	595.5	19.94
2010	3317.553	953.1	20.15
2011	4330.513	661.8	19.58
2012	3973.153	559.4	19.54
2013	4125.300	1531.4	19.83
2014	4366.593	778.6	19.77
2015	3668.113	1123.1	19.96
2016	3366.773	1208.3	21.28
2017	4479.447	990.9	19.72
2018	4278.432	992.2	19.56

Table 3, The Average Production of Oilseeds ('000 tonnes), Weather(c), and Rainfall (mm) from 2006 to 2018 in India Figure 7 shows variations depending on the oilseeds production growth, rainfall at the time of monsoons. It adds to the growth of oilseeds production or reduces the oilseeds production growth.

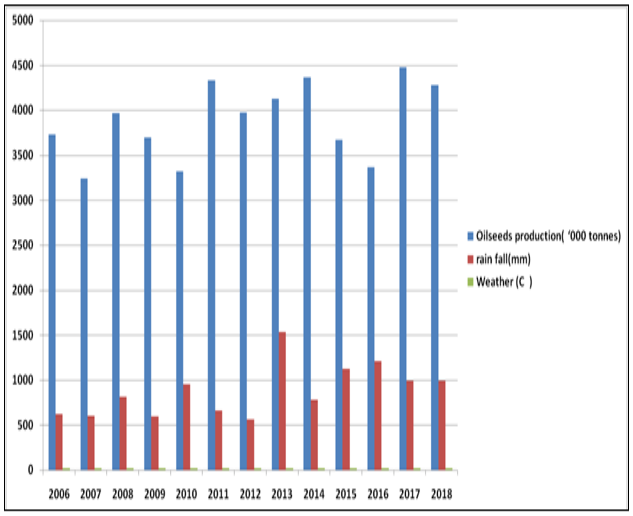

Fig. 7. Variations Depending on the Oilseeds Production Growth, Rainfall at the time of Monsoons

III. FUNCTIONAL DESCRIPTION

Adaboost algorithm and predictive technique are used in analyzing big data. The sequence steps for processing Adaboost algorithm. Big Data must be checked in the two-dimensional form of representing information as a variable $D_k = 1,2,3,\ldots, k$.

$$T_k = [(x_1,y_1),(x_2,y_2),\ldots(x_m,y_m)] \quad (1)$$

In equation (1), it is shown that T_k Where is training dataset, $[(x_1,y_1),(x_2,y_2),\ldots(x_m,y_m)]$ a two dimensional information assigned to training data. If there is a weak learning present in class of big data, assign sequence number i.e., training data is set $T_k = 1$.

$$W_1(i) = D_1(i) = 1/m, \; i=1,2,\ldots,m \quad (2)$$

In equation (2) is shown the weak learning dataset as a D_1 (i) is assigned to weight factor W_1 (i) which is helpful as many times as the original dataset is split into weak datasets. This paper provides numbers to weak dataset as the number one $t=1,2,3,\ldots, T_k$.

$$Di = \frac{w_i}{\sum_{i=1}^{m} wt(i)} \quad (3)$$

In equation (3) is shown TR_t subset from D_i and this instance assign a sequence number to every dataset. Further, if we have to evaluate the weak learning dataset then its result will be unbelievable. If this paper has obtain any hypothesis value from stronger dataset ht : x -> y and computes the error Ɛt which is shown in equation (4).

$$\varepsilon t = \sum_{i:ht(xi)\neq yi} D(i) \quad (4)$$

If there is strong dataset, find the error and resolve the equation (3) if aggregation of error value Ɛt > 0.5 the go to equation (1) which is tested in Figure 8. Otherwise evaluate normalized error is shown in equation (5)

$$Bt = \varepsilon t \; / (1 - \varepsilon t) \quad (5)$$

Finally, obtains the weighted majority voting by hypothesis which is shown in equation (6).

$$H_t = \sum_{t:ht(x)=y} \log\left(\frac{1}{\beta t}\right) \quad (6)$$

Provide a normalized weight for information on class is shown in equation (7).

$$H_{Final}(x) = \sum_{k=1}^{K} * \sum_{t:ht(x)=y} \log\left(\frac{1}{\beta t}\right) \quad (7)$$

Adaboost algorithm work steps are given:

Step 1: Create Training Data Set from the information provided by the normalized weight.

Step2: Naming a party for the subclass created from this.

Step 3: Finding that the information in this class is a weak learning or strong learning one.

Step 4: It is ascertained that the error value from strong learning is lower than 0.5.

Step 5: If not, go to the second step.

IV. RESULT AND DISCUSSION

The value of strong learning dataset is computed for all classes of dataset with the value previously obtained from trained data. Rapidminer software is used to analyze big data. This software provides the learning and testing procedure on dataset. Adaboost algorithm is used with predictive technology and used for additional data analysis that indicates when training data set gives high error rate which is shown in Fig. 8. These are both processed and produced from the available weak train data set that are examined whether the oilseeds produce higher yields and the current market price is less. Decision tree Technique is used for giving suggestion to the farmers, the new subset of sensor technology available in the farmer, and then what kind of feasible results are obtained. Data analysis tools are important to foretell how much profit or loss is achieved when farmers use sensor technology.

Adaboost algorithm uses it to divide the entire information into classes. The class which is divided into classes is examined to determine whether specific information has been fixed. This means that data processing is determined to know the value of the material in the class. Analyzing this way, make a structured decision in the individual data components. As this paper analyzes the information of Big Data, it is very useful for agriculture. If the information analyzed in this way requires more accurate knowledge, to get the most accurate knowledge from the analysis information, this paper has a long term stored information.

The dataset consists of words or numbers and size of dataset is very large, which is directly used in a computer system. It will take long time for processing. So this paper has converted this information to integer format, which is an easy machine understandable format and using the

rule-based system on the information to make those numbers.

A. Weak learning process

0-> null, 1-> low, 2-> average, 3-> high

[0] IF Profitable Status is Yes THEN Weather is average and rainfall is low (raw error = 0.00)

[1] IF profitable Status is No THEN Weather is low and rainfall is low (raw error =0.66)

[2] IF profitable Status is No THEN Weather is low and rainfall is high (raw error =0.77)

[3] IF profitable Status is No THEN Weather is high and rainfall is low (raw error =0.88)

[4] IF profitable Status is No THEN Weather is high and rainfall is low (raw error =0.55)

[5] IF profitable Status is No THEN Weather is low and rainfall is low (raw error =0.66)

[6] IF Profitable Status is Yes THEN Weather is null and rainfall is low (raw error = 0.33)

Fig. 8. Adaboost Apply on Weak Dataset and to Make Strong Dataset ClusterThe outcome is stored in M-dimensional format and then it is reconstructed, which is shown in Table 4. This study has a track of the data which we have collected as per the years shown in the Table 4

Year	Oilseeds Production ('000 tonnes)	Rainfall (mm)	Weather (°C)	Profitable Status	Profitable Status
2006	3730.387	618.7	20.07	Yes	1
2007	3238.587	601.6	19.69	No	-1
2008	3967.373	815.0	19.60	Yes	1
2009	3695.867	595.5	19.94	No	-1
2010	3317.553	953.1	20.15	No	-1
2011	4330.513	661.8	19.58	Yes	1
2012	3973.153	559.4	19.54	No	-1
2013	4125.300	531.4	19.83	Yes	1
2014	4366.593	778.6	19.77	Yes	1
2015	3668.113	1123.1	19.96	No	-1
2016	3366.773	1208.3	21.28	No	-1
2017	4479.447	990.9	19.72	Yes	1
2018	4278.432	992.2	19.56	Yes	1

Table 4: Raw training data of oilseeds production with rainfall, weather, and profitable status with Integer Format from 2006 to 2018 Predictive analysis focuses on the principles of the following information:

Rapped in place	Type of sensor with Save information type	Value of saved information	I / C #	Value					
				1	2	3	4	5	6
Argyle - Thunder 1: MK-III Weather Station Temp	MK-III Weather Station Temp (Temperature)	65509 °C	I	39178	73548	2.52E+18	3.92E+23	41.97	-87.66
Argyle - Thunder 1: MK-III Weather Station RH	MK-III Weather Station RH (Relative Humidity)	75%	I	39180	73550	2.52E+18	3.92E+23	41.97	-87.66
Argyle - Thunder 1: MK-III Weather Station Pressure	MK-III Weather Station Pressure (Differential Pressure)	10066 Pa	I	39182	73552	2.52E+18	3.92E+23	41.97	-87.66
Argyle - Thunder 1: MK-III Weather Station Wind Speed	MK-III Weather Station Wind Speed (Wind Speed)	24m/s	I	39184	73554	2.52E+18	3.92E+23	41.97	-87.66
Argyle - Thunder 1: MK-III Weather Station Wind Direction	MK-III Weather Station Wind Direction (Wind Direction)	112 °N	I	39186	73556	2.52E+18	3.92E+23	41.97	-87.66
Argyle - Thunder 1: MK-III Weather Station Rainfall	MK-III Weather Station Rainfall (Cumulative Precipitation)	145 count	C	39188	73558	2.52E+18	3.92E+23	41.97	-87.66
Argyle - Thunder 1: SM-I Soil Moisture Sensor	SM-I Soil Moisture Sensor (Soil Moisture)	453 mV	I	39174	73544	2.52E+18	3.92E+23	41.97	-87.66

#Instantaneous / Cumulative

Source of information from : Atmosphere Date and Time of Data : 03/14/2017 10:23:57 a.m.

Table 5: Data Gathered Through the Sensors in Agriculture Land

Whether the analyzed information is properly trimmed to fit or not. Is information too simple? Whether Information is sufficiently trained. Based on the above principles the study has been conducted.

B. Rainfall

How much rainfall in every year, which is considered as a normal rainfall and / or a marginal level. If we take any particular rainfall data which is tested with normal rainfall and if its result is greater, then this paper takes value between from 0 to (+1) to attributes otherwise this paper takes value between from 0 to (-1).

C. Oilseeds

How much oilseeds production in every year, which is considered as a normal production and /or a marginal level. If this paper takes any particular any oilseeds production data which is testes with normal production and if its results are greater, then this paper assigns value between from 0 to (+1) to attributes otherwise we assign value between from 0 to (-1).Similarly, assign rules to the remaining attributes. This algorithm is written to analyze excess information. It will be analyzed as it is operating in machine technology and will be taken from the existing information which is shown in Table 5. The Adaboost algorithm is used to convert the weaker class to a stronger class. The following process is divided into two parts, one of which is the training model and the other, the test model. The training model is created as a first model of information. This model is being tested to determine if its information has the same characteristic or collective information. It is considered a second model to analyze if it is known to be collective information. If this paper finds the information available to us to be stronger information, it is stored as for strong information, since it only analyzes the weak information and re-analyzer it.

V. CONCLUSION AND FUTURE WORK

The information collected from the Indian agricultural department and the source of the sensor devices from the ground which stored in big data. Two different technologies used to analyze big data information like Adaboost, predictive technologies. In additionally, it has added more knowledge for learning processes into exiting technologies, such as the machine learning techniques. The solution available from this process is at the small scale farmer, who can get the information he needs from this project. This research is useful for small and marginal farmers, and this research stands before predicting how much yields can be obtained by using any type of crop in any season. When farmers used sensor technology in agriculture farming then they obtain smart works and more benefits. This approach ensures that the two different type's farmer will obtain the benefit. Finally, this paper gives an idea, how farmer can get a surplus profit.

REFERENCES

1. "Microsoft word synthesis report", How to Feed the World in 2050
2. Gopinathan, Mambully Chandrasekharan, and Rajasekaran Sudhakaran. "Biofuels: opportunities and challenges in India." In Vitro Cellular & Developmental Biology-Plant 45.3 (2009): 350-371.
3. Rathmann, Régis, Alexandre Szklo, and Roberto Schaeffer. "Land use competition for production of food and liquid biofuels: An analysis of the arguments in the current debate." Renewable Energy 35.1 (2010): 14-22.
4. Aimin, Hao. "Uncertainty, risk aversion and risk management in agriculture." Agriculture and agricultural science procedia 1 (2010): 152-156.
5. Sharma, Vijay Paul. "Problems and prospects of oilseeds production in India." (2014).
6. Pretty, Jules. "Agricultural sustainability: concepts, principles and evidence." Philosophical transactions of the Royal Society B: biological sciences (2007).
7. De Ponti, Tomek, Bert Rijk, and Martin K. Van Ittersum. "The crop yield gap between organic and conventional agriculture." Agricultural systems 108 (2012): 1-9.
8. Trenberth, Kevin E. "Changes in precipitation with climate change." Climate Research 47.1-2 (2011): 123-138.
9. Deaton, Angus, and Jean Drèze. "Food and nutrition in India: facts and interpretations." Economic and political weekly (2009): 42-65.
10. Thornton, Philip K. "Livestock production: recent trends, future prospects." Philosophical Transactions of the Royal Society B: Biological Sciences 365.1554 (2010): 2853-2867.
11. Jayaraman, Prem Prakash, et al. "Internet of things platform for smart farming: Experiences and lessons learnt." Sensors 16.11 (2016): 1884.
12. Department of Agriculture and Cooperation Annual Report 2017-18. Link: http://agricoop.nic.in/sites/default/files/Krishi%20AR%202017-18-1%20for%20web.pdf
13. Fan, Shenggen, Ashok Gulati, and Sukhadeo Thorat. "Investment, subsidies, and pro-poor growth in rural India." Agricultural Economics 39.2 (2008): 163-170.
14. Sharma, Vijay Paul. "Problems and prospects of oilseeds production in India." (2014).
15. Blunden, Jessica, and Derek S. Arndt. "State of the Climate in 2015." Bulletin of the American Meteorological Society 97.8 (2016): Si-S275.
16. Jha, Girish Kumar, et al. "Edible oilseeds supply and demand scenario in India: Implications for policy." Report of the Indian Agricultural Research Institute (2012).
17. Ramakumar, R. "Indian Agriculture under Economic Reforms: A Preliminary Review." Retrieved May 10 (2009): 2011.
18. Jeble, Shirish, et al. "Impact of big data and predictive analytics capability on supply chain sustainability." The International Journal of Logistics Management 29.2 (2018): 513-538.
19. Antle, John M., James W. Jones, and Cynthia E. Rosenzweig. "Next generation agricultural system data, models and knowledge products: Introduction." Agricultural systems 155 (2017): 186-190.
20. Kache, Florian, and Stefan Seuring. "Challenges and opportunities of digital information at the intersection of Big Data Analytics and supply chain management." International Journal of Operations & Production Management 37.1 (2017): 10-36.
21. Mintert, James, et al. "The challenges of precision agriculture: is big data the answer." Southern Agricultural Economics Association Annual Meeting, San Antonio, Texas. No. 230057. 2016
22. Kamilaris, Andreas, Andreas Kartakoullis, and Francesc X. Prenafeta-Boldú. "A review on the practice of big data analysis in agriculture." Computers and Electronics in Agriculture 143 (2017): 23-37.
23. Devalkar, Sripad K., et al. "Data science applications in Indian agriculture." Production and Operations Management(2018).

AUTHORS PROFILE

K.Baskar, Research Scholar, pursuing Ph.D in the Department of Computer Science and Engineering at Annamalai University, Chidambaram, Tamilnadu State,India. He is currently working as an Associate Professor in the Department of Computer Science and Engineering at Sri Venkateswaraa College of Technology, Sriperumbudur.Tamilnadu, India.
Email: baskar_ka@yahoo.com

Dr. S. Arivalagan is working as an Assistant Professor in the Department of Computer Science and Engineering at Annamalai University, Chidambaram, Tamilnadu State, India. His area of research includes Image Processing, Data Structures, Computer Networks and DBMS.
Email: arivucseau@gmail.com

Dr. P. Sudhakar is working as an Associate Professor in the Department of Computer Science and Engineering at Annamalai University, Chidambaram, Tamilnadu State, India. His area of research includes Software Engineering, Image Processing, Wireless Networks and Data Mining.
Email: kar.sudha@gmail.com