

Analysis of Agriculture Data using Data Mining Techniques

Ashwini Navghare¹, Shreya Deo², Apurva Padwad³, Mansi Agrawal⁴, Prof. A.D. Shah⁵

¹Ashwini Navghare, Dept. of Computer Engineering, Sipna College of Engineering & Technology, Amravati, Maharashtra, India

²Shreya Deo, Dept. of Computer Engineering, Sipna College of Engineering & Technology, Amravati, Maharashtra, India

³Guide: Prof. A.D. Shah, Dept. of Computer Engineering, Sipna College of Engineering & Technology, Amravati, Maharashtra, India

Abstract:- Agriculture is the backbone of Indian Economy. In meteorology, weather forecasting is a crucial application. In the sector of agriculture where farmers and agri-businesses have to make innumerable decisions every day and intricate complexities involves the various factors influencing them. An essential agricultural planning is needed for the accurate yield guesstimate for the various crops and its quality factor involved. But due to intervention by the human by their artificial industries and constant destruction of natural resources, the climate has been changing rapidly and as a consequence, we are endangering the crops and soil fertility. This results in rapid seasonal changes in rainfall pattern, summer and winter and hence, agriculture and farmers have been suffering from serious destruction of both crops and their livelihoods. For prevention of crop and economic damages, Data Mining tools and techniques can be used. Agriculture is a noticeable earmark for Big Data. Conditions of Environmental, Variability of Soil, levels of input of farmers, prices of used and needed material have made it all more admissible for farmers to use the climate prediction information and get help to make critical farming decisions. Data Mining of large amount of crop, soil used for their cultivation and changing climate data, and non-experimental data optimization makes agriculture more irrepressible to climatic change. Predictive data mining technique can be used for predicting in future use of crops, pesticides and fertilizers, revenue will be generated for growth and function of crops in agriculture.

Keywords: - Weather Forecasting, Data Mining, Agricultural Climate Change, Big Data, Cultivation Soil, Crops, Fertilizers, Pesticides.

1. INTRODUCTION:

This paper is a survey about the unhealthy environmental changes in the climate regarding the agriculture and cultivation of crops. This shows how these climatic changes are adversely affecting the people's eating habits and health. Here, data mining is used for the analyzation, pattern searching and data processing to understand the variation in climate and environment more predominantly. This needs lot of literature survey on one hand of climate change data mining and agricultural data mining as both of these are the most important factors for productive profitable results with the economic growth.

1.1 Data Mining:

Data mining is a process in which patterns are recorded based on findings from data by using machine learning algorithms, statistics and database systems. Data mining has multiple steps; first, we take large chunks of data and organize it into a database, then we sort the data by using any of the machine learning

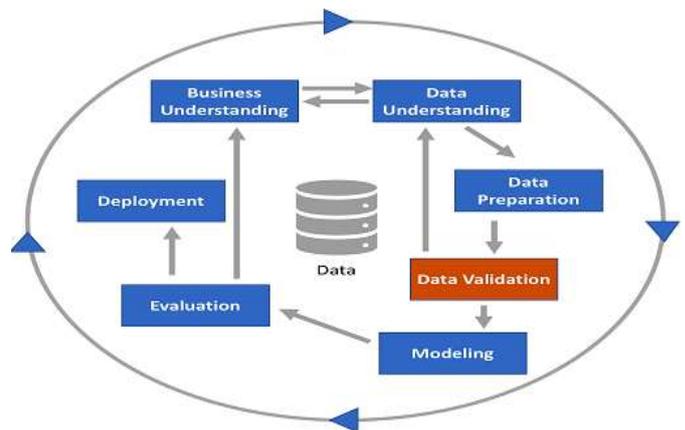


Fig -1: Data Mining Process Diagram

Algorithms depending on the problem, after the algorithm has solved performing the sorting, we can find patterns related to the problem and providing solutions to dealing with it. A good examples of data mining can include grocery stores, where the data for mining is gathered depending on how people are purchasing the goods based on their needs and the store's services. This data can also vary on the basis of various other factors such as area of the store, people living in that area.

Data mining techniques go as follows: Tracking patterns, Classification, Clustering, Association, Outlier Detection, Regression and Prediction. By using some of the above techniques we can perform highly effective data mining process. Data mining has proven highly useful in providing customers with wide range of products with multiple demands that are not heard directly. Only by data analysis we could perform actions for great profits.

1.2 Data Mining in Agriculture:

Agriculture is the backbone of Indian economy. Over 80% households in rural areas are dependent on agriculture. It is Agriculture is contemplated as a standard occupation and plays an important role in everyone's lifestyle. Agriculture is an important sector in Indian economy as it contributes to about 17% of GDP and provides employment to over 60% of the population. Due to human intervention, the climate has been changing like never, situations like droughts, untimely rainfall, hotter climates create problems for the farmers, especially poor farmers who cannot cope with the continuous changes in climate and due to which suicide commits of farmers are increasing day by day. Hence, to reduce this situational death of farmers, it is very necessary to predict the rainfall and other climate changes for the betterment of the farmers as well as the economy of the country.

Worldwide India has been ranked second in the farming cultivation & food products. Demographically, this is the broadest economic sector and plays a significant role in the overall socio-economy of India. It is dependent on various climatic and economic characteristics such as soil, climatic conditions, crop cultivation, irrigation, fertilizers used, temperature, rainfall, harvesting, pesticides weeds, etc. Therefore, data mining in agriculture s very important. In very recent research, the applications for the agricultural data mining techniques are included for better and effective results. These recent technologies nowadays are providing a lot of information data on agricultural-related research activities, which can be later used to analyzation to find important results.

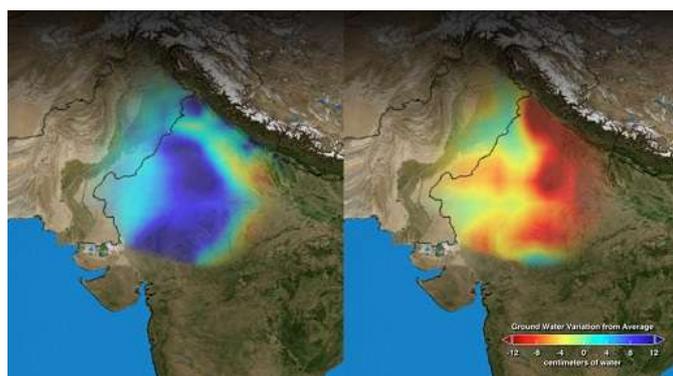


Fig -2: Deceasing Water Levels for Agricultural Lands

Here, Fig-2 shows how the levels of water are decreasing day by day which is for sure is not a good symptom for the agricultural fields which yields the crops. Whereas, Fig-3. shows that these tremendous changes in weather and levels of ground water the fertility of soil is getting disturbed and destroyed which has taken a drastic turn in agricultural field. And this will keep affecting if there are no solutions found on it.



Fig -3: Infertility of Land due to Water Loss

1.3 Data Mining in Climate Change:

Several breakthroughs have been provided by the Big data and Data Mining in fields for instance such as health informatics, smart cities and marketing, etc. However, these same techniques have not delivered consistent key findings for climate change. There are reasons, why the main previous data mining work in climate science, and in the analysis of climate teleconnections, has count on methods that renders oversimple answers in "yes or no". It is not that simple in climate that weak connections from all the regions globally to give results from an underlying physical phenomenon. And if these threshold and weak connections are thrown out then it will all halt. Substitutently, millions of data points spread out all over globally.

Therefore, a new methodology for data mining on climate data is developed by the Georgia Tech Team. This methodology is more self-contained than the previous tools. It finds out the commonalities of data sets without manual user expertise. Hence, this method can be trusted due its transparency, robustness and clear results.

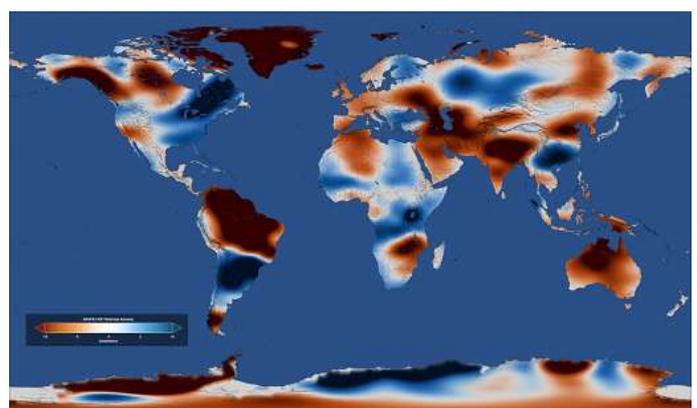


Fig -4: Changes in Climate caused due to Human Activities Harmful to Environment

Here, the Fig-4 shows the drastic climate changes caused due to the large amount of harsh human activities which results in the harm to the environment. This is made possible due the new methods and technologies developed by the scientists which generates the proper results in order to give accurate climate predictions and their timings.

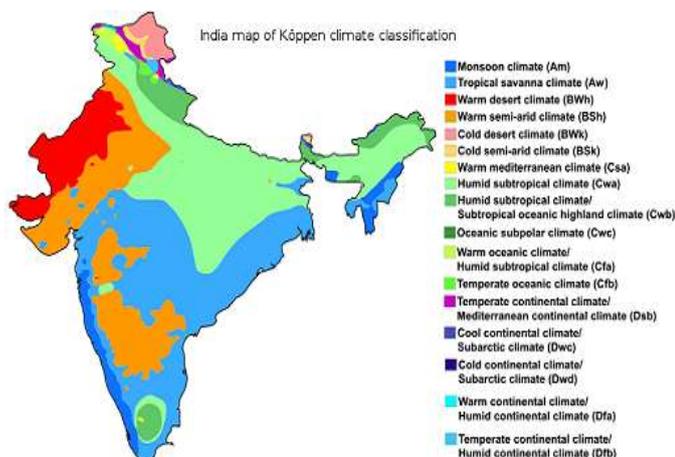


Fig -5: Classification of Climate in India

Here, Fig-5 shows the classification of climate in India at various locations as Temperate, Warm, Cool, Blossom, Tropical, etc. This gives an scenario of how it changes rapidly with change in weather and ultimately activities by humans.

2. Literature Survey:

There are number of methods of data mining, some of which we have surveyed. The main intentions are to survey various methods and see the Analysis of the Agricultural Data using Data Mining Techniques. We have done a literature survey about these methods as such now, Clustering is an unsupervised classification process which is the main part in data mining. These clustering techniques are categorized into Partitioning Clustering, Hierarchical Clustering, Density-Based Methods, Grid-Based Methods and Model-Based Clustering Methods. Partitioning Clustering algorithms, for instance, K-means, K-medoids PAM, CLARA and CLARANS has assign predefined cluster number i.e. 'k' objects into clusters. Hierarchical Clustering algorithms assign a cluster that can have data point's which are representatives of low-level cluster objects in tree structure. Similarly, there are certain methodologies used agricultural data mining. Now, lets see those methods useful for data mining for our survey regarding the Analysis of the Agricultural Data using Data Mining Techniques and also for climate change and its predictions

2.1 Which Data Mining techniques can be useful for predicting the change in Climate will be helpful for Agriculture?

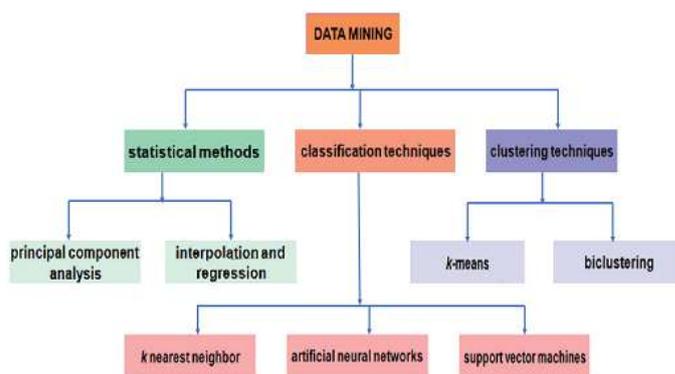


Fig -6: Data Mining Techniques

(a) Data Mining Techniques in Agriculture:

There are various different weather forecasting methodologies which researchers have developed and evaluated in the field of agriculture all over the world. Some of the methods used in implementing to solve the problem related to the agriculture are the K-Means (KM), K-Nearest Neighbors (KNN), Artificial Neural Network (ANN), Support Vector Machine (SVM). Also, Multiple Linear Regression (MLR) method is used for modelling the linear relationship among the dependent variable and the independent variables. Here, rainfall is the dependent variable and production, area of sowing and year are the independent variables.

Methodology	Applications
K-means	Forecasts of pollution in atmosphere Classifying soil in combination with GPS
k-nearest Neighbor	Simulating daily precipitations and other weather variable
Support Vector Machine	Analysis of different possible change of the weather scenario
Decision Tree Analysis	Prediction soil dept
Unsupervised Clustering	Generate cluster and determine any existence of pattern
WEKA Tool	Classification system for sorting and grading mushrooms.

Table -1: Data Mining Methods

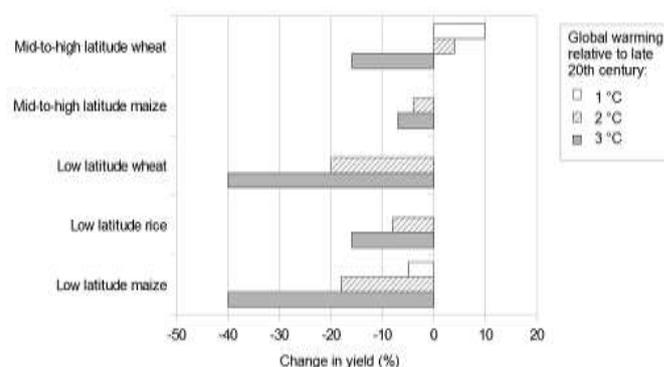


Chart -1: Changes in some Important Yields

Here, Chart.-1 shows the changes in the major crop yields of India. These all changes are happened due to the change in climate. If the rainfall is not proper and on proper time, then the yields get highly affected by it. Therefore, there is a high need of measuring the climate predictions and weather broadcastings so that the farmers can decide the best time to cultivate their crops and hence, their situations can be helped and made good so far.



Fig -7: Loss of Yield caused by Rainfall Variation

Here, Fig.-7 shows the loss in yields in agricultural fields due to rainfall variations. This loss is a major drawback for the farmers in regards of their hard work, energy, and fund they have invested. And, ultimately this results in the decrease of economy.

Also, Chart.-2 shows that India is the country of major agricultural crop cultivations over worldwide and tropical areas. This reveals that if the agricultural yields are given proper help then not only the country's but over the world it will be a profitable share to the world economy and hence, in the growth.

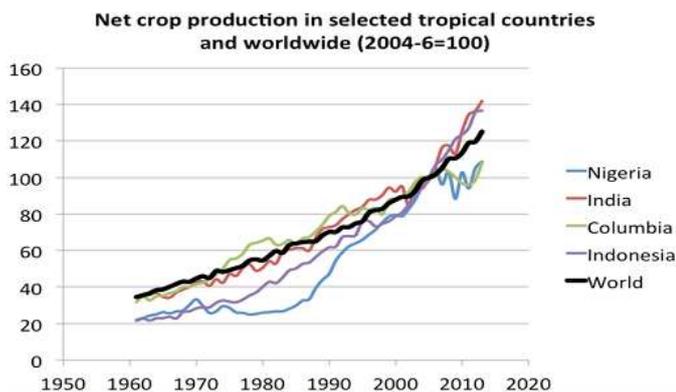


Chart -2: Crop Production in Tropical & Worldwide Countries

(b) Data Mining Techniques in Climate Change:

For understanding the data mining of the climate change or weather forecasting predictions, we need to understand what exactly the climate is consist of, a Scale Disparity among the Atmospheric, Hydrological and Land Process Models and Water, Energy, Environmental impacts and, purpose of Knowledge Discovery in climate is to build a connection

among these disparate scales. Therefore, this outputs to a lower resolution impacts which are needed to be higher level. These extreme events and stresses are caused by the regional change in climate. The process of extracting useful knowledge from databases is known as Knowledge Discovery.

There are mainly three methods used for the climate change predictions such as:

1. *Synoptic Weather Prediction*
2. *Numerical Weather Prediction*
3. *Statistical Weather Prediction*

1 *Synoptic Weather Prediction:*

The study of observational data and different data collection are observed from thousands of weather stations.

2 *Numerical Weather Prediction:*

Weather prediction capabilities of the computer from the gained numerical data.

3 *Statistical Weather Prediction:*

Statistical past records of weather are used to predict the future occurrences.

3. CONCLUSION

In this paper, we have surveyed the various climate change methodologies in data mining and agricultural methodologies in data mining. Climate change affects on agriculture and its fields and yields, crop cultivations and its effect on economy. This survey suggests that the more predictions are accurate about climate the more helpful is it for the farmers. Therefore, the Analysis of Agriculture Data using Data Mining Techniques is done,

REFERENCES

- [1] Jharna Majumdar, Sneha Naraseeyappa and Shilpa Ankalaki "Analysis of agriculture data using data mining techniques: application of big data" Majumdar et al. J Big Data (2017) 4:20 DOI 10.1186/s40537-017-0077-4
- [2] Deepak Sharma, Priti Sharma "Rain Fall Prediction using Data Mining Techniques with Modernistic Schemes and Well-Formed Ideas" International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-9 Issue-1, November 2019
- [3] Aishwarya Dhore1, Anagha Byakude2, Bhagyashri Sonar3, Mansi Waste4 "Weather prediction using the data mining Techniques" International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 - 0056 Volume: 04 Issue: 05 | May -2017 www.irjet.net p-ISSN: 2395-0072
- [4] P.Kalaiyarasi [1], Mrs.A.Kalaiselvi [2] "Data Mining Techniques Using To Weather Prediction" International Journal of Computer Science Trends and Technology (IJCTST) - Volume 6 Issue 3, May - June 2018
- [5] [5] Folorunsho Olaiya "Application of Data Mining Techniques in Weather Prediction and Climate Change Studies " Application of Data Mining Techniques in Weather Prediction and Climate Change Studies" I.J. Information Engineering and Electronic Business, 2012, 1, 51-59 Published Online February 2012 in MECS

(<http://www.mecs-press.org/>)
DOI:10.5815/ijieeb.2012.01.07

- [6] Shalin Paulson "A Survey on Data Mining Techniques in Agriculture" Special Issue - 2015 International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181 Published by, www.ijert.org RTPPTDM-2015 Conference Proceedings
- [7] K. KRISHNA KUMAR^{a,*}, K. RUPA KUMAR^a, R. G. ASHRIT^a, N. R. DESHPANDE^a and J. W. HANSEN^b "CLIMATE IMPACTS ON INDIAN AGRICULTURE" INTERNATIONAL JOURNAL OF CLIMATOLOGY Int. J. Climatol. 24: 1375–1393 (2004) Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/joc.1081
- [8] Auroop R Ganguly and Karsten Steinhäuser "Data Mining for Climate Change and Impacts" 2008 IEEE International Conference on Data Mining Workshops
- [9] Jeetendra Prakash Aryal, Tek Bahadur Sapkota, Dil Bahadur Rahut, M.L. Jat "Agricultural sustainability under emerging climatic variability: the role of climate-smart agriculture and relevant policies in India" International Journal of Innovation and Sustainable Development . January 2019 DOI: 10.1504/IJISD.2019.10020869
- [10] Tek B. Sapkota ^{a,✉}, Sylvia H. Vetter ^b, M.L. Jat ^a, Smita Sirohi ^c, Paresh B. Shirsath ^d, Rajbir Singh ^e, Hanuman S. Jat ^f, Pete Smith ^b, Jon Hillier ^g, Clare M. Stirling ^h "Cost-effective opportunities for climate change mitigation in Indian agriculture" Science of the Total Environment 655 (2019) 1342–1354