International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
http://www.ijcsjournal.com Volume 5, Issue 1, No 7, 2017 ISSN: 2348-6600
Reference ID: IJCS-186 PAGE NO: 1132-1136
Sri Vasavi College, Erode Self-Finance Wing 3rd February 2017
National Conference on Computer and Communication NCCC'17
http://www.srivasavi.ac.in/ nccc2017@gmail.com

A SURVEY OF SOIL TEMPERATURE AND MOISTURE PREDICTION TO RECOMMEND PLANTING DIFFERENT CROPS IN AGRICULTURE USING DATA MINING

D.CHANDRALEGA
Assistant Professor in BCA
Sri Vasavi College (Self Finance)
Erode.
chandralegaemail@gmail.com

ABSTRACT- To obtain information or knowledge which can be helpful to farmers and government organizations for making better decisions and for make better policies which help to increased production. In this paper, our focus is on application of data mining techniques which is use to extract knowledge from the agricultural data to estimate better crop yield for major crops in major districts of India. It is very hard to acquire the information what really want with the accumulation of large number of data. In this paper our focus is on the applications of Data mining techniques in agricultural field. Generally for doing agriculture land, labor, capital and organization are required without that cannot produce with a new agriculture technology. Fuzzy sets are suitable for handling the issues related to understandability of pattern of incomplete data, and human interaction, mixed media information and can provide approximate solutions faster for given pattern of data. Yield prediction is a very important agricultural problem to farmer that remains to be solved based on the previous available data. The yield prediction problem can be solved by employing Data Mining techniques. This work aims to find suitable data models that achieve a high accuracy and a high generality in terms of yield prediction capabilities.

INTRODUCTION

India is a country of rural economy and it is predominately agriculture oriented. In India more than 82 % are farmers belongs to small and large marginal farmers. Crop yield prediction is a very important area of research till which is use to ensure the food security all over the world. India's economy is almost depends on the crop yield. This will help to select the best crop for planting in the different district.

All the crops are depend on the various factors like temperature, rain, sun light, humidity, moisture and carbon dioxide (CO2) concentration to produce grains and other crop products that are so essential to our nutrition and health. However the amount of rainfall is vary between one districts to other and also vary in one year to another. Crop management is about managing climate risk so as to have financially viable and sustainable agricultural systems.

International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
http://www.ijcsjournal.com Volume 5, Issue 1, No 7, 2017 ISSN: 2348-6600
Reference ID: IJCS-186 PAGE NO: 1132-1136
Sri Vasavi College, Erode Self-Finance Wing 3rd February 2017
National Conference on Computer and Communication NCCC'17
http://www.srivasavi.ac.in/ nccc2017@gmail.com

The most important part of farming is a pesticide. Without pesticides crops would die significantly due to insects and other pests which lead to sudden drop in the crop yield. And also the too much pesticides may affect the crop, while too little may not useful for crop. So the amount of pesticides required by crop is very important parameter. In our project research, we have considered the various environmental, biotic (soil salinity) and the areas of production are the factors for crop in India. Taking this into consideration we developed a database for various districts, for this we applies clustering techniques to divide regions, and then we apply suitable classification techniques to obtain crop yield predictions.

RELATED WORK

Bangladeshi student purposes Data mining techniques to predict annual yield of major crops and recommend planting different crops in different districts in Bangladesh. They considered the effects of environmental (weather), biotic (pH, soil salinity) and area of production as factors towards crop production in Bangladesh. Taking these factors into consideration as datasets for various districts, they applied clustering techniques to divide regions; and then they apply suitable classification techniques to obtain crop yield predictions. In David H White and S Mark Howden, Research paper they focus on the climates determinants of crop productivity.

They considered how the climate envelopes of different crops based on light, temperature and moisture influence the distribution of cropping and other land uses around the world. They also discuss how these and other climatic variables influence the growth and yield of crops. Adaption strategies are also discussed that helps a lot to assist the crop producers to cope with the rising global temperatures and carbon dioxide (CO2) levels, along with the often reduced rainfall, soil, moisture and water availability. In M.C.S Geetha, study different data mining techniques in agriculture Research paper aims of finding suitable data models which achieves a high precision and a high generality with respect to parameters namely rainfall, year and production. In research paper discusses applications of data mining techniques used in agriculture domain.

DATA MINING

Knowledge Discovery and Agriculture: The term KDD refers to the overall process of knowledge discovery in databases. In this process, Data mining is a particular step involving the application of specific algorithms for extracting patterns (models) from data set. The additional steps in the knowledge discovery process, such as data preparation, data selection, data cleaning, incorporation of appropriate knowledge of data set, and proper interpretation of the results of mining, ensures that useful knowledge is derived from the data set. KDD systems incorporate theories, algorithms, and methods from all these fields. Many successful applications have been reported from varied sectors such as marketing, finance, banking, manufacturing, telecommunications and Agriculture.KDD is used in agriculture to show the statistical information about soil condition, climate

International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
http://www.ijcsjournal.com Volume 5, Issue 1, No 7, 2017 ISSN: 2348-6600
Reference ID: IJCS-186 PAGE NO: 1132-1136
Sri Vasavi College, Erode Self-Finance Wing 3rd February 2017
National Conference on Computer and Communication NCCC'17
http://www.srivasavi.ac.in/ nccc2017@gmail.com

conditions, past crop yield, Government strategies, all the information about pesticides, fertilizer.

Fuzzy Set: The modeling of imprecise and qualitative knowledge, as well as the transmission and handling uncertainty of data at various stages are possible through the use of fuzzy sets. Knowledge discovery in databases is mainly concerned with identifying interesting patterns and describing them in a concise and meaningful manner. Fuzzy models can be said to represent a prudent and user-oriented sifting of data, qualitative observations and calibration of commonsense rules in an attempt to establish meaningful and useful relationships between system variables. Despite a growing versatility of knowledge discovery systems, there is an important component of human interaction that is inherent to any process of knowledge representation, manipulation, and processing.

Clustering: Data mining aims at sifting through large volumes of data in order to reveal useful information in the form of new relationships, patterns or clusters, for decision making by a user farmer. So that, user can finds appropriate data for the crop development and their management.

Association Rules: An important area of data mining research deals with the discovery of association rules. An association rule describes an interesting association relationship among different attributes. In Agriculture insecticide association involves disease related to that insect.

Functional Dependencies: Fuzzy inference generalizes both imprecise (set-valued) and precise inference. Similarly, fuzzy relational databases generalize their classical and imprecise counterparts by supporting fuzzy information storage and retrieval. In agriculture field functional dependency describe between yield and their price. Also it provides information about soil and their productivity.

Data Summarization: Summary discovery is one of the major components of knowledge discovery in databases. Data summarization provides the user with comprehensive information for grasping the essence from a large amount of information in a database. Fuzzy set theory is also used for data summarization. Agriculture is very big field so comprehensive information much required to grasping the data about agriculture.

APPLICATION

There are several applications of Data Mining techniques in the field of agriculture. Some of the data mining techniques are related to weather conditions and Short term forecasting of air pollution in the atmosphere. Data mining techniques are applied to study sound recognition problems. Data mining techniques are often used to study soil characteristics. Classifying soils in combination with GPS-based technologies. The K Nearest Neighbor (KNN) is applied for simulating daily precipitations and other weather variables.

International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
http://www.ijcsjournal.com Volume 5, Issue 1, No 7, 2017 ISSN: 2348-6600
Reference ID: IJCS-186 PAGE NO: 1132-1136
Sri Vasavi College, Erode Self-Finance Wing 3rd February 2017
National Conference on Computer and Communication NCCC'17
http://www.srivasavi.ac.in/ nccc2017@gmail.com

MOTIVATION

Agriculture is the backbone of the Indian nation. In spite of the fact that large areas in India have been brought under irrigation, only one-third of the cropped part is irrigated. The productivity of agriculture is very low as compare to land which farmers has So as the demand of food is increasing, the researchers, farmers, agricultural scientists and government are trying to put extra effort and techniques for more production. And as a result, the agricultural data increases day by day. As the volume of data increases, it requires involuntary way for these data to be extracted when needed. Still today, a very few farmers are actually using the new methods, tools and technique of farming for better production. Data mining can be used for predicting the future trends of agricultural processes. Indian farmer must have a good understanding of the soil type, the biotic factors governing the soil, environmental factors and also a thorough knowledge about the traditional agricultural practices to gain maximum crop yield.

ANALYSIS DATA

Excel software are use to conduct qualitative analyses and to create a benchmark for the analysis of the research dataset. The benchmark allowed current statistical methods for the dataset to be established and any limitations for dataset to be identified. The dataset was then analyzed using a clustering process within the data mining software. That result were compared against the benchmark for a number of factors that included ease of application, speed, time and accuracy of results to determine if data mining was superior to current methods. The results of statistical and data mining experiments may still require some expertise to be understood and utilized of given dataset. The dataset used in this project has been collected from Mahatma Phule Krushi Vidyapith Rahuri. From the dataset, we pre-processed and selected only the attributes which are important for our project rainfall, temperature humidity etc. And also cultivated area for every crop considered according to the districts. We also considered the two biotic attributes – soil salinity and soil pH for our project.

METHODOLOGY

The method of our project is initially divided into two major parts: (1) Clustering (2) Classification. A clustering of the selected districts in our project, we have considered a total of 20 districts of India. In order to group the districts into distinct clusters, the assumption that we had to use was that the districts containing the similar values of relevant attributes should belong to the same cluster. According to this assumption the method of our project is initially divided into two major parts: (1) Clustering (2) Classification. A clustering of the selected districts in our project, we have considered a total of 20 districts of India. In order to group the districts into distinct clusters, the assumption that we had to use was that the districts containing the similar values of relevant attributes should belong to the same cluster. According to this assumption, we categorized our selected attributes for the consideration of clustering the districts as follows:

International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
http://www.ijcsjournal.com Volume 5, Issue 1, No 7, 2017 ISSN: 2348-6600
Reference ID: IJCS-186 PAGE NO: 1132-1136
Sri Vasavi College, Erode Self-Finance Wing 3rd February 2017
National Conference on Computer and Communication NCCC'17
http://www.srivasavi.ac.in/ nccc2017@gmail.com

1) Cluster Type-1 is based on the following attributes: Rainfall, minimum temperature, maximum temperature, humidity and sunshine. These are the environmental or climatic attributes considered for our research. The degree of similarity of the collection of these attributes should indicated distinct clusters for the selected districts.

2) Cluster Type-2 is based on the following attributes: soil pH and soil salinity. As discussed earlier, these biotic factors contribute largely towards the prediction of the crops.

CONCLUSION

With the improvement of data mining technologies, especially those without any premises or humans subjective, data mining can be applied in many areas. In this paper some Data Mining techniques were adopted in order to estimate crop yield analysis with existing data and their use in data mining. Some data mining technique have not yet applied to agriculture problem, as an example GPS techniques may be employed for discovering important information from agricultural-related like soil identification.

REFERENCES

[1] A. Mucherino, P. Papajorgji, P.M. Pardalos References:

[1] A. Mucherino, P. Papajorgji, P.M. Pardalos, Data Mining in Agriculture, Springer, 2009.

[2] U. Fayyad, P. Stolorz, Data mining and KDD: Promise and challenges, Future Generation Computer Systems,

Vol.13, pp. 99-115, 1997.

[3] I. Jagielska, C. Matthews, T. Whit fort, an investigation into the application of neural networks, fuzzy logic, genetic

Algorithms, and rough sets to automated knowledge acquisition for classification problems, Neurocomputing, Vol. 24,

pp. 37-54, 1999

[4] U. M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy, Eds., Advances in Knowledge Discovery and

Data Mining. Menlo Park, CA: AAAI/MIT Press, 1996.

[5] "Special issue on knowledge discovery in data- and knowledge bases," Int. J. Intel. Syst., vol. 7, 1992.