IJIEMR Transactions, online available on 20th July 2020. Link

:http://www.ijiemr.org/downloads.php?vol=Volume-09&issue=ISSUE-07

Title: CROP YIELD PREDICTION USING DATA MINING

Volume 09, Issue 07, Pages: 139 - 143

Paper Authors

CHOKKAKULA DEVI, SARIPALLI PRASUNNA, S VASAVI

USE THIS BARCODE TO ACCESS YOUR ONLINE PAPER

To Secure Your Paper As Per UGC Guidelines We Are Providing A Electronic Bar Code

CROP YIELD PREDICTION USING DATA MINING

CHOKKAKULA DEVI[1*], SARIPALLI PRASUNNA[2], S VASAVI[3***]**

[1, 2, 3] Department of CSE, Wellfare Institute of Science Technology And Management, Pinagadi, Pendurthi, Visakhapatnam ,Andhra Pradesh, India

[*] chokkakula.devi@gmail.com , [**] prasunasaripalli43@gmail.com , [***] vasavimogalapalli@gmail.com .

Abstract: India is generally an agricultural country. Agriculture is the most important contributor to the Indian economy. Agriculture crop production depends on season, biological, and economic cause. Now-a-days, farmers are struggling to produce the yield because of unpredictable climatic changes and also they are unable to predict which crop is suitable for their lands according to their soil conditions. They are also unaware of pesticides which are to be used against different insect attacks. Due to the lack of enough knowledge on fertilizers, farmers are using chemical fertilizers. The purpose of any fertilizer is to increase the fertility of the soil. However, using chemical fertilizers do have their hidden dangers about which most farmers may not know. The chemical fertilizers and pesticides have important role in crop production but recently appeared many problems and caused environmental pollution and human health. And the essential issue for agricultural intention is the accurate yield estimation for the numerous crops involved in the planning .To overcome these problems, we are proposing a solution that, finding suitable crop based on weather and soil conditions by using data mining techniques. This will help the farmers in increasing the yield production of their crops. We also suggest the farmers which pesticides and fertilizers are to be used for better crop yielding. By this, the farmer will be more benefited by getting high yield production with more income.

1. Introduction

The Indian agriculture is mostly depend in terms of climate and soil conditions .Agriculture is facing a problem called sudden change in their climatic changes that are directly affecting to the crop yielding. for the sustainable result in agriculture, these resources are need to be estimated and analysed. Precise and dependable information about crop yield prediction is important for taking decisions for agricultural risks management. Farmers are leaving agriculture because they are not getting

the right reward for their harvest, they are not knowing about which crop is suitable for their soil, and also not knowing about the pesticides and fertilizers to be used for their crop yielding .Crop yield prediction is also important for supply chain operation of companies engaged in industries that use agricultural produce as a raw material. Data mining techniques are used in agriculture for data clustering, predictive and forecasting purpose. Different techniques have been used for mining data over the years. the most used data mining techniques are k-means and ID3 algorithm in the

agriculture field. By using these algorithms we can predict the soil and forecast the weather conditions. k-means cluster algorithm is an iterative algorithm that tries to partition the datasets i.e.,(soil clustering).ID3 algorithm create a decision tree based on collection of sample data. this algorithm predict the weather by using historical data as a training data.

2. Literature Review

In recent years many farmers are leaving their farming and migrated from the villages because of they don't getting expected yield for their crops . They don't predict weather conditions. Many of the farmers are using chemical fertilizers so that they get more profit but crops and vegetables are affected with chemical. Those are harmful to human health.Now a daythere many modern agricultural techniques are available but the most of the farmers are unaware of those techniques. To avoid these problems we are used the some data mining techniques.

Data mining

Data mining is defined as a process used to extract usable data from a larger set of any raw data. In data mining we can cluster the similar type of data and classify the data. There are pre-processing techniques

Data-Cleaning

The data can have many irrelevant and missing parts. To handle this part, data cleaning is done. It involves handling of missing data, noisy data etc.

Data-Transformation

This step is taken in order to transform the data in appropriate forms suitable for mining process.

Data-Reduction

Since data mining is a technique that is used to handle huge amount of data. While working with huge volume of data, analysis became harder in such cases. In order to get rid of this, we use data reduction technique. It aims to increase the storage efficiency and reduce data storage and analysis costs.

Clustering

Clustering is the process of making a group of abstract objects into classes of similar objects.Clustering analysis is broadly used in many applications such as market research, pattern recognition, data analysis, and image processing.

K-mean Algorithm

K-Means clustering algorithm is defined as a unsupervised learning methods having an iterative process in which the dataset are grouped into k number of predefined non- overlapping clusters. It follows a simple procedure of classifying a given data set into a number of clusters, defined by the letter "k," which is fixed beforehand. The clusters are then positioned as points and all observations or data points are associated with the nearest cluster, computed, adjusted and then the process starts over using the new adjustments until a desired result is reached.

The k-meansis used for k-means algorithm for clustering the different types of soils.

By using this algorithm we can predict the which crop is suitable for their soil conditions.

Algorithm

STEP-1:Select the number of cluster center.

STEP-2:Set initial cluster center randomly.
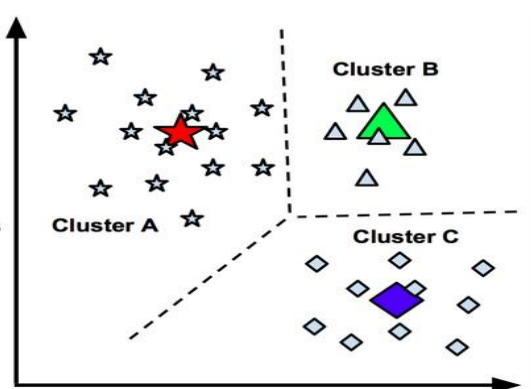
STEP-3 :Put object to closest cluster center.

STEP-4 :Recalculate the new cluster center.

STEP-5 :Create cluster based on smallest distance.

Formula for k-mean

$$J(V) = \sum_{i=1}^{c} \sum_{j=1}^{c_i} (\|x_i - v_j\|)^2$$

Where,'$\|x_i - v_j\|$' is the Euclideandistance between x_i and v_j. 'c_i' is the number of data points in i^{th} cl uster. 'c' is the number of cluster centers.

DECISION TREE

A Supervised Machine Learning Algorithm, used to build classification and regression models in the form of a tree structure.

A decision tree is a tree where each -

Node - a feature(attribute) Branch - a decision(rule)

Leaf - an outcome(categorical or continuous)

ID3 Algorithm

ID3 stands for Iterative Dichotomiser 3.It is a classification algorithm that follows a greedy approach by selecting a best attribute that yields maximum InformationGain(IG) or minimum Entropy(H). The ID3 algorithm is used for predict the weather conditions based on previous gathered data. We make decision tree for weather data then predict and forecast the weather conditions.

STEP -1 : Calculate the training data set gain of every attribute.

STEP -2 : At least two most important attributes with the same value of trained data.

STEP -3:randomly choose one from the set of equally important attribute.

STEP - 4 Make a decision tree node using select attribute.

STEP -5 : Expand decision tree node with selected attribute value.

3. Existing System

Farmers manually do their work of cultivation according to their experienced knowledge. But they are unable to predict which crop is suitable for their land.As they are not having enough knowledge on the soil which is having more fertility and also unable to forecast the climatic

International Journal for Innovative Engineering and Management Research
A Peer Reviewed Open Access International Journal
www.ijiemr.org

conditions, they seek advice from the research experts. This will take more time for the farmers.

And they are using chemical fertilizers for better yielding .Those chemicals are harmful to our health.
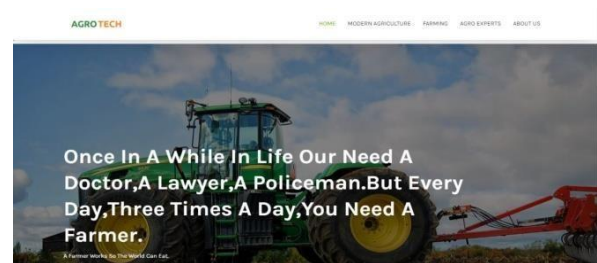
4. Proposed System

By our application farmer can find the better crop for their fields depending on the soil and weather conditions.We also guide the farmer for their modern way of farming, effects due to the usage of chemical fertilizers that leads to issues on health of human beings.Also suggests the organic fertilizers required for their soil situation and the pesticides which are to be used against different insect attacks.

4.1System Architecture:

fig: System architecture

Our system architecture consists of three modules.

1. The farmer requests for crop details. Then the request will be sent to cloud through our web application. By comparing the soil and weather conditions through soil and weather API's which is integrated to our application the required information will be provided. Then the farmer gets the crop details for his field.
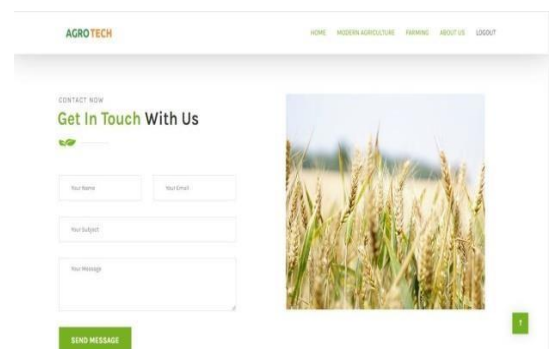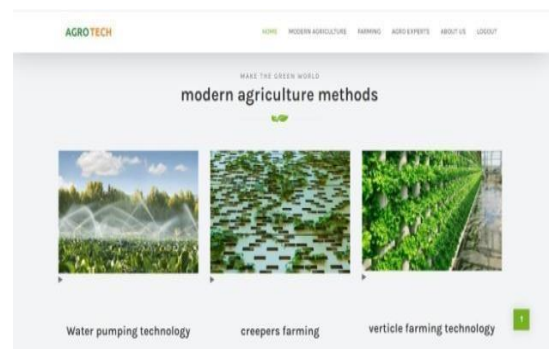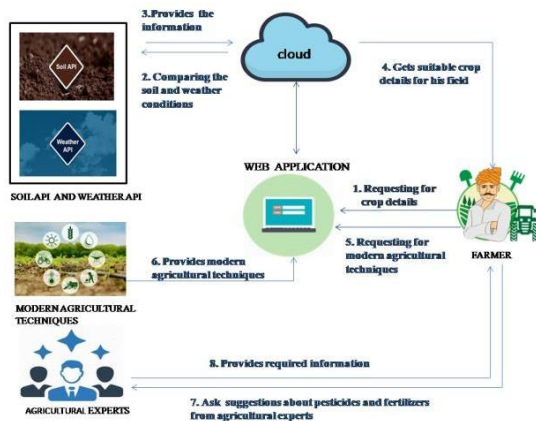
2. The farmers can also requests for modern agricultural techniques, and the information will be provided by our web application.

3. Farmers can also ask suggestions for the agricultural experts about which pesticides and fertilizers are to be used for his fields and the required information will be provided.

5 . Screens

UI SCREEN FOR HOME

UI SCREEN FOR MODERN AGRICULTURE,UI SCREEN FOR AGRO EXPERTS

6. Conclusion

As you all know, during the last decade the climate has become very uncertain .Due to this the farmers who were planting trading traditional crops are now facing problems , the yield of their crops is getting reduced. Therefore, the rate of suicide done by farmers has increased specifically in india . In our study we found out that if farmers knew the the yield of the crop that they are planting before hand,they would choose the crop that will produce better yield based on the type of soil and climate of that region.

7. Future Enhancement

By using algorithms like apriori to count the frequently occuring features which helps for predicting the crops in different locations. Implement a system which will predict the crop name along with apprropriate yield in a farm.

8. References

[1]Abello, P.M. Pardalos, M. Resende, Handbook of massive data sets,Kluwer,New York, 2002

[2]W. Klosgen, J.M. Zytkow, Handbook of datamining andknowledge discovery,Oxford University Press, 2002.

[3]P.M.Pardalos, L.V. Boginski, A. Vazacopoulos, Data mining in biomedicine, Springer, New York, 2007.

[4]P.M. Pardalos, P. Hansen, Data mining and mathematical programming, American Mathematical Society, USA, 2008.

[5]Crop cultivation season retrieved from Agriculture Market Information System (AMIS) crop calender

[6]Aymen E Khedr, Mona Kadry, Ghada Walid(2015), "Proposed framework for Implementing Data Mining Techniques to enhance Decisions in Agriculture sector Applied case on Food Security Information Center Ministry of Agriculture, Egypt", international conference on Communication, management and Information technology (ICCM)

[7]Afan Galih Salman, Bayu Kanigoro, Yaya Heryadi, "Weather forecasting using deep learning Techniques", 2015 International Conference on Advance Computer science and Information Systems(ICACSIS)

[8]Niketa Gandhi, Owaiz Petkar, Leisa J. Armstrong, "Predicting Rice crop yield using Bayesian Networks", 2016 Intl. Conference on Advances in Computing, Communications and Informatics (ICACCI), Sept. 21-24, 2016, Jaipur, India.

[9]Niketa Gandhi, Leisa J. Armstrong "Rice Crop Yield Forecasting of Tropical Wet and Dry Climatic Zone of India using data mining Techniques", 2016 IEEE International Conference on Advances in Computer Applications (ICACA)

[10]Kuljit kaur, Kalnwalpreet singh Attwal "Effect of Temperature and Rainfall on Paddy Yield using data mining ", 2017 7th international conference on Cloud Computing, Data Science and EngineeringConfluence

[11]Rupanjali D. Baruah, R.M. Bhagat, Sudipta Roy, L.N. Sethi "Use of Data Mining Technique for Prediction of tea yield in the face of Climate Change of Assam, India ", 2016 International Conference on Information Technology

[12]Scikit-learn, Machine learning in python retrieved from http: //scikitlearn. org