IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 7, July 2017

Regression Analysis with Cloud computing Technology in the field of Agriculture

Mrs. V. Ramya, MCA.[1], Mr. K. Mohamed Amanullah, MCA., M.Phil.,[2]

Computer Science, Bishop Heber College, Bharathidasan University[1,2]

Abstract: Nowadays, in India transformation tools are necessary in agronomic field sector to increase the economy. Internet of Things (IoT) is the efficient way to growth the productivity of agriculture. Development in internet has leads to the growth in data which results in emergence of cloud and IoT environment. In recent decades, pulling out of useful knowledge based content and recognizing the patterns in dataset are understood. Meaningful data need to be investigate and applied various self-restraints for identifying the significant features in agricultural domain. Understanding of suitable techniques in cloud computing and IoT requires for examining large datasets. The aim of this research is to build an information system for educating interaction between farmer and purchaser. The focus of this paper is to analyze and use of agri-cloud and IoT particularly by using Multiple Linear Regression (MLR) analysis to predict the crop manufacture to have decision making process easier.

Keywords: Regression analysis, Internet of Things (IoT), cloud computing.

I. INTRODUCTION

Cloud computing is a tool to make statistics knowledge related services available with simple manner and hiding the complications in the services. The user don't really need to know who are providing the facilities. It reduces the cost of availing those facilities and offers preservation. It provides facilities to user whenever and anywhere they needed. Agriculture cloud is the superior cloud that is composed of founded on MAD cloud architecture. Most of the agriculture based developing country like India, the agricultural sector donates are 20% of the country's GDP and around 65% of total population working in the sector. the main impartial of our work is to define land potential explicitly and energetically for unique and continuously changing soil and climatic condition the cloud solicitation helps the farmer to growth their agricultural yield .this work will also enable more rapid and complete integration and disseminate of local and scientific knowledge. This request provides high quality service and efficient data to user in anywhere anytime. It delivers width coverage area. The scope of this paper to concern the description and analysis of agricultural data with varying climatic and soil condition.

Cloud support numerous services to farmers to interact with cloud by using any low-priced ways like sensors, mobile devices, scanners etc. The query to the workers can be asked via internet connection. The investigation application is entirely based on MAD-cloud architecture, data are stored rendering to the co-ordinate and physical and chemical requirements of the crop. The data also defines soil texture, humidity, wind speed, rain amount. The user can obtain detain material about related crop which is require to increase the production. The user can select the co-ordinate location and define individual detail like name, place, etc. It also describes crop ailment and method to cure. Cloud provides objective way like required quality, dependability and security by using the cloud in agriculture helps the farmers and also used to growth our economic level. The information system in our proposed work delivers customer and farmer a proper channel of announcement where information can access, recovered and integrated. The quality of crop is key feature in agriculture domain, the quality can be validated by using limitation such as price, fertilizer used soil productivity etc. The quality analysis and comparison made by customer for different crops available in portfolio give competitive advantages to farmers and customers. Forecasting future manufacture of crops requires ancient data analysis and finding useful pattern in the datasets, certain set of dependent and self-governing variable can be used to forecast the future production of respective crops. The relationship between dependent and independent variables can predict the future manufacture of crops; the results may vary with change in values of reliant on and independent variables. Strong association between different dependent and autonomous variable signify that proposed model fit into measured agriculture system. There are many data mining techniques in presence.

Agriculture's very important problematic is yield prediction that can be solved using the current data and using that is regression [9]. They discuss that by put on data mining techniques the yield prediction problem can be resolved. They work at finding data models which is appropriate to achieve a high level accurateness also high level generality for predicting yield capabilities. So many types of techniques were projected on to the different obtainable data sets. They conclude that when it associate to the actual average production approximation of average production using Multiple Linear Regression Technique is given as 98% correctness with respect to limits.

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 7, July 2017

II. PROPOSED METHODOLOGY

It transforms collected information into comprehensible format. As we know that real world data is noisy, inconsistent, incomplete and having lack of certain trends. This unruly can be resolved by using data preprocessing task. Data preprocessing includes spring-cleaning, alteration, standardization, feature abstraction and selection. Data Collection allows users to take regulator of the entire data gathering lifecycle, from survey creation, through survey deployment and management to examination reporting. Users can create anything from humble web-based surveys to sophisticated data collection projects managed in multiple languages and numerous modes (on the web, on the phone, on paper, face-to-face, etc.). Users can enforce data quality through survey logic, real-time validation and calculations, sample management and quotas. Data Gathering uses an author once, deploy anywhere and in any language approach to survey expansion ensuring that you can easily leverage manifold means of reaching an audience without compromising review excellence. For web-based, phone-based and achieved face-to-face surveys, Data Collection includes robust survey organization and administration tools online to ensure discernibility into the survey process. Survey results can be made obtainable at any stage of data collection through flexible reportage facilities.
Statistical Package for the Social Sciences (SPSS) is software in which data can be arrived directly, or it can be smuggled from a number of dissimilar sources. Making data for analysis is one of the most significant steps done by SPSS software.

2.1 Agriculture Virtualized Information Database

Agriculture information resource such as multimedia files inaccessible sensing images and watching data streams are characterized with disseminated and massive storage. The development of the modern agriculture information techniques and the group and the utilization of farming information resources and the sharing of the data is urging problems. By taking remote sensing duplicate data which is widely used for pest controller, disease, yield forecast crop quality analysis etc. and the active monitoring of the crop growth also been understood.
There are different agriculture requests data, pastures monitoring data, climate data and soil enquiry data to the isolated association, cloud computing, distributed computing and grid calculating which features for virtualization, highly reliable, common ability and serve to the request. Hadoop, as an open source framework, is appropriate for the distributed data storing and management in cheap computers. The cloud storage platform is constituted by the central server the cloud storage stage provides the interface like resource access platform management and status monitor etc. after submitted to the cloud platform, it delivers evidence like remote sensing images, videos, and text can be stored in resource server.
Information service understands the management of fundamental services like registration information, legal validation and prominence monitor. Normal user also can register, regain and access to the agriculture evidence. The admin user can energetically monitor the node of resource server checking. Cloud storage is a kind of virtualized resource storage pool. Collective with one or more software or hardware, it has a capability of dynamic allocation, smooth extension storage and announcement. Cloud storage deals with the incorporation of resources distributed separately. The source of storage cloud is stored in dissimilar resource node. The high well-organized data transmission is achieved by dispersed storage algorithm. When users add farming information, whose schema has been registered, the category should be quantified.
A distributed node of dynamic nursing platform based on Hadoop cloud storing was developed by open source software ganglia. At present, the agriculture virtualized material database on efficiently developed and implement in many countries. It solve the problem of blocking of massive farming information and multi-copy storage and concurrency transmission of agriculture resource information. The server checking was done dynamically and agriculture source security in each investigation organization should be deliberated.

2.2 Cloud Computing

Cloud computing designates a multiplicity of computing impressions which include huge amount of computers which are associated through a real-time announcement network like the Internet. In science, cloud computing is for disseminated computing over the associated network, and the way of ability to run a package or the existing application on to the many connected processers at the same time. The phrase also refers to the facilities based on network, which again appear to be provided by real server hardware, and these services also served up by virtual hardware and simulated by software which are running on one or more real apparatuses. Actually virtual servers substantially do not exist and so it can be moved around and scaled up or down on the fly and it will not distress the end user possibly, rather like a cloud. Cloud computing provide the resources through which the whole thing like from computing the power to computing the infrastructure, requests required, business processes to individual collaboration. It can be delivered to anyone as a service wherever and whenever they need. Main advantages of cloud computing is centralized database, readiness for accessing data, local language announcement, guidance to farmers and others to get correct and related solution, data security and mainly dropping load.

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 7, July 2017

The analyzed data is then used as an input in the SPSS software to perform prediction method for predicting outcome. This instrument professionally simplifies the data mining process using qualitative data of the unique dataset, to produce most accurate outcome.
However, the qualitative data investigation process has to follow some steps to have positive effect on regression equation and to increase the extrapolative accuracy of results. The linear regression comparison in regression method is modified accordingly using efficient dissimilarity of equation measures between self-governing features, this provide better separation of pattern classes, which provide all the imaginable patterns and improve the result accuracy.

Cloud computing in agriculture

For updating agricultural field, cloud calculating in agriculture is the efficient way to deliver most updated information. Also the boundaries of producer that is inappropriate agricultural knowledge get fulfilled, restating same process is reduced and helps to recover utilization of a current resources. Nowadays, don't have sufficient technical support for the predicting weather. Most of growers are in a blind conformity state. Organizational form of production in agriculture is very simple, regressive, and low degree of concentration of agricultural production areas, combined agriculture is problematic to achieve. In addition, due to the boundaries of the farmers such as at marketplace predicting, decision-making commercial, gathering useful material and capacity of logistics management is more lacking, which leads to a divergence between the supply and request.
However, cloud computing in cultivation will provide guidance regarding modern equipment's, machineries, methods, methods to farmers and others to revolutionize agricultural field. The farmer in India should meet the request of latest information, tendencies and news if any. This demand can be serve by cloud computing expertise. It can increase farmer's own interests and provide healthy expansion of market supply and request.

III. EXPERIMENTAL RESULT

By illustrating the process of Multiple Linear Regression using an example adapted from Chaterjee, Hadi and Price from on estimating the performance of supervisors in a large financial organization.
The data shown in Table 1.1 are from a survey of clerical employees in a sample of departments in a large financial organization. The dependent variable is a performance measure of effectiveness for supervisors heading departments in the organization. Both the dependent and the independent variables are totals of ratings on different aspects of the supervisor's job on a scale of l to 5 by 25 clerks reporting to the supervisor. As a result, the minimum value for each variable is 25 and the maximum value is 125. These ratings are answers to survey questions given to a sample of 25 clerks in each of 30 departments. The purpose of the analysis was to explore the feasibility of using a questionnaire for predicting effectiveness of departments thus saving the considerable effort required to directly measure effectiveness.

Table 1.1 Predictions on the validation data

Case	Prediction	Error
21	44.46	-5.54
22	63.98	-0.02
23	63.91	10.91
24	45.87	5.87
25	56.75	-0.62
26	65.22	-0.78
27	73.23	-4.77
28	58.19	10.19
29	76.05	-8.95
30	76.10	-5.90

It can be applied for real time monitoring and control of controlled environment in greenhouses, and can be used for real time monitoring and decision support systems for field parameters, e.g. soil parameters (moisture, conductivity, etc.), environmental parameters (e.g. temperature, humidity, light, wind). The data received from remote sensing satellites and aerial imageries (through satellite/drones) can further complement the decision making process.

IV. CONCLUSION AND FUTURE SCOPE

This approach, planned agricultural information system which completes qualitative analysis of agricultural data. Mainly it bonds the gap amongst farmer and customer so that customer can distinguish among different farmers and the compulsory products. The cloud computing and IoT techniques that is regression analysis has been implemented to

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 7, July 2017

predict the crop production and analyze the patterns. SPSS software is used to perform multiple linear regression on to the large dataset to have precise results. Many investigators are involved in this scheme they can access the data and perform examination and provide analyzed files to users. The cloud computing technology has been implemented so that data would be protected on cloud. Data could be safe and easy to make any time at any location. In future scope, the system can be enhanced by using other techniques to have better consequences.

The future scope of this system will include the intelligent system which will take the decisions or actions according to the conditions prevailing. So that the farmer's interaction with the system will be minimized which will lead to less human efforts for the monitoring. This will allow farer to vilipend the nominal warnings as system will take care of it, which will be a lucrative deal for the end user.

REFERENCES

[1] Tuli, A, Hasteer, N, Sharma M, Bansal A., "Framework to leverage cloud for the modernization of Indian agriculture system", IEEE International Conference on Electro/Information Technology, 2014.
[2] K.Venkataramana, Dr.M.Padmavathamma, "A Design of Framework for AGRI-CLOUD",IOSR Journal of Computer Engineering (IOSRJCE), 2012.
[3] Shitala Prasad, Sateesh K. Peddoju, and Debashis Ghosh, "AgroMobile: A Cloud-Based Framework for Agriculturists on Mobile Platform", International Journal of Advanced Science and Technology, 2013.
[4] Sagar B.Jadhav Dr. Rajesh Prasad, Shantanu S.Panhale ,Chetan S. Mohture, "Review of Cloud Computing and Its Application", International Journal of Advanced Research in Computer Engineering and Technology(IJARCET), 2013.
[5] Yanxin Zhu, Di Wu and Sujian Li, "Cloud Computing and Agricultural Development of China: Theory and Practice", IJCSI International Journal of Computer Science Issues, 2013.
[6] Seena Kalghatgi, Kuldeep P. Sambrekar, "Review: Using Cloud Computing Technology in Agricultural Development", IJISET International Journal of Innovative Science, Engineering and Technology, 2015.
[7] Mr. Mahesh D. S1, Ms. Savitha S2, Dr. Dinesh K. Anvekar3., " A Cloud Computing Architecture with Wireless Sensor Networks for Agricultural Application", International Journal of Computer Networks and Communications Security, 2014.
[8] Prashant Satpute, Omprakash Tembhurne, "A Review of: Cloud Centric IoT based Framework for Supply Chain Management in Precision Agriculture", International Journal of Advance Research in Computer Science and Management Studies, 2014.
[9] D Ramesh, B Vishnu Vardhan, "Data Mining Techniques and Applications to Agricultural Yield Data", International Journal of Advanced Research in Computer and Communication Engineering, 2013.
[10] Geraldin B. Dela Cruz, Member, IACSIT, Bobby D. Gerardo, and Bartolome T, "Agricultural Crops Classification Models Based on PCA-GA Implementation in Data Mining", International Journal of Modeling and Optimization, 2014.
[11] Prof. Chandrakanth. Biradar1, Chatura S Nigudgi,"Statistical Based Agriculture Data Analysis", International Journal of Emerging Technology and Advanced Engineering, 2012.
[12] Farah Khan, Dr. Divakar Singh, "Association Rule Mining in the field of Agriculture: A Survey", International Journal of Scientific and Research Publications, 2014.
[13] Hetal Patel, Dharmendra Patel, "A Brief survey of Data Mining Techniques Applied to Agricultural Data", International Journal of Computer Applications, 2014.
[14] Dehua Zhang,"Analysis on the Influencing Factors of Farmers Income in Heilongjiang", International Conference on Civil, Materials and Environmental Sciences, 2015.
[15] Georg Ru and Rudolf Kruse, "Regression Models for Spatial Data: An Example from Precision Agriculture", ICDM, 2010.
[16] Aditya Shastry , Sanjay H A and Madhura Hegde, "A Parameter based ANFIS Model for crop yield prediction", IEEE, 2015.
[17] Feng Yu, Qian Zhang, RuPeng Luan, JunFeng Zhang, Xin Liu, "Application and Improvement of Intelligent Recommendation for Agricultural Information", Ninth International Conference on Natural Computation (ICNC), 2013.
[18] Jaiwei Han and Micheline Kamber, "Data Mining Concepts and Techniques", Second Edition,Morgan Kaufmann Publishers, 2006.
[19] Jinyu Chen, Wenxiu Zhang,"Analysis for Regional Differences and Influence Factor of Rural Income in China", Scientific Research, 2012.
[20] Liu Haime, Chen Yun, "Linear Regression Analysis of Gross Output Value of Farming, Forestry, Animal Husbandry and Fishery Industries", 2013.