ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

Regression method and Cloud computing Technology in the field of Agriculture

Pallavi V. Jirapure[1], Prarthana Deshkar[2],

P.G Student, Dept. of Computer Technology, Yeshwantrao Chavan College of Engineering, Nagpur, Maharashtra, India[1]

Prof., Dept. of Computer Technology, Yeshwantrao Chavan College of Engineering, Nagpur, Maharashtra, India[2]

ABSTRACT: Nowadays, in India modernization tools are necessary in agricultural field sector to increase the economy. Information and Communication technology (ICT) is the efficient way to increase the productivity of agriculture. Development in internet has lead to the increment in data which results in emergence of data mining. In recent decades, extraction of useful knowledge based content and identifying the patterns in dataset are comprehended. Meaningful data need to be analyze and applied various disciplines for identifying the important features in agricultural domain. Understanding of appropriate techniques in data mining requires for analyzing large datasets. The aim of this paper is to build an information system for improving interaction between farmer and customer. The focus of this paper is to analyze and use of data mining technique specially regression analysis to predict the crop production to have decision making process easier.

KEYWORDS: Regression analysis, Information and Communication technology (ICT), cloud computing, SPSS software

I. INTRODUCTION

Data mining is a multidisciplinary field in which data is analyzed from different perspectives and finds useful patterns from large dataset, transform information into knowledge by using sophisticated data analysis tool. Data mining is the process of extraction of information from large databases and it is a powerful new technology having a great potential to help researchers as well as others on the most important information in their data warehouses [1]. Data mining tools are used to predict the future trends and behaviors thus allowing businesses to make knowledge-driven decisions. The prediction and classification are two key aspect of data mining. Due to the advancement in various data mining Techniques and methodologies for agricultural domain, they are characterized in such a way that researchers can get in depth details of their work area. Data mining process can be broadly classified into two approaches explorative data mining approach and predictive data mining approach for proposed project. Explorative technique evaluates the big data and formulates the high utility datasets pattern of database in a brief descriptive way. Predictive considers major fields, variables or keys of database to predict the other dependent variables and this is the one of the most important technique of Data Mining.

The information system in our proposed work provides customer and farmer a proper channel of communication where information can access, retrieved and integrated. The quality of crop is key aspect in agriculture domain, the quality can be validated by using parameter 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 production of crops requires historical data analysis and finding useful pattern in the datasets, certain set of dependent and independent variable can be used to forecast the future production of respective crops. The relationship between dependent and independent variables can predict the future production of crops; the results may vary with change in values of dependent and independent variables. Strong correlation between different dependent and independent variable signify that proposed model fit into considered agriculture system. There are many data mining techniques in existence. Like clustering, classification, prediction i.e. regression, association rule mining.

Agriculture’s very important problem is yield prediction that can be solved using the existing data and using data mining technique that is regression [9]. They discuss that by applying data mining techniques the yield prediction problem can be solved. They work at finding data models which is suitable to achieve a high level accuracy also high

DOI: 10.15680/IJIRCCE.2016. 0404068

level generality for predicting yield capabilities. So many types of Data Mining techniques were estimated on to the different available data sets. They conclude that when it compare to the actual average production estimation of average production using Multiple Linear Regression Technique is given as 98% accuracy with respect to parameters.

Considering some assumptions data is preprocessed and then analyzed properly. Data preprocessing is important task of prediction. Agricultural sector is relatively an emerging research field where lot of work is to be done. Many researchers are involved in this field.

II. RELATED WORK

Related to cloud deployment model, work has been done earlier [1], it provide agriculture related information to the farmer mainly to those who have problem of getting updated information on time using cloud computing technology. Exploitation of the latest technologies like cloud computing and internet would make direct contribution to the productivity of agriculture.

During the development of other countries like China, they used cloud computing which provide early-warning and policy-making based on the agricultural products market, the tracing management of agricultural products quality [5]. In this paper they discussed for China's agricultural development what is the impacts of cloud computing and accordingly presented suitable application and also the cloud computing technology were considered as a long term system works. Cloud computing is a technology which build all data at center, enhance the service capabilities for different usage, integrate all the resources to perform special task, and make the data secured. Based on the analysis of single factor from previous studies, [14] the paper attempts to study more factors comprehensively and finds the main factors which influence analysis by using multiple linear regression method. Its drawback is that the factors influencing farmers' income is very complex, the paper can only choose some relatively important factors to analyze.

The Spss software is best software to analyze statistical data. The applied Spss software, how statistics of china agriculture are analyzed and established Linear regression model which estimates the effect of machinery and planting area that is total output of forestry, farming, fishery industries and animal husbandry [20]. They try to find the factors which will influence the total output of farming, animal husbandry, forestry and fishery industries

In [21], they focus on the scope for e-powering the people who live in the area like rural India and those who work for their benefits. The latest developments in Information Technology that facilitate effective Information Technology penetration to rural India, effective changes in pattern of information requirements & role of Information Technology, type of systems required in the post WTO (World Trade Organization) environment, first is the problems in e-powering rural India and second is that the possible solutions are examined.

III. PROPOSED METHODOLOGY

3.1 The Data Mining Process

There are many data mining techniques available. For prediction purpose regression analysis method is present. Regression analysis is one of the most important statistical techniques for business applications. It's a statistical methodology that helps estimate the strength and direction of the relationship between two or more variables. The analyst may use regression analysis to determine the actual relationship between these variables by looking at a corporation's sales and profits over the past several years. The regression results show whether this relationship is valid. The first phase is the data collection from different sources and data preprocessing using SPSS software to find if all the assumptions fulfilled by the data, this process can be considered as an optimum technique. The second phase is to understand the data and find appropriate regression method for generating model of prediction for forecasting of crop production. This prediction model is mainly utilized for making decision making process easier.

3.2 Methods and Procedures

3.2.1 Data preprocessing

Data preprocessing is an important task in data mining process. It transforms collected information into understandable format. As we know that real world data is noisy, inconsistent, incomplete and having lack of certain trends. This problem can be resolved by using data preprocessing task. Data preprocessing includes cleaning, transformation, normalization, feature extraction and selection. Data Collection allows users to take control of the entire data collection lifecycle, from survey creation, through survey deployment and management to survey reporting. Users can create anything from simple web-based surveys to sophisticated data collection projects administered in multiple languages and multiple 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 Collection uses an author once, deploy anywhere and in any language approach to survey development ensuring that you can easily leverage multiple means of reaching an audience without compromising survey quality. For web-based, phone-based and managed face-to-face surveys, Data Collection includes robust survey management and administration tools online to ensure visibility into the survey process. Survey results can be made available at any stage of data collection through flexible reporting facilities.

SPSS is software in which data can be entered directly, or it can be imported from a number of different sources. Preparing data for analysis is one of the most important steps done by SPSS software.

3.2.2 Regression

This is a task performed to know the correlation between different fields and a way of knowing a function in which all the data items are analyzed and provided a real valued prediction outcome. There are two methods for prediction, they are linear regression (LR) and non linear regression (NLR).

3.2.3 Linear regression

In linear regression, an equation is generated which explains the relationship between more than one data fields especially between .independent and dependent variable. These variables are responsible for prediction of any given observation. There are again two methods depending on to the type of independent variables, they are simple linear regression and multiple linear regression. If the prediction is done on the basis of single independent variable i.e. if only one independent variable is responsible for prediction of single dependent variable then it is referred as simple linear regression and if two or more than two independent variables are responsible for predicting a single dependent variable then a method is referred as multiple linear regression.

3.2.4 Equation

In simple linear regression, we predict score of one variable from score of other variable. The variable we are predicting is referred to as y. The variable we are basing our prediction on is referred to as x. The equation is given as $y=bx+a$

Where, y - value to be predict
b - Slope or rate of increase or decrease of y at each unit increase in x
x – Explanatory variable

In multiple linear regression, we predict score of one variable from score of more than one variables. The variable we are predicting is referred to as y. The variables we are basing our prediction on are referred to as x1,x2,…,xn.

The equation of multiple linear regression.

$y=b+a0x1+a1x2+a2x3+a3x4+a4x5$

Where, y - dependent variable i.e. variable to be predict
b - Regression weight
x1, x2, x3, x4, x5 – independent variables
a0,a1,a2,a3,a4 – correlation coefficients

3.2.5 Analysis

Regression analysis is an indispensable tool for analyzing relationships between financial variables. For example, it can: Identify the factors that are most responsible for prediction. Determine how much a change in dependent variable will impact a prediction. Develop a forecast of the future value of the dependent variable.

3.3 Cloud Computing

Cloud computing describes a variety of computing concepts which involve huge amount of computers which are connected through a real-time communication network like the Internet. In science, cloud computing is for distributed computing over the connected network, and the way of ability to run a program or the existing application on to the many connected computers at the same time. The phrase also refers to the services 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 machines. Actually virtual servers physically do not exist and so it can be moved around and scaled up or down on the fly and it will not affect the end user possibly, rather like a cloud. Cloud computing provide the means through which everything like from computing the power to computing the infrastructure, applications required, business processes to personal 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 communication, guidance to farmers and others to get correct and related solution, data security and mainly reducing load.

IV. IMPLEMENTATION OF INFORMATION SYSTEM

Figure 4.1 shows the architecture of information system. The idea is to implement the application of providing qualitative analyzed data to the users to have decision making process easier.

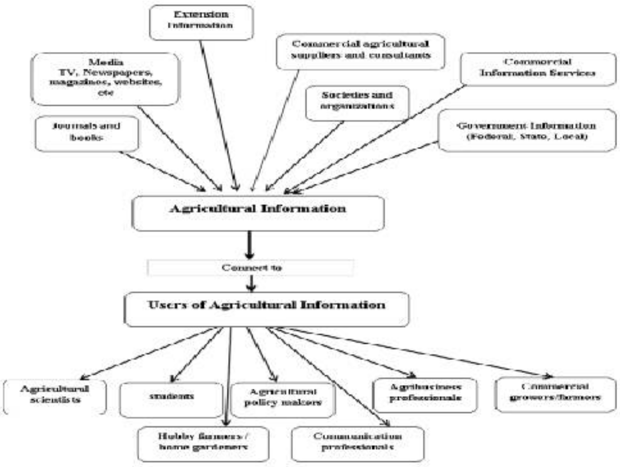

Extension Information
Media TV, Newspapers, magazines, websites, etc
Commercial agricultural suppliers and consultants
Commercial Information Services
Societies and organizations
Journals and books
Government Information (Federal, State, Local)
Agricultural Information
Connect to
Users of Agricultural Information
Agricultural scientists
students
Agricultural policy makers
Agribusiness professionals
Commercial growers/farmers
Hobby farmers / home gardeners
Communication professionals

Fig 4.1 Information system model

The analyzed data is then used as an input in the SPSS software to perform prediction method for forecasting outcome. This mechanism efficiently simplifies the data mining process using qualitative data of the original dataset, to produce most accurate outcome.

However, the qualitative data analysis process has to follow some steps to have positive effect on regression equation and to increase the predictive accuracy of results. The linear regression equation in regression method is modified accordingly using efficient variation of equation measures between independent features, this provide better separation of pattern classes, which provide all the possible patterns and improve the result accuracy.

Information system provides all the details of farmer and customer and others to have direct communication with each other without using middle man. Bridging the gap between farmers and customer by providing products online analyzing features like production location, warehouse location, quantity , etc,. E-farming and modernizing the agricultural field is the main aim of this information system. Using data mining technique, regression analysis all the process is possible and easier.

A. Cloud computing in agriculture

For modernizing agricultural field, cloud computing in agriculture is the efficient way to provide most updated information. Also the limitations of producer that is improper agricultural knowledge get fulfilled, repeating same process is reduced and helps to improve utilization of an existing resources. Nowadays, we don’t have enough technical support for the forecasting weather. Most of farmers are in a blind conformity state. Organizational form of production in agriculture is very simple, backward, and low degree of specialization of agricultural production areas, integrated agriculture is difficult to achieve. In addition, due to the limitations of the farmers such as at market forecasting, decision-making business, gathering useful information and capacity of logistics management is more lacking, which leads to a mismatch between the supply and demand.

However, cloud computing in agriculture will provide guidance regarding modern equipments, machineries, techniques, methods to farmers and others to modernize agricultural field. The farmer in India should meet the demand of latest information, trends and news if any. This demand can be serve by cloud computing technology. It can increase farmer’s own interests and provide healthy development of market supply and demand.

V. EXPERIMENTAL RESULTS

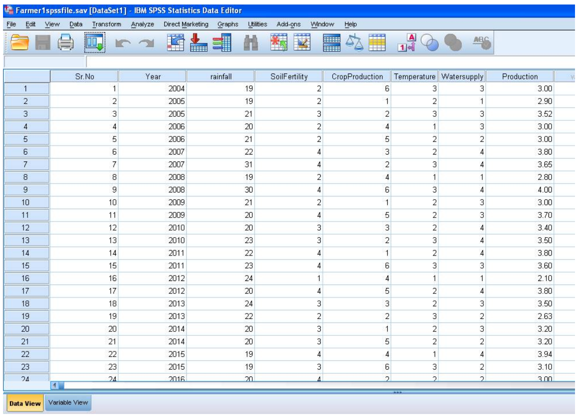
Now under this section, extensive experiments on real-world agricultural data are present. Multiple Linear Regression approach of regression analysis with respect to prediction factors is analyzed. The linear regression equation is estimated after following data preprocessing process. The multiple linear regression method was applied and results are compared accordingly.

The experiment used the SPSS version 20 software in the implementation and utilization. A computer with 2 Gigabyte memory, equipped with 2.80 Ghz Processor, and a proprietary 32 bit operating system was utilized. Microsoft Azure cloud computing platform to perform online data storage.

The default settings in the SPSS software and in the multiple linear regression method configurations, was used in the experiment.

A. Data collection

Real world data is collected from different surveys. The dataset is collection of previous year historical data including the 7 parameters; they are year, rainfall, soilfertility, cropProduction, Temperature, watersupply and production.

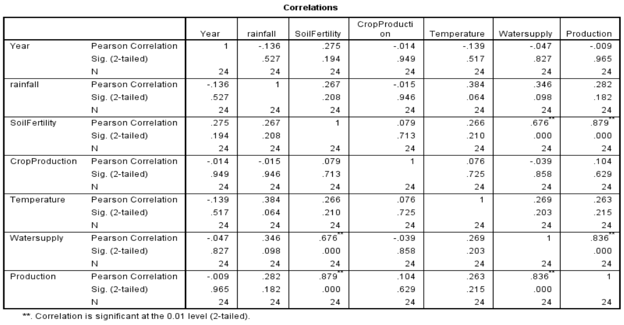


	Sr.No	Year	rainfall	SoilFertility	CropProduction	Temperature	Watersupply	Production
1	1	2004	19	2	6	3	3	3.00
2	2	2005	19	2	1	2	1	2.90
3	3	2005	21	3	2	3	3	3.52
4	4	2006	20	2	4	1	3	3.00
5	5	2006	21	2	5	2	2	3.00
6	6	2007	22	4	3	2	4	3.80
7	7	2007	31	4	2	3	4	3.65
8	8	2008	19	2	4	1	1	2.80
9	9	2008	30	4	6	3	4	4.00
10	10	2009	21	2	1	2	3	3.00
11	11	2009	20	4	5	2	3	3.70
12	12	2010	20	3	3	2	4	3.40
13	13	2010	23	3	2	3	4	3.50
14	14	2011	22	4	1	2	4	3.80
15	15	2011	23	4	6	3	3	3.60
16	16	2012	24	1	4	1	1	2.10
17	17	2012	20	4	5	2	4	3.80
18	18	2013	24	3	3	2	3	3.50
19	19	2013	22	2	2	3	2	2.63
20	20	2014	20	3	1	2	3	3.20
21	21	2014	20	3	5	2	2	3.20
22	22	2015	19	4	4	1	4	3.94
23	23	2015	19	3	6	3	2	3.10
24	24	2016	20	4	2	2	2	3.00

Fig 5.1 Data collection in SPSS software

B. Data Pre-processing

From the given data set selecting Production as dependent variable and all other are independent variables. However, only those variables are selected as independent variables that have linear relationship with dependent variables.



Correlations

		Year	rainfall	SoilFertility	CropProduction	Temperature	Watersupply	Production
Year	Pearson Correlation	1	-.136	.275	-.014	-.139	-.047	-.009
	Sig. (2-tailed)		.527	.194	.949	.517	.827	.965
	N	24	24	24	24	24	24	24
rainfall	Pearson Correlation	-.136	1	.267	-.015	.384	.346	.282
	Sig. (2-tailed)	.527		.208	.946	.064	.098	.182
	N	24	24	24	24	24	24	24
SoilFertility	Pearson Correlation	.275	.267	1	.079	.266	.676**	.879**
	Sig. (2-tailed)	.194	.208		.713	.210	.000	.000
	N	24	24	24	24	24	24	24
CropProduction	Pearson Correlation	-.014	-.015	.079	1	.076	-.039	.104
	Sig. (2-tailed)	.949	.946	.713		.725	.858	.629
	N	24	24	24	24	24	24	24
Temperature	Pearson Correlation	-.139	.384	.266	.076	1	.269	.263
	Sig. (2-tailed)	.517	.064	.210	.725		.203	.215
	N	24	24	24	24	24	24	24
Watersupply	Pearson Correlation	-.047	.346	.676**	-.039	.269	1	.836**
	Sig. (2-tailed)	.827	.098	.000	.858	.203		.000
	N	24	24	24	24	24	24	24
Production	Pearson Correlation	-.009	.282	.879**	.104	.263	.836**	1
	Sig. (2-tailed)	.965	.182	.000	.629	.215	.000	
	N	24	24	24	24	24	24	24

**. Correlation is significant at the 0.01 level (2-tailed).

Fig 5.2 Correlation table

The figure shows the correlation table, specifying correlation of each variable with each of other variables. The value varies between -1 to +1, negative value specify negative relationship and positive value specify positive relationship. If the value is close to 1 gives more correlation of that variable with respected variable. From above table, year has negative relation while others have positive relation with production variable. Hence, independent variables are decided as rainfall, SoilFertility, CropProduction, Temperature and waterSupply. The model is estimated using least square method and tested the fit of model using coefficient of variation (R²) in SPSS software. The hypothesis test estimates if all the coefficients in equation are statistically significant as shown in above table.

C. Assumptions

Before performing multiple linear regression data should fulfill some assumptions. First is, observations should be independent i.e. not autocorrelated. Using Durbin Watson method autocorrelations are detected.

$$d = \frac{\sum_{t=2}^{T}(e_t - e_{t-1})^2}{\sum_{t=1}^{T} e_t^2}$$

If d is between 1.5 and 2.5 means observations are independent. Using SPSS software value estimated as,

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.942^a	.887	.856	.17657	1.609

a. Predictors: (Constant), Watersupply, CropProduction, Temperature, rainfall, SoilFertility

b. Dependent Variable: Production

Fig. 5.3 Durbin Watson value

Second is, Data need to show homoscedasticity i.e. variance among the regression line is same for all the values of x.

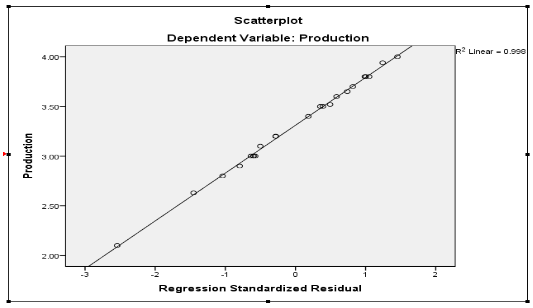
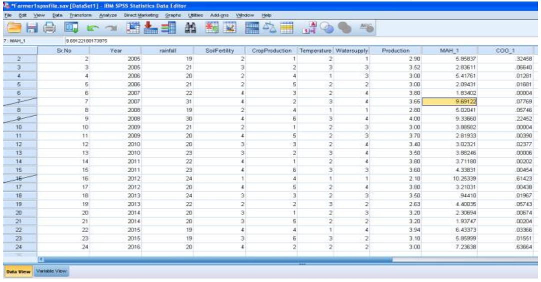


Fig. 5.4 Scatterplot

Next is, Data should not show Multicollinearity i.e. more than two independent variables are not highly correlated with each other.

Data must not have any outliers, using Mahalanobi's Distance Outliers can be detected. Finding critical value, the outliers can be checked. Then ignoring these observations for prediction gives accurate solution.



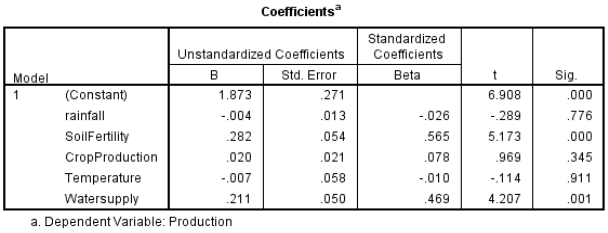
7 : MAH_1 9.69122180173976

	Sr.No	Year	rainfall	SoilFertility	CropProduction	Temperature	Watersupply	Production	MAH_1	COO_1
2	2	2005	19	2	1	2	1	2.90	6.85837	.32458
3	3	2005	21	3	2	3	3	3.52	2.83611	.06640
4	4	2006	20	2	4	1	3	3.00	5.41761	.01281
5	5	2006	21	2	5	2	2	3.00	2.09431	.01681
6	6	2007	22	4	3	2	4	3.80	1.83402	.00004
~~7~~	7	2007	31	4	2	3	4	3.65	9.69122	.07769
8	8	2008	19	2	4	1	1	2.80	5.02041	.05746
~~9~~	9	2008	30	4	6	3	4	4.00	9.33660	.22452
10	10	2009	21	2	1	2	3	3.00	3.86982	.00004
11	11	2009	20	4	5	2	3	3.70	2.81933	.00390
12	12	2010	20	3	3	2	4	3.40	3.02321	.02377
13	13	2010	23	3	2	3	4	3.50	3.86246	.00006
14	14	2011	22	4	1	2	4	3.80	3.71180	.00202
15	15	2011	23	4	6	3	3	3.60	4.33831	.00454
~~16~~	16	2012	24	1	4	1	1	2.10	10.25339	.61423
17	17	2012	20	4	5	2	4	3.80	3.21031	.00438
18	18	2013	24	3	3	2	3	3.50	.94410	.01967
19	19	2013	22	2	2	3	2	2.63	4.40035	.05743
20	20	2014	20	3	1	2	3	3.20	2.30694	.00674
21	21	2014	20	3	5	2	2	3.20	1.93747	.00204
22	22	2015	19	4	4	1	4	3.94	6.43373	.03366
23	23	2015	19	3	6	3	2	3.10	5.85999	.01551
24	24	2016	20	4	2	2	2	3.00	7.23638	.63664

Fig 5.5 Preprocessed Data

D. Performed multiple linear regression

On preprocessed data, multiple linear regression is performed. Coefficient table gives the value of coefficient to be placed in multiple linear regression equation.



Coefficients^a

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	1.873	.271		6.908	.000
	rainfall	-.004	.013	-.026	-.289	.776
	SoilFertility	.282	.054	.565	5.173	.000
	CropProduction	.020	.021	.078	.969	.345
	Temperature	-.007	.058	-.010	-.114	.911
	Watersupply	.211	.050	.469	4.207	.001

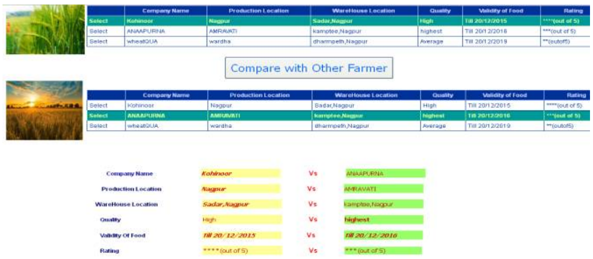
a. Dependent Variable: Production

Fig 5.6 Coefficient table

The equation produced is,
y= 1.873-0.004 * rainfall + 0.282 * SoilFertility + 0.020 * CropProduction - 0.007 * Temperature + 0.211 *Watersupply

E. Interaction among customer and farmer

For qualitative analysis of product, the product is analyzed and compared with each available product and then making decision.



	Company Name	Production Location	WareHouse Location	Quality	Validity of Food	Rating
Select	Kohinoor	Nagpur	Sadar,Nagpur	High	Till 20/12/2015	****(out of 5)
Select	ANAAPURNA	AMRAVATI	kamptee,Nagpur	highest	Till 20/12/2016	***(out of 5)
Select	wheatQUA	wardha	dharmpeth,Nagpur	Average	Till 20/12/2019	**(outof5)

	Company Name	Production Location	WareHouse Location	Quality	Validity of Food	Rating
Select	Kohinoor	Nagpur	Sadar,Nagpur	High	Till 20/12/2015	****(out of 5)
Select	ANAAPURNA	AMRAVATI	kamptee,Nagpur	highest	Till 20/12/2016	***(out of 5)
Select	wheatQUA	wardha	dharmpeth,Nagpur	Average	Till 20/12/2019	**(outof5)

Company Name	Kohinoor	Vs	ANAAPURNA
Production Location	Nagpur	Vs	AMRAVATI
WareHouse Location	Sadar,Nagpur	Vs	kamptee,Nagpur
Quality	High	Vs	highest
Validity Of Food	Till 20/12/2015	Vs	Till 20/12/2016
Rating	****(out of 5)	Vs	***(out of 5)

Fig. 5.6 Customer analyzing each product

VI. CONCLUSION AND FUTURE SCOPE

In this paper, we proposed agricultural information system which performs qualitative analysis of agricultural data. Mainly it bridges the gap between farmer and customer so that customer can differentiate among different farmers and the required products. The data mining technique that is regression analysis has been implemented to predict the crop production and analyze the patterns. Spss software is used to perform multiple linear regression on to the large dataset to have accurate results. Many researchers are involved in this system they can access the data and perform analysis and provide analyzed files to users. The cloud computing technology has been implemented so that data would be saved on cloud. Data could be safe and easy to fetch any time at any location. In future scope, the system can be improved by using other data mining techniques to have better results.

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.

[21] S.C. Mittal, "Role of Information Technology in Agriculture and its Scope in India", Fertilizer Association of India (FAI) Forum, 2002.