International Academic Institute for Science and Technology

International Academic Journal of Science and Engineering
Vol. 4, No. 2, 2017, pp.190-197.
ISSN 2454-3896

International Academic Journal of Science and Engineering

www.iaiest.com

Regression Techniques Using Data Mining in Flowering Plant

I , Leila MORTAZAVI FAR [2],

Department of ICT, Computer Faculty, Zand Daneshgostar Institute of higher Education, Shiraz, Iran.

Abstract

Nowadays, data is considered as the heart of every business. Regardless of being macro/micro industry Data is vital in wholesales, retailer, production, communication, facilities, transportation, insurance, credit cards, booking through operational systems. Therefore, a device is needed to process the stored data and then provides users with the same. Data mining in recent years, have increasingly developed and provided users several applications limitlessly.In previous decades, large number of data has been accumulated and stored in data Unit .This research analyzes data through regression and discussed the effect of dry leaf weight on the period of flowering. The data base has been collected from different plants of Kish Island in various areas in order to determine, classify and predict effective growth factors on blooming. The analysis of these agricultural data base with data mining methods may have some advantages in agriculture.

Keywords: agriculture, data mining, regression, dry leaf weight, flowering

Introduction:

In fact, organizations have been bombarded by data. Still are looking for knowledge. It shows that they could not use the stored data properly. On the other had They are obliged to use appropriate methods for finding and grouping practical and effective data from a vast range of data available , for correct decision – making where there are various markets , services, Job environments, demands and users. Data mining is a prompt solution in organizations and institutes. The more ranges of data result in more complexity among them. And it would be harder to access the data. So, Data is a method to discover knowledge [1].

Today, the traditional methods, used in data collection and data analysis and not feasible due to the economy and Data mining can be used as a good substation. There are different definitions for the term "data mining". Some consider it as a tool/ device which makes users capable of direct connection to a vast range of data. Others suggest an accurate definition for Data mining i.e., data investigation. Data mining is a process for discovery the hidden knowledge of data through connection to useful patterns. Data mining the analysis of a vast range of data in order to discover patterns and meaning full law, mostly

through modeling and algorithm of input data, based on a specific purpose. Data-mining is a semi-automatic procedure of data unit analysis to find useful patterns. It is also defined as a procedure in which different techniques of modeling and analysis are used to discover. Patterns and proper connections practical for prediction. As shown above, data mining indicates knowledge extraction, analysis and patterning among data. The main idea in data-mining is that the old data have pieces of information, feasible and useful in future. It is hard because the patterns are not clear and signals received from data and sometime ambiguous. The purpose of Data-mining is to find patterns in previous data which make needs, preferences, and tends clear. Therefore. One of the host functions of data-mining is to distinguish signals and recognize basic patterns in seemingly accidental variable. It is generally related to modeling. Basically each model is an algorithm or a group of laws which like a group of inputs based on specific purposes. One can classify the environmental problems into six groups. In other words, in order to convert a business problem in to a data-mining problem it should change the same into one of data-mining activities, namely, classification, estimate, prediction, similarity Categorization, clustering, and description. Classification, estimate, and prediction are examples of controlled data-mining in which the purpose is to find the value of a goal variable. Similarity Categorization. And clustering and Uncontrolled data-mining in order to find data structure without focusing on any specific variable, Description may be controlled or uncontrolled. Data mining models are as follows [2-8]:

Nervous systems: these system, inspired by Human brain pattern store related information inside systems. Nervous system is capable of learning and also can teach or decide based on input data, similar to Bio system.

Decision Tree: a structure made in order to facilitate decision-making is W5 which is to deride large group of gathered data into smaller ones based on a series of simple decision making laws. In a sequential division (brunching) members are move similar hierarchically.

Low Deduction: It is a method in a system used to categorize items based on law deduction. Although a decision tree can deduce a group of laws, Law deduction method can also conclude in depend ant Laws which are hot necessarily generating a tree. Law deduction is a group of observations. The laws deduced can show a complete scientific model of data or a limited (local) model of data.

Regression: regression (logistic model), decimal Analysis and regression (loges tic model) is used when it is to predict the membership in a category, not the values of a quantitative variable. When a category (response variable) has more than two strata, decimal analysis is required and when the response variable has only two strata, regression (logistic model) can be applied.

Regression technique:

Regression is a data mining (machine learning) technique used to fit an equation to a dataset. The simplest form of regression, linear regression, uses the formula of a straight line ($y = mx + b$) and determines the appropriate values for m and b to predict the value of y based upon a given value of x. Advanced techniques, such as multiple regression, allow the use of more than one input variable and allow for the fitting of more complex models, such as a quadratic equation [9].

Regression technique can be used in agriculture. reported the results of a regression analysis of the relationship between energy use and agricultural productivity. The study was based on the analysis of the yearbook data for the period 1971-2003[10].

[11] Showed regression technique can predict real demand water in Indian agriculture. A major task in agriculture production is water and fertilizer management. Deficit application, however, may limit the growth of crop.
This paper discusses a process model for analyzing data and describes the support that SPSS provides for this model. SPSS-based analysis and application construction process is illustrated through a case study in the agricultural domain-ornamental plants.

1. Methodology

1.1. Location of study

The study was conducted in Kish Island which is located in the Persian Gulf on the mainland Iran. Kish's climate is considered to be a very dry, semi equatorial. Four different regions of Kish Island were selected to do the experiment. They were Sanae, Sadaf, Pavion and Saffein.

1.2. Data collection

Data were collected from different ornamental plants which are in selected regions. The ornamental plants included *Nerium oleander*, *Tecoma stans*, *Thevetia nerrifolia* and *Delonix regia*. The data were related to growth factors such as number of flowers, florets, leaves, duration of flowering, leaf area, chlorophyll and proline contents, fresh and dry weight of leaves.

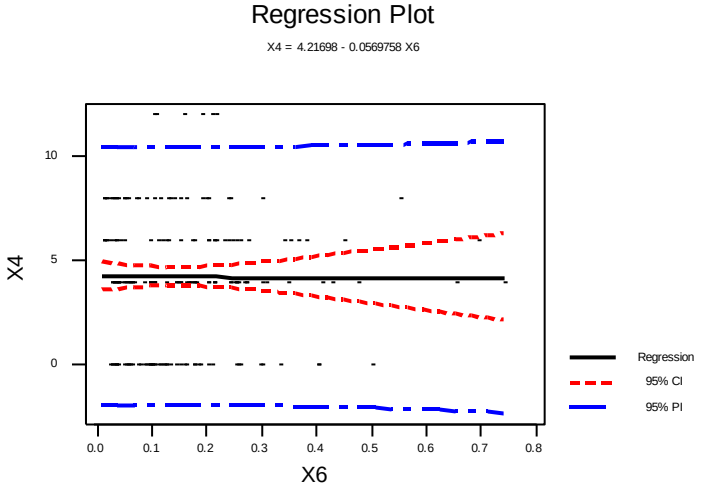
2. The effect of dry leaf weight on flowering period through regression

In this section we discussed the effect of dry leaf weight on flowering [12,13]. Factors for investigation the effect of number of leaves in floral plants on flowering period are:
X4: Flowering period
X6: Dry leaf weight
X10: Flowering period in floral plants in season (1)
X12: Dry leaf weight in floral plants in winter
X16: Flowering period in floral plants in zone (1), cold season (1)
X18: Dry leaf weight in floral plants in winter, zone (1), cold season (1)
P1: Predicted interval
C1: Reliance interval

Table 1: The effect of dry-leaf weight on flowering period in floral trees through regression Analysis

Number of leaves observed	Exact value (per mount)	error	Reliance interval	Predicted value(future)
0/1	4/22	0/24	3/73-4/69	1/98-10/40
0/3	4/21	0/34	3/52-4/87	2/01-10/41
0/22	4/20	0/25	3/70-4/70	1/99-10/40
0/33	4/19	0/3	3/44-4/95	2/01-10/41
0/5	4/18	0/6	2/91-5/46	1/99-10/40

Figure1: The effect of dry-leaf weight on flowering period in floral trees through regression Analysis.



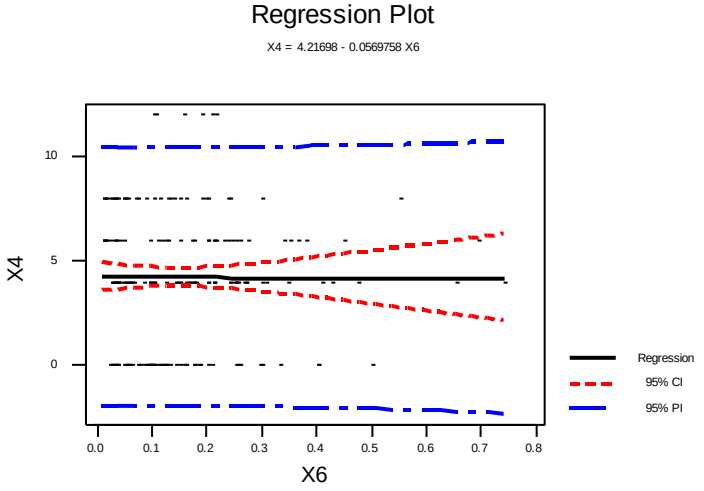
According to table 1 and figure 1 dry-leaf weight has no effect on the flowering period in floral trees, because p>0.5.

2.1. The effected of dry-leaf weight on flowering period, season (1) through regression Analysis.

Table 2. The effect of dry-leaf weight on flowering period, season (1), through regression Analysis.

Number of leaves observed	Exact value	error	Reliance interval	Predicted value(future)
0/1	4/19	0/34	3/49-4/88	0-10/48
0/3	4/26	0/51	3/23-5/28	0-10/58
0/22	4/23	0/37	3/49-4/97	0-10/52
0/33	4/27	0/58	3/11-5/42	0-10/62
0/5	4/33	0/99	2/35-6/30	0-8810

Figure 2. The effect of dry-leaf weight on flowering period, season (1), through regression Analysis.



According to table 2, and figure.2, dry-leaf weight has no effect on the flowering period in cold season.

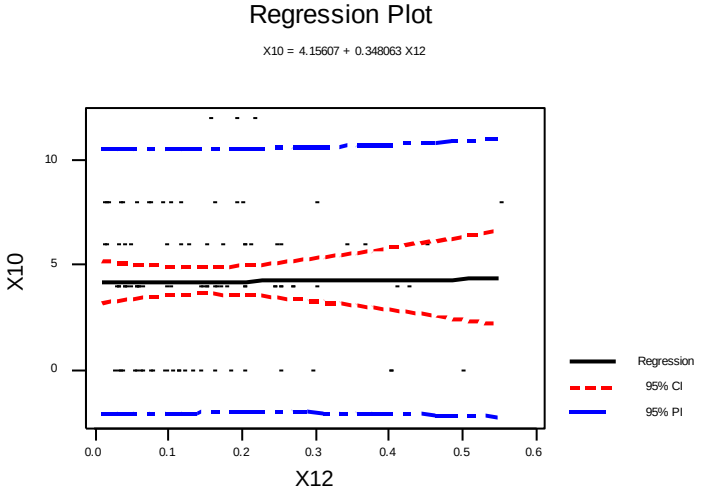
2.2. The effected of dry-leaf weight on flowering period, season (1), zone (1) through regression Analysis.

Table 3. The effect of dry-leaf weight on flowering period, season (1), zone (1) through regression Analysis.

Number of leaves observed	Exact value	error	Reliance interval	Predicted value(future)
0/1	5/11	0/91	.3/23-7/00	0-13/49
0/3	4/52	1/52	2/14-6/19	0-13/03

0/22	4/76	0/86	2/96-6/56	0-13/11
0/33	4/44	0/29	1/75-7/12	0-13/02
0/5	3/94	2/23	0-8/56	0-13/32

Figure 3. The effect of dry-leaf weight on flowering period, season (1), zone (1) through regression Analysis.



According to table 3, and figure. 3, the dry-leaf weight has no effect on the flowering period.

2.3. The effected of dry-leaf weight on flowering period, season (1), zone (1), type (1) through regression Analysis.

3. Table 4. The effect of dry-leaf weight on flowering period, season (1), zone (1), type (1) through regression Analysis.

Number of leaves observed	Exact value	error	Reliance interval	Predicted value(future)
0/1	8/95	2/23	2/77-15/14	0/81-18/75
0/3	9/05	2/29	2/70-15/40	0/84-18/96
0/22	9/01	1/20	5/69-12/33	0/71-17/31
0/33	9/06	2/82	1/25-16/88	1/84-19/97

0/5	9/15	6/02	0/58-25/87	9/23—27/52

Figure 4. The effect of dry-leaf weight on flowering period, season (1), zone (1), type (1) through regression Analysis.

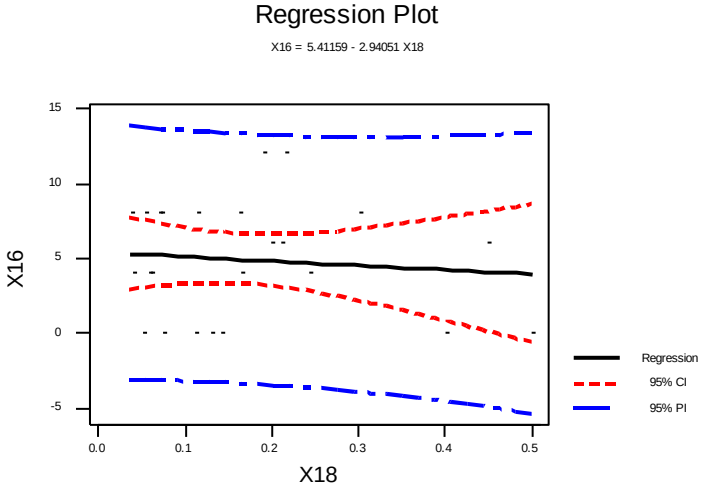


Table 4, and Figure 4, shows that dry-leaf weight has no effect on flowering period, season (1), and high zone in ornamental plants.

Conclusion:

The collection of information and data has increased with the advent of new computing technology, but establishing patterns within this data has become more difficult and requires new approaches and tools if it is to be undertaken. The advent of this problem has provided an opportunity from which data analysis has started to take over from current methods. Since, a great deal of data was collected in this research some effective outcome gained. This paper studied the effect of dry-leaf weight on flowering period. The results also showed that there is no correlation between these two variables because p>0.5. Tables and figures show that dry-leaf weight has no effect on flowering period.

References

Ashok Kumar D, Kannathasan N. A Survey on Data Mining and Pattern Recognition Techniques for Soil Data Mining. IJCSI 2011; 8(3): 20-32

Au W H, Chan K C C. An effective algorithm for discovering fuzzy rules in relational databases. In: Proceedings of IEEE International Conference on Fuzzy Systems FUZZ IEEE 1998. Alaska, USA; 1998. P. 1314-1319

Chapple M (2011) Regression. , Accessed 2 Feb 2012

Fayadd U, Piatesky–Shapiro G, Smyth P. Data Mining to Knowledge Discovery in Databases. USA: AAAI Press / the MIT Press; 1996. p.50-67

H. Fetanat, l. Mortazavifar, N. Zarshenas(2015)“ The analysis of agricultural data with regression data mining technique” Ciência e Natura, v. 37, p. 102−107, COMCONF01

H. Fetanat, l. Mortazavifar, N. Zarshenas(2015)“ THE APPLICATION OF DATA MINING TECHNIQUES IN AGRICULTURAL SCIENCE”, Ciência e Natura, v. 37, p. 108−116 ,COMCONF01

Jackson J. Data mining: a conceptual overview, communications of the association for information systems 2002; 8: 267-296

Kacprzyk J, Zadrozny S. Data mining via linguistic summaries of data: an interactive approach. In Proceedings of IIZUKA 1998. Fukuoka, Japan; 1998. p. 668-671

Lee R S, Liu J N K. Tropical cyclone identification and tracking system using integrated neural oscillatory leastic graph matching and hybrid RBF network track mining techniques. IEEE Transactions on Neural Networks 2000; 11: 680-689

Mai J (2006) Data mining application journal of amercing no 2

Mitra S, Mitra P, Pal SK. Evolutionary modular Design of rough knowledge-based network using fuzzy attributes. Neurocomputing 2001; 36: 45-66

Osman K (1997) Arterial hypertension in Saudi Arabia,” Annals of Saudi Medicine, Vol. 17, No. 2, 1997, pp. 170-174

Zhang Y Q, Fraser M D, Gagliano R A, Kandel A. Granular neural networks for numerical-linguistic data fusion and knowledge discovery. IEEE Transactions on Neural Networks 2000; 11: 658-667