

**Research Article**

## **DISCUSSING HOW TO IMPROVE THE QUALITY OF INFORMATION TECHNOLOGY SYSTEMS IN ELECTRONIC LEARNING THROUGH DATA MINING IN SCHOOLS OF SARDASHT**

**Abdollah Pour Ahmad and \*Mehran Molavi**

*Department of Public Administration, Mahabad Branch, Islamic Azad University, Mahabad, Iran*

*\* Author for Correspondence*

### **ABSTRACT**

The purpose of this research is to discuss how to improve the quality of information technology systems in e-learning through data mining in schools of Sardasht. The method of research was descriptive-correlative. The population is consisted of entire computer teachers in schools of Sardasht as 300 individuals. The sample size was determined through Cochran formula as 169 individuals. For data collection, a questionnaire was used. For determining the validity, the questionnaire was checked by supervising professor and management professors. Also for determining the reliability of questionnaires, the Cronbach's alpha was used. For data analysis, the Pearson's correlation and Regression tests are used. Results indicated that data mining and data mining techniques (neurological-systems, decision tree, rules inference, and summarizing, clustering and heterogeneity detection) are effective on the quality of information technology systems and e-learning in schools of Sardasht.

**Keywords:** *Data Mining, Decision Tree, Clustering, Electronic Learning*

### **INTRODUCTION**

Development of educational system in the era of information technology and communications is one of the main challenges of educational policy makers. As a result of intense impacts of information technology and its increasing penetration in different pillars of educational system, reviewing the educational system becomes one of the most important affairs to attend in order to create a coordinated and synchronized system for dynamic presence in field of information and guaranteed durability (Zarei, 2005). Electronic learning is a method for providing a more flexible learning, more opportunities for learners, facilitated progress follow up for learners and also an opportunity for creating new and effective learning environments (Khalifa and Razavi, 2013). Results of research conducted by Fallah *et al.*, indicated that e-learning has more impacts on cognitive, psychological, social and moral growth of students who have learned through this method compared to students who have not received this method of learning. Results of research conducted by Hung *et al.*, (2010) indicated that student's perspective towards knowledge in normal schools and smart schools are in significant differences.

On the other hand, with respect to the fact that schools have always been faced with several information regarding schools, students, teachers, personnel, materialistic sources and etc. and these information are most often susceptible of carrying valuable information and examples; school principals are forced to obtain information regarding data, how to achieve knowledge, strategies, policies and respective solutions. One of the apparatuses that is used by organizations who are faced with a huge amount of data is data mining. Data mining means extracting knowledge from huge amounts of data and is also considered as the most important step in the process of discovering knowledge. This apparatus simultaneously benefits from several scientific fields including databases, statistics, artificial intelligence, machine learning, neurological networks, pattern recognizing, knowledge based systems, knowledge, information recovery, high speed calculations and visual presentation of data. It is also known as the most efficient technique in discovering organizational beneficial knowledge (Ahmadvand & Akhoundzadeh 2011). Payandeh *et al.*, (2009) conducted a research titled as comparison of data mining techniques for better personalization in electronic learning. This research was concerned with predicting the educational yields of participants of electronic learning systems through machine learning in which the methods of decision tree, neurological systems, Biz algorithm and rational regression were used. Among these

## Research Article

methods, rational regression (80.69%) had the most precise predicting ability among applied methods. The methods of decision tree and simple biz also had a similar precision of 80.113% and in the following level, neurological network method with 79.261% of precision had the least prediction precision. Minaei *et al.*, (2013) discussed the realization of elements effective on educational fall through associational rules and cluster analyzing. In this research, authors had tried to deploy predicting data mining models for prediction of University student's educational status based on personal characteristics and educational history. With respect to statistical results of predicting models of student's status, the student's educational future can be known with high assurance based on previous data. Kahn *et al.*, (2008) used data mining techniques for predicting and avoiding social network felonies on the internet.

With respect to stated content and school's definite need for re-designing and with respect to the issue that currently no other element is able to change the redesign of educational structures and processes of schools except information technology, and also with further respect to important applications of information technology in schools, research in this contexts seems necessary. Since unlike most production and service providing organizations, schools have had a little share of new technologies, scientific and pragmatic activities in this context can be extremely beneficial. On this basis, the main purpose of this research is to discuss the improvement of quality of information technology and electronic learning through data mining.

## MATERIALS AND METHODS

### Methods

The present research is an applicable study conducted in a survey-correlation manner. The population of this research includes all teachers who are familiar with computers and data mining as 300 individuals. The sample size was determined as 169 individuals via Cochran's formula and respectively 169 questionnaires were distributed via simple random sampling. For data collection, a questionnaire was used which includes two segments. First segment is consisted of six questions for determining the level of implementing data mining and also in order to examine information technology and electronic learning, 11 items were arranged. Validity of questionnaire was approved by a supervising and some of management field's professors and in order to check the reliability of questionnaires, the Cronbach's alpha was used. The value of calculated Cronbach's alpha shows a desirable reliability of questionnaires. For the purpose of analyzing data, the tests of Kolmogorov-Smirnoff, Pearson's correlation and Regression were used.

**Table 1: Validity of questionnaires**

| Measures      | Cronbach's alpha | N  |
|---------------|------------------|----|
| Data mining   | 0.86             | 5  |
| Quality of IT | 0.91             | 10 |

## RESULTS AND DISCUSSION

### Results

#### Descriptive Findings

Demographic findings revealed that 85.7% of the respondents were male and 14.3% percent were female. 28.2% of the respondents held a semi B.A, 64% held a B.A and 7.8% held M.A and higher degrees. 36% had less than five years, 32.3% had 6-10 years, 17.8% had 11-1 years, 9.6% had 16-20 years and 4.3% had more than 20 years of experience.

#### Normality of Data Distribution

For determining the normality of data distribution the Kolmogorov-Smirnoff test was used and the results are shown in table 2. As you can see, the calculated significance level for all variables is more than 0.05; as a result they are not significant. With taking this issue into account, it can be said with 95% assurance level that data are distributed normally and the parametric test of Pearson's correlation can be implemented.

**Research Article**

**Table 2: Smirnov-Kolmogorov test results for normality of data**

| variables               | Data mining | IT and E-learning |
|-------------------------|-------------|-------------------|
| Kolmogorov-Smirnov test | 1.14        | 0.906             |
| Significance level      | 0.148       | 0.385             |

In this section, research hypotheses are tested

Main Hypothesis: Data mining has effects on the quality of IT systems and E-learning in schools of Sardasht

**Table 3: Person's correlation results for main hypothesis**

| Authentic sample | Significance | Quality of IT and E-learning | Data mining | sources                      |                       |
|------------------|--------------|------------------------------|-------------|------------------------------|-----------------------|
|                  | /000         | /51                          | 1           | Data mining                  | Pearson's correlation |
| 169              | /000         | 1                            | /51         | Quality of IT and E-learning |                       |

As you can see in table 3, the level of calculated significance is 0.000 which is significant at 0.05. With respect to this content, it can be claimed with 95% assurance that there is a statistic significant relation between the variables of data mining and quality of IT systems and E-learning. The explanation is that as the level of implementing data mining increases, the level of quality of IT systems and E-learning increases. The results of regression model are provided in the table. With respect to table 6, we can see that the regression coefficient is significant ( $P= 0.00 < 0.05$ ). The independent variable determines 26% of variance of dependent variable (quality of information technology systems and electronic learning). With considering these criteria, the main hypothesis is accepted.

**Table 6: Results of regression model for the main hypothesis**

|             | Value of R | Square R | Beta | Significance |
|-------------|------------|----------|------|--------------|
| Constant    |            |          |      | 0.000        |
| Data mining | 0.51       | 0.2601   | 0.51 | 0.000        |

**First Subsidiary Hypothesis:** the method of neurological networks is effective on the quality of information systems and electronic learning in schools of Sardasht.

**Table 7: Results of Pearson's correlation test for first subsidiary hypothesis**

| Authentic sample | Significance | Quality of IT and E-learning | Neurological systems | sources                      |                            |
|------------------|--------------|------------------------------|----------------------|------------------------------|----------------------------|
|                  | /000         | /40                          | 1                    | Neurological systems         | Pearson's correlation test |
| 169              | /000         | 1                            | /40                  | Quality of IT and E-learning |                            |

As you can see in table 7, the level of calculated significance is 0.000 which is significant at 0.05. With respect to this content, it can be claimed with 95% assurance that there is a statistic significant relation between the variables of neurological networks and quality of IT systems and E-learning. The explanation is that as the level of implementing neurological networks increases, the level of quality of IT systems and E-learning increases. The results of regression model are provided in the table. With respect to table 7, we can see that the regression coefficient is significant ( $P= 0.00 < 0.05$ ). The independent variable determines 16% of variance of dependent variable (quality of information technology systems and electronic learning). With considering these criteria, the first subsidiary hypothesis is accepted.

**Research Article**

**Table 8: Results of regression model for the first subsidiary hypothesis**

|                       | Value of R | Square R | Beta | Significance |
|-----------------------|------------|----------|------|--------------|
| Constant              |            |          |      | 0.000        |
| Neurological networks | 0.40       | 0.160    | 0.51 | 0.000        |

**Second Subsidiary Hypothesis:** the method of decision tree is effective on the quality of Information systems and electronic learning in schools of Sardasht.

**Table 9: Results of Pearson's correlation test for the second subsidiary hypothesis**

| Authentic sample | Significance | Quality of IT systems and E-learning | Decision tree | sources                              |                            |
|------------------|--------------|--------------------------------------|---------------|--------------------------------------|----------------------------|
|                  | /004         | /29                                  | 1             | Decision tree                        | Pearson's correlation test |
| 169              | /004         | 1                                    | /29           | Quality of IT systems and E-learning |                            |

As you can see in table 9, the level of calculated significance is 0.004 which is significant at 0.05. With respect to this content, it can be claimed with 95% assurance that there is a statistic significant relation between the variables of decision trees and quality of IT systems and E-learning. The explanation is that as the level of implementing Decision trees increases, the level of quality of IT systems and E-learning increases. The results of regression model are provided in the table. With respect to table 9, we can see that the regression coefficient is significant ( $P= 0.004 < 0.05$ ). The independent variable determines 8.41% of variance of dependent variable (quality of information technology systems and electronic learning). With considering these criteria, the second subsidiary hypothesis is accepted.

**Table 10: Results of regression model for the second subsidiary hypothesis**

|               | Value of R | Square R | Beta | Significance |
|---------------|------------|----------|------|--------------|
| Constant      |            |          |      | 0.000        |
| Decision tree | 0.29       | 0.841    | 0.29 | 0.004        |

**Third Subsidiary Hypothesis:** inference of rules is effective on the quality of Information technology systems and electronic learning in schools if Sardasht.

**Table 11: Results of Pearson's correlation test for the third subsidiary hypothesis**

| Authentic sample | significance | Quality of IT systems and e-learning | Inference of rules | sources                              |                            |
|------------------|--------------|--------------------------------------|--------------------|--------------------------------------|----------------------------|
|                  | /001         | /63                                  | 1                  | Inference of rules                   | Pearson's correlation test |
| 169              | /001         | 1                                    | /63                | Quality of IT systems and e-learning |                            |

As you can see in table 11, the level of calculated significance is 0.01 which is significant at 0.05. With respect to this content, it can be claimed with 95% assurance that there is a statistic significant relation between the variables of inference of rules and quality of IT systems and E-learning. The explanation is that as the level of implementing inference of rules increases, the level of quality of IT systems and E-learning increases. The results of regression model are provided in the table. With respect to table 11, we can see that the regression coefficient is significant ( $P= 0.01 < 0.05$ ). The independent variable determines

**Research Article**

39.66% of variance of dependent variable (quality of information technology systems and electronic learning). With considering these criteria, the third subsidiary hypothesis is accepted.

**Table 12: Results of regression model for the third subsidiary hypothesis**

|                    | Value of R | Square R | Beta | Significance |
|--------------------|------------|----------|------|--------------|
| Constant           | 0.63       | 0.396    |      | 0.000        |
| Inference of rules |            |          | 0.63 | 0.001        |

**Fourth Subsidiary Hypothesis:** the method of summarizing is effective on the quality of information systems and e-learning in schools of Sardasht.

**Table 13: Results of Pearson's correlation test for fourth subsidiary hypothesis**

| Authentic sample | Significance | Quality of IT systems and e-learning | Method of summarizing | sources                              |                            |
|------------------|--------------|--------------------------------------|-----------------------|--------------------------------------|----------------------------|
|                  | /000         | /43                                  | 1                     | Method of summarizing                | Pearson's correlation test |
| 169              | /000         | 1                                    | /43                   | Quality of IT systems and e-learning |                            |

As you can see in table 13, the level of calculated significance is 0.000 which is significant at 0.05. With respect to this content, it can be claimed with 95% assurance that there is a statistic significant relation between the variables of summarizing method and quality of IT systems and E-learning. The explanation is that as the level of implementing summarizing method increases, the level of quality of IT systems and E-learning increases. The results of regression model are provided in the table. With respect to table 13, we can see that the regression coefficient is significant ( $P= 0.000 < 0.05$ ). The independent variable determines 18.4% of variance of dependent variable (quality of information technology systems and electronic learning). With considering these criteria, the fourth subsidiary hypothesis is accepted.

**Table 14: Results of regression model for the fourth subsidiary hypothesis**

|                    | Value of R | Square R | Beta | Significance |
|--------------------|------------|----------|------|--------------|
| Constant           |            |          |      | 0.000        |
| Summarizing method | 0.43       | 0.184    | 0.43 | 0.000        |

**Fifth Subsidiary Hypothesis:** The clustering method is effective on the quality of information technology systems and e-learning in schools of Sardasht.

**Table 15: Results of Pearson's correlation test for the fifth subsidiary hypothesis.**

| Authentic sample | Significance | Quality of IT systems and e-learning | Clustering method | sources                              |                            |
|------------------|--------------|--------------------------------------|-------------------|--------------------------------------|----------------------------|
|                  | /000         | /49                                  | 1                 | Clustering method                    | Pearson's correlation test |
| 169              | /000         | 1                                    | /49               | Quality of IT systems and e-learning |                            |

As you can see in table 15, the level of calculated significance is 0.000 which is significant at 0.05. With respect to this content, it can be claimed with 95% assurance that there is a statistic significant relation

**Research Article**

between the variables of clustering method and quality of IT systems and E-learning. The explanation is that as the level of implementing clustering method increases, the level of quality of IT systems and E-learning increases. The results of regression model are provided in the table. With respect to table 15, we can see that the regression coefficient is significant ( $P= 0.000 < 0.05$ ). The independent variable determines 24% of variance of dependent variable (quality of information technology systems and electronic learning). With considering these criteria, the fifth subsidiary hypothesis is accepted.

**Table 16: Results of regression model for the fifth subsidiary hypothesis**

|                    | Value of R | Square R | Beta | Significance |
|--------------------|------------|----------|------|--------------|
| Constant           |            |          |      | 0.000        |
| Summarizing method | 0.49       | 0.240    | 0.49 | 0.000        |

**Sixth Subsidiary Hypothesis:** the method of identification of heterogeneity is effective on the quality of information technology systems and e-learning in schools of Sardasht.

**Table 17: Results of Pearson's correlation test for sixth subsidiary hypothesis**

| Authentic sample | significance | Quality of It systems and e-learning | Heterogeneity identification method | sources                              |                            |
|------------------|--------------|--------------------------------------|-------------------------------------|--------------------------------------|----------------------------|
| 169              | /000         | /35                                  | 1                                   | Heterogeneity identification method  | Pearson's correlation test |
|                  | /000         | 1                                    | /35                                 | Quality of It systems and e-learning |                            |

As you can see in table 17, the level of calculated significance is 0.000 which is significant at 0.05. With respect to this content, it can be claimed with 95% assurance that there is a statistic significant relation between the variables of Heterogeneity identification method and quality of IT systems and E-learning. The explanation is that as the level of implementing Heterogeneity identification method increases, the level of quality of IT systems and E-learning increases. The results of regression model are provided in the table. With respect to table 17, we can see that the regression coefficient is significant ( $P= 0.000 < 0.05$ ). The independent variable determines 16% of variance of dependent variable (quality of information technology systems and electronic learning). With considering these criteria, the sixth subsidiary hypothesis is accepted.

**Table 18: Results of regression model for sixth subsidiary hypothesis**

|                                     | Value of R | Square R | Beta | Significance |
|-------------------------------------|------------|----------|------|--------------|
| Constant                            |            |          |      | 0.000        |
| Heterogeneity identification method | 0.35       | 0.1225   | 0.35 | 0.000        |

**Discussion and Conclusions**

Nowadays, implementing traditional methods of collecting and analyzing data is not appropriate since they are time consuming and extremely costly. On this basis, implementing new methods of data analyzing such as data mining seems very necessary. During the recent years, there have been several researches in the context of implementing the process of data mining in the field of education. This new research field is called educational data mining which is concerned with development of knowledge

### **Research Article**

exploring methods from data of educational environments, especially university students (Romero and Ventura, 2007). Results of data analyses indicated that there is a significant and positive relation among data mining and quality of information systems and electronic learning. Also it was revealed that data mining has a significant effect on quality of information systems and electronic learning. It means that, corresponding to the level of implemented data mining technique in schools, the quality of information systems and educational systems increase as well. Gathered data about students can be personal or educational which are both collected through data bases established in schools. Also these types of data are accessible through electronic educational systems. With implementing data mining techniques on educational information, beneficial data can be extracted from them. This knowledge can also be implemented in order to understand student's behaviors, aiding the process of education, evaluation and optimization of electronic learning systems, improving curriculums, improving efficiency and improving capability of students as well as other purposes. With respect to the results of this research, it may be concluded that related authorities and subordinates should be concerned with application of data mining in schools and in this sense, optimize the quality of information systems and ultimately lead to student's progression and prosperity.

Results of data analyses revealed that methods of neurological systems, decision tree, rule's inference, summarizing, clustering and the method of identifying heterogeneity are effective on the quality of information technology systems and electronic learning. As a result, we can conclude that authorities and related subordinates must take into account the application of methods of neurological systems, decision tree, rule's inference, summarizing, clustering and the method of identifying heterogeneities in schools and in this sense improve the quality of information technology systems and electronic learning and ultimately lead to student's progression and prosperity.

### **REFERENCES**

- Agricultural Zavarak and Babol Ismail (2008).** Measurement and evaluation of e-learning Education. *Higher Education* **3** 88-73.
- Ahmadvand Ali and Mohammad Akhoundzadeh (2010).** Application of data mining techniques in modeling framework crime. *Police Journal of Human Development* **7**(30) 22-12.
- Epstein BS, Myrafzl Samia and Hani Hassan (2012).** Identification of factors affecting academic failure using association rules and clustering analysis. *Data Mining Conference, Tehran*.
- Khalifa G and Rizvi Syed A (2012).** Evaluation and quality assurance in e-learning model to analyze the importance – performance. *Journal of Media* **3**(1) 44-33.
- Khan JI and Shaikh SS (2008).** Computing in Social Networks with Relationship Algebra. *Journal of Network and Computer Applications* **31** 862–878.
- Lasting Curls H, Sydrzy Faith, Thereby M and Farah A (2008).** Comparison of data mining techniques for better personalization in e-Learning. *National Conference of Computer Engineering, Electrical and Information Technology, Sama, Hamadan*.
- Ong ET, Foo L-K and Lee SM (2010).** Smart schooling and its impact on students' attitudes towards science. *World Conference on Innovations in Science and Technology Education, University of Tartu, Estonia*.
- Romero C and Ventura S (2007).** Educational data mining: A survey from 1995 to 2005. *Expert Systems with Applications* 135–146.