A Data Mining Approach To Evaluate The Cognizance Of Teaching Faculty In Kingdom Of Saudi Arabia

Hisham M. Alsaghier

College of Computer and Information Sciences, Majmaah University, Saudi Arabia

Abstract
This paper presents an approach to evaluate the perception of teaching faculty using data mining techniques. A real data set of students' database is processed to examine hidden relationship between the attributes in the course registration database. The results showed that data mining techniques can be an effective tool to predict the cognizance of teaching faculty when used with database of student information systems.

Keywords :
Data Mining, Educational Systems, Classification-Based Techniques, Faculty Perception

1. Introduction

The role of faculty in higher educational institutions are very vital and is important in deciding the future of the country. Evaluating the cognizance of teaching faculty members in higher educational institutions in Kingdom of Saudi Arabia is a challenging task as it involves more complex and deeper analysis. Education systems such as Students Information Systems and Learning Management Systems include a big data of students' records, course registration criteria, curriculum plans, course assessment results, and academic performance. This data can be considered as a treasure if considered to analyze various trend, phenomena and to support the academic strategic decision. Data mining techniques can be used to analyze the large set of data of education systems. There by, this paper will review the literature of the present approached to evaluate the cognizance of teaching faculty in Saudi higher education system by considering different aspects. This paper is an exclusive attempt in apply suitable data mining techniques in the database maintained by the Universities and exploring the hidden relationships between the attributes in the course registration database.

Review of Literature

In order to analyze the present context, the review of literature has been carried in the aspects such as evaluation of teaching effectiveness by the student, faculty members perception of e-learning in higher education system in Kingdom of Saudi Arabia, Saudi higher education reality and prospects etc.,

Alshahrari et al. (Alsharari, Ratyan, & Mohammad, 2016) stated that there is no significant statistical relationship between current realty of university careers and the three axes within all examined variables, but there is an impact of job title for deans only with university teaching and scientific research, but no impact in community service. This paper talks only about the administrators and not the faculty members. The role of faculty members or students are not addressed at any level in this paper.

Al-Kuwaiti (Al-Kuwaiti, 2014) discusses the role of students in evaluating the teaching effectiveness in the Kingdom of Saudi Arabia. The author carried out a study on seven medical colleges Government universities in the Kingdom of Saudi Arabia. The results show that a group of randomly drawn final year students and a group of their teaching faculties were studied. There were statistically significant differences between instructors' and students' perception of SETE. Whereas, students registered disapproval in three of the four areas studied, the pattern of instructors' response was a mirror image of the students. No attempt is being made to evaluate the cognizance of teaching faculty as proposed in this paper.

Kaur et al., (Kaur, Singh, & Josan, 2015) adopted data mining to investigate the educational institutes in school databases. Their research aimed to identify the slow learners among students using data mining approach, i.e. classification-based algorithm. Using WEKA and Open Source Tool, a real data set from a high school is examined to filter the protentional variables. The results showed the importance and effectiveness of classification-based data model in the context of education.

Srivatsava and Sirvastava (Srivastava & Sirvastava, 2016) present the application of different data mining tools and techniques which proven to be effective when used to predict the students' performance and address their academic profiles. The author discussed the advantage's and techniques of using data mining in educational institutional databases in general.

Bhullar (Bhullar, 2012) reviewed the advantages, differences, and classification of data mining classification mode in order to understand students' data sets using WEKA data mining tool. This paper uses J48 algorithm to predict the result of the student.

Ali (Ali1, 2013) examines the role of data mining in an education sector. The author lays emphasis on application

Manuscript received November 5, 2018
Manuscript revised November 20, 2018

of data mining that contribute to offer competitive courses and improve their business. The paper also highlights the benefits of data mining.

Priyadarshini and Ray (Priyadarshini & Ray, 2017) presents the use of data mining techniques to enhance the process performance in higher education systems. Different types of clustering, classification and association techniques are used in this paper. Using these techniques has shown enhancement of the student performance, their life process management, selection of courses, to measure their reservation rate and allow the fund management of the organization.

Sharma and Harmanjit (Sharma & Harmanjit, 2013) investigated various effective applications of data mining on education systems. The paper showed how data mining can assist in improvement of education system by enabling better understanding of the students. The extra information can give a great help for the teachers to effectively manage their classes and to give a proactive feedback to the students.

Johina and Kamra (Johina & Kamra, 2015) represent the data mining technique like classification and clustering using educational dataset to manage the information of education system. This analysis will be of good use to the students to take effective decisions.

Agarwa et al., (Agarwa, Pandey, & Tiwari, 2012) presents the decision tree approach which is considered as an important base of student selection for any course program. The paper developed a faith on Data Mining techniques that can be used as a strategic management tool in education and business system

It is very clear from the above reviewed literature that much research is not being carried out to evaluate the cognizance of teaching faculty by applying data mining as proposed in this research work.

R language is a statistical language for generating graphs and calculating statistical quantities. Later it is added with the packages of machine learning. R is an open source language available under GNU General Public License.

3. Need of using R-tool for knowledge discovery

R is an open source language and environment that is used to perform statistical and graphics computing. Similar to the S language and environment, R is a GNU project developed by John Chambers and Collogues at Bell Laboratories (formerly AT&T, now Lucent Technologies). R can be considered as a different implementation of S. Though there are some significant differences, but most of the code written for S runs unaltered under R.

R offers a various variety of statistical (linear and nonlinear modelling, classical statistical tests, time-series analysis, classification, clustering, …) and graphical techniques, and is highly scalable. The S language is considered as the vehicle of choice for research in statistical methodology, and R provides an Open Source route to participation in that activity (R-Project, 2017).

Some of the machine language algorithms which are implemented in R are:

a) Regression
    i) Linear regression
        - Least Square regression
        - Principal Component regression
        - Partial Least square regression
    ii) Non-linear regression
        - Support Vector Machine
        - k-Nearest Neighbor
        - Neural Network

b) Classification
    i) Linear classification
        - Logistic regression
        - Linear discriminant analysis
    ii) Non-linear classification
        - Neural network
        - Support Vector Machine
        - k-Nearest Neighbor
        - Naïve Bayes
    iii) Decision trees for classification
        - Classification and Regression trees
        - Random forest
        - PART

R language can be used for data analytics. Hotel booking prediction

4. Objective and Scope of this Paper

The objective of the research is to use machine learning approaches for evaluating the cognizance of teaching faculty in the context of Kingdom of Saudi Arabia. This research work uses R as a tool for evaluation. Various features in R is used to extract peculiar patterns from the chosen dataset.

5. Dataset

For the defined objective, the research work uses the following database where all the previous course details are stored as semantic data sources. R language is used for implementation.

Table 5a: Student Details

S.No	Field Name	Data Type	Size	Description
1	Reg No	Text	Varchar 2(10)	Student Registration Number
2	Specialization	Text	Varchar 2(25)	Subject Specialization
3	Name	Text	Varchar 2(30)	Name of the Student
4	House No	Number	Varchar 2(5)	House Number
5	Street No	Number	Varchar 2(5)	Street Number
6	Way No	Text	Number (10)	Sub Street Number
7	Place	Text	Varchar 2(25)	Name of the Place
8	Wilayah	Text	Varchar 2(25)	Name of the Location
9	Manthaka	Text	Varchar 2(25)	Name of the city
10	Country	Text	Varchar 2(25)	Name of the Country
11	E-mail	Text	Varchar 2(30)	Student E-mail Address
12	Telephone	Number	Number (20)	Student Phone Number

Table 5b: Internal and External Assessment

S.No	Field Name	Data Type	Size	Description
1	Semester	Text	Varchar 2(15)	Name of the Semester
2	Year	Number	Number (4)	Academic Year
3	Reg No	Text	Varchar 2(10)	Student Register Number
4	Module Code	Text	Varchar 2(20)	Subject Code
5	Assessment parameter	Text	Varchar 2(30)	Assessment Parameter
6	Max Mark	Number	Number (5)	Maximum Mark for each Subject
7	Marks Scored	Number	Number (5)	Mark Secured
8	Percentage	Number	Number (5)	Total Percentage
9	Group	Number	Number (5)	Student Group

Table 5c: Exam Summary

S.No	Field Name	Data Type	Size	Description
1	Semester	Text	Varchar 2(15)	Name of the Semester
2	Year	Number	Number (4)	Academic Year
3	Module Code	Text	Varchar 2(20)	Subject Code
4	Tutor	Text	Varchar 2(20)	Name of the Tutor
5	Session	Text	Varchar 2(2)	Exam Session
6	Avg_Internal	Number	Number 2(5)	Average Internal Marks
7	Avg_External	Number	Number 2(5)	Average External Marks
8	Avg_Total	Number	Number 2(5)	Average Total
9	Pass_Percentage	Number	Number 2(5)	Student Pass Percentage

6. Role of Students

Students are the prime observers for evaluation of their cognizance of the teaching faculty. Ultimately end results carry major weightage in overall teaching learning process. It is obvious that the pass percentage will give an idea of how far the students understood the subject from the faculty and their level of confidence in the subject. Thus, we focused on the instructor performance in a specific module over the years. We also observed the variation of pass percentage over several years in a specific module.

7. Role of the Department/College

Any department/college focuses on teaching learning process and the quality of teaching and the cognizance of the subject from the students' point of view. The department should take the responsibility of continuously motivating the faculty members to see that the decision taken by the system is being implemented.

8. Case Analysis

The database discussed in section 6 is analyzed with R.

```
$ Sem          : chr  "Fall" "Fall" "Fall" "Fall" ...
$ Year         : int  2010 2010 2009 2009 2011 2006 2010 2011 2007 2010
$ Modulecode   : chr  "ELEC 0001" "ELEC 0001" "PROJ 0002" "DSNT 0301" ...
$ Tutor        : chr  "Muhammad Bashir" "Muna Al Rahbi" "Rosalio Pana" "Michelle D Souza" ...
$ Session      : chr  "Session-D" "Session-D" "Session-A" "Session-A" ...
$ Avg_Internal : chr  "12.266666" "12.266666" NA "22.25" ...
$ Avg_External : chr  "5.216666" "5.216666" "20.433333" NA ...
$ Avg_tot      : int  17 17 20 22 22 23 24 24 25 25 ...
$ Pass_percentage: num  0 0 33.3 0 46.7 ...
```

Database Snapshot for the Faculty Reem for the Module COMP0003

```
>head(df2)
  Sem Year Modulecode          Tutor Session Avg_Internal Avg_External
1: Fall 2006  COMP 0003 Reem Mohammed Al Siyabi Session-A  35.596153  13.307692
2: Spring 2007  COMP 0003 Reem Mohammed Al Siyabi Session-C  31.444444  19.333333
3: Fall 2006  COMP 0003 Reem Mohammed Al Siyabi Session-K  39.319444  15.111111
4: Fall 2006  COMP 0003 Reem Mohammed Al Siyabi Session-J  38.847222  18.333333
5: Fall 2007  COMP 0003 Reem Mohammed Al Siyabi Session-L  31.424242  26.212121
```

6: Spring 2009 COMP 0003 Reem Mohammed Al Siyabi Session-F 31.3 27.133333

Avg_tot Pass_percentage

1: 48 50.00000
2: 51 58.33333
3: 54 55.55556
4: 57 69.44444
5: 57 60.60606
6: 58 56.66667

Fig 1 shows that the faculty Reem was taking the module COMP0003 from 2006 onwrds. And the pass percentage was varying based on the sessions. For the same instructr the results are varying based on the level of students in that session.

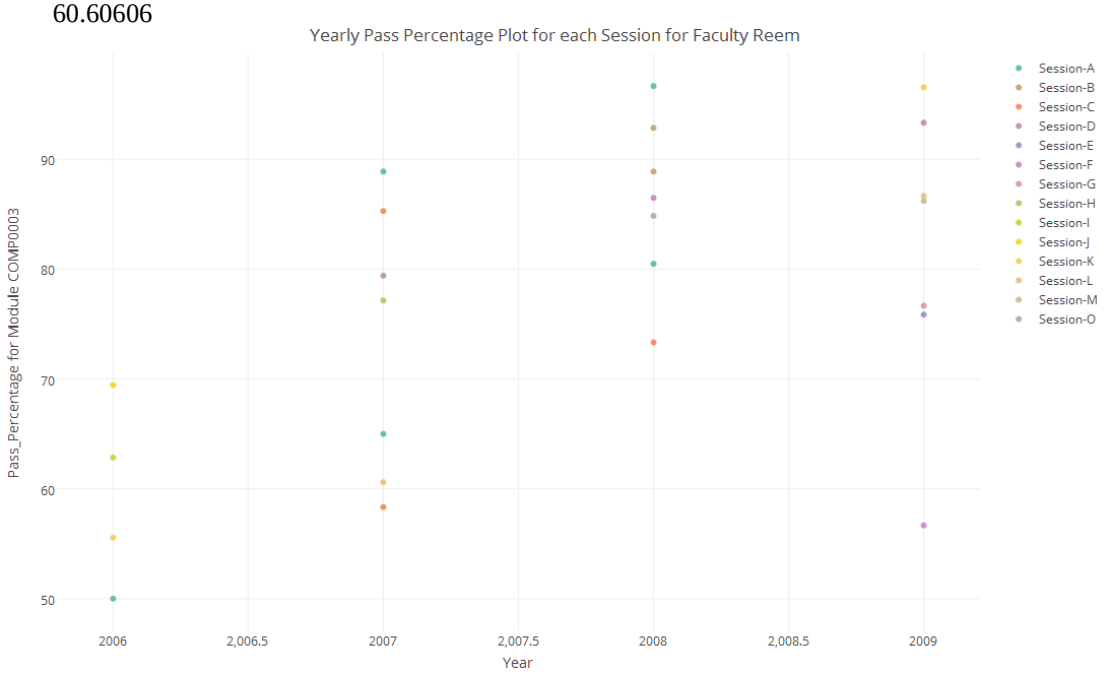

Fig. 1

Average external score for Module Comp0003 increased over the years irrespective of the instructors. In 2011 the maximum Avg_external score is 34.62. In 2005 the maximum Avg-external score is 21.00. Figure 2 shows this relationship.

Fig. 2

9. Conclusion

Data Mining could be used to improve business intelligence process including education system to enhance the efficacy and overall efficiency by optimally utilizing the resources available. This paper primarily focuses on evaluating the cognizance of teaching faculty members on their history of handling the courses. The discovered knowledge will be of good use to the department / college for successful implementation.

References

[1] Agarwa, S., Pandey, G. N., & Tiwari, M. D. (2012). Data Mining in Education: Data Classification and Decision Tree Approach. International Journal of e-Education, e-Business, e-Management and e-Learning, 140-144.

[2] Ali1, D. M. (2013). Role of Data Mining in Education Sector. International Journal of Computer Science and Mobile Computing, 374-383.

[3] Al-Kuwaiti, A. A. (2014). Students evaluating teaching effectiveness process in saudi arabian medical colleges: A comparative study of students' and faculty members perception. Saudi Journal of Medcine and Medical Sciences, 166-172.

[4] Alsharari, M., Ratyan, M., & Mohammad, A. A. (2016). Saudi Higher Education Reality and Prospects: Evaluating Careers' Dimensions of University Teaching, Scientific Research and Community Service Northern Border University as a Model. Indian Journal of Science and Technology, Vol 9(4), 1-16.

[5] Bhullar, M. S. (2012). Use of Data Mining in Education Sector. Proceedings of the World Congress on Engineering and Computer Science (pp. 1-4). San Francisco, USA: WCECS.

[6] Johina, & Kamra, V. (2015). A Review: Data Mining Technique Used In Education Sector. International Journal of Computer Science and Information Technologies, 2928-2930.

[7] Kaur, P., Singh, M., & Josan, G. S. (2015). Classification and prediction based data mining algorithms to predict slow learners in education sector. 3rd International Conference on Recent Trends in Computing 2015(ICRTC-2015) (pp. 500-508). Elsevier.

[8] Priyadarshini, N., & Ray, M. (2017). A Review: Data Mining Techniques in Education Acadmia. SEI.

[9] R-Project. (2017, August 16). What is R? Retrieved from https://www.r-project.org: https://www.r-project.org/about.html

[10] Sharma, R., & Harmanjit, S. (2013). Data Mining in Education Sector. International Journal of Electronics & Data Communication, 2(2), 4-8.

[11] Srivastava, J., & Sirvastava, A. (2016). Data Mining in Education Sector: A Review. Special Conference Issue: National Conference on Cloud Computing & Big Data, (pp. 184-190).