Investigating Students' Temporal Behaviors and Evaluation of Faculty through Educational Data Mining

Fatemeh Orooji  
Dept. of Electrical & Computer Engineering  
University of Tehran  
Tehran, Iran  
forooji@ut.ac.ir

Fattaneh Taghiyareh  
Dept. of Electrical & Computer Engineering  
University of Tehran  
Tehran, Iran  
ftaghiyareh@ut.ac.ir

Received: October 5, 2010- Accepted: November 27, 2010

Abstract—Educational data mining (EDM) extracts implicit and interesting patterns from large data collections to provide a more effective learning environment. Introducing EDM concepts and techniques, this paper aims to discover existing behavioral pattern of students in course selection and faculty evaluation as well as educators policies in grading. This study has been carried out to determine the correlation between Student Evaluation of Teachers (SET) ratings and their gained scores in University of Tehran, department of Information Technology engineering. Dividing students based on their grades, a weak direct relationship has been demonstrated among weak and good students (0.086, 0.108) meaning that the more students’ grades were, the higher teacher evaluation scores were observed. Insufficient students’ awareness of SET importance and the existence inappropriate questions beyond students’ knowledge in the questionnaire may cause these results. Recognizing effective factors in SET ratings can reveal the strengths and weaknesses of this kind of faculty evaluation and provide the possibility for better planning and obtaining authentic results.

Key words- educational data mining, student behavior, teacher evaluation, sequence pattern

I. INTRODUCTION

Nowadays, most of educational institutions apply automated web-based mechanisms for students’ academic activities such as course enrolments and faculty evaluation. These systems can record all students and teachers activities accumulating large data collections, which need utilizing data mining techniques in order to extract implicit and interesting patterns automatically. According to [1] educational data mining (EDM) may be defined as “the process of converting raw data from educational systems to useful information that can be used to inform design decisions and answer research questions”. EDM is an emerging interdisciplinary research area facilitates some educational user-oriented applications[2]. EDM normally begins with data preprocessing, then applying appropriate data mining techniques, followed by evaluating discovered patterns[3]. Automatic data clustering and classification, discovery of associations and sequence patterns between data, and identification of outliers may be enumerated as well known approaches in EDM[4].

Due to the significant effect of students evaluations on instructors promotion process, some recent research are focused on exploring affective factors such as course, student, and instructor characteristics
In this respect, there are a lot of investigations indicating that Student Evaluation of Teacher (SET) ratings are affected by students' expected grades which enables faculty to 'buy' higher SET ratings through lowering their grading standards[17]. Despite the vast amount of researches in most of the countries, there are a negligible number of studies reporting the situation of Iranian academic centers through utilizing EDM techniques. This paper studies the correlation between SET ratings and students' grades in the University of Tehran, department of Information Technology Engineering. Contrary to the literature there was no significant relation among all students but clustering students based on their grades, has revealed some weak direct correlation between students' final grades and SET ratings. In addition, overall picture of students grades and evaluation rating as well as educators grading policies have provided an understandable depiction of students and educators behaviors. Research findings can be applied in recommender systems in order to provide some more flexible educational maps in adaptation to learners' characteristics and capabilities. In addition, results can be considered as important evidence implying the vital demand for essential verification and possibly essential change in evaluation methods and questionnaire.

The paper is organized as follows. Section 2 overviews the most relevant surveys carried out in EDM to date. Introducing EDM concepts, it describes EDM applications for different groups of users, recommended steps, and some of most common techniques. It then goes on the description of faculty evaluation reviewing the affecting parameters, in section 3. Section 4 is devoted to the research information and goals. Section 5 explains the research findings complemented by a detailed discussion in section 6. Then, we have concluded the paper and outlined areas for future research.

II. EDUCATIONAL DATA MINING

Recently web-based educational systems are being utilized more and more by universities and schools to supplement their conventional teaching environments where neither students nor teachers are bound to a specific location or time. They can collect large amounts of daily-generated information about teaching–learning interactions recorded in their own databases or related log files. These large data collections require data mining techniques to be applied in order to discover desired knowledge automatically[8]. Mining educational data known as EDM may be considered as a valuable mean in order to understand learners' activities and learning process as a whole. EDM can provide computational approaches to analyze educational data in order to study educational questions[4]. Using advanced techniques, EDM tries to extract interesting and potentially useful patterns from a set of large data sources that have been cleaned and integrated before. EDM techniques are oriented to the users who will use their discovered knowledge to comprise students, educators and administrators. These advanced techniques include characterization and summarization of data, the discovery of associations and correlation between data, automatic classification of data, clustering, discovery of discriminated features, identification of outliers, etc. Also EDM needs the complementary steps consists of the evaluation of discovered patterns and possibly using interactive visualization.

A. EDM Applications

Application of educational data mining can be represent as an iterative cycle, which processes and refines gathered educational data (see Fig. 1). The discovered knowledge may be entered the data mining loop again as an informative resource in order to improve the e-learning process.

There is a lot of information about all system educational resources, activities, and users' interactions. Different data mining techniques can be applied in order to reveal useful knowledge, which will be used, by both students and providers. Indeed EDM provided a variety of applications due to different actors with each particular point of view described in Table 1[9,10].

<table>
<thead>
<tr>
<th>EDM users</th>
<th>EDM applications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Students</td>
<td>Suggest activities, resources and learning tasks based on student capabilities and characteristics, tasks already done by the learner and his successes, and on tasks made by other similar learners in order to develop an effective customized education system</td>
</tr>
<tr>
<td>Educators</td>
<td>Provide more objective feedback for instruction reporting the course content and effectiveness on the learning process, the efficiency of different learning activities, the most frequently made mistakes and students who need more guidance and monitoring</td>
</tr>
<tr>
<td>Administrators and academics</td>
<td>Give a more detailed view of the system situations and user interactions interpreting how to better organize institutional resources (human and material) and their educational offer</td>
</tr>
<tr>
<td>responsible</td>
<td>Determine the effectiveness of the new approach offered in computer mediated distance learning</td>
</tr>
</tbody>
</table>

Figure 1. The cycle of applying data mining in educational systems[9].

Table 1, some of EDM user-oriented applications
B. EDM Steps

Generally an EDM, like any other data mining process[8], has the steps of first gathering and preprocessing data, then applying the appropriate data mining tasks and at last post-processing as follows[11]:

- **Data gathering.** Current LMSs mostly use their own relational databases that store all user profiles, their interaction data, and academic results, instead using needed textual log files. LMSs also keeps transactional log files which record all details about users activities (clicks, elapsed time, etc.).

- **Data pre-processing.** In contrast to the typically observational datasets in data mining, EDM data preprocessing is simpler since it uses structural data stored in databases for analysis purposes while typically observational datasets were generated to support the operational setting rather than analysis. LMS database keeps data usually very clean and the values are correct, normally contain no noise from measuring devices as it is in the other fields of data mining.

- **Applying the mining algorithms.** In this step it is necessary to choose some useful data mining algorithm and implementation, set their special parameters and other constraints such as minimum support and confidence, specify their output files and what specific attributes can be present as resultant of the discovered rules. Figure 2 shows the percentage of using different techniques in a sample country along one year.

- **Data post-processing.** One of the most important phases in EDM is evaluating and interpreting the obtained rules and representing them understandable in order to put the results into use for further actions. Discovered information may be used by the teachers for making decisions about the course activities and the students to improve the learning processes.

![Image](image-url)

Figure 2. The poll of Data mining analytic methods used frequently in a specific duration [12]

It may be useful to mention that the size of educational datasets is normally small in comparison to other similar fields such as ecommerce applications. EDM may deals with users less than 1000 students while traditional DM may involve thousands of clients. Of course, the amount of users will be increased if the purpose of EDM demands to pool data from several years or from several similar courses. On the other hand, each student may interact in each individual process, which means the total number of instances or transactions can be quite large depending on the amount of information the LMS stores about the interactions and the level of granularity.

C. User Modeling through EDM

Generally, there are two learner modeling approach as follows[13]:

- **Collaborative:** In this approach, users explicitly provide information about themselves via filling out a questionnaire. Although this approach may provide some more reliable information, most of the learners do not like spending a considerable amount of time to fill corresponding questions.

- **Automatic:** In this approach, the system implicitly gathers any information needed while users are working/learning in the system. Class discussion/student reflection, instructor observations, cognitive sciences or from artificial intelligence are all utilized in this approach, no need to ask implicit questions. However, there are some problems in information accuracy provided by this direct approach, which make it a necessary to get as much reliable information as possible to build a more robust student model.

EDM techniques may be considered as some valuables tools, which can gather precious information about each user and groups of users constructing an up to date automatic user profiles. Besides, it reports some group behavioral features in order to provide better understanding of system performance. In this respect, each user behaviors may be compared to its reference group regarding predefined standards, too.

III. STUDENT EVALUATION OF TEACHER (SET)

It is mostly accepted that students can be considered as the immediate consumers of any educational systems whose satisfaction must have a significant role in any academic activities. In this respect, students are often demanded to evaluate their educators through some predefined questionnaire complemented by some open-end questions. Student ratings information can be used by educators to know the effectiveness of their teaching strategies and to make some needed changes as well as by administrators to make personnel and program decisions.

In spite of the support for the validity, reliability, and usefulness of student ratings provided by researchers, faculty often have questions about the influence of some other factors beyond their teaching performance on the ratings students give. Responding to those questions, a considerable number of researches have investigated the relationship between student SET ratings and a number of variables. [5] has determined
related and non-related factors, summarized in Table 2, comprising some course, student, and instructor characteristics.

There is also some other variables taking into account in examining whether they relate to SET ratings, such as students' personality traits [14] and the instructors' ability to participate in a highly interactive world [6]. It may be beneficial to mention that some researches demonstrate the opposite aspects of the relation between SET rating and some variables. [15] concluded that if learning is measured more objectively (rather than by just a single grade), it will less likely be related to the evaluations.

Even utilizing different methods to analyze SET ratings can lead to different conclusions [16]. Using EDM various techniques [6] has examined students' rating of instructors explaining the suitability of categorization and regression tests the evolving nature of higher education.

<table>
<thead>
<tr>
<th>Factors related</th>
<th>Factors not related</th>
</tr>
</thead>
<tbody>
<tr>
<td>Course characteristics</td>
<td>Class size</td>
</tr>
<tr>
<td></td>
<td>Discipline</td>
</tr>
<tr>
<td></td>
<td>Reason for Taking Course</td>
</tr>
<tr>
<td></td>
<td>Course Level</td>
</tr>
<tr>
<td></td>
<td>Difficulty Level of Class</td>
</tr>
<tr>
<td>Student characteristics</td>
<td>Expected Grade</td>
</tr>
<tr>
<td></td>
<td>Motivation</td>
</tr>
<tr>
<td></td>
<td>Major</td>
</tr>
<tr>
<td></td>
<td>Gender</td>
</tr>
<tr>
<td>Instructor characteristics</td>
<td>Faculty Rank</td>
</tr>
<tr>
<td></td>
<td>Personality</td>
</tr>
<tr>
<td></td>
<td>Research productivity</td>
</tr>
<tr>
<td>Administration of rating</td>
<td>Instructor's Presence in Room</td>
</tr>
<tr>
<td></td>
<td>Time of Administration</td>
</tr>
<tr>
<td></td>
<td>Student Anonymity</td>
</tr>
</tbody>
</table>

A. Relation between SET ratings and Grades

One of the most interesting items in this literature is the relationship between student SET ratings and expected grades. For more than 80 years researchers have investigating the relationship between student evaluation of teaching (SET) ratings and relative grades and over 1,300 articles had been published trying to explain how students evaluate faculty [17]. This vast quantity of research reflects that student evaluation of teaching (SET) ratings are influenced by expected grades, revealing the fact that faculty are able to 'buy' higher SET ratings by giving higher grades and lowering grading standards [7]. 'Higher expected grades do lead to significantly better SET scores among both principles and upper division [economics] classes' [18]. Also it is shown that 'individual students reward teachers with higher evaluations as both their own grade and the grades of their peers increase' [19].

Besides it is revealed that students' midterm grades are positively and significantly related to educators' evaluations [20]. Furthermore it is discovered that students who have a poor idea of their own understanding evaluate their teacher weak [21].

However [22] indicates that instead viewing the relationship between grades and student satisfaction simply as evidence of contamination due to grading leniency, it is better to see it as a result of important causal relationships among other variables.

Investigating the relation among Iranian students, [23] has reported a negative weak relationship demonstrated between SET and students' grade average points (GPA) ($r = -0.091$) which means that students who have higher GPA had evaluated the teacher lower. Regarding different level of difficulty, a direct relation has been reported just about difficult course which has an acceptable distribution of grades [24].

IV. EDM CASE STUDY

This study considers data collected during 2004-2010 at the University of Tehran, department of electrical and computer engineering, ECE, which consist some information about all students who study the field of Information Technology. Datasets comprises more than 7000 records consist of course name, educator name and the final grades.

In addition, students are always demanded to fill a web based evaluation form in order to express their levels of satisfaction about the educators. There are more than 7200 records describing student evaluation of teacher (SET) values in the last 3 semesters. It is worth to mention that before utilizing the new web-based evaluation mechanism, all evaluations had being performed handly, so there is no evidence available to the researchers.

We first describe our specific problem and then show the experimental results obtained in the execution of the different data mining algorithms. Finally, we explain our results and discuss about some of them.

A. Research Questions

This research is the second and complementary phase of our previous research [25] trying to answer some more detailed questions especially investigating the relationship between SET rating and students' grades. Our main objective is using educational data mining techniques in a web based learning environment to discover some temporal patterns of students and educators behaviours.

Through EDM techniques, we wanted to discover students’ sequence pattern of courses selection. Although there are some perquisite between course consequences, some selective sequence of course selection are available which are usually under the guide of the supervisors. Furthermore some overall view of educators’ grading policies or students level of success are demanded to be illustrated.
The main question goals interactions between students and teachers. Students are demanded to fill an evaluation form that asks students’ approve of each educator. There is an important issue that states there would be a reasonable relation between the satisfaction of a student and his/her level of success. Researchers are going to analyze the relation between final mark obtained for a course and student evaluation of teacher (SET) rating.

V. RESULTS

A. Sequence patterns of students’ course selection

Students’ sequence pattern of course selection was the first purpose of this research. In this respect, using Clementine and different valubles for minimum support and confidence, students’ consequences of courses was investigated, as it is represented in Figure 1. Discovered knowledge will be more informatics if some reconsideration are applied and courses which are completely adapt the faculty academic chart and indeed have no new information are put aside.

![Figure 3. Students sequence pattern of course selection with minsup=40 and minconf=50](image)

These sequences can be utilized in a recommender system, which help learners to choose their course sequences based on the behaviour of successful learners.

B. Behavioral pattern of students groups

After cleaning and preprocessing the research data, some summery table have been constructed in order to be utilized in the next steps of analyzing processes. Figure shows grades points averages (GPA) for all students grouped by their entrance year. There is a little decrease in the last semesters, which may be explained as the result of graduating hard study students. In addition, it shows a greater GPA values for two groups (82, 85) almost in all main semesters (1-8) and a changing behavior for two groups (83, 84) showing up and down GPA average in respect to the others. The analysis and presentation of data at a group level provide stakeholders with a method for distinguishing differences in students’ behaviours across entrance year and a benchmark of current activity for future comparison.

![Figure 4. Overall picture of student groups](image)

Through aggregating the grades given by each educator, it has been facilitated to have a comprehensive picture of his/her behaviours as it depicted in Figure . It may enable administrators and educators to discover the trends of behavioral policies. For example, it shows some decreases in two fall semesters which can be investigated in the future semesters in order to whether it happen again or not. These types of measurements and reports are of particular interest when attempting to understand one educator behaviour consistency through his/her consequent years of teaching. Regarding Figure against Figure , it will be possible to seek any correlation between educators and student groups’ behaviours whether they match each other or not.

![Figure 5. Overall picture of an educator’s grading behavior](image)

C. Relation between SET rating and students grades

The datasets include more than 2300 web-based SET ratings for just for 3 semesters. A summary report of all evaluations is depicted based on the semester name and grouped by students’ entrance year in Figure 6(a). Since students participated in this investigation have been spread in all different number of semester, (first, second and so on), we put them in a row (based on the number of semester) in order to estimate students’ temporal behaviors. Representing average SET ratings of all groups as a timeline (depicted in the right side), it may be concluded that although students have rated their teachers highly at the beginning semester, the ratings have been decreased along the consequent semesters lead to the least SET ratings in semester 8. As it mentioned before, ratings are available just for 3 semesters, but if the SET ratings are available for all semesters in the future, more behavioral patterns will be discovered.
In this paper, we use SET data to investigate the relationship between SET ratings and relative grades. First step, we investigated all courses and their students as a whole. Contrary to recent literature that suggests that instructor is able to purchase high SET scores when relative grades are considered, we have found no relation among our data set. Then we eliminated some strange data that seemed to be outliers: some top students had evaluated their instructors very poor while some weak students had evaluated them excellent.

Figure 7 represents the results for bachelor and master students separately after the elimination of those outliers. Because of the insufficient number of master students, we have considered the bachelors in the rest.

Furthermore, to analyzing educators’ behaviors, each teacher had been considered individually. Figure shows the relation between SET ratings and students grade for five courses all had been taught by the same educator.
VI. DISCUSSION

Analyzing students behaviors in teacher evaluating (Figure ), it seems that students evaluate their teachers optimistically at the beginning but as they become juniors in the university, they were more familiar with the evaluation measurements and expected their educators more and more which leads to a decrease in their evaluations. Of course since there are more complex courses in the last semesters, it demands more students efforts which may cause a little less satisfaction especially when the students are not interested in the courses as it mentioned in [5].

Considering educators' behaviors depicted in Figure , makes it hard to conclude any reasonable relation between the observed parameters since some direct and inverse relations are discovered. It shows that there are more parameters affecting the relations, which must be considered. These factors are related to characteristics of students and courses as well as the educator's teaching strategy in each course[5, 6, 14] [15, 16].

We have investigated the relationship between SET ratings and students grades. Contrary to the literature that faculty are able to 'buy' higher SET ratings by giving higher grades[7, 18], no important correlation recognized. Categorizing students in three groups make it possible to discover some interesting behavioral patterns for each group. Some information about the GPA and SET ratings in addition to their correlation coefficients are reported briefly in Table 3. some computational results about students Table 3.

<table>
<thead>
<tr>
<th>Grades</th>
<th>Records</th>
<th>Grade Ave</th>
<th>Grade Var</th>
<th>SET Ave</th>
<th>SET Var</th>
<th>Pearson</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>2283</td>
<td>14.63</td>
<td>12.08</td>
<td>17.99</td>
<td>17.58</td>
<td>0.037</td>
</tr>
<tr>
<td>Weak</td>
<td>&lt;14.5</td>
<td>941</td>
<td>11.42</td>
<td>6.02</td>
<td>18.10</td>
<td>14.61</td>
</tr>
<tr>
<td>Moderate</td>
<td>14.5 ≤ A&lt;17</td>
<td>601</td>
<td>15.75</td>
<td>0.369</td>
<td>17.62</td>
<td>23.56</td>
</tr>
<tr>
<td>Good</td>
<td>&gt;17</td>
<td>741</td>
<td>18.41</td>
<td>1.00</td>
<td>18.14</td>
<td>16.42</td>
</tr>
</tbody>
</table>

Pearson's correlation coefficient between two variables reveals that there is a direct correlation among weak and good students (0.086, 0.108) and a negligible indirect correlation among moderate ones (-0.015). It seems that weak students evaluate their educators better than the moderate one. It brings about this conclusion that they may know their own low level of efforts or insufficient information and have less expectation, which means that they accept their own responsibility. We think that this situation may be considered as a cultural issue.

On the other hand, since moderate SET average is less than the other ones, it seems that good and weak students evaluate their teacher better than moderate students do. In addition, standard variances of SET ratings show a serious difference between the categories: weak and good students have more similar ratings values than moderate students. Trying to explain these results, moderate students may be considered as some different types of students, who have unequal intelligence, efforts, and even free time, which leads to a wide range of opinions, and successes they have. Results would be more knowledgeable if we considered other students related factors such as how many semesters they selected the course later than their classmates did. It seems that different clusters of students will be more precisely detected if other important information is considered.

Investigating SET rating, it seems that there is no negative question in the questionnaire as it is common in other survey mechanisms[17]. Therefore, it is impossible to check whether the students have filled the form spending enough time or they have quickly selected some choices. Besides, students may have insufficient experiences and maturity to answer some questions related to their educator's abilities such as his/her knowledge in the field. In addition, all questions are just related to the teacher with no concern for students. For example, questionnaire asks "the educator ability in explaining the course main concepts", which would be more efficient if it asked: "How much have you understood the course main concepts through the educator's lectures?" Furthermore, there is a real need to add some questions investigating students' personality and their self-expectations so that we will be able to give meaning to the results in comparison to each person characteristics. It may be realized through some questions like: "The pace of the course[6] which tries to know the student sense of the course teaching speed. Current questionnaire consider no course related factors, which if have been evaluated we could categorize all courses in adoption to their demanded students efforts and other important factors.

VII. CONCLUSION

Educational Data Mining (EDM) has emerged as a research area in recent years that deals with large amounts of data provided in different types of educational environments by different groups of users[2]. Utilizing various advanced techniques, EDM has the ability to provide appropriate services oriented to students, educators, and administrators. Reviewing more common EDM applications and techniques, this research has investigated institutional data gathered via more than 5 years in the University of Tehran for the students who study in Information Technology engineering.

Extracting the sequence pattern of course selection in consequent semesters, different patterns of students' behaviors have been recognized. These results can be used in a recommender system in order to facilitate the consulting process and providing a more precise view of each student individually as well as in comparison to other students.

The second goal of this research has been realized though providing an understandable picture of all students, grouped by their entrance year, to recognize some of their characteristics and capabilities. In this
Discovering students' behaviors in different levels of detail will enable us to setup a better individual and group model of all students. We will use this knowledge to enrich our simulation and modeling interfaces introduced in [26] and to provide more precise models of a learner in a simulated educational environment.

This research will be continued on investigating the relations between SET ratings and other factors related to students, educators, and course characteristics as well as rating administration. In addition, some weighting parameters will be considered to determine effects of questionnaire items in its total rating. In addition, it really needs to add some new negative questions, to check the questionnaire internal consistency, as well as some students and new technology relevant questions like some other questionnaires [6, 14, and 17]. We think it will be highly beneficial if researcher have enough information to compare the demanded relationship among some Iranian universities in order to investigate the effects of their special academic policies on satisfaction level of students.

ACKNOWLEDGMENT

We thank Dr. Mocini, the head of University of Tehran Informatics Center who made the research data set available for us.

REFERENCES

[8] Han, J., Data Mining: Concepts and Techniques. 2006: University of Illinois at Urbana-Champaign, Micheline Kamber.
[16] Croushore, D. and R.M. Schmidt, RIDIT ANALYSIS OF STUDENT COURSE EVALUATIONS. University of Richmond 2010

Fatemeh Orooji received her B.Sc. degree in Electrical Engineering (Electronics) from Sharif University and her M.Sc. degree in Software Engineering from Shahid Beheshti University of Tehran. She is now a Ph.D. Candidate in University of Tehran, Iran. She is interested in web-based educational systems (WBES) and technology-enhanced learning. Her research activities include Agent-Based Architectures & Applications, Educational Data Mining, Personalization & Adaptation, and Information Retrieval in WBES.

Fattaneh Taghiyareh is Assistant Professor of Computer Engineering- Software & Information Technology, at the University of Tehran, where she has served since 2001. She received a Ph.D. in Computer Engineering- Parallel Algorithm Processing from the Tokyo Institute of Technology in 2000. Her research interests include Human-Centered Computing applied to Learning Management Systems and based on Multi-Agent Systems. She is currently working on "Group Collaborative Learning" as well as "Educational Data Mining (EDM)" with a pedagogical approach. In addition, Web Services are also another concern of her research to solve the composition problem from an ontology-based perspective. She was ex manager of the Department of Information Technology at the School of Electrical & Computer Engineering, University of Tehran, and former Director of Technology Incubator, the University of Tehran. She is a member of Editorial Board of International Journal of Information & Communication Technology. Currently she is serving both Software Engineering and Information Technology Departments as a member and she is supervising more than 10 post-Graduate students of the Multi-Agent System Lab and HCC Lab has founded by her. Moreover, she is supervising IT Foundation lab which was established also by her 5 years ago. She is the author of more than 50 conference and journal papers.