An Application of Educational Data Mining Techniques at Faculty of Technical Sciences in Novi Sad

Vladimir Ivančević[#1], Milan Čeliković[#2], Slavica Aleksić[#3], Ivan Luković[#4]

[#]*Department of Computing and Control, Faculty of Technical Sciences, University of Novi Sad*
Trg Dositeja Obradovića 6, 21000 Novi Sad, Serbia

[1]dragoman@uns.ac.rs

[2]milancel@uns.ac.rs

[3]slavica@uns.ac.rs

[4]ivan@uns.ac.rs

***Abstract*— As the Educational Data Mining gets more prominence, so rise the expectations of what is considered standard in modern education process. By using an already proposed data warehouse system as a basis, we analyse in the paper three issues typical for modern education: discovery of actual relationship between different assignment types and final grades, formation of student project groups, and discovery of atypical students. We elaborate the appropriate solutions and propose a way of improving the education process.**

***Keywords*— Data Mining, Data Warehouse, Educational Data Mining, Education Process.**

I. Introduction

In recent years, Educational Data Mining (EDM) has become more prominent in both research and practice. Some of the principal reasons for this are the fast-paced development of technology, the rapid growth of human knowledge, and the increase in the number of educational institutions. At the same time, at the Faculty of Technical Sciences of University of Novi Sad (FTS), there is also a continuous increase in the number of students enrolled, as well as a greater selection of curricula and courses. Such circumstances call for a fast response and the improvement of faculty infrastructure including further development of the faculty information system. Some ideas regarding this issue have already been proposed in [1]. Data concerning students, their assessments, and grades should be stored in an OLTP database, along with information about courses, programmes, and instructors, while a separate data warehouse (DW) system is to be populated through ETL process with data necessary for the generation of various reports. However, much of potential of the DW system would lie dormant without the use of data mining (DM) techniques.

In this paper, we propose a usage of DM techniques in a DW system dedicated to education process, so some of the typical educational problems present in academic institutions could be mitigated. The main goal of the paper is to demonstrate some useful results that can be quickly obtained through DM and exploited for the improvement of the education process. We illustrate the application of selected DM techniques through these cases:

- analysis of dependencies between student assessment scores and final grades;
- formation of student project teams; and
- discovery of atypical students.

Also, we discuss main motivation factors behind these issues, and give a short overview of the DW system and data used. Our conclusions are based on analyses performed on a data sample obtained from several courses performed at the Chair of Applied Computer Science of FTS.

II. Related Work

Although EDM is a young discipline, considerable research effort is put into developing new approaches intended to further improve education process. Beside that, many of the existing DM techniques are being adapted and applied at all the levels of various education systems.

At the School of System Engineering at Universidad de Lima, a recommendation system helps students in deciding which courses to choose [2]. It utilizes C4.5 algorithm [3] to discover rules from previous records which include student course selections and grades. Classifiers then predict whether a student will be successful in selected courses. This relies on assumption that current students who share similar course choices and grades with previous students, will have similar success. Similarly, various algorithms are used to predict final student grades in a web-based education system from features present in logged data [4]. Combination of multiple classifiers is used together with a genetic algorithm (GA) [5] which optimizes the process and helps to achieve greater accuracy.

Two approaches to computer-supported group formation (CSGF) which include semantics are proposed in [6]. One of them treats group formation as a constraint satisfaction problem where the semantics of the formation criteria is modelled as a set of constraints. The other approach relies on clustering, while semantics is added to data using a Semantic Web domain ontology so a more extensive description of a person is available for mining. In [7], a framework for measuring student motivation in online learning environment is proposed. It assumes that motivation includes three dimensions (engagement, energization, and source) to which various variables computable from logged data can be related. These relationships were speculated using results obtained by hierarchically clustering the variables. An example of the use of another non-predictive DM technique is given in [8], where association rule mining is performed with the objective of discovering hidden relationships between failed courses.

When comparing the aforementioned works with the approach that we propose, a noticeable difference is the specification of a data source used. In the referenced works, necessary data were directly fetched from general-purpose databases and then preprocessed in order to be used in analyses. On the other hand, our approach used data from the DW system possessing already cleaned and transformed data, thus reducing duration of the data preparation phase. Another important distinction is the granularity of analysed data. While we used data at the level of a single course assignment, the others utilized course grades [2], [8]; grades and student personal information (e.g., age, gender, nationality, interests, and roles) [6]; and details concerning student use statistics, success rate, and behaviour in online systems [4], [7].

III. User Requirements

One of the main objectives of FTS is to strengthen its position in the international education and research community. In order to achieve this goal, the quality of education process needs to be continuously improved. Some of the possible actions for improving current conditions include modernization of course programme content, additional investments in faculty facilities together with increased attendance to student needs and problems concerning their studies. The last group of measures has potential to yield significant results because it is expected that influencing students, which are central to education process, will produce more benefits with less expenditures when compared to other methods mentioned. Furthermore, when these measures are coupled with proper software and available data, the final effect should be even more pronounced. Still, some problems cannot be solved with a report produced directly from a DW system. By applying DM techniques and algorithms to such a system, the examination process can be shortened, and hidden relationships and dependencies between different data sets uncovered.

During the past year, at the Chair of Applied Computer Science at FTS, students made complaints on disproportionate levels of difficulty of certain assignments and expressed concerns regarding the types of assignments actually emphasized in some courses. Some of the students wanted more balanced project teams not randomly formed, while others were worried about the unfairness which could arise when groups were created through self-selection. Also, several teachers noticed that often there were certain students whose success with given assignments varied considerably throughout semester. Taking everything into consideration, the Head of the Study Programme identified three key issues that needed to be addressed in a special study: discovery of actual dependencies between single assignment type scores and final grades; formation of student groups for project assignments; and identification of atypical students. The study was expected to include analyses of data containing student scores and grades, offer interpretation of its results, and suggest procedures that could resolve issues identified in the education process.

IV. Data Warehouse System

In the context of advanced analyses, the main role of the data warehouse system is to store homogenized and adequately organized data. The actual DW system used for DM analyses was developed using Oracle Warehouse Builder (OWB) and included a single database implemented with Oracle RDBMS. This database was not a part of the OLTP system and was exclusively used for data warehousing.

A. *Data Warehouse Schema*

Database schema of the already proposed DW system was firstly presented in [1]. Since then, however, it has evolved so the new user requirements could be met (see Fig. 1).

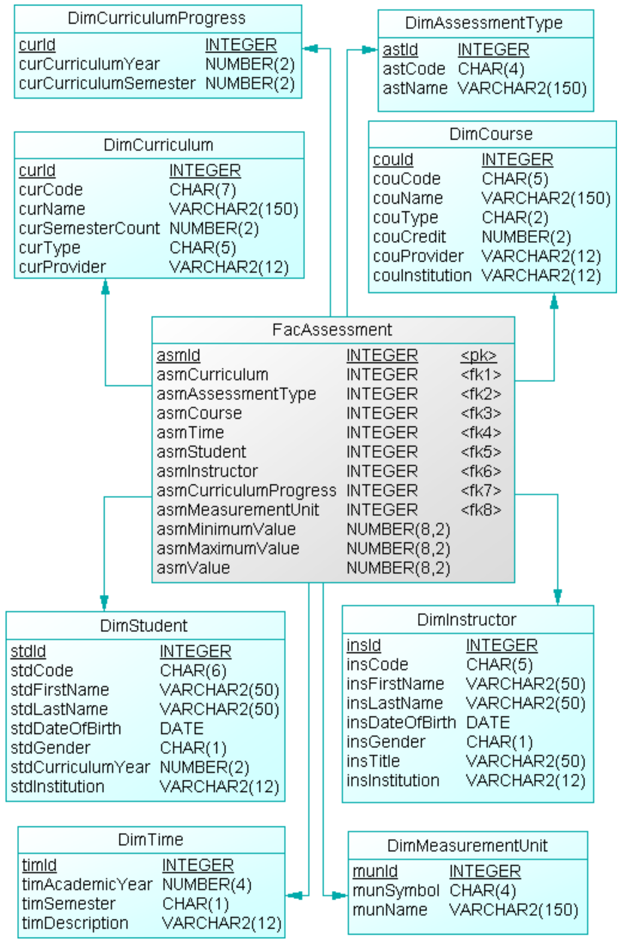

Fig. 1 A Diagram of the DW Schema

Some of the dimensions were reorganized and new measure attributes were added. The only fact table used is *FacAssessment* which stores data about assignments including passed, failed and skipped student assignments. Dimensions which are referenced by *FacAssessment* fact table are:

- *DimTime* – time context of the assessment. Includes academic year (e.g., 2010 for academic year 2010/2011) and semester (winter or summer) during the year when the assessment was done;
- *DimCurriculum* – a curriculum to which the assessment belongs, e.g., *Computing and Automation BSc (Hons)*;
- *DimCurriculumProgress* – time in context of a curriculum. Includes curriculum year (e.g., 4 for 4th year of curriculum) and curriculum semester (e.g., 7 for 7th semester of curriculum) in which assessment was to be carried out;
- *DimCourse* – a course for which the assessment was done (e.g., *Databases* or *Data Warehouse Systems*);
- *DimAssessmentType* – a category of assessment measure (e.g., test of theoretical or practical knowledge, project, oral or written exam, or final grade);
- *DimInstructor* – a teacher responsible for the assessment;
- *DimMeasurementUnit* – a measurement unit used for the assessment (e.g., points, percents, or numeric grades); and
- *DimStudent* – a student who was assessed.

Only a part of the DW schema that is relevant for the aforementioned DM analyses is presented.

B. A Description of DW Data

The sample data were selected from five courses (including both mandatory and elective courses) conducted by Chair of Applied Computer Science at FTS during the past two years. Four of them are bachelor level courses, deliberately selected only from the 4th study year, while the fifth one is a master level course. The selected courses are:

- *Database Structures and Organization* – mandatory 4th year bachelor course in *Power, Electronics and Telecommunications Engineering* curriculum, for year 2008/2009 with 28 students, and 2009/2010 with 37 students;
- *Databases* – mandatory 4th year bachelor course in *Computing and Control* curriculum, for year 2008/2009 with 70 students and 2009/2010 with 83 students;
- *Information Systems and Management* – elective 4th year bachelor course in *Computing and Control* curriculum, for year 2009/2010 with 23 students;
- *Database Systems* – elective 4th year bachelor course in *Computing and Control* curriculum for academic year 2009/2010 with 22 students; and
- *Data Warehouse Systems* – elective 1st year master course in *Computing and Control* curriculum, for year 2008/2009 with 54 students and 2009/2010 with 46 students.

The sample data include student scores for all pre-exam assignments, attendance scores and final exam grades. The data were extracted from proprietary *xls* files and imported into the OLTP database. This database was then used as a source in ETL process, while the DW system was the destination.

Through this process, we obtained 2020 fact records in total, stored in *FacAssessment* table. The number of records in dimension tables is as follows:

- *DimTime* – 8 records,
- *DimCurriculum* – 6 records,
- *DimCurriculumProgress* – 3 records,
- *DimCourse* – 5 records,
- *DimAssessmentType* – 10 records,
- *DimInstructor* – 3 records,
- *DimMeasurementUnit* – 3 records , and
- *DimStudent* – 270 records.

This amounts to 308 dimension records in total.

V. DATA ANALYSES AND RESULTS

In this section, we present three practical examples of DM usage which offer solutions to the issues identified in requirements analysis together with short discussions of the obtained results. In the first example, we analyze a contribution of various course assignments to students' final grades. In the second example, we apply a clustering technique to provide formation of students' project groups, while in the third example we apply the anomaly detection technique to discover atypical students. All algorithms used in our DM analyses were implemented under Oracle Data Mining environment [9]. Results were acquired in a mining process conducted through Oracle Data Miner (a GUI data mining tool) and stored in database tables, so they could be managed using regular SQL.

A. Contributions to Final Grades

There are several assessment types regularly used in education to calculate final grades. These include, but are not limited to, mid-term assessments of theoretical and practical knowledge, project assignments, written and oral exams, attendance statistics, etc. Often, a teacher assigns a weight to every course requirement defining the impact of an assignment on a final grade. For all the courses being analysed, assignment scores had already been assigned by the teachers in such a way that each sum of all assignment scores for a single course is a number from the interval [0..100] which directly corresponds to the final grade. In this case, the weight of a single assignment is the maximum score possible for that assignment divided by 100. Therefore, every weight is a number between 0 and 1. Since various types of assignments may have a larger or lesser impact on a final grade, a correlation between variables representing scores for one assignment type and a final grade may be greater or lower.

Course teachers are responsible for creating assessment types and defining weights of all assessment types included in their courses so as to keep the courses in line with the promoted course goals. For example, a course intended to be mostly of a theoretical character, should have greater weights assigned to assessments focused on an evaluation of theoretical knowledge, while the teachers should carefully tailor the principal assignments so a better and more complete view of student's theoretical knowledge could be obtained. In order to make proper decisions, teachers should be provided with historical data and tools that would allow them to identify contributions of individual assessment types to final grades and confirm whether the course goals are met.

In order to acquire information concerning the significance of individual assessment types, various statistical techniques can be used. To discover the dependence between assignments and grades, one possibility is to calculate correlation between the data values of aggregated scores for each assessment type and the values of the final grade, and then to compare these correlation values to corresponding weights of the assessment types. The other option is to apply a DM technique known as *Attribute Importance* i.e. *Feature Selection* [9] to identify the significance of attributes for the prediction of a selected target attribute. This technique utilizes the minimum description length (MDL) principle which states that the most compact representation of data is also the most probable explanation of the data. As we associate a predictive model of the target class to each attribute, the models are ranked according to the compression levels that can be achieved on the basis of these models. Since the technique usually benefits from binning technique in data preparation, the typical data preparation process includes the application of decision trees for the identification of optimal bin boundaries for both categorical and numerical attributes. Because this technique is mostly automated, optimized for larger data sets and easily executable for data stored in a DW database, it is more suitable for solving the identified problem than the calculation of correlation.

We applied the technique on a sample data that include student grades and scores aggregated to different assessment types defined for two courses: *Databases* in 2008/2009 year with $N = 70$ students and *Database Structures and Organization* in 2008/2009 and 2009/2010 with $N = 65$ students. Final grade for both courses was calculated using the scores belonging to one of the four categories: theoretical, practical, attendance, and oral, while their weights (w) are: 0.3 (30% of final grade), 0.4 (40%), 0.05 (5%), and 0.25 (25%), respectively. Pearson correlation coefficients were calculated for pairs of data sets – grades paired with scores for each assessment type. Calculated values are presented in Table 1.

From these results, it follows that for both courses the greatest correlation exists between the oral assessment and the grade, although the corresponding weight is not the highest one. Similarly, there is a high correlation for the pair theoretical-grade. However, for the *Database* course, the correlation value for the pair practical-grade is relatively low despite the fact that practical assessments have the greatest weight, while the attendance correlates moderately with the final grade, although it only has a 5% impact. This discrepancy is not present in the results obtained for the *Database Structures and Organization* course, where correlation values for practical assignments and attendance follow their weights more closely.

TABLE I
PEARSON CORRELATION COEFFICIENTS FOR THE GRADES AND SINGLE ASSESSMENT TYPES

Databases (N=70)				
	Theoretical (w=30%)	**Practical (w=40%)**	**Attendance (w=5%)**	**Oral (w=25%)**
Grade	0.77	0.46	0.64	0.87
Database Structures and Organization (N=65)				
	Theoretical (w=30%)	**Practical (w=40%)**	**Attendance (w=5%)**	**Oral (w=25%)**
Grade	0.79	0.56	0.31	0.86

After the application of *Attribute Importance* technique on the same data, we may notice that the order of the assessment types when sorted by their importance (see Fig. 2) coincides with the order obtained by sorting the assessment types according to their respective correlation values. We have actually expected such a behaviour, since correlation describes dependence between two data sets.

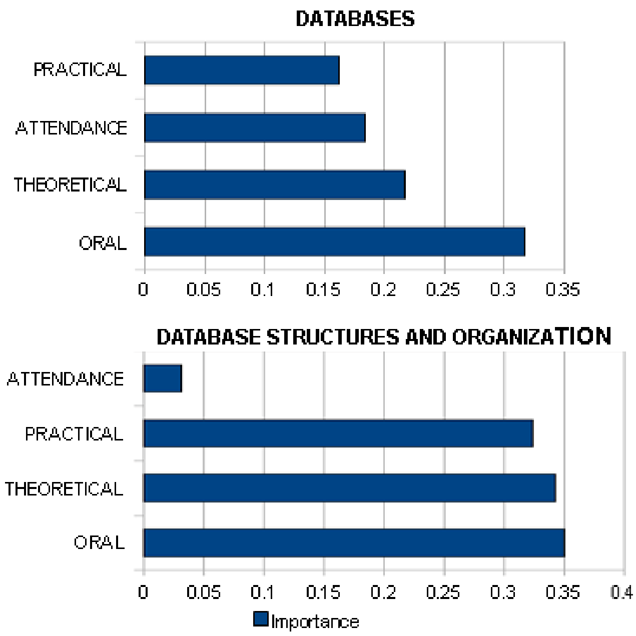

Fig. 2 Importance of all the types of assessments

Similar conclusions follow as we analyse the values from Table 2, which shows calculated importance values for all the assessment types together with the corresponding assessment type weights and ranks determined by these importance values.

TABLE II
IMPORTANCE OF EACH ASSESSMENT TYPE IN THE PREDICTION OF THE FINAL GRADE

Databases (N=70)			
Rank	**Name**	**Weight**	**Importance**
1	oral	0.25	0.32
2	theoretical	0.30	0.22
3	attendance	0.05	0.18
4	practical	0.40	0.16
Database Structures and Organization (N=65)			
Rank	**Name**	**Weight**	**Importance**
1	oral	0.25	0.35
2	theoretical	0.30	0.34
3	practical	0.40	0.32
4	attendance	0.05	0.03

For the *Databases* course, it is interesting to notice that the final grade can be slightly better predicted by the attendance scores which have by far the lowest weight than by the practical assessment scores which are favoured in grade calculation by the greatest weight. Although the attendance score has only a minor impact on the final grade, the information it holds is much more valuable than it was initially expected. The lowest importance of the practical assessments in the prediction of the final grade can be attributed to the fact that the scores for practical assignments mostly lie in a narrow range and have a low deviation, while, at the same time, there is a greater dispersion of final grades that cannot be explained well by these scores. We believe that one of the main reasons for this anomaly is the fact that many students view practical assignments as a chance to secure a passing grade. This would allow them to avoid an oral exam that is generally not required to pass the course. For all the students who are more interested in making a higher grade, scores on other assignments would determine the final grade.

B. *Formation of Student Project Groups*

Teamwork has gained increasing prominence in education process over the last decades, particularly in engineering courses where larger project assignments are given to groups of students who then have to solve a problem relying on individual skills of team members. However, one of the biggest problems in this kind of assessment is forming student groups. A simple solution is to randomly assign students to teams, but this leads to inferior group dynamics, less positive attitudes about the group experience, and slightly lower group outcome ratings when compared to self-selected teams where students themselves choose their own teams [10]. Similarly, student groups formed through criterion-based selection, e.g., groups assembled based on mixed academic ability, common career aspirations of members, and gender dispersion, are preferred to random groups [11]. Although students favour self-selected teams, that approach may produce homogeneous groups which are not generally considered as effective on the whole as more diverse groups. When it comes to the choice between mixed-ability groups and streamed groups where members have similar academic ability, the former are preferred [12]. Another important issue to be considered is the group size. While there is no definite answer appropriate for every domain, it is widely accepted that there should be 4±1 members.

The process of forming student groups can be supported by a clustering technique. All the recommendations for the assembly of groups should be taken into account when executing one of the algorithms that creates clusters of similar students. The calculation of a similarity between two students can be based on the difference between their respective course scores and grades made on one of the previous mandatory courses related to course for which groups are formed. The closer the scores of two students, the more similar they are considered to be and the greater is the chance of them becoming members of a same cluster. If the size of a team is selected to be 4, then 4 such clusters need to be formed. For this purpose, k-means [13] clustering method is suitable, since the enhanced version of this algorithm [9] requires a number of clusters as the input and works best when there are not many attributes, which was the case in this analysis, since there were only 4 attributes. This method is distance-based, i.e. a distance function is used to calculate similarities between entities and to assign them to clusters. A top-down model is formed hierarchically using binnary splitting and growing one cluster at a time. In the usual data preparation phase for this method, outlier-sensitive normalization is performed. Once the clusters have been generated, final student groups can be created by randomly selecting one team member from each cluster.

Data on which this approach was tested included student scores aggregated to different assessment types defined for the *Databases* course in 2008/2009 with N = 70 students. Final grade was determined by the scores belonging to one of the four categories: theoretical with max. score 30, practical 40, attendance 5, and oral 25, which together gave a max. sum of 100. Using these scores, four student clusters were formed. Their case counts, rules approximating cases contained, together with confidence values (the percentage of all the cases which satisfy the left side of the rule that also satisfy the right side of the rule) and support counts (the number of cases that satisfy both sides of the rule), are shown in Table 3.

For each cluster, there is a rule which describes the majority of cases in that cluster. It can be noticed that clusters labelled *1* and *2* generally consist of students with good scores, while clusters *3* and *4* mostly include students who have lower scores. Students have been assigned to clusters more or less evenly with the exception of cluster *3* which has a bit more students than the other clusters. Still, this can be solved by moving some of the students from cluster *3* to the similar cluster *4*. Using this data, project teams can be formed for future courses in the same curriculum which are dependant on the *Databases* course.

If the same procedure is applied on student scores on the *Databases* course in 2009/2010 year with $N = 83$ students, the 4 clusters constructed vary much more in their size. However, if the clustering is based solely on scores from theoretical and oral assessments, created clusters are more balanced in size. This is actually a consequence of the fact that these scores have a greater variance and are more evenly distributed. Only one cluster, mostly comprised of students who passed the course without taking oral exam, has a significantly greater case count than the rest. Two of the clusters centre mainly around students with better scores, while students belonging to the remaining cluster had the lowest scores.

TABLE III
STUDENT CLUSTERS

<table>
<tr><th colspan="4">Databases (N=70)</th></tr>
<tr><th>Cluster</th><th>Cases</th><th>Confidence (%)</th><th>Support Count</th></tr>
<tr><td rowspan="2">1</td><td>18</td><td>83.33</td><td>15</td></tr>
<tr><td colspan="3">IF attendance in (5.0) and 19.9<=oral<=25 and 36.67<=practical<=40 and 8.4<=theoretical<=28 THEN cluster=1</td></tr>
<tr><td rowspan="2">2</td><td>16</td><td>75.00</td><td>12</td></tr>
<tr><td colspan="3">IF attendance in (5.0) and 0<=oral<=2.05 and 33.35<=practical<=40 and 8.4<=theoretical<=25.2 THEN cluster=2</td></tr>
<tr><td rowspan="2">3</td><td>25</td><td>84.00</td><td>21</td></tr>
<tr><td colspan="3">IF attendance in (0.0) and 0<=oral<=12.25 and 33.35<=practical<=40 and 0<=theoretical<=22.4 THEN cluster=3</td></tr>
<tr><td rowspan="2">4</td><td>11</td><td>90.91</td><td>10</td></tr>
<tr><td colspan="3">IF attendance in (0.0) and 0<=oral<=2.05 and 6.73<=practical<=30.02 and 0<=theoretical<=5.6 THEN cluster=4</td></tr>
</table>

If clustering algorithms not requiring a number of clusters as an input are used in this process, a number of natural clusters in a data set may not be equal to the one needed. We can easily overcome the problem by joining the similar clusters and dividing the larger ones. Also, the clustering approach need not be restricted to the formation of project teams. It can be extended to the formation of class or laboratory groups. In order to create more diverse groups, besides student scores, other attributes for clustering can also be included, such as personality traits.

C. Discovery of Atypical Students

Modern educational institutions attempt to adapt their curricula and courses to individual needs and abilities of their students. One of the reasons for this is the fact that different students require different teaching styles. Some of the students do not react well to teaching methods suitable for the majority and become disinterested. On the other hand, there are those who have little difficulty in completing their course assignments. However, all of them need some form of additional support.

In order to achieve this, one of the first steps is to identify students who significantly differ from an average student. When all of the usual measures of location and dispersion are calculated for student scores and grades at the end of a semester, atypical students and their results remain largely overlooked. One of the reasons for this is that most of the students' scores lie in a narrow interval centred around the mean value. However, through a closer examination of scores and grades data, we can detect atypical students using the DM technique *Anomaly Detection* [9]. This technique is based on a classification algorithm that relies on the Support Vector Machine (SVM) method [14] and works only with one class. The method uses a kernel function to transform entity representations into points in a high-dimensional space and then searches for a hyperplane that optimally separates two sets of points belonging to different classes. Once the support vector machine has been trained, i.e. the hyperplane has been identified, entities can be classified. This technique recognizes data deviation cases, in which attribute values differ significantly from those of other cases. A final output of the algorithm is a prediction that indicates for each record if it is classified as a regular case or not.

To illustrate this approach, we selected data containing student scores aggregated to assessment types for the courses: *Databases* in 2008/2009 and 2009/2010; *Database Structures and Organization* with integrated data for 2008/2009 and 2009/2010; *Data Warehouse Systems* with integrated data for 2008/2009 and 2009/2010; *Information Systems and Management* in 2009/2010; and *Database Systems* in 2009/2010. All the scores were aggregated for each student and belonged to one of the five assessment categories: theoretical, practical, project, attendance, or oral. To identify atypical students, a classification model was built with the outlier rate set to 10% and applied to this data set. Because there were not many attributes in this analysis (only 5 attributes), the Gaussian kernel function was used.

The overall results are shown in Table 4. Among the students labelled as atypical, three subgroups were identified after a closer inspection of the results: exceptional students; students with inconsistent scores over different assessment types; and students with consistently very low scores. For the courses with a fewer number of students, atypical students make a larger percentage of the population. In such cases, anomaly detection tends to be less reliable because there is not enough data to make such predictions.

TABLE IV
ATYPICAL STUDENTS BY COURSE

Course	Total	Atyp.	%
Databases 2008/2009	70	5	7
Databases 2009/2010	83	13	16
Database Structures and Org.	65	7	11
Data Warehouse Systems	100	13	13
Inf. Systems and Management	23	4	17
Database Systems	22	7	32

The anomaly detection technique is a good starting point for the identification of atypical students. However, the results do not always have a high level of accuracy and the algorithm requires a larger data set. Furthermore, if a student with consistently good or bad scores is not in a very small minority, the technique will not detect an atypical case because that record does not fit to the definition of an anomaly. To overcome this, the usual approach of setting a threshold and then comparing it to student scores can be adopted. Still, this DM technique is not suitable for the discovery of students who are atypical but whose scores are close to typical cases. Such students could be discovered through questionnaires, interviews, or psychological assessments. Only when all of the different groups of atypical students have been identified, it will be possible to start personalizing course assignments according to individual students' profiles.

VI. Conclusion

In the paper, we discuss three issues typical for modern educational institutions. We also propose the solutions relying on the use of a DW system and various DM techniques. Conclusions that can be reached with the help of DM algorithms have potential to improve overall academic performance of students as well as to increase the amount of acquired students' knowledge during education process. In order to perform the DM analyses, we prepared a data sample and stored it in our DW system. We believe that application of DM techniques on a DW system specially designed for EDM can considerably improve education process in academic institutions.

Further work may involve the development of a software tool which would allow inexperienced users to easily perform some of the predefined analyses including those presented in the paper. Such a software tool should be problem-oriented with an ability to clearly present the results together with their interpretations and useful suggestions. For such a software tool to be effective, carefully prepared data stored in a DW system, are to be available. Therefore, apart from applying a wider set of DM techniques and algorithms, our future work will be targeted to the formulation of the appropriate data acquisition and preparation methods to systematically obtain a broader and more accurate sample of data from different types of data sources.

Acknowledgment

The research was supported by Ministry of Science and Technological Development of Republic of Serbia, Grant III-44010 Title: *Intelligent Systems for Software Product Development and Business Support based on Models*.

References

[1] M. Čeliković, S. Aleksić and I. Luković, "A data warehouse system for monitoring university education process," in *Proceedings of the 4th International Conference on Information Technology (ICIT 2009)*, 2009, Al-Zaytoonah University of Jordan, ISBN 9957-8583-0-0, pp. 1-7, in CD ROM.

[2] C. Vialardi, J. Bravo, L. Shafti and A. Ortigosa, "Recommendation in higher education using data mining techniques," in *Proceedings of the 2nd International Conference on Educational Data Mining,* 2009, pp.190-199.

[3] J. R. Quinlan, *C4.5: Programs for Machine Learning*, Morgan Kaufmann Publishers, San Mateo, California, USA, 1993.

[4] B. Minaei-Bidgoli, D. A. Kashy, G. Kortemeyer and W. F. Punch, "Predicting student performance: an application of data mining methods with the educational web-based system LON-CAPA," in *Proceedings of the 33rd ASEE/IEEE Frontiers in Education Conference,* 2003, Boulder, CO, USA.

[5] J. H. Holland, *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*, The MIT Press, USA, 1975.

[6] A. Ounnas, "Enhancing the automation of forming groups for education with semantics," PhD thesis, University of Southampton, UK, 2010.

[7] A. Hershkovitz and R. Nachmias, "Developing a log-based motivation measuring tool," in *Proceedings of 1st International Conference on Educational Data Mining,* 2008, pp. 226-233.

[8] O. O. Oladipupo and O. J. Oyelade, "Knowledge discovery from students' result repository: association rule mining approach," *International Journal of Computer Science and Security (IJCSS)*, vol. 4 (2), pp. 199-207, May 2010.

[9] *Oracle Data Mining Concepts, 11g Release 1 (11.1)*, B28129-04, Oracle, 2008.

[10] K. J. Chapman, M. Meuter, D. Toy and L. Wright, "Can't we pick our own groups? The influence of group selection method on group dynamics and outcomes," *Journal of Management Education*, vol. 30 (4), pp. 557-569, Aug. 2006.

[11] G. A. Lawrie, K. E. Matthews and R.L. Gahan, "Forming groups to foster collaborative learning in large enrolment courses," in *Proceedings of the 16th UniServe Science Annual Conference*, 2010, University of Sydney, ISBN Number 978-0-9808597-1-3, pp. 66-71.

[12] G. Cumming, "The introductory statistics course: mixed student groups preferred to streamed," *Teaching of Psychology*, vol. 10 (1), pp. 34-37, 1983.

[13] S. P. Lloyd, "Least squares quantization in PCM," *IEEE Transactions on Information Theory*, vol. 28 (2), pp. 129–137, 1982.

[14] B. E. Boser, I. M.Guyon and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proceedings of the 5th Annual ACM Workshop on COLT*, 1992, Pittsburgh, PA, ACM Press, pp. 144-152.