XVII International Scientific Conference on Industrial Systems (IS'17)
Novi Sad, Serbia, October 4. – 6. 2017.
University of Novi Sad, Faculty of Technical Sciences,
Department for Industrial Engineering and Management
Available online at http://www.iim.ftn.uns.ac.rs/is17

Development of Business Intelligence systems to predict Behaviour Patterns of Students

Slobodan Petrovic, M.Sc.Eng.
(Teacher, Business Technical College of Vocational Studies Uzice, Serbia, slobodan.petrovic@vpts.edu.rs)
Darko Stefanovic, Ph.D.
(Assistant Professor, University of Novi Sad, Faculty of Technical Sciences, Department of Industrial Engineering and Management, Serbia, darkoste@uns.ac.rs)
Teodora Lolic, M.Sc.
(Assistant, University of Novi Sad, Faculty of Technical Sciences, Department of Industrial Engineering and Management, Serbia, teodora.lolic@uns.ac.rs)
Milan Mirkovic, Ph.D.
(Assistant Professor, University of Novi Sad, Faculty of Technical Sciences, Department of Industrial Engineering and Management, Serbia, mmirkov@uns.ac.rs)

Abstract

This paper presents the methodology for definition of behaviour patterns, success criteria and formation of clusters of students. The objective is the formation of clusters at an early phase, while a student has not even started studies yet. With the formed clusters, improvement activities are implemented as a part of support system in decision making towards better results of higher education. With the presented research, sustainability and the originality of the approach are proved, by which the support system in decision making in the information system is improved at higher education institutions. For definition of behaviour patterns, Business Intelligence tools were used: On-Line Analytical Processing (OLAP) and statistical analyses.

Key words: *Behaviour patterns, Business intelligence, Decision support, Higher education*

1. INTRODUCTION

The vast amount of data is collected and located both on the Internet and in the field of information systems (IS) of companies, institutions, institutions of higher education. From a variety of data one should extract the most useful information [1], [2] for timely (decisions and corrective) activities. There are a number of software [3], [4], techniques, tools, known as Data Warehouse (DW), data mining, OLAP, etc., all within the Business intelligence technology – BI [5]. All of them should enable the quick and efficient way to get the right information relevant for decision-taking. However, the use of BI in higher education is not used extensively enough (e.g. in Croatia [6]) or is not automated sufficiently [7]. At higher education institutions, the indicators are: a pass examination rate in the examination period, an average mark at exams, and a pass rate from one semester to the next one and so on.

1.1 The objective of the research and hypotheses

The objective of this paper is to identify students (at an early phase, while a student has not even started studies yet) those who shall abandon (drop out studies) further studies (they shall not study in the next academic year or next semester) or those who shall not study in continuity (they shall attend again the same academic year, i.e. semester) - cluster of students, hereinafter referred to as: CoS-year student. The set hypotheses (H designations and item number) in terms research objectives are:

– H1: correct design of OLAP reports and modelling of them, trends can be determined by a pass rate of students per study programs and semesters;

– H2: by a definition of the criterion $C0_{limit}$ dependent on the achieved success of students in secondary school - the previous level of education and its implementation, it is possible to identify students from a generation (Gyear – e.g. Generation of 2007 will be designated G07) who are CoS-year students;

– H3: by a definition of the criterion $C1_{limit}$ dependent on the achieved success of students in one of the subjects at the Admission test or complete Admission test and its application, it is possible to identify students from a generation who are CoS-year students;

– H4: to define a combination of criterions $C0_{limit}$ and $C1_{limit}$ and by applying the combined criterions, it is possible to identify students of a generation who are CoS-year students.

2. REFERENCES AND PREVIOUS RESEARCH

The importance of the application of OLAP tools and BI technologies in processing of data relevant to a particular segment of organisation's business is shown in many papers [9], [10], technologies that are used by BI [2], [4], and application multiple criteria analysis in higher education[11]. However, the number of studies that research the possibility of using BI technology at a higher education organizations (for example, Thailand [12]) to improve the educational process is small (the same conclusion was made by [7] for developing countries). The study [8] presents BI technologies (the eUNIV and e-EdU-Quality Projects) for quality assurance at some higher education institutions in Romania.

A survey in [3] shows that a relatively small number of IT students have an idea what BI is and they „see" BI as a possible field of employment. The problem for the successful implementation of BI tools is how to utilize data efficiently (ETL process - Extract, Transform and Load [6]) from IS higher education institutions (or from a Student Relationship Management – SRM [13] or DW) for identifying and predicting a students' behaviour.

Research studies [14], [15] describe the techniques of data mining from an e-learning system and form students' behaviour patterns. Research studies [7], [12] presents a conceptual model and tools of BI and OLAP-based DW to support management in decision-making at higher education institutions, as well as other studies, ignore the data available at higher education institutions when enrolling students at the institution itself (the importance of research) and that should be the start of data mining in the e-learning system. A certain time is necessary to elapse for forming the data in the e-learning system (studying time). Some educational activities that are to improve the process can be realized without waiting for formation of the data in the e-learning system.

Luan[16] describes a case study in the U.S., how data mining techniques are used for monitoring and identifying the profile of community college students (described with significant attributes; in this paper: H1-H4) those students who continue towards higher levels of studies (four-year institutions), which provides significant benefits for all the stakeholders. It also shows the measures of help and support to students in the process of continuing the studies (including financial assistance as well) towards a higher level of studies (four-year institutions). A similar study [17] was conducted in Spain, on Secondary Schools, in order to helping about the academic orientation of student.

The study [18] presents the methodology and developed model for anticipating which students will abandon studies and that is for the purpose to start up certain improvement activities. The anticipation is performed at the end of the first semester, while in this paper it is anticipated earlier, when a student enrolees a higher education institution.

Heise [19] points out that the reason for his study is a small number of research studies: how the data from the DW are used to improve the support to a complex decision-making for process (a similar conclusion by [7] in the United States Higher Education). The study [13] emphasizes the problem of Higher Education of Portugal, that students abandon usually in the first year of studies and they recommended the development of Student Relationship Management (SRM) system. The data collected from the e-learning system are used for data analysis with the help of OLAP and data mining tools for undertaking improvement activities.

The study [6] presents an example of a successfully developed national information system (Croatia) for higher education where an important item is the Data Warehouse system and its own-developed BI tool for data analysis: a pass rate, the average mark of students at exams, pass rate and average mark considering the position of a teacher, etc.

The paper [12] presents how BI was implemented in higher education, also at the national level, in Thailand, which BI technology is used and key factors for successful implementation are identified: 1) completeness and the data integrity and 2) the intelligent OLAP tools to support fully the a decision support system (DSS).

3. METHODOLOGY

During this research, which has resulted in formation of reports, graphs, criteria, numerous methods were used: the method of analysis (OLAP), statistical analysis (measures of dispersion and measures of distribution form) and the method of synthesis as components of BI.

If students want to start their studies at higher education institutions in Serbia, they have to take Admission tests. An Admission test consists of a knowledge check in one or more subjects/fields which prospective students, applicants-candidates for studying, had studied at the secondary school (the previous level of education). Results, respectively the demonstrated knowledge of candidates in the Admissions test is measured in terms of points (a correct answer to a question or a task brings a previously-defined number of points) for each subject/field individually. Together with the results from the Admission test, candidates' result/success from their previous level of education is evaluated (Secondary School) by this rule: the sum of average marks of a candidate in all four years of Secondary School is multiplied by number two. The minimum number of points that candidates can get for the result/success from Secondary School is 16 and the maximum is 40. Based on these two criteria (Points from Secondary School and Points at Admission test) candidates can achieve: the minimum of 16 and maximum of 100 points.

The progress of students during the research is monitored and expressed with the parameter: an enrolled semester. With the semester enrolled, a student acquires the right to attend classes from that semester. Data are collected through a survey which is carried out during the application of candidates for

taking an Admissions test. Data available at all higher education institutions in Serbia during the enrollment of students in the first semester are:
1) achieved an average mark in Secondary School (link to H2), which is expressed in number of points, hereinafter referred to as: Points from Secondary School or C0 and
2) achieved number of points at the Admission test (link to H3).

In order to track and describe the indicators of point's allocation, the authors used the dispersion measures: μ - arithmetic mean, and σ- standard deviation. In accordance with the set goals of the research (item 1.1) and the hypotheses H2 and H3, a general expression for the criteria (quality indicators) of all the levels of higher education is defined:

$Ck_{limit} = \mu_k - \sigma_k$ (1)

where is:
k = 0, 1, 2, 3 ... 10 – a criterion designation; μ_k - arithmetic mean of points of students as per *k* criterion; σ_k - standard deviation of points of students as per *k* criterion.

Concretely, for the hypothesis H2 (C0 criterion dependent on: *Points from Secondary School*, i.e. achieved result/success of students in Secondary School), the expression (1) has the form:

$C0_{limit} = \mu_0 - \sigma_0$ (2.0)

where is: μ_0 - arithmetic mean for C0, $\mu_0 = \sum_{i=1}^{N} \frac{C0_i}{N}$;

N - the total number of students at a study program and

σ_0 - standard deviation for C0, $\sigma_0 = \sqrt{\frac{1}{N}\sum_{i=1}^{N} C0_i^2 - \mu_0^2}$.

The criterion $C0_{limit}$ is obtained on the basis of data for the generation of students at study programs who enrolled in the first semester (hereinafter referred to as: start Gyear) - higher education institutions have available these data at the enrollment at a higher education institution. If the i-th student has $C0_i$ (number of *Points from Secondary School*) less than $C0_{limit}$, i.e. $C0_i < C0_{limit}$, such a student is identified (hypothesis H2) as a student CoS-year. Hypothesis H2 can be expressed:

H2: $C0_i < C0_{limit}$ => CoS-year (2.1)

Also, in accordance with the hypothesis H3 (a criterion dependent on: *Points at the Admission test* or C1, i.e. from the achieved result/success in one of the subjects at the Admission test or a complete Admission test), the expression (1) has the form:

$C1_{limit} = \mu_1 - \sigma_1$ (3.0)

where is: μ_1 - arithmetic mean for C1 and σ_1 - standard deviation for C1.
The criterion $C1_{limit}$ is obtained based on the data as well as the criterion $C0_{limit}$. If the i-th student meets the condition $C1_i < C1_{limit}$, a student is identified (hypothesis H3) as a CoS-year student. Hypothesis H3 can be expressed:

H3: $C1_i < C1_{limit}$ => CoS-year (3.1)

By using adequate tables and graphs from any OLAP application that provide OLAP analysis, one can get to new and useful information in a fast, efficient and adequate way. The whole process of OLAP analysis is simple and intuitive for the top management. Forming a data file in a table format, analyzing these data by presenting the same data in different views, one obtains the information necessary for DSS based on facts, i.e. data.

4. A CASE STUDY - OLAP REPORTS FOR ANALYSES IN HIGHER EDUCATION

The entire research, and the proofs for the hypotheses H1-4, were carried out on a case study: Business Technical College of Vocational Studies (BTC) in Uzice, Serbia. The research that was conducted, comprised the population of students of BTC study programs:
– per five study programs (hereinafter referred to as: SP), the generation in 2007 (G07 - 414 students) and in 2008 (G08 - 501 students) and
– per six SPs, the generation in 2009 (G09 - 486 students) and in 2010 (G10 - 478 students), the total of 22 study programs, 1879 students.

In the tables and graphs below, a part of the data obtained from research for reasons of available space and clarity of presentation of On-line analysis is separated. Roman numerals (I, III, V, etc.) represent enrolled the first, third, fifth semester and so on, while designations with indices "a" (I-a, III-a) represent re-enrolled (the letter "a" has a meaning: the student did not satisfy the criteria for admission to the next semester: CoS-year student) the first, third, fifth semester.

4.1 Trends in passing rate of students: H1

Based on the analysis of data and histograms (Fig.1), in accordance with the hypothesis H1, it is possible to determine the trend: in all SPs there is a decline in the number of students at a transition (enrolment) of students from first semester towards the next semesters. A similar phenomenon exists in the Higher Education of Portugal [13]. This trend is somewhat less expressed in the transition of students from the second / third into the fifth semester (Fig. 1).

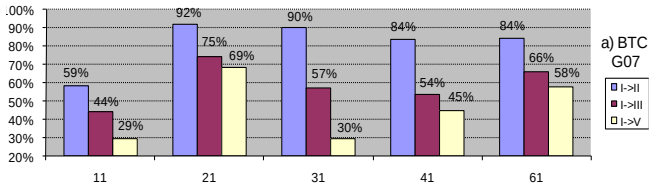

Figure 1. Percentage of students by SP in II, III and V semester compared to the start generation

In Figure 1, there are graphs, histograms for students of BTC generation in 2007, and they are similar to the generations of 2008, 2009 and 2010. Based on these analysis and identified trends (hypothesis H1), for further research data for I, II, III and V will be valid semester generation of students that will be compared with each other and the regularity and legality of students' behaviour will be established.

By analyzing the data from Fig. 1, it can be concluded that a further research is possible in order to anticipate and identify CoS-year students.

4.2 **Trend (H1) dispersion measure and distribution form of C0i and C1i for H2 and H3**

In accordance with the set hypotheses H2 and H3, the data are analyzed for C0i - *Points from Secondary School* and C1i - *Points at Admission test* at the level of study programs.

4.2.1 Trend of arithmetic mean for C0 and C1

In Figure 2 there are given histograms for the generations of students from 2007 to 2010, for the study program 21 (SP 21). From the histogram, it is concluded that there is a trend that students who have a higher number of points C0 (they achieved a higher average mark in their Secondary school and thus were better students) they study in the continuity (they enrol the final semesters at SP) and do not interrupt studies (a confirmation of hypothesis H1). The same trend is present for other SPs.

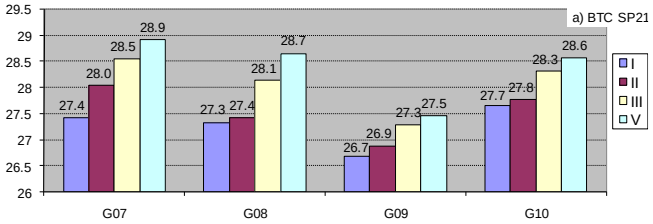

Figure 2. The arithmetic mean of C0 (μ_0): per semesters

By analysing the histograms in Figure 3, the conclusion is that there is a trend (hypothesis H1) as well as for the criterion C0 and the same character of the trend. Students who have a higher value of C1 (they showed greater knowledge at the Admission test and achieved a higher number of points) study in continuity and do not interrupt their studies. The same trend is present for the remaining SPs at BTC and their histograms are not given for the reasons of available space.

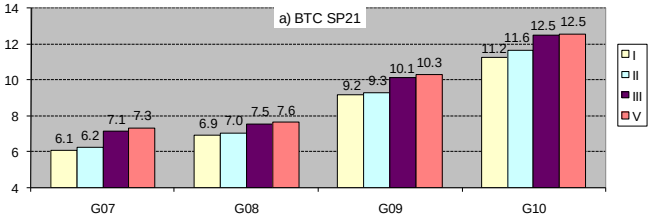

Figure 3. The arithmetic mean of C1 ($\mu 1$): per semesters

In accordance with the hypotheses H2 and H3, and the methodology used, the standard deviation σ_k is analyzed for criteria Ck, k = 0, 1.

4.2.2 Trend of standard deviation for C0 and C1

From the histograms in Figure 4, it can be seen that the standard deviation has a declining trend (hypothesis H1) per enrolled semesters of generations of students.

This means that the grouping of C0 data is closer to the arithmetic mean, and in practice, it allows identifying students who drop out further studies and proving for hypotheses H2 and H3.

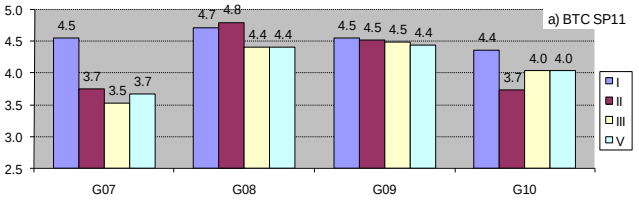

Figure 4. The standard deviation of C0 ($\sigma 0$): per semesters

Figure 5 shows that σ_1 (standard deviation for C1) has a tendency of slight increase (H1) in relation to the enrolled semester of generation of students. In Figure 5, there is an example for a study program 41 and the trend is the same as for other study programs at BTC.

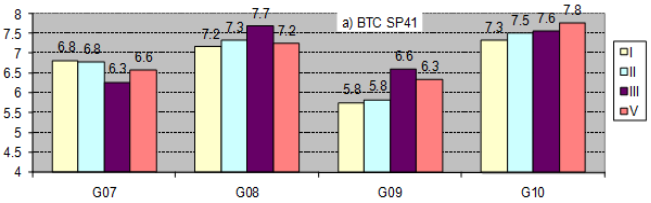

Figure 5. The standard deviation of C1 ($\sigma 1$): per semesters

5. IDENTIFICATION OF THE CoS-YEAR STUDENTS AND DISCUSSION

OLAP reports represent a start-up for identification of CoS-year students needed for a decision support system within the business intelligence of the higher education institution. In accordance with the set objective of the research and hypothesis H2 and H3 the number of students by study programs was determined (Table 1) from the generation of students who met the conditions 2.1 and 3.1.

5.1 **Hypothesis H2: CoS-year students**

The established criterion $C0_{limit}$ identified for G10, on average, 64% of students who are CoS-2010 (enrolled I, II, or III semester but not the V semester) out of the total number of students who met the expression 2.1 (a column (I+II+III)/ Total, H2 - Table 1). The data are similar for G07, G08 and G09: 76%, 76% and 63%. At the same time, the identified students by the expression 2.1 are, on average, 24% of students of CoS-2010 (Table 1, a column (I+II+III)/ ΣCoS2010). Because the percentage of successfully identified students CoS-2010 based on the criterion $C0_{limit}$ is high, it can be started with the identified students, with the implementation of additional, educational activities in order to reduce abandonment of the studies.

5.2 **Hypothesis H3: CoS-year students**

The same analysis, as for C0 data, was also conducted for C1 data and hypothesis H3. Based on the trends for C1 data, $C1_{limit}$ criterion defined by the expression 3.0 was modified for two reasons:

1) the value of standard deviation σ_1 (Figure 5) is, in some cases, higher compared to the arithmetic mean μ_1 (Figure 3), i.e. $\sigma_1 > \mu_1$, so that the proposed criterion $C1_{limit}$ can have a negative value (minimum is 0 points),
2) the value of C1 may be a whole number or value = an integer (whole number) + 0.5.

In a further analysis, the criterion $C1_{limit}$ was initially calculated by the expression 3.0 and corrected:

1) if it is $C1_{limit} \leq 1$, the value of the criterion $C1_{limit} = 1$,
2) if it is $C1_{limit} > 1$, the value of the criterion is rounded to the first higher value by the rule: the integer or integer + 0.5, depending on which value is closer to the calculated value.

Table 1. Number and percentage of students CoS-2010: hypotheses H2 and H3

Gene-ration	Study pro-grams	Semester I		II		III		V		**Total**		(I+II+III)/ Total		ΣCoS 2010	(I+II+III)/ ΣCoS2010	
		H2	H3	H2	H3	H2	H3	H2	H3	**H2**	**H3**	H2	H3	H2 H3	H2	H3
2010	11	4	3	2	1	-	-	1	3	**7**	**7**	86%	57%	30	20%	13%
	21	1	2	5	4	3	1	5	13	**14**	**20**	64%	35%	31	29%	23%
	31	3	4	3	2	-	-	2	7	**8**	**13**	75%	46%	24	25%	25%
	41	3	4	2	3	1	-	4	7	**10**	**14**	60%	50%	23	26%	30%
	52	2	5	1	4	2	3	10	19	**15**	**31**	33%	39%	28	18%	43%
	61	7	6	6	8	2	3	8	17	**23**	**34**	65%	50%	53	28%	32%
										Average		**64%**	**46%**		**24%**	**28%**

Such a corrected criterion is the beginning for hypothesis H3 and represents the criterion $C1_{limit}$ in expression 3.1. On the basis of the expression 3.1, the number of students from the generation (CoS-2010) was determined, the students who met it and it is given in Table 1. The established criterion $C1_{limit}$ identified, on average, 46% of students for G10 who are CoS-2010 (they enrolled I, II, or III semester but not the V semester) out of the total number of students who met the expression 3.1 (a column (I+II+III)/Total, H3 - Table 1). Similar are the data for G07, G08 and G09: 69%, 52% and 57%. At the same time, the identified students are, on average, 28% of students of CoS-2010 (Table 1, a column (I+II+III)/ΣCoS2010). The percentage of successfully identified students CoS-2010 based on the criterion $C1_{limit}$ is lower than the criterion $C0_{limit}$, but still significant, so that on the basis of this criterion also, one can start with improvements with the identified students and with the implementation of additional, educational activities.

5.3 **Hypothesis H4: CoS-year students**

The research continues in regard to the hypothesis H4: combinations (union and intersection) of hypotheses H2 and H3, i.e. expressions 2.1 and 3.1 (Table 2).

Table 2. Number and percentage G10, H4: Union H2+H3 and Intersection H2∩H3

Gene-ration	Study pro-grams	Semester I		II		III		V		**Total**		(I+II+III)/ Total		ΣCoS 2010	(I+II+III)/ ΣCoS2010	
		H2+ H3	H2∩ H3	H2+ H3	H2∩ H3	H2+ H3	H2∩ H3	H2+ H3	H2∩ H3	**H2+ H3**	**H2∩ H3**	H2+ H3	H2∩ H3		H2+ H3	H2∩ H3
BTC 2010	11	7	-	3	-	-	-	4	-	**14**	**-**	71%	0%	30	33%	0%
	21	3	-	7	2	4	-	17	1	**31**	**3**	45%	67%	31	45%	6%
	31	5	2	3	2	-	-	8	1	**16**	**5**	50%	80%	24	33%	17%
	41	5	2	5	-	1	-	10	1	**21**	**3**	52%	67%	23	48%	9%
	52	5	2	4	1	4	1	23	6	**36**	**10**	36%	40%	28	46%	14%
	61	8	5	12	2	5	-	21	4	**46**	**11**	54%	64%	53	47%	13%
										Average		**52%**	**53%**		**42%**	**10%**

The percentage of students who can be identified if one of the hypothesis is met, H2 or H3 (2.1 or 3.1) or both hypotheses together, H2 and H3 (hereinafter referred to as: Union H2 + H3), does not represent the sum of the previously listed percentages because there are students who were identified by both hypotheses (both H2 and H3). The union H2+H3 (Table 2) identifies, on average, 52% of students for G10 who are CoS-2010 out of the total number of students who met one and/or both hypotheses (a column (I+II+III)/Total, Table 2). At the same time, the identified students are 42% of the students of CoS-2010, (Table 2, a column (I+II+III)/ ΣCoS 2010), while the percentages for G07, G08 and G09 are: 48 %, 39% and 44%. Percentage of students who were identified by both hypotheses together, H2 and H3 (expressions 2.1 and 3.1, hereinafter referred to as: Intersection H2∩H3), for G07-G10: 87%, 81%, 80% and 53% while the percentages compared to the total number CoS-2010 are: 9%, 10%, 8% and 10%.

5.4 **The number of CoS-year students**

It is necessary to determine the percentage of identified students as per the above hypotheses (H2, H3, H4) in relation to the Start Gyear (Fig. 6). If this percentage is high, then it is better to implement in practice more educational activities over the entire generation or to look for another, a better quality criterion meaning an adequate number of identified students.

Figure 6 shows that, when the identification students (H4) is applied: Union H2 + H3, this percentage for G10 is in a range of 27.6% to 37.9%, an average of 33.9%. The average for G07, G08 and G09 is: 37.2%, 32.9% and 32.1%. The above percentages are high (about and above 1/3) and it is not rational for perform additional educational activities.

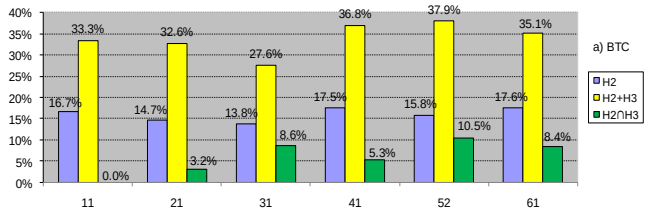

Figure 6. Percentage of identified CoS-2010 compared to the Start G10: H2, H3, H4 (H2+H3 and H2∩H3)

Percentage of identified students for the hypothesis H2 compared to the start G10 is within the limits: 13.8% to 17.6%, on average 16.0%. The average for G07, G08 and G09 is 19.4%, 15.5% and 17.5%. Percentage of the identified students by the criterion: intersection H2∩H3, compared to the start G10 for is within the limits: 0.0% to 10.5%, on average, 6.0%. The lowest percentage of identified CoS-2010 compared to the Start G10 is obtained by applying the hypothesis H4, the criterion: Intersection H2∩H3 and the highest percentage by using the criterion: Union H2+H3.

6. CONCLUSION

Previous analyses (Chapters 4 and 5) contain a part of the discussion on the results because it is impossible to create reports without emphasizing the meaning and significance of certain results of the report.

The set objective of the research: to define clusters of students it was achieved with the help of proven hypotheses H1-H4: identified CoS-year students. With the identified students, one can start the implementation of additional educational activities in order to reduce the abandonment from the studies. A limiting factor is in the number of identified CoS-year students. Hypothesis H4, a criterion: Intersection H2∩H3 has the lowest percentage of identified students in comparison to the Start Gyear and it is important to point out, the percentage (%) of successfully identified CoS-year students is high, about 1/2 (H2 has a higher average). Criterion: Intersection H2∩H3 is the basis on which it is possible and necessary with the identified students to start up additional educational activities. Out of the remaining hypotheses, the analysis in Section 5.4 shows that the hypothesis H2 is sustainable for additional educational activity. The number of identified CoS-year students as per hypotheses H3 and H4, the criterion: Union H2+H3 suggests starting up educational activity or modification of the criteria established in hypotheses H3 and H4. Also, the derived proofs of the hypotheses H1-4 suggest that it is necessary to include the data with which the higher education institution disposes about the students during the enrollment process in the first semester in the process of data mining from the e-learning system.

The results of the research represent a significant percentage of identified students to whom the higher education institution must pay an additional attention and to direct some additional educational activities. The methodology presented in Section 3 - Methodology, together with the presented criteria is independent of higher education institutions and the nation, and is applicable to any higher education institution.

7. REFERENCES

[1] Hsinchun, C., Chiang, R. H. L. and Storey, V. C. (2012), *"Business Intelligence And Analytics: From Big Data to Big Impact*", MIS Quarterly, Vol. 36, No. 4, pp. 1165-1188.

[2] Vavpotic, D. and Vasilecas, O. (2012), *"Selecting a Methodology for Business Information Systems Development: Decision Model and Tool Support"*, Computer Science and Information Systems, Vol. 9, No. 1, pp. 135-164.

[3] Brandon, A. (2011), "*The Future of Business Intelligence is in Education. Trends and Outliers"*, available at: https://www.tibco.com/blog/2011/06/16/the-future-of-business-intelligence-is-in-education/ (accessed: 30 june 2017)

[4] Chaudhuri, S., Dayal, U. and Narasayya V. (2011), *"An Overview of Business Intelligence Technology*", Communications of the ACM, Vol. 54, No. 8, pp. 88-98.

[5] Negash, S. (2004), *"Business Intelligence",* Communications of the Association for Information Systems, Vol. 13, pp. 177-195.

[6] Mekterovic, I., Brkic, Lj. and Baranovic, M. (2009), *"Improving the ETL process and maintenance of Higher Education Information System Data Warehouse",* WSEAS Transactions on Computers, Vol. 8, No. 10, pp. 1681-1690.

[7] Javed, A., Shaikh, M. U. and Bhatti, B. M. (2008), *"Conceptual Model for Decision Support System Based Business Intelligence OLAP Tool for Universities in Context of E-Learning",* In Proc. of the World Congress on Engineering and Computer Science 2008, San Francisco/USA.

[8] Pah, I., Oprean, C., Moisil, I. and Kifor, C. (2008), *"Technology to Support Education Software Solutions for Quality Assurance in e-learning*", International Journal of Computers, Communications & Control, Vol. 3, No. 3, pp. 433-436.

[9] Camilovic, D., Becejski-Vujaklija, D. and Gospic, N. (2009), *"A Call Detail Records Data Mart: Data Modelling and OLAP Analysis*", Computer Science and Information Systems, Vol. 6, No. 2, pp. 87-110.

[10] Golfarelli, M., Mandreoli, F., Penzo, W., Rizzi, S. and Turricchia, E. (2012), *"OLAP query reformulation in peer-to-peer data warehousing*", Information Systems, Vol. 37, No. 5, pp. 393–411.

[11] Bana e Costa, C. A. and Oliveira, M. D. (2012), *"A multicriteria decision analysis model for faculty evaluation*", Omega, Vol. 40, No. 4, pp. 424–436.

[12] Kleesuwana, S., Mitathab, S., Yupapin, P. P. and Piyatamrongd, B. (2010), *"Business Intelligence in Thailand's Higher Educational Resources Management*", Procedia - Social and Behavioral Sciences, Vol. 2, No. 1, pp. 84–87.

[13] Piedade, M. B. and Santos, M. Y. (2009), *"Business Intelligence supporting the teaching-learning process*", In Proc. of the 9th WSEAS International Conference on Simulation, Modelling And Optimization, Budapeast/Hungary 2009, pp. 292-297.

[14] Blagojevic, M. and Micic, Z. (2012), *"A web-based intelligent report e-learning system using data mining techniques*", Computers & Electrical Engineering, Vol. 39, No. 2, pp. 465-474.

[15] Lara, J. A., Lizcano, D., Martínez, M. A, Pazos, J. and Riera, T. (2014), *"A system for knowledge discovery in e-learning environments within the European Higher Education Area – Application to student data from Open University of Madrid, UDIMA"*, Computers & Education, Vol. 72, pp. 23–36.

[16] Luan, J. (2002), *"Data Mining and Its Applications in Higher Education",* New Directions for Institutional Research, Vol. 2002, No. 113, pp. 17–36.

[17] Castellano, E. J. and Martínez, L. A. (2009), *"Web-Decision Support System based on Collaborative Filtering for Academic Orientation. Case Study of the Spanish Secondary School"*, Journal of Universal Computer Science, Vol. 15, No. 14, pp. 2786-2807.

[18] Delen, D. (2010), *"A comparative analysis of machine learning techniques for student retention management",* Decision Support Systems, Vol. 49, No. 4, pp. 498–506.

[19] Heise, D. L. (2005), *"Data warehousing and decision making in higher education in the United States*", available at: http://www.andrews.edu/leaderpart/roundtable/2005/joust2/hd/dissprop.pdf, (accessed: 30 june 2017)