CORRELATION ANALYSIS OF EDUCATIONAL DATA MINING BY MEANS A POSTPROCESSOR'S TOOL

Georgi Teodorov, Oktay Kir, Irina Zheliazkova

Abstract: *The paper deals with the correlation analysis as educational data technique that is easy to interpret and simple to implement. Two datasets respectively from environment for knowledge testing and for exercise tasks modelling testing are gathered. Programming of tasks for test parameters relationships, test reliability, cheat recognition, and test validation in a specialized postprocessor tool is discussed .*

Keywords: *data mining, correlation analysis, dataset, test reliability, test validity, cheap recognition, postprocessor tool*

ACM Classification Keywords: *Computer and Information Science Education, Knowledge Representation*

Introduction

In the recent review of Romero & Ventura [6] correlation analysis has been pointed out as one of the Educational Data Mining (*EDM*) techniques for extracting useful information to support reasonable decisions making in the educational environments. In table 1 all kinds of tasks (from A to K) to which this technique has been applied are listed. It is not surprising that teachers widely use the correlation analysis as it's easy to interpret similar to the descriptive statistics and simpler for computation in comparison with other known techniques as neural, Bayesian, and Kohenen networks, rule-based systems, cluster and regression analysis. The main requirements for design of the *EDM* tools also are formulated in the same survey and concern: the user interface, visualization task, integration of the tool with an educational environment, standardization of data and models, as well as algorithms for data mining.

An earlier paper of Hernandez et al. [1] deals in depth with the task F and more precisely with cheat in online testing. There a questionnaire study is cited concerning the students' cheat and conducted by Donald McCabe. In a representative sample of 1,800 students from nine state universities in USA, seventy percent of students admitted to cheate on exams. As a result five reasons for this undesirable student's behavior were discovered, namely: lazy or didn't study or prepare, to pass a class or improve a grade, external pressure to succeed, didn't know answers, time pressure or too much work. In the above-mentioned paper Genderman, who founded four main factors associated with academic dishonesty (individual characteristics, peer group influences, instructor influences, and institutional policies) also is cited. In the everyday teaching practice some students even become masters in the art of cheating.That is why it is interesting to analyze the student's abnormal behavior and compare it with the normal one.

Table 1. The groups of tasks for EDM using correlation analysis technique

	Objective
A. Analysis and Visualization of Data	to highlight useful information and support decision making. In the educational environment, for example, it can help educators and course administrators to analyze the students’ course activities and usage information to get a general view of a student’s learning.
B. Providing Feedback for Supporting Instructors	to provide feedback to support course authors/teachers/administrators in decision making (about how to improve students’ learning, organize instructional resources more efficiently, etc) and enable them to take appropriate proactive and/or remedial action.
C. Recommendations for Students	to be able to make recommendations directly to the students with respect to their personalized activities, links to visits, the next task or problem to be done, etc, and also to be able to adapt learning contents, interfaces, and sequences to each particular students.
D. Predicting	to estimate the unknown value of a variable that describes the student. In education, the values normally predicted are performance, knowledge, score, or mark. This value can numerical/continuous value (regression task) or categorical/discrete value (classification task).
E. Student Modeling	to develop cognitive models of human users/students, including a modeling of their skills and declarative knowledge. Data mining has been applied to automatically consider user characteristics (motivation, satisfaction, learning styles, affective status, etc.) and learning behavior in order to automate the construction of student models.
F. Detecting Undesirable Student Behaviors	to discover/detect those students who have some type of problem or unusual behavior such as: erroneous actions, low motivation, playing games, misuse, cheating, dropping out, academic failure, etc.
G. Grouping Students	to create groups of students according to their customized features, personal characteristics, etc. Then the clusters/groups of students obtained can be used by the instructor/developer to build a personalized learning system to promote effective group learning, to provide adaptive contents, etc.
H. Social Network Analysis	Social networks analysis, aims at studying relationships between individuals, instead of individual attributes or properties. A social network is considered to be a group of people, an organization or social individuals who are connected by social relationships like friendships, cooperative relations, or informative exchange.
I. Developing Concept Maps	to help instructors/educators in the automatic process of developing/constructing concept maps. A concept map is a conceptual graph that shows relationships between concepts and expresses the hierarchal structure of knowledge.
J. Constructing Courseware	to help instructors/development process of courseware and learning contents automatically. On the other hand, it also tries to promote the reuse/exchange of existing learning resources among different users and systems.
K. Planning and Scheduling	to enhance the traditional educational process by planning future courses, helping with student course scheduling, planning resource allocation, helping in the admission and counseling processes, developing curriculum, etc.

For detecting students cheats in on-line exams Hernandez et al. [1] proposed to use Knowledge Discovery in Databases (*KDDs*) a non-trivial process of identifying valid, novel, potentially useful, and ultimately understandable patterns in databases. They viewed the *EDM* simply as an essential step in the process of *KDDs* and use *WEKA* as a Data Mining Engine *(DME*). *WEKA* contains tools for data preprocessing, classification, regression, clustering, association rules, and visualization. The machine learning algorithms can either be applied directly to a dataset or be called from the researcher own Java code. It is also well-suited for developing new machine learning algorithms. At the same time it's easy to use and understandable and provides a comprehensive environment for testing methods against other existing methods.

To solve this *DM* task Jelev et al. applied the descriptive statistics and visualization techniques on a dataset with test results. The test was extracted by a database containing 150 questions covering the lecture material for the course “Programming structures” and created by means of the popular during the 90's multi-media environment ToolBook. The test included 35 multiple-choice questions each one with 5 alternatives only one of which is the correct answer. The goal of the first study [4] was to determinate the test validity using one-factor analysis with Fisher's criterion. Each of 6 test versions was analyzed according to the correct answers students gave to 7 randomly selected questions (one from each topic taught). The examination of the test difficulty through one-factor analysis shown that the different test versions do not pose a significant influence on the grades, e.g. the test is a good means for the knowledge testing. The intervals of correct answers were determined to correspond to a six-point marking scale as 10 scores to correspond to the mark “poor”. That threshold presents 0.35 from the total test scores is close to that one (0.40) accepted by Zheliazkova's group. In the second study [5] the scores distribution in percentage of the experimental curve of the student's marks and the theoretical curve representing a normal Gauss distribution were visualized in a common coordinate system. The conclusion was made that the two curves have approximately the same distributions with mean equal to the “good” mark (4.00). The probability for a student to give a definite number of correct answers under full lack of knowledge or random choice of answers was calculated applying the Bernoulli formula. The obtained graphical result of this probability shown that the case of 7 correct answers has the highest probability. For analysis of the dispersion measures the sigma derivation method was applied to assess if the sample of 35 questions is representative. The obtained result was approximately 33 questions as this number covers with 90% guarantee the lecture material.

Authors' Team Previous Studies

Since 2006 Zheliazkova's research group has been using different *DM* techniques for postprocessing students' tests and exercise results. Two experimental studies had been conducted to assess the effectiveness of the intelligent computer-based tests in comparison with the traditional ways of testing such as multiple-choice tests, and exams [8,9]. Objects of the studies were students-bachelor (1-st year, 2-nd semester), specialties Computer Systems and Technologies (CST) and Communication Technique and Technologies (CTT) at Rousse University. A multiple-choice test (*T1*) and an intelligent test (*T2*) covering the theme “Algorithms” from the subject “Programming 1” were generated by means of a specialized environment for knowledge testing. As a tool for postprocessing the gathered datasets was used Excel. The relationship between the exam mark given previous semester by the lecturer (*M3*) and the intelligent test mark (*M2*) in traditional six-grade scale was found to be high while the relationship between *M1* and *M3*, as well as between *M1* and *M2* was moderate. Another interesting relationship between the time undertaken and *M1* and time and *M2* was calculated as lower. According to *M3* the experimental data was divided in two tables respectively with data for the students with mark “5” or “6” and for the students with mark “3” or “4”. That was made because at the end of the previous semester the students from the first group were assessed by the lecturer and released from the exam. Note that, the values of *r (time, M1*) close to 0 confirm the statement that is not objective due to some well-known reasons in contrast to *r*(*time*,*M2*) had a positive value, greater for the first group of students than for the second one.

In a more recent study of Zheliazkova’s group [10] another interesting relationship that between the test mark (*TM*) and exercise mark (*EM*) in the traditional scale also had been investigated. Two datasets were gathered from two specialized environments respectively for knowledge testing and modeling dynamic systems used in the authors’ team teaching practice. Again Excel was used as a tool for postprocessing. The value of *r* (*TM, EM*) was 0.44 that means a moderate relationship. The conclusion made was that both environments are feasible and yield to sustainable and valid results. Probably the reasons why this value was not higher is that the exercise tasks for modeling were more complex and the students used a new software environment to perform them. It took most of students 2-3 times longer to complete the first task for modeling than to complete the following ones. So, unavoidably, the exercise performed within the above-mentioned environment beside the subject specific knowledge partly measures also technological skills to use this environment.

A teacher’s tool implemented by Zheliazkova’s group for the *EDM* called postprocessor was reported from design, implementation, and user’s points of view. For ensuring the tool’s intelligence and its adaptation to the teacher a power and expressive script language called *SessionScript* was implemented. Programming of descriptive statistics, visualization, and correlation analysis techniques was demonstrated using two output data sets respectively from both above-mentioned environments. Application of the linear methods of prediction using this tool is reported in another paper submission for the present conference [3]. The description of the experimental dataset and technology of the tool using can be found there.

The present paper deals with the same experimental dataset and the same tool that’s why their descriptions are omitted here. The next paper sections focus on the correlation analysis application respectively for the following *DM* tasks: test parameters relationships, test reliability, cheat recognition, and test validation. Conclusion summarizes the methodology proposed for their application using the postprocessor.

Correlation Analysis for the Test Parameters Relationships

The coefficient of correlation analysis r_{XY} can serve as a qualitative indicator for the relationship between two statistical test’s parameters, for example, X and Y with the number of the questions n and the mean indicators respectively m_x, m_y $r_{XY} = \sum_{i=1}^{n}(x_i - m_x).(y_i - m_y) / \sqrt{\sum_{i=1}^{n}(x_i - m_x)^2 . \sum_{i=1}^{n}(y_i - m_y)^2}$. The value of r_{XY} shows how strong is the relationship between the considered parameters and changes in the range from -1.00 to +1.00. In order to move from its concrete value to more clear for a non-skilled teacher (T) a five-intervals scale with a linguistics value is used. For example, if r_{XY} is in the range 0.0 ÷ 0.3 then the relationship is low; 0.3 ÷ 0.5 – moderate; 0.5 ÷ 0.7 – significant; 0.7 ÷ 0.9 – high; 0.9 ÷ 1.0 – very high. If two parameters are moving in the same way r = +1.0 and if in the opposite r = -1.0. The value 0 means that there is no relationship between the considered parameters.

The test mark in the traditional six-grades scale was computed only on the base of correct knowledge scores (P) as follows: $0 \leq P \leq 0.4^* P_{max}$ – “2”; $0.4^* P_{max} < P \leq 0.55^* P_{max}$ – “3”; $0.55^* P_{max} < P \leq 0.70^* P_{max}$ – “4”; $0.70^* P_{max} < P \leq 0.85^* P_{max}$ – “5”; $0.85^* P_{max} < P \leq 1.0^* P_{max}$ – “6” where P_{max} = 352 was the total test scores. The experience accumulated during the last decade by Zheliazkova’s research group has pointed out that such a non-linear scale gives a Gauss distribution of the student’ marks and is acceptable by both teachers and students.

Correlation between the student’s mark and time undertaken for test performance is one of the very interesting relationships and at the same time not well studied. The time planned for the test performance was T_{max} = 120 min but the students were told that the time for the test performance actually is unlimited and together with wrong and missing knowledge will be used as assessment indicators only for research purpose. The input table 2 contains test mark (M) and time undertaken (T) for the “very good” students, e.g. with mark “6” or “5”on the base of their correct knowledge. The coefficient of correlation calculated was 0.21, e.g. that means low correlation between the considered parameters.

Table 2. The test mark and time undertaken for the “very good” students

TRANS_STUDENTS

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
M	6	6	6	6	6	6	6	5	5	5	5	5	5	5	5	5	5	5
T	150	160	109	150	144	146	140	120	130	122	160	160	110	151	120	140	145	160

	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
M	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
T	115	108	115	120	110	130	120	160	150	150	155	160	150	112	119	110	118	129

Table / Description

The input table 3 contains the values of M and T for the “good” students, e.g. received mark “4” or “3”. For this part of the students the coefficient of correlation was 0.02 that means no correlation between the considered parameters. The total coefficient of correlation computed was equal to 0.36 that means low rather than moderate relationship.

Table 3. The test mark and time undertaken for the “good” students

TRANS_STUDENTS

	1	2	3	4	5	6	7	8	9	10	11	12	13
M	4	4	4	4	4	4	4	4	4	4	4	3	3
T	105	96	125	106	140	111	150	150	90	116	135	120	119

Table / Description

Note, that the total number of students decreased from 63 to 49 due to unwilling of some students to register T in their Word documents. Among them was the single student received the mark “poor”. It is seen from the table 3 that only two students received the mark “satisfactory”. These test results also confirm the finding that the students believe in the objective and precise test assessment and go to test only if they assessed themselves at least with mark “satisfactory”.

Correlation Analysis for the Test Reliability

The idea for application of the correlation analysis for the test reliability belongs to Savelev et al., 1986 [7] under the assumption that all other factors are constant, and a longer test will be probably more reliable than a shorter one the indicators of the answers of the even and odd questions and the coefficients of their rank correlation have to be computed.

Table 4. Test results of the ST4

STUDENT 063145

	QUESTIONS	ODD	EVEN	PERCENT1	PERCENT2	RANK1	RANK2	D	D2
1	Q1 Q2	10	15	100	100	10	9	1	1
2	Q3 Q4	10	7	100	100	10	9	1	1
3	Q5 Q6	5	13	100	100	10	9	1	1
4	Q7 Q8	4	16	80	100	2.5	9	6.5	42.25
5	Q9 Q10	10	10	100	100	10	9	1	1
6	Q11 Q12	3	10	100	100	10	9	1	1
7	Q13 Q14	10	15	100	100	10	9	1	1
8	Q15 Q16	3	12	80	100	2.5	9	6.5	42.25
9	Q17 Q18	10	12	100	100	10	9	1	1
10	Q19 Q20	12	6	85	60	4	1.5	2.5	6.25
11	Q21 Q22	10	16	100	100	10	9	1	1
12	Q23 Q24	10	10	100	100	10	9	1	1
13	Q25 Q26	10	11	100	100	10	9	1	1
14	Q27 Q28	12	10	100	100	10	9	1	1
15	Q29 Q30	5	6	33	60	1	1.5	0.5	0.25

Table / Description

Table 5. Test results of the ST1

STUDENT 063156

	QUESTIONS	ODD	EVEN	PERCENT1	PERCENT2	RANK1	RANK2	D	D2
1	Q1 Q2	10	15	100	100	12.5	10	2.5	6.25
2	Q3 Q4	8	7	80	100	9	10	1	1
3	Q5 Q6	3	13	60	100	4.5	10	5.5	30.25
4	Q7 Q8	3	16	60	100	4.5	10	5.5	30.25
5	Q9 Q10	10	10	100	100	12.5	10	2.5	6.25
6	Q11 Q12	2	10	67	100	7	10	3	9
7	Q13 Q14	10	2	100	13	12.5	1	11.5	132.25
8	Q15 Q16	2	12	50	100	2	10	8	64
9	Q17 Q18	10	12	100	100	12.5	10	2.5	6.25
10	Q19 Q20	10	6	71	60	8	2.5	5.5	30.25
11	Q21 Q22	6	14	60	87	4.5	4	0.5	0.25
12	Q23 Q24	10	10	100	100	12.5	10	2.5	6.25
13	Q25 Q26	6	11	60	100	4.5	10	5.5	30.25
14	Q27 Q28	12	10	100	100	12.5	10	2.5	6.25
15	Q29 Q30	5	6	33	60	1	2.5	1.5	2.25

Table / Description

Table 6. Test results of the ST2

STUDENT 063160

	QUESTIONS	ODD	EVEN	PERCENT1	PERCENT2	RANK1	RANK2	D	D2
1	Q1 Q2	10	15	100	100	13	12	1	1
2	Q3 Q4	6	7	60	100	8	12	4	16
3	Q5 Q6	2	7	40	53	6	2	4	16
4	Q7 Q8	2	16	40	100	6	12	6	36
5	Q9 Q10	10	8	100	80	13	6	7	49
6	Q11 Q12	3	2	100	20	13	1	12	144
7	Q13 Q14	8	15	80	100	9	12	3	9
8	Q15 Q16	4	10	100	83	13	8	5	25
9	Q17 Q18	10	12	100	100	13	12	1	1
10	Q19 Q20	12	6	85	60	10	3.5	6.5	42.25
11	Q21 Q22	4	16	40	100	6	12	6	36
12	Q23 Q24	0	8	0	80	2	6	4	16
13	Q25 Q26	2	11	20	100	14	12	2	4
14	Q27 Q28	0	8	0	80	2	6	4	16
15	Q29 Q30	0	6	0	60	2	3.5	1.5	2.25

Table / Description

Table 7. Test results of the ST3

STUDENT 063951

	QUESTIONS	ODD	EVEN	PERCENT1	PERCENT2	RANK1	RANK2	D	D2
1	Q1 Q2	10	15	100	100	12.5	12.5	0	0
2	Q3 Q4	10	7	100	100	12.5	12.5	0	0
3	Q5 Q6	3	6	60	46	7	6	1	1
4	Q7 Q8	0	12	0	75	3.5	7	3.5	12.25
5	Q9 Q10	10	8	100	80	12.5	8	4.5	20.25
6	Q11 Q12	10	10	100	100	12.5	12.5	0	0
7	Q13 Q14	8	15	80	100	9	12.5	3.5	12.25
8	Q15 Q16	3	12	75	100	8	12.5	4.5	20.25
9	Q17 Q18	10	0	100	0	12.5	3	9.5	90.25
10	Q19 Q20	0	0	0	0	3.5	3	0.5	0.25
11	Q21 Q22	0	0	0	0	3.5	3	0.5	0.25
12	Q23 Q24	0	0	0	0	3.5	3	0.5	0.25
13	Q25 Q26	0	9	0	81	3.5	9	5.5	30.25
14	Q27 Q28	10	10	100	100	12.5	12.5	0	0
15	Q29 Q30	0	0	0	0	3.5	3	0.5	0.25

Table / Description

Table 8. The generated table

STUDENTS

	STUDENT	E	CORR	H
1	ST1	6	0.742	0.889
2	ST4	3	0.609	0.798
3	ST2	5	0.006	0.524
4	ST3	4	-0.045	0.267

Table / Description

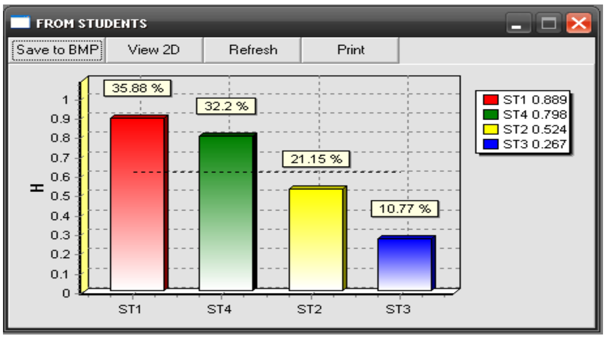

Fig. 1. The bar diagram of the Spearman-Brown coefficient

The reliability of test is measured applying the formula of Spearman-Brown: $H = 2.r_{XX} / (1 + r_{XX})$, where $r_{XX} = 1 - 6.\sum_{i=1}^{n} (x_i^{'} - x_i^{"})^2 / (n^3 - n)$. It is accepted that the test reliability is enough if $H > 0.8$. For a student’s test performance the r_{XX} between the odd and even test questions (CORR) and the Spearman-Brown coefficient (H) one are given in table 8. For the first and fourth student the test is reliable enough. For the second and third students the test turned to be non-reliable. This can be explained with the fact that the test was oriented to the “4” students supposed to be a substantial part of all students. Obviously, for the “5” and “6” students, as well as the “3” students the test is more likely unreliable. The average of the H (approximately 0.6) depicted with a dotted line on fig. 1 shows that the test can be accepted as reliable.

Correlation Analysis for the Cheat Recognition

For cheat discovering Hernandez et al. [1] used a complex patterns recognition approach using test correct and inccorrect answers as dataset. The approach based on correlation analysis has three stages of proving the cheat - datasets with correct (table 9) , missing (table 10), and wrong knowledge (table 11). The observation during the test performance showed at least four students possibly attempted to cheat.

A visual comparison of the correlation between ST1, ST2, ST3 and ST4 is shown on tables 12, 13, and 14. The correlation of the ST4 with ST1, ST2, and ST3 is close to moderate that is a normal student. The highest correlation of 1.00 is between ST2 and ST3 which means one of them could be the test answers source.

Table 9. The test questions correct knowledge of four students possibly attempted to cheat

TRANS_STUDENTS

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
ST1	10	15	10	7	4	13	4	16	10	10	3	10	8	8	4	12	8	8	14	8	0	16	10	10	0	11	0	0	0	0
ST2	10	15	10	7	4	13	4	16	10	10	3	10	8	13	4	12	10	12	14	8	0	16	0	0	0	11	0	0	0	0
ST3	10	15	10	7	4	13	4	16	10	10	3	10	8	13	4	12	10	12	14	8	0	16	0	0	0	11	0	0	0	0
ST4	10	9	8	7	3	13	3	16	10	10	3	10	8	15	3	10	10	12	0	8	6	16	10	10	10	11	12	10	5	6

Table 10. The test questions missing knowledge of four students possibly attempted to cheat

TRANS_STUDENTS

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
ST1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0
ST2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0
ST3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ST4	0	6	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	14	2	4	0	0	0	0	0	0	0	10	4

Table 11. The test questions wrong knowledge of four students possibly attempted to cheat

TRANS_STUDENTS

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
ST1	0	0	0	0	0	0	1	0	0	0	0	0	2	2	0	0	0	10	0	2	10	0	0	0	10	0	12	10	15	10
ST2	0	0	0	0	1	0	1	0	0	0	0	0	2	2	0	0	0	0	0	0	10	0	10	10	10	0	12	10	15	10
ST3	0	0	0	0	1	0	1	0	0	0	0	0	2	2	0	0	0	0	0	2	10	0	10	10	10	0	12	10	15	10
ST4	0	0	0	0	2	0	2	0	0	0	0	0	2	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0

The ST3 has zero missing knowledge for all questions answers (table 10) and due to the error "division by zero" when calculating the correlation so he/she is omitted in table 13. To confirm this finding the lecturer had a talk with these students and assessed them with the test mark "3" for their honesty.

Table 12. Correct knowledge case

CORRELATION ANALYSIS

	ST1	ST2	ST3	ST4
ST1	1	0.8566	0.8566	0.4156
ST2		1	1	0.413
ST3			1	0.413
ST4				1

Table 13. Missing knowledge case

CORRELATION ...

	ST1	ST2	ST4
ST1	1	-0.0345	-0.0805
ST2		1	0.0345
ST4			1

Table 14. Wrong knowledge case

CORRELATION ANALYSIS

	ST1	ST2	ST3	ST4
ST1	1	0.7744	0.7768	-0.1996
ST2		1	0.9972	-0.2002
ST3			1	-0.2076
ST4				1

Correlation Analysis for the Test Validation

For the test validation correlation between the test mark (*M1*) and exercise mark (*M2*) given by the assistant at the end of the semester was computed. He/she had been told to assess each student's activity during each exercise (their number was 15) in the traditional six-grade scale. That information was brought in an Excel table and *M2* of each student was computed as an average mark with accuracy 0.25. It is assumed that the difference between both marks hasn't to exceed one interval for the students with normal behavior. The difference between both marks would exceed one interval for the students with low *M2* probably used the common device for cheat.

The results for both groups of students are shown in the input tables 15 and 16 respectively in which the last column (30 and 15) contains the corresponding average mark. Note, that the total number of the students decreased from 63 to 43 as for some students *M2* missing. A positive tendency also is confirmed that for the intelligent tests the average mark is shifted from "good" to "very good". The mean of the test mark close to "excellent" (5.35) against that for the exercise (3.71) confirms that these students have abnormal behavior.

Table 15. The test and exercise mark for the students with normal behavior

TRANS_STUDENTS

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
M1	6	6	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	4	4	4	4	4	4	4	4	3	3	4.65
M2	5.57	5.07	6	5.5	5.54	5.17	5.29	4.33	4.29	5.15	4.57	5.86	4.57	4.43	4	5.57	4.57	5	4	4	5.67	5.33	4.33	3.5	4.14	4.93	4.93	3.86	4.14	4.8038

Table Description

Table 16. The test and exercise mark for the students with abnormal behavior

TRANS_STUDENTS

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
M1	6	6	6	6	6	5	5	5	5	5	5	5	5	5	5.35
M2	4	4.5	4.8	3.86	4.43	3.5	3.67	3.5	3	3.75	3	3.9	3.6	2.4	3.7073

Table Description

There is a small difference (0.15) between *M1* and *M2* for the first group of students while as this difference for the second group is substantial (1.65). The coefficient of correlation for the first group *r (M1,M2)* = 0.43 is close to those in our previous study [11] and for the second group is 0.74 is much greater.

Conclusions

The correlation analysis for the test parameters relationships, test reliability, cheat recognition, and test validation has been applied using a specialized tool based on two data sets respectively for knowledge testing and exercise performing. Programming these tasks is simpler and easy to interpret by the educators. Some findings in this study are in line with some previous studies.

The following methodology for using the tool for such group of tasks is proposed: 1) Constructing the input table with columns equal to the students with normal behaviour and rows to their test and exercise marks; 2) Adding new column with the average values of both mark; 3) Calculating the correlation between these marks; 4) Repeating 1, 2, and 3 for the group of students with abnormal behaviour (if it exists);

The authors are grateful to all students participated in this risky study with help of which it becomes possible.

Bibliography

[1] Hernandez J. A., Ochoa A.,, Munoz J., Burlak G., Detecting Cheats in online students assessments using Data Mining, Conference on Data Mining, 2006, pp 204-210.

[2] Kir O., Zheliazkova I. I., Teodorov G., Educational Data Mining by Means of a Power Instructor’s Tool, Proceedings of International Conference on Entrepreneurship, Innovation, and Regional Development (ICEIRD), 2011 (Accepted).

[3] Kir O., Zheliazkova I. I., Prediction of Educational Data Mining by Means of a Postprocessor Tool, International Conference “Modern E-Learning”, Varna, 2011 (Accepted).

[4] Jelev G., Minkovska D., Approaches for Definition the Validity of the Results of the Test for Knowledge Mastering, International Conference on Computer Science, 2004, Sofia, pp. 268-273.

[5] Jelev G., Minkova Y., Determination of Representative Sample Size and Knowledge Assimilation Test Results Processing. Problems and Discussion, International Conference on Computer Science, 2004, Sofia, pp. 274-279.

[6] Romero Cr., Ventura S. Educational Data Mining: A Review of the State of the Art, IEEE Transactions on Systems, Man, and Cybernetics – Part C: Applications and Reviews 2010; 40/6: 601-18.

[7] Savelev A. Ya., Novikov V.A., Lobanov Yu. I., Preparing Information for Automated Teaching Systems, Moscow, 1986.

[8] Zheliazkova I. I., Andreeva M. H., Kolev R. T., Knowledge Testing in Algorithms – An Experimental Study, International Conference “Modern E-Learning”, Varna, 2006, pp.55-62.

[9] Zheliazkova I. I., Kolev R. T., Andreeva M. H., Knowledge Testing in Algorithms by Means of a Word Technology, Proceedings of the Second National Conference on E-Leaning in the Higher Education, Kiten, Bulgaria, 2006, pp. 21-25.

[10] Zheliazkova I. I., Valkova P. A., Georgiev G. Todorov, A Computer-Based Technology for Processing and Visualization of Session’s Data, Int. Journal of Information Technologies and Control, 2011 (Accepted).

Authors' Information

Georgi Teodorov – PhD student, University of Rousse, Studentska street 8, Rousse 7017, Bulgaria; e-mail: georgi.t.georgiev@gmail.com

Oktay Kir – PhD student, University of Rousse, Studentska street 8, Rousse 7017, Bulgaria; e-mail: kir.oktay@gmail.com

Irina Zheliazkova – Associate Professor; University of Rousse, Studentska street 8, Rousse 7017, Bulgaria; e-mail: irina@ecs.ru.acad.bg