Knowledge Prediction of Different Students' Categories Trough an Intelligent Testing

Irina Zheliazkova[1], Oktay Kir[1], Adriana Borodzhieva[1]

[1]University of Ruse, Ruse, Bulgaria

Abstract – Student's modelling, prediction, and grouping have remained open research issues in the multi-disciplinary area of educational data mining. The purpose of this study is to predict the correct knowledge of different categories of tested students: good, very good, and all. The experimental data set was gathered from an intelligent post-test performance containing student's correct, missing, and wrong knowledge, time undertaken, and final mark. The proposed procedure applies consequently correlation analysis, simple and multiple liner regression using a power specialized tool for programming by the teacher. The finding is that the accuracy of the procedure is satisfactory for the three students' categories. The experiment also confirms some findings of other researchers and previous authors' team studies.

Keywords –Correct Knowledge; Missing Knowledge; Wrong Knowledge; Student's Prediction; Intelligent Testing.

1. Introduction

During the last decade a new multi-disciplinary research area called Educational Data Mining (EDM) has been established to support reasonable educator's decisions though extracting useful information from raw experimental data. Romero et al. [19] described step by step this process and applied the main data mining techniques in Moodle – one of the most popular course management systems. Later in an exhaustive survey Romero & Ventura [20] identified eleven kinds of EDM tasks among which student's modeling, prediction, and grouping. According to this survey correlation and regression analysis [17] have been applied to almost kinds of EDM tasks. This fact can be explained with easy interpretation, simple computation, table and graphical visualization of these techniques in comparison with more complex ones such as association rules, neural and Bayesian networks.

Researchers as Hernandez et al. [7] view the EDM only as an essential step of the knowledge discovery in databases. When data mining techniques are applied to educational data sets semantic information, e.g. knowledge can be used and/or extracted. This requires more effective than Microsoft Excel tools integrating domain knowledge into data mining algorithms. The design of tools for EDM has to meet specific requirements concerning user interface, visualization task, integration with an e-learning environment, standardization of data and models, as well as the algorithms for data mining. For detecting the students' cheats in on-line exams the above-cited authors used a specialized data mining engine WEKA in which machine learning algorithms can be applied directly to a data set or called by the teacher's Java code.

Student's modeling is one of the oldest EDM tasks used mainly for the system's adaptation to the individual students and/or for modeling different categories of students, e.g. basic, proficient, advanced [15]. Intelligent Tutoring Systems (ITSs) [22] spread in the 80's sup-ported a detailed and rich Student Model (*SM*) storing his/her knowledge, skills, attitudes, style of learning, and so on. In the last two decades different types of the SM appeared such as parametrical [21], overlay [4], fuzzy temporal [6], probabilistic [25], based on machine learning [14], temporal and permanent models [24], individual and group models [8].

Nowadays the most popular EDM techniques applied to check the student's knowledge improvement are descriptive statistics and visualization. The raw data more often used are the student's performance, knowledge, scores, and final mark. Usually they are gathered by means of a subject pretest and posttest for an experimental and control group.

Web put the beginning of a new educational technology that hasn't parents among the non-web teaching systems most of which support a domain-independent SM integrated in a relational database. Brusilovsky [2] called this technology Learner Model Matching (LMM) due to its opportunity to analyze and match at once many Students' Models (SMs). He also identified two kinds of this technology and respectively called them Adaptive Collaboration Support (ACS) and Intelligent Class Monitoring (ICM). The goal of the ACS technology is to use system's knowledge about different learners to form a matching group for different kinds of collaboration, for example, to form students' teams for practical exercises, work-shops, course projects, and so on. The goal of the ICM technology is to identify the learners who are essentially different from the others

due to variety of reasons, for example, progressing too fast/slow or passing much/less lecture material. The LMM technology requires shifting the research from e-learning to e-teaching environments offering power tools for support the teacher's pedagogy knowledge base [9,16].

The present study attempts to solve the task of Correct Knowledge (CK) prediction for three categories of students at three stages applying respectively correlation analysis, simple and three-parametrical regression analysis. In section 2 the theoretical background of intelligent testing put in the teaching practice during more than a decade is discussed. In section 3 the methods and materials of the current study are presented including: involved participants, gathered data set, applied methods, and a specialized tool for processing and visualization. In the next section the received results are presented in tables and graphically. The final section presents a discussion of these results and some conclusions from this author's team study.

2. Theoretical Background of Intelligent Testing

The theoretical background in intelligent testing described here has being applied during more than a decade in the teaching practice of Zheliazkova's research group [11, 12, 23, 27, 28]. Some findings in cognitive and pedagogy sciences concerning statistical analysis of the test results were confirmed and new ones were found experimentally.

The questions answers are modeled as graph with the following visualization accepted (Table 1). The graph nodes correspond to the sub-answers, e.g. atomic knowledge, and graph arcs to the default logical relationships (sequential, combinational, relational and so on) between them. The nodes/arcs visible during the student's answering are presented with solid circles/lines and the hidden nodes/arcs – with dotted circles/lines. Such graph models allow a precise and objective measurement of the following test question parameters: volume of knowledge, e.g. number of scores (Q_t), degree of system prompt (C_p), and the weights w_i of sub-answers (sa_i) ($i = 1, \dots, n$), where n is the number of the sub-answers. So, it is reasonable to compute Q_t as the weighted sum of the nodes number and the arcs number.

According to the notation the classical multiple-choice answer (row 1) presents an isolated graph with solid nodes equal to the checked answers (n), and the rest (m) nodes dotted. It is reasonable to accept for this simplest type of answer that Q_t is equal to n, $w_i = 1$ ($i = 1, \dots, n$), and C_p to 1.00.

In case of unordered keywords answer (row 2) chosen from the test dictionary with much more number of keywords m the corresponding graph also is isolated, so Q_t and w_i retains the same, but C_p can be accepted equal to 0.

The next type of question answer (row 3) presents n ordered keywords missing in a text, table, programming code, etc. and chosen by the student from the test dictionary. The adequate graph model is linear, $Q_t = n + (n - 1) = 2.n - 1 \cong 2.n$, $w_i = 2$ ($i = 1, \dots, n$) and C_p accepted equal to 0. Hereinafter the approximation formulae for Q_t are recommended in order to compute the student's scores as an integer number.

Row 4 corresponds to n learning objects numbered arbitrary and mentioned in the question text formulation. The student has to put these numbers in their correct, e.g. logical sequence. The adequate graph model also is linear, so, the formulae for Q_t and w_i retain as in the previous type but $C_p = n/(2n - 1)$.

Row 5 corresponds to ordered numbers, e.g. binary, decimal, heximal each one with an arbitrary length. The student has to enter these numbers from the keyboard. The corresponding graph model also is linear, but Q_t is computed as a weighted sum of the numbers length plus number of arcs, and C_p also can be accepted equal to 0.

Row 6 requires from the student to enter from the keyboard a linear formula having on the left side a dependent variable equal to an arithmetical or logical expression including m independent variables, numbers and brackets on the right side. The symbols for the variables are supposed to be mentioned in the question text formulation. It is well-known from the theory of compilers that the adequate graph model presents a binary tree. According to the graph theory if the number of the tree nodes is n then the number of its arcs is $n - 1$. The formulae for Q_t and w_i retain the same, but $C_p = (m + 1)/Q_t$.

The answer type computing a formula (row 7) requires from the student to determine the values of the independent variables (for example, from some free text, symbolic text or picture) associated with the question formulation, and to replace them in the formulae. If the formula itself has not been included in the test it is reasonable to accept $Q_t = 3.n$ and $C_p = 0$.
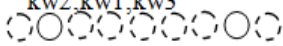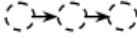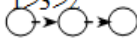
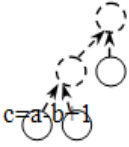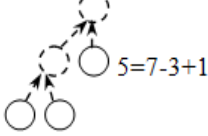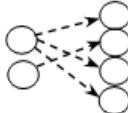Row 8 is for question answers called unordered pairs. The first pair's element belongs to n learning objects denoted with numbers and its second element to m objects' attributes denoted with letters. The objects and attributes are mentioned in the question text formulation. The student has to point out all k binary relations between objects and attributes. So, it is reasonable to accept $Q_t = 3.k$; $C_p = (n+m)/Q_t$, $w_i = 3$ $(i = 1, …, k)$, where $k \geq n$.

Row 9 corresponds to unordered triples answers copied from the dictionary. This kind of model is appropriate for testing knowledge about matrixes, table functions of variables X and Y, and so on with dimension $n \times m$. The first element of a triple presents the current matrix row or function X-value,

the second one – the current matrix column or function Y-value and the third element – the value of the corresponding matrix cell or function itself. So, in case of such type of answers $Q_t = 3.m.n$, e.g. $w_{ij} = 3$ $(i = 1, …, n; j = 1, …, m)$ and $C_p = 0$.

The weights (1, 2, and 3) inferred from the underlying graph models allow to divide the intelligent test questions into very easy, easy, and moderate. This finding is not only in concordance with some previous studies of Zheliazkova's group but also is shared by other researchers [3, 5]. The study carried out shown that this background yields to valid and stable results acceptable by both students and teachers for testing the lecture material at the level of comprehension.

Table 1. Graph Models and Formulae for Different Answers Types

Answer type	Example graph model	Formulae for parameters	Values of parameters
Multiple-choice	b;e	$Q_t = n$; $w_i = 1$; $C_p = 1.00$	$n=2$; $Q_t=2$; $w_i=1$; $C_p = 1.00$
Unordered keywords	kw2;kw1;kw3	$Q_t = n$; $w_i = 1$; $C_p \approx 0.00$	$n=2$; $Q_t=2$; $w_i=1$; $C_p \approx 0.00$
Ordered keywords	kw1>kw2>kw3	$Q_t = 2.n-1$; $n=3$; $w_i=2$; $C_p = (n-1)/Q_t$	$n=3$; $Q_t=5$; $w_i=2$; $C_p = 2/5 = 0.40$
Numbered objects	1>3>2	$Q_t = 2.n-1$; $n=3$; $w_i=1$; $C_p = n/Q_t$	$n=3$; $Q_t=5$; $w_i=2$; $C_p = 0.60$
Ordered numbers	10>110>1	$w_i = length(sa_i)$ $(i=1,…,n)$; $Q_t = sum(w_i) + (n-1)$; $C_p \approx 0.00$	$n=3$; $w_1=2$; $w_2=3$; $w_3=1$; $Q_t = (2+3+1)+(3-1)=8$;
Enter formula	c=a\b+1	$Q_t = 2n+1$; $w_i=2$; $C_p = (n+2)/(2.n+1)$	$n=5$; $Q_t=11$; $w_i=2$; $C_p = 0.28$
Compute formula	5=7-3+1	$Q_t = 3n+1$; $w_i=3$; $C_p = n/(3n+1)$	$n=5$; $Q_t=16$; $w_i=3$; $C_p = 3/16 = 0.19$
Unordered pairs		$Q_t = 3.k$; $w_i=3$; $k=4$; $C_p = (n+m)/(3k)$	$k=4$; $n=2$; $m=3$, $Q_t=12$; $w_i=3$; $C_p = (2+4)/3.4 = 0.50$
Unordered triples	4 3 1 8 2 6 2 5 4	$Q_t = 3.k$; $w_i=3$; $k=9$; $C_p \approx 0.00$	$k=9$; $n=3$; $m=3$; $Q_t = 3.9 = 27$; $w_i=3$; $C_p \approx 0.00$

3. Materials and Methods

3.1. Participants and Data Collection

The participants involved in the knowledge testing were 65 bachelor degree students from the specialty Computer Systems and Technologies at University of Ruse. The test session was carried out within the framework of the course Software Engineering and covered 30 hours taught lecture material.

The test was created by the course lecturer as an intelligent post-test in order to evaluate the Correct Knowledge (*CK*), Missing Knowledge (*MK*), Wrong Knowledge (*WK*), as well as the Time (*T*) undertaken for the test performance. The parameter *CK* can be defined as a part of the student's answer that coincides with the teacher's one. The parameter *MK* is a part of the teacher's answer that does not present in the student's one, while the parameter *WK* exists in the students answer, but is missing in the teacher's one.
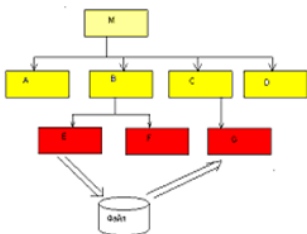
Each student was accessible through a common device to the test Microsoft Word document finishing with the teacher's keywords dictionary.

The number of questions was 30 with total scores $P_{max} = 352$ and time planned for the test performance $T_{max} = 120$ min. The questions answer types were four, namely: multiple choice, unordered keywords, ordered keywords, numbered keywords, and unordered pairs [26]. Table 2 contains one example question per each of the listed types of answers.

Depending on its type each question brought different number of scores $p_{max\,j}$. The student's answering was reduced to: filling an empty edit field, copying and pasting dictionary keywords and using the special symbols ";" and ">" as separators of non-ordered and ordered sub-answers respectively. In order to make difference between *MK* and *WK* usage of question answer or sub-answer "no" was recommended.

Note, that the above-presented external knowledge description of the lecture material content differs from the most popular term of knowledge representation in ITSs. Such knowledge description "refers to how information content from the external realm is mentally organized, structured, conveyed and made available for use when needed." [10].

Table 2. Examples of the answers types used in the test

Answer type	Text formulation	Correct answer	Parameters values
Multiple-choice	**Check the features of the program architecture on the figures: a)** direct recursion, **b)** indirect recursion, **c)** reading from file, **d)** hierarchical structure, **e)** writing to file.	c;d;e	$w=1$; $Q_i=3$; $C_p=1.00$
Unordered keywords	**List the principles of the object-oriented programming.**	abstraction; restricted access; modularity; hierarchy; typicality; parallelism; stability	$w=1$; $Q_i=7$; $C_p=0.00$
Ordered keywords	**Fill in the missing keywords:** *"If the ... MyAnimal, belonging to ... class, points out to the class ..., the virtual method TAnimal.Voice() will be invoked. In contrary, if it points out to an object from the class TDog, ... of ..., the method ....Voice() will be invoked."*	object>TAnimal>child>TAnimal>TDog	$w=2$; $Q_i=10$; $C_p=0.50$
Numbered objects	**Order in the right sequence the number of the keywords in the fragment in the right sequence.** 1) Procedures & Functions (titles + bodies) **2)** Implementation **3)** Unit **4)** Procedures & Functions (titles) **5)** Begin **6)** Interface	3>6>4>2>1>5	$w=2$; $Q_i=12$; $C_p=0.50$
Unordered pairs	**Point out the correspondence between** the keywords: **4)** private; **3)** public; **2)** protected; **1)** published; and the statements: **a)** Attributes and methods of the class are accessible from each other part of the program; **b)** they are accessible from the current class and its children; **c)** they are not accessible out of the module in which the class is declared; **d)** accessible not only during the run time, but during the design time too.	4>c;3>a;2>b; 1>d	$w=3$; $Q_i=12$; $C_p=0.66$

Before uploading the filled in document back on the common device, the students had to answer two additional questions "Whether you think the mark will be objective and precise? Yes/No" and "How many minutes the test performance took you?" The students were also told that the time for the test performance actually is unlimited and that this parameter together with the MK and WK will be used as assessment indicators only for research purpose.

3.2. Raw Data

On the next day the lecturer computed by EXCEL the *CK*, *MK*, and *WK* for each test question. In case of the *CK* the student's sub-answers taken into account were equal to the teacher's ones, for the *MK* as the total scores of the missing student's sub-answers, and for the *WK* as the total scores of the student's sub-answers different from the teacher's ones. The total scores for the test also were written down in the student's WORD document. The student's final Mark (*M*) was computed depending on his/her test *CK* scores (*P*) in a traditional for Bulgarian assessment scale. More precisely the assessment scale is: $0 \leq P \leq 0.4* P_{max}$ – "Poor" (2); $0.4*P_{max} < P \leq 0.55*P_{max}$ – "Satisfactory" (3); $0.55*P_{max} < P \leq 0.70*P_{max}$ – "Good" (4); $0.70*P_{max} < P \leq 0.85*P_{max}$ – "Very good" (5); $0.85*P_{max} < P \leq 1.0*P_{max}$ – "Excellent" (6). The experience accumulated during the last decade by Zheliazkova's research group has pointed out that such a non-linear scale leads to marks acceptable by both teachers and students.

3.3. Methods

It was reasonable to accept the *CK* as predicted, e.g. dependent variable because the student's mark *M* was promised to be computed only on its base. The rest variables, e.g. *MK*, *WK*, and *T* served as independent variables in the regression analysis. The corresponding type of the regression equation $CK = f(MK, WK, T)$ depending on the data set was preliminary unknown. To find it the correlation analysis was applied before the CK prediction itself. As it is known [21] the computed value of the linear correlation coefficient $r(X,Y)$ (a real number in [-1.00,+1.00]) shows how strong is the relationship between two statistical parameters (*X* and *Y*). For better understanding by the teacher this quantitative value was transformed to a linguistics one by means of a five-point Likert's scale. More precisely, if $r(X,Y)$ is in the range $0.0 \div 0.3$ then the relationship is "low"; $0.3 \div 0.5$ – "moderate"; $0.5 \div 0.7$ –

"significant"; $0.7 \div 0.9$ – "high"; $0.9 \div 1.0$ – "very high". If the two parameters are moving in the same way $r = +1.0$ and if in the opposite one $r = -1.0$. The value $r = 0$ means no correlation between the considered parameters. It was expected that $r(CK, MK)$ will have the highest value among all possible combinations of the test parameters. In this case the simple regression equation [13, 18] for the *CK* prediction will have the form: $CK = B0 + B1.MK$. Formula (1) presents the computed values of B0 and *B*1 for "all" tested students. At this stage the corresponding regression equation will have the general form: $CK = B0 + B1.WK + B2.T$. The impact of the *WK* and *T* was supposed to be the first order that means the corresponding changes are additive. The regression coefficient $B1/B2$ represents change in the *CK* per one unit increase in *WK*/*T* when the other variable *T*/*WK* is constant.

3.4. Teacher's Tool

The used teacher's tool especially developed for the EDM purpose offers a power programming language called SESSIONSCRIPT classified as "script", "visual", "very high-level", and "mark-up". A program in this language is reduced to a parallel arithmetic's and table processing (Figure 1). The visualization allows choosing different kinds of diagrams (bar, pie, line, and point) viewed in separated windows. The common structure of programs in SESSIONSCRIPT described in the meta-language of Bascus-Naur and the technology scheme for the tool's using can be found in [11]. To perform a step of this scheme the teacher has to know the syntax and semantics of the corresponding group of commands. In practice some steps can be omitted others repeated or interpreted as a subtask solving. The teacher can save programs in a standard text file serving as a common catalogue of the tasks already solved by him/her and/or the members of the course team. To enhance the language learning a special color coding scheme is applied to the program script. Additionally a lot of messages (information and diagnostics) are generated by the tool. The teacher can choose menu-commands to view the names and values of both system and program variables and before to exit to see a detail report, print and/or save it in .xls format (figure 2). The extracted parameters: number of the program variables, rate of task solving, invalid commands frequency, and so on can be used for comparison of one and the same task performance by different teachers.
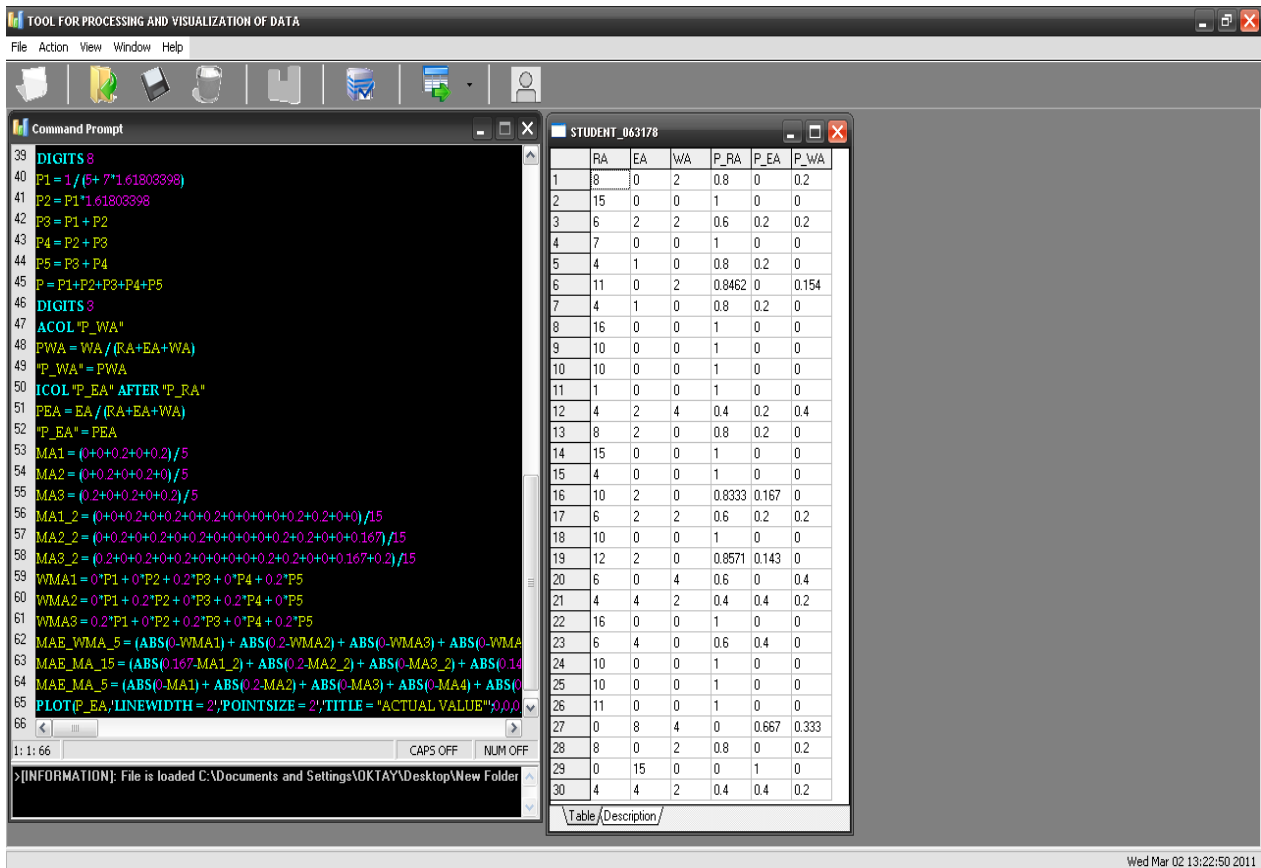
Figure 1. The Tool's Screen

Figure 2. The Teacher's Report

4. Results

4.1. Raw Data Filtering

After the raw data filtering due mainly to the cheat recognition [23], 47 students rest. Only one student received mark "2" that confirms our previous observations on the student's belief in the objective and precise test assessment. The students go to test only if they assessed themselves at least with mark "3". Each row of Table 3 can be treated as parametrical temporal SM that differs from those

mentioned in the introduction in dividing the student's knowledge into three kinds, namely: CK, MK, and WK. The whole table 3 (not given here) presents the category of "all" students, the first 35 students of "very good", and the last 12 ones of "good" students. Another positive tendency also confirmed that for the intelligent tests the average mark AVG(M) is shifted from "4" to "5" mark. Perhaps, a few students with mark "6" from the second category and a few students with mark "3" from the third category needs additional teacher's attention, for example, more complex/simple lecture material.

Table 3. The Raw Data Set

4.2. Correlation Analysis

Table 4, 5, and 6 contain the computed values of all possible coefficients of correlation between the *CK*, *MK*, *WK*, and *T* for "all", "very good", and "good" students respectively. The comparison of the values of these coefficients shows that the correlation coefficient between the *CK* and *MK* has the greatest value. That was the main argument for application of simple linear regression at the next stage of prediction.

Table 4. Correlation Coefficients for "All" Students

CORRELATION ANALYSIS			
	CK	MK	WK	T
CK	1	-0.881	-0.164	0.297
MK		1	-0.321	-0.296
WK			1	0.031
T				1
\Table \Description /			

Table 5. Correlation Coefficients for "Very Good" Students

CORRELATION ANALYSIS			
	CK	MK	WK	T
CK	1	-0.733	-0.114	0.165
MK		1	-0.592	-0.177
WK			1	0.074
T				1
\Table \Description /			

Table 6. Correlation Coefficients for "Good" Students

CORRELATION ANALYSIS			
	CK	MK	WK	T
CK	1	-0.911	0.316	-0.046
MK		1	-0.68	-0.064
WK			1	0.228
T				1
\Table \Description /			

4.3. Simple Regression Analysis

Both experimental data line of the *CK* and its prediction on Figure 3 show a good prediction. The similar formula (2) and Figure 4 were received for the category of "very good" students, and formula (3) and Figure 5 for that of "good" students. Comparison of the three equations shows that the quantitative results of the simple regression are in the line of the qualitative results of the correlation analysis. The positive and negative symbols of *B*1 are the same as of the corresponding correlation coefficients. Visually for each category of students the prediction line is very close to the experimental one. For comparison the traditional graphical interpretation of

the simple linear regression model $Y = F(X)$ is a line (*Y*) with an angle to the axis *X* [2].

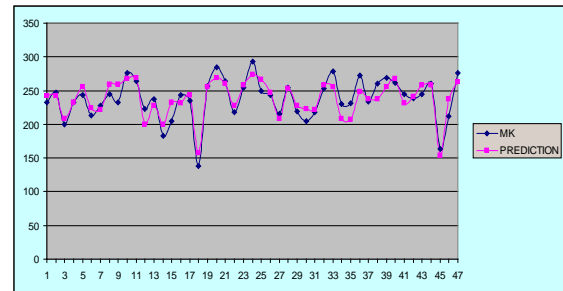$$CK = 276.9506811 - 0.844308*MK \qquad (1)$$

Figure 3. CK Prediction from MK for "All" Students
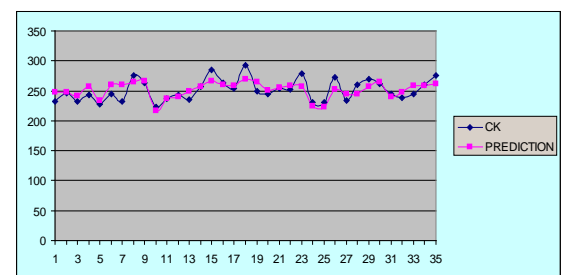
$$CK = 271.8887111 - 0.591738862*MK \qquad (2)$$

Figure 4. CK Prediction from MK for "Very Good" Students

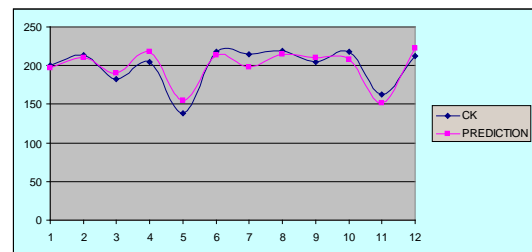$$CK = 254.8956634 - 0.70425653*MK \qquad (3)$$

Figure 5. CK Prediction from MK for "Good" Students

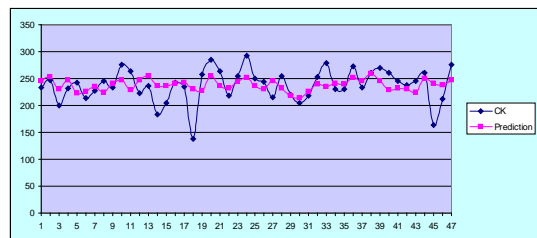$$CK = 186.6356518 - 0.343801627*WK + 0.469223454*T \qquad (4)$$

Figure 6. CK Prediction from WK and T for "All" Students
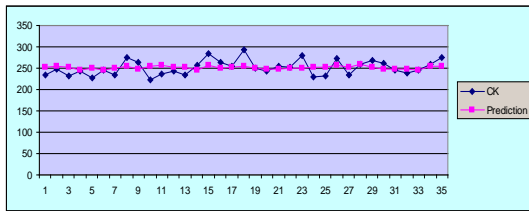
$$CK = 232.707 - 0.148*WK + 0.168*T \qquad (5)$$

Figure 7. CK Prediction from WK and T for "Very Good" Students

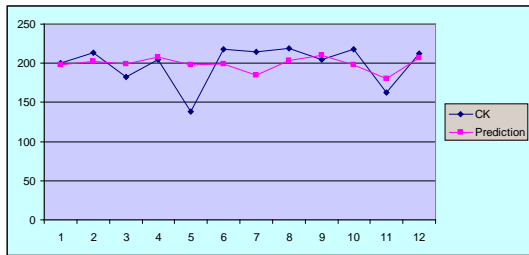$$CK = 196.3442384 + 0.613112891*WK - 0.16248322*T \qquad (6)$$

Figure 8. CK prediction from WK and T for "Good" Students

4.4. Three-Parameters Regression Analysis

Formula (4) presents the calculated values of *B*0, *B*1, and *B*2 for all students. On Figure 6 the experimental line of the *CK* and its prediction for this category are shown. The similar formula (5) and Figure 7 concerns the category of the "very good" students, and formula (6) and Figure 8 the category of the "good" students. For "all" tested students the impact of the *WK* (*B*1 = -0.344) is less and negative in comparison with that for *T* (*B*2 = +0.469). For the "very good" students the signs and values of B1 and *B*2 are equal but for the "good" students the impact of the *WK* (*B*1 = +0.613) is much greater than that of *T* (*B*2 = -0.162) and the signs are opposite. Note that a coincidence exists in both signs and values with the corresponding coefficients of correlation (Table 4). For comparison the traditional graphical interpretation of three-parametrical linear regression model *Y* = *F*(*X*1, *X*2) is several parallel lines (*Y*) with an angle to the axis *X*1 under different *X*2 = const. [13].

In this study the accuracy indicators Coefficient of Determination (CD) and Adjusted Coefficient of Determination (ACD) [1] were added to the Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) used in the previous study concerning the short-term prediction of the *CK*, *MK*, and *WK* [12]. In Table 7 the values of all four indicators are given together with the Average (AVG) of the dependent and independent variables.

Table 7. Accuracy Indicators of Simple Regression

Category	AVG(CK)	AVG(MK)	MAE	MAPE	CD	ACD
All	237,91	46,23	11,82	0,05	0,78	0,77
Very good	251,29	34,91	9,69	0,04	0,54	0,52
Good	199,08	79,25	8,76	0,04	0,83	0,81

The value of the MAPE for the three categories of students is constant (4-5%) that is acceptable for the testing purpose. The CD and ACD for "all" tested students indicates that 77-78% of the variance in the *CK* is predicable from the *MK*. This percentage is substantially less (52-54%) for the "very good" students, and much (81-83%) for the "good" students. Perhaps, the rest percentage could be predicted from the other independent variables WK and T for which the correlation coefficients were much more less. Therefore, there is a need of the next stage multiple linear regression.

In Table 8 the corresponding accuracy indicators are given together with the AVG of the dependent and independent variables. The MAPE value for the three students' categories is approximately constant (1%). For all tested students the values of the CD indicates that only 4% of the *CK* variance is predicable from the *WK* and T. This percentage increases to 11-12% for the "very good" and "good" students.

Table 8. Accuracy Indicators of Multiple Regression

AVG (CK)	AVG (WK)	AVG (T)	MAE	MAPE	CD	ACD
237,91	30,77	131,83	14,10	0,06	0,04	-0,02
251,23	28,74	135,37	21,74	0,1	0,12	0,078
199,08	36,67	121,5	17,00	0,1	0,11	-0,08

4. Discussion and Conclusions

An original procedure for correct knowledge prediction of three categories of tested students was proposed applying consequently correlation, single and multi-parametrical regression analysis. The experimental data set included also missing knowledge, wrong knowledge and time from an intelligent test performance. By means of several indicators the prediction accuracy was assessed as satisfactory although the applied models were linear.

The actual results for correct knowledge of different categories of students can significantly differ from the predicted ones. The main reasons for this can be some side effects and/or external factors not taken into consideration in this study. Regarding students, they could be: attempt for fraud, unfamiliar type of questions, insufficient attention or low motivation. Other external factors, related to the teacher, could be: poor test session planning, organization, and/or delivery. So, the expectation is that the proposed procedure can improve knowledge management, instructional design, and pedagogical decisions.

The procedure was implemented by means of a specialized EDM tool with a powerful, expressive and flexible script language. The programs' script can be reviewed and reused by the teacher thus enhancing his/her productivity. Under the same structure of the data set the correct skills also can be predicted for a test-like exercise including more complex tasks than test questions. In such a way the teacher's pedagogy base, containing facts, rules, and accumulated experience can be easily and quickly extended and shared with other teachers from a course team or even from different course teams.

The precise and objective students' test assessment has drawn the students' attention and increased their motivation not only during the carried out experiment. The observation of intelligent testing in other subjects has shown that the students go to an intelligent test performance only if they were self-assessed at least with the mark "3". Another positive tendency of intelligent testing to shift the average test mark from "good" to "very good" also was confirmed. Across the years both students and teachers became aware the terms "correct knowledge", "missing knowledge", and "wrong knowledge", as well as the terms "individual" and "group student's models".

References

[1]. Adjusted Coefficient of Determination. http://www.r-tutor.com/content/adjusted-coefficient-determination

[2]. Brusilovsky, P. (1999). Adaptive and Intelligent Technologies for Web-Based Education. *Special Issue on Intelligent Systems and Teleteaching, Rollinger, C., &Peylo, C. (Eds). KinstlicheIntelligenz, Vol. 4*, 9-25.

[3]. Karpova, I. P. (2001). Analysis of Learner's Answers in Automated Teaching Systems.*Russian Journal on Information Technologies,* 49-55.

[4]. Freuch, P. A. (1990). Domain-Independent Learner Model for AI-Based Training System.*Int. J. Computers & Education, Vol. 15(1-3),* 49-61.

[5]. Goh, G. M.,Quek, C.,Maskell, D. L. (2010).EpiList II: Closing the Loop in the Development of Generic Cognitive Skills.*IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans, Vol. 40(4),* 676-685.

[6]. Hawkes, L. W., Derry, S. J.,Rundensteiner, E. A. (1990). Individualized Tutoring Using an Intelligent Fuzzy Temporal Relational Database.*Int. J. on Man-Machine Studies,* 409-429.

[7]. Hernandez, J. A., Ochoa, A.,Munoz, J.,Burlak, G. (2006). Detecting Cheats in Online Students Assessments Using Data Mining.*Proceedings of Conference on Data Mining,* 204-210.

[8]. Hong et al., Adaptivity in Web-Based Educational System. Massey University, New Zealand, http://www10.org/cdrom/posters/1052.pdf

[9]. Jenny,L.,B. Moon, (Eds)(2000). Learners and Pedagogy, SAGE Publications Ltd.

[10]. Kavita, L. S., R. J. Mislevy, (2008). *Design Patterns for Assessing Internal Knowledge Representations. PADI technical report 22 (Draft).*

[11]. Kir, O., Zheliazkova, I. I. (2011). *A Script Language for Data Mining, Proceedings of the Second Scientific Conference Dedicated to Yuri Gagarin's Flight, Shumen,* 91-101.

[12]. Kir, O.,Zheliazkova, I. I. (2011).Prediction of Educational Data Mining by Means of a Postprocessor Tool.*Int. J. of Information Technologies & Knowledge, Vol. 5(2),* 190-199.

[13]. Lecture25: Multiple Linear Regression, 1990, courses.washington.edu/qsci483/Lectures/25.pdfs/25.pdf

[14] Lichnaropoulos, N,Y. Glinos, (1999). Student Modeling in an Inductive Learning Environment.*Bulgarian Journal of "E+E"*, *Vol. 5-6,* 58-62.

[15]. Lissitz, R. W., H. Wei (2008). Consistency of Standard Setting in an Augmented State Testing System*. J. of Educational Measurement: Issues and Practice*, pp. 46-56.

[16]. Paschin, E.,Mitin, A. (1985). Automated Teaching System EKSTERN", Moscow State University, (in Russian).

[17]. Rawlings, J. O.,Pantula, S. G.,Dickey, D. A. *Applied Regression Analysis: a Research Tool*. Second edition, Springer.

[18]. http://library.northsouth.edu/Upload/Applied%20Regression%20Analysis.pdf

[19]. Romero, C., Ventura, S., &García, E. (2008). Data Mining in Course Management Systems: Moodle Case Study and Tutorial. *Int. J. Computers & Education, Vol. 51,* 368-384.

[20]. Romero, Cr.,Ventura, S. (2010). Educational Data Mining: a Review of the State of the Art. *IEEE Transactions on Systems, Man, and Cybernetics – Part C: Applications and Reviews, Vol. 40(6),* 601-618.

[21]. Savelev, A. Ya.,Novikov, V. A.,Lobanov, Yu. I. (1986). Preparing Information for Automated Teaching Systems, Moscow (in Russian).

[22]. Sleeman, D., Brown, J. S. (Eds) (1982). Intelligent Tutoring Systems, Academic Press: London.

[23]. Teodorov, G.,Kir, O.,Zheliazkova, I.I. (2011).Correlation Analysis of Educational Data Mining by Means a Postprocessor Tool.*Int. J. of Information Theories and Applications, Vol. 18(3),* 289-299.

[24]. Vadillo, J. A. &Ilarraza, A. D. (1995). The Role of the User Model for Explanation in Intelligent Tutoring Systems for Training. Proceedings of the 3rd *Int. Conference on Computer Aided Engineering Education, Bratislava, Slovakia,* 101-106.

[25]. Villano, M. (1992). Probabilistic Learners Models: Bayesian Belief Networks and Knowledge Space Theory.*Proceedings of the Second Int. Conference ITS'92, Montreal, Canada,* 491-498.

[26]. Zheliazkova, I. I., Andreeva, M. H. (2004). An Intelligent Multimedia Environment for Knowledge Testing. *E- learning and the Knowledge Society, Gent & Brussels, Belgium,* 3.13.1-3.13.24.

[27]. Zheliazkova, I.I., Andreeva, M.H., Kolev, R.T. (2006). Knowledge Testing in Algorithms – an Experimental Study. *Proceedings of Int. Conference Modern E-Learning, Varna, Bulgaria,* 55-62.

[28]. Zheliazkova, I.I.,Kolev, R.T. (2008). Task Results Processing for the Needs of Task-Oriented Design Environments. *Int. J. Computers & Education, 51,* 86-96.

Corresponding author: Irina Zheliazkova
Institution: University of Ruse, Bulgaria
E-mail: Irina@ecs.uni-ruse.bg