Intelligent Modelling for Predicting Students' Final Grades

Monika Simjanoska, Marjan Gusev and Ana Madevska Bogdanova
University Ss. Cyril and Methodius, Faculty of Computer Science and Engineering,
Skopje, Macedonia
Email: m.simjanoska@gmail.com, {marjan.gushev, ana.madevska.bogdanova}@finki.ukim.mk

***Abstract*—The main objective of this paper is producing an intelligent virtual teacher who will be able to predict the students' final grades at the end of the semester. Our approach is based on continual observation of the student's activities on the particular course during the semester. In order to achieve realistic modelling of the students' devotion to the given lectures and also the degree of how much the student has learned from the given lecture, we take into account both the e-Learning and the e-Assessment results through the semester. In our previous work we did an intelligent students' Profiling to classify the students into a pass, or, fail category. In this paper we go deeper into the problem, achieving more precise modelling according to which we will be able to determine the student's most likely final grade, using multi classification methodology. The advantage of our model is in its ability to take into account all the assessments during the semester, not relying only on the results from the last student's assessment. It can be a good indicator whether the teacher needs to perform additional testing of the student's knowledge in order to derive an overall conclusion on the most appropriate grade.**

***Index Terms*—e-Learning, e-Assessment, Machine Learning**

I. Introduction

E-Learning, or Computer Assisted Learning, refers to the learning by using sophisticated information and communication technologies, whereas, the e-Assessment, or Computer Assisted Assessment, aims to assess the learner's knowledge. Vonderwell et al. [1] make a clear distinction between assessment of learning (assessment for the purposes of grading and reporting with its own established procedures) and assessment for learning (assessment whose purpose is to enable students, through effective feedback, to fully understand their own learning and the goals they are aiming for). A good overview of e-Assessment systems is presented by Gusev and Armenski [2]. In our research the assessment for learning is referred to as e-Learning and the assessment of learning is referred to as e-Assessment.

Rovai [3] in his research on online and traditional learning identifies the major principles of general assessment theory, examines how these principles can be applied to an online environment, and identifies and describes several assessment issues that have special significance in a virtual classroom. Practical suggestions are included to assist both first-time and experienced online instructors develop assessment components for their online courses.

Various conflicting attitudes about distance learning have been discussed by Hannay et al. [4] in order to present the perceptions of both the students and the teachers. Sun et al. [5] developed an integrated model with six dimensions: learners, instructors, courses, technology, design, and environment to investigate the critical factors affecting learners' satisfaction in e-Learning. The results showed that learners' computer anxiety, instructors' attitude toward e-Learning, e-Learning course flexibility, e-Learning course quality, perceived usefulness, perceived ease of use, and diversity in assessments are the critical factors affecting learners' perceived satisfaction.

The research that we conduct in this paper is a result of a long term e-Learning and e-Assessment process during the course on Computer Architecture and Organization (CAO). We believe that besides the benefit of the intelligent prognosis of the students' final grades, the results will be a useful indicator for the teachers of whether the given lectures for the e-Learning process are well formed and represent the real students' knowledge, and also whether the students have adapted to this kind of learning a can fully express their knowledge.

Beside the surveys on the students' satisfaction, universities have been given a lot of advices to turn their interests towards cloud computing, since the cloud offers good resources flexibility and scalability, storage, computational requirements, network access, reduced power consumption and lower costs [6], [7], [8]. As an upgrade to the e-Learning and e-Assessment system that we use in this paper, Ristov et al. [9] propose a SOA architecture of a cloud hosted e-Assessment system which uses scalability and elasticity in order to achieve sustainable performance. Their solution reduces the overall costs since it uses minimum resources utilized only during the e-Assessment.

However, the focus of this research is not to analyse the existing system's performance and cost, but to introduce an intelligence in the e-Learning and e-Assessment processes which will result in producing an intelligent grading agent (IGA) whose decisions will be as similar as possible to the teacher's opinion on the student's final grades. For that purpose we use the results from the online learning system that has been implemented at our university since 2002. Based on 10 years of e-Learning and e-Assessment experience, we have developed a strategy by which we proved that the student has achieved all learning objectives (LOs) by providing correct answers for relevant knowledge items during the online learning process. During the strategy development we confronted issues of guessing the correct answers in a random manner, which we solved by introducing negative scoring for wrong answers. This is how we expect the students to answers the given questions only when they know the correct answer. During the development of the strategy, we also determined the minimal number of questions the system should ask to get relevant assessment of student's knowledge. After analysing several strategies, we agreed that the least vulnerable strategy that offers satisfactory results is 3R (three correct answers in a row). The probability analysis showed that the probability to randomly choose three consequently correct answers is smaller than the probability to choose any three correct answers from a given set. The overall contribution of this research is that we developed strategy by which the students are asked as minimum questions as 3 and expect 3 correct answers in a row, in order to allow them to continue the online learning process. This strategy somehow confirms the student's knowledge on a particular learning objective.

Therefore, the resulting online learning system is with the following features:

- *Corrective measure* - whenever the student gives a wrong answer the system suggests corrections.
- *Adaptive learning path* - better students can move faster towards the more complex learning objectives.
- *Clear objectives and goals* - the navigation explains what else has left to be learned.

However, this approach does not estimate the final grade of the e-Assessments, nor evaluates the student efficiency, and this is the motivation to conduct a research whose results are reported in this paper.

The work reported in this paper is based on our previous research [10] where we achieved modelling of the students overall work during the semester, taking into account the results from both the e-Learning and the e-Assessment, to determine whether the student will pass or fail the CAO course. We believe that this kind of grading that depends on the students' overall work through the whole course will additionally motivate them to work continuously during the semester and present their real knowledge during the e-Learning process, since it directly affects their final grade.

Our research mainly consists of two parts. At first we verify the existing methodology for classifying the students into a pass or fail class where we use most common Support Vector Machines (SVM) kernels. The second part of the research reported in this paper focuses on building an IGA able to estimate the students' final grade.

Figure 1 presents the interaction between the online learning system and our intelligent agent. As depicted, the online learning system continuously assigns e-Learning exams from the current LOs that have to be learned before the student proceeds to the next learning level. The results from the e-Learning are then stored into the online learning system. Periodically, on teacher's request, the system assigns e-Assessment to the students in order to test their knowledge. All the results are again saved by the online learning system. As soon as the course is finished, the IGA collects all students e-Learning data which then it uses to create individual students' Profiles. The Profiles refer to the students' manner of passing the learning objectives of the particular lecture. On the other hand, the e-Assessment results are used to update those Profiles. After that, the Profiles are used in the ML process, i.e. the SVMs to produce a classifier able to determine the students' final grades. Since the teacher is allowed to monitor the results and each student's knowledge background, the IGA's decision is verified by a human factor. Hereupon, we propose a system that simulates a real teacher and is able to estimate the real achievements of its students.

The rest of the paper is organized as follows. In Section II we present the use of machine learning (ML) methods in education. Our methodology for multi-class classification of the students' knowledge is presented in Section III. In Section IV we present the experiments and the results from the presented approach and we derive a conclusion and present the plans for future work in Section V.

II. Related Work

In this section we give a review of the latest ML methods used for educational problems.

Castro et al. [11] aim to provide an up-to-date snapshot of the current state of research and applications of Data Mining methods in e-learning. They provide a taxonomy of e-learning problems to which Data Mining techniques have been applied, including: students' classification based on their learning performance, detection of irregular learning behaviours, e-learning system navigation and interaction optimization, clustering according to similar e-learning system usage and systems' adaptability to students' requirements and capacities.

Zaiane et al. [12] in their paper suggests the use of web mining techniques to build an agent that could recommend on-line learning activities or shortcuts in a course web site based on learners' access history to improve course material navigation as well as assist the online learning process. These techniques are considered integrated web mining as opposed to off-line web mining used by expert users to discover on-line access patterns.

The paper by Fei et al. [13] describes their work in exploring automatic question classification tests which can be used in e-learning system. Such tests can take the form of multiple-choice tests, as well as fill-in-the-blank and short-answer tests. They proposed a text categorization model using an artificial neural network trained by the backpropagation learning algorithm as the text classifier. Their test results showed that the system achieved the performance of nearly 78%.

Chang et al. [14] state that with the growing demand in e-learning, numerous research indicated that adaptive learning is a critical requirement for promoting the learning performance of students. The first step for achieving adaptive learning environments is to identify students' learning styles. Therefore, in their paper they propose a learning style classification mechanism to classify and then identify students' learning styles. The proposed mechanism improves k-nearest neighbour classification and combines it with genetic algorithms. To demonstrate the viability of the proposed mechanism, the

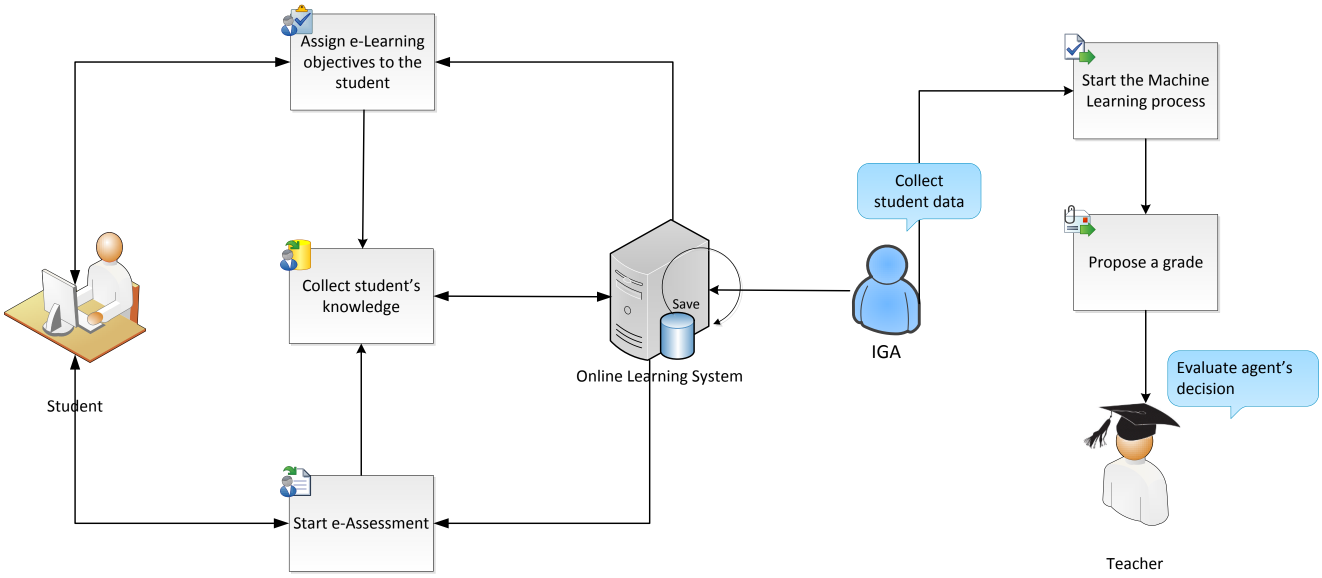

Fig. 1. Interaction between different roles in the machine learning process

proposed mechanism is implemented on an open-learning management system. The experimental results indicate that the proposed classification mechanism can effectively classify and identify students' learning styles.

Applying intelligence in e-Learning services has been the focus of several recent research projects and papers, and with the explosive growth of information sources available on the World Wide Web, the existing e-Learning services requires pedagogical overhauling for keeping pace with the trends in e-Learning. One of several possibilities can be its application to Online Assessment. In this Paper, we present the system architecture for applying intelligent methodologies to online assessment that adapts to the examinee's ability level and the other side we describe that web mining can be applied to Learning Content Management Systems to build knowledge about e-Learning and has potential to help and improve learner's performance.

The design of online assessment tools has been in the focus recently [15], [16], as well as applying intelligence in e-Learning services. For example, Uday et al. [17] present a system architecture for applying intelligent methodologies to online assessment that adapts to the examinee's ability level and the other side we describe that web mining can be applied to Learning Content Management Systems to build knowledge about e-Learning and has potential to help and improve learner's performance.

III. Methods and Methodology

In this section we present the methods and the methodology that we developed for building the students' individual Profiles as well as the classification process itself.

A. *Online Learning System's Structure*

Before we describe the students' Profiling, we must discuss the structure of the online learning system. The knowledge database for each course can be described in a tree-like manner. If we assume that one course represents the root of the tree, than the course lectures are its branches, and the learning objectives are branches' branches. Each learning objective consists of few questions for which there is one, or, several correct answers. In order to make the 3R strategy applicable for any learning objective, there are at least five questions provided.

B. *Student's Profiling*

In order to show how the results from the long term e-Assessment of a particular student can affect the previous e-Learning impression, we must provide each student an individual Profile, which can be described as follows.

The first step is extracting their knowledge from the online learning process in a n-dimensional vector, where n is the number of LOs. The students knowledge of a particular LO is represented as a number of trials to pass the given LO and the number of questions answered before the 3R pattern occurred in the passing trial. In order to express the student's efficiency for each LO that has been assigned, we proposed new metric in (1) to measure the student's *Initial Success*, IS. In the beginning, the IS for every student enrolled in the course is a null vector. The dimensionality of the IS vector is the same as the number of all LOs in the knowledge database. Therefore, $IS_i = (LO_1, LO_2, ..., LOj)$, where the highest value of i is the total number of students and j is the total number of LOs in the database. Considering the fact that the students are assigned each week with new questions from the LOs, the IS is calculated as sum of the number of trials (NT) necessary for the student to pass the given LO, and the product of the scalar 0.01 and the number of questions (NQ) in the last trial. We use the value 0.01 in order to scale the number of questions in range $0.03 \leq NQ < 1$, since the maximum number of answered questions per learning objective is a double digit number. Therefore, the lower the IS score is, the more the student has learned the taught lectures.

$$IS_{ij} = NT_{ij} + (0.01 * NQ_{ij}) \tag{1}$$

Once we built the IS vector, our IGA uses them to conceive preliminary impression on the students' devotion to the course. However, taking into account only the ISs is not enough to derive conclusion of the students' most probable final grade. Therefore, we propose another approach which takes into consideration the students' e-Assessment results before concluding the grade. Thus, the students' success is revised every time the student has taken an e-Assessment. By introducing this approach, we simulate a real surroundings where the teacher evaluates the students' activities during the whole course to determine their grades. Hereupon, we calculate the *Revised Success*, RS, as:

1) *Punishment:*

$$RS_{ij} = IS_{ij} + 1 \tag{2}$$

2) *Reward:*

$$RS_{ij} = IS_{ij} - 1 \tag{3}$$

where i denotes the currently analyzed student and j denotes the particular LO.

The RS fully depends on the student's result from answering the questions from the e-Assessment. If the student's answer on a question from a particular LO is $false$, then the student is punished with adding one more trial on his IS of the particular LO from the e-Learning knowledge background. On the opposite, when the student's answer on a question from a particular LO is $true$, the student is rewarded with reducing the number of trials by 1 in the e-Learning history. That is how we use the e-Assessment results to affect the student's impression during the semester.

C. Machine Learning Analysis

After creating the students' individual Profiles, in this section we show how to use them to train a SVM classifier to perform multi-grade classification. We consider SVM to be an appropriate classification technique since we know the desired outputs during the training process. In order to obtain the best training, we compared the performance of three types of kernels: linear kernel (LK), polynomial kernel (PK) and radial basis function (RBF). In order to avoid overfitting in the process of training the classifier, we used hold-out cross-validation technique which does not permit overlap between training data and test data, yielding a more accurate estimate for the generalization performance of the algorithm [18].

The SVM technique is mainly used for binary classification problems. Since, we want to adapt it to classify the students into six classes, we need to organize the classification across a multiple-machines strategy. Before we discuss the training procedure, we must describe the classes we want our IGA to learn from:

- C_{fail} - contains all the students who have failed the course, i.e., the average score from their e-Assessments is $\leq 50\%$;
- C_6 - contains the students with on average 51 - 60% success;
- C_7 - contains the students with on average 61 - 70% success;
- C_8 - contains the students with on average 71 - 80% success;
- C_9 - contains the students with on average 81 - 90% success;
- C_{10} - contains the students with on average 91 - 100% success;

Considering the classes, we defined two test cases for the training process:

- *Training Case 1*: In the first training case we organize the students Profiles into 2 classes; the first contains the Profiles from C_{fail} and the second contains the Profiles in the remaining $C_{6:10}$ classes. Thus, the first SVM machine will assume that $C_1 = C_{fail}$ and $C_2 = C_{6:10} = C_{pass}$. Hereafter, we need to train the classifier to recognize each grade distinctively. Therefore, we defined five additional SVM classifying machines (CM). The training process is performed in 5 iterations, which means that in each iteration we set the class C_1 to be the class C_i for $i = 6 : 10$, and the class C_2 to be a set from the classes $C_{6:10} \setminus C_i$. The whole process is depicted in Figure 2.
- *Training Case 2*: In the second training case, instead of excluding the Profiles that have failed the course when determining the correct grade, we assume $C_{fail} = C_5$. Thus, the training process is developed as follows. We define six SVM machines and each of them sets the class $C_1 = C_i$ for $i = 5 : 10$, and the class $C_2 = C_{5:10} \setminus C_i$. Figure 3 visually explains the process.

In order to evaluate the classifier's performance we use $Precision$ (4) and $Recall$ (5) measures. The classifier's $Precision$ is a measure of true positives (TP) among all Profiles classified as positive (C_1), which include the true positives and the false positives (FP). The classifier's $Recall$ is a measure of its ability to select the Profiles that are positive from all the Profiles in the positive class.

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

$$Recall = \frac{TP}{TP + FN} \tag{5}$$

The experiments and the results from the analyses are presented in Section IV.

IV. Experiments and Results

In this section we present the results from the methodology described in Section III.

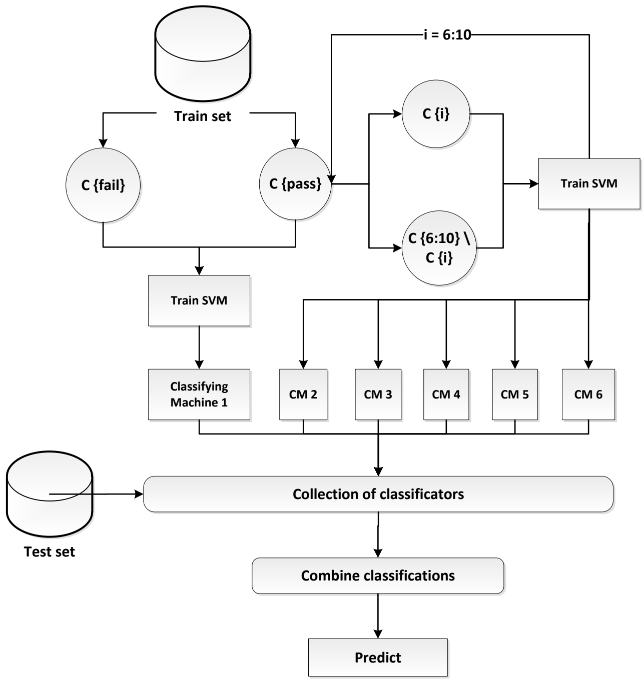

Fig. 2. Training Case 1

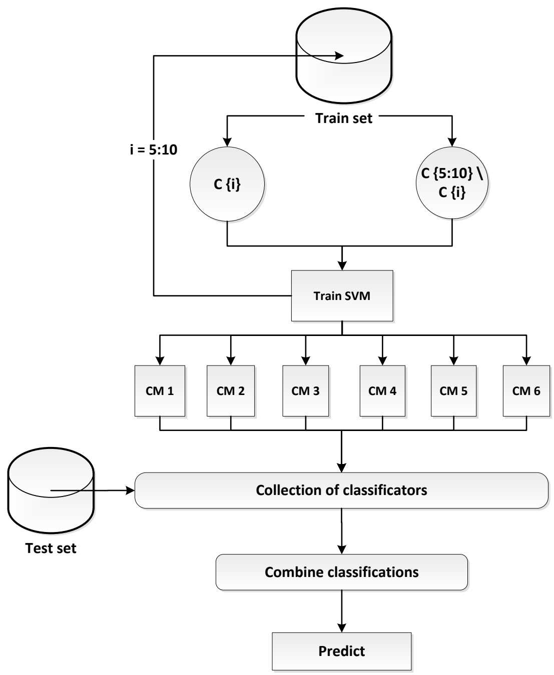

Fig. 3. Training Case 2

In order to obtain reliable results, we analysed the Profiles of 1273 randomly selected students that have attended the CAO course in the past years. The analyses were performed according to the training cases described in Section III-C, by using the *SVMlight* [19] tool. Since the results from the RBF were not promising, we present only the results from the classification with LK and PK.

Considering the fact that in both training cases, the training and the classification procedure is the same for the class of Profiles that have failed the course, $C_{fail} = C_5$, at first we describe the classification results from this analysis. Thus, in order to train the classifier to recognize the Profiles categorized as failed, we created two classes:

- C_5 which contains 629 Profiles that have failed
- $C_{6:10}$ which contains 644 Profiles that have passed

However, the initial results from the classification were not satisfying and we decided to use the bootstrapping method in order to obtain increased number of Profiles that have failed and passed the course. The method resulted in a 1000 Profiles in each class. The results from the binary classification are presented in Table I.

TABLE I
RESULTS FROM THE BINARY CLASSIFICATION

SVM Results	LK		PK	
Class	Precision	Recall	Precision	Recall
C_5 vs $C_{6:10}$	100%	100%	100%	100%

Since there are variable number of Profiles in each of the classes, we also used the bootstrapping method to enlarge the data in the first class so the number of Profiles will be the same as in the second class. Therefore, the number of Profiles before and after applying the bootstrapping method in each Training Case 1 and 2 are presented in Table II.

TABLE II
NUMBER OF PROFILES BEFORE AND AFTER THE BOOTSTRAPPING METHOD

Number of Profiles	Before	Training Case 1	Training Case 2
C_6	247	324	956
C_7	192	379	1011
C_8	101	470	1102
C_9	30	541	1173
C_{10}	1	570	1199

The preparation of the Training Case 1 and 2 according to the methodology defined in Section III-C produced the classification results in tables III and IV.

TABLE III
RESULTS FROM TRAINING CASE 1

SVM Results	LK		PK	
Class	Precision	Recall	Precision	Recall
C_6 vs $C_{6:10} \setminus C_{5,6}$	99.39%	100%	98.18%	100%
C_7 vs $C_{6:10} \setminus C_{5,7}$	99.47%	100%	98.95%	100%
C_8 vs $C_{6:10} \setminus C_{5,8}$	99.16%	100%	98.74%	100%
C_9 vs $C_{6:10} \setminus C_{5,9}$	100%	100%	100%	100%
C_{10} vs $C_{6:10} \setminus C_{5,10}$	100%	100%	98.96%	100%

The analyses from the both training cases showed that the Training Case 2 produced barely improved results, but both of

TABLE IV
RESULTS FROM TRAINING CASE 2

SVM Results	LK		PK	
Class	Precision	Recall	Precision	Recall
C_6 vs $C_{5:10} \setminus C_6$	99.58%	100%	98.96%	100%
C_7 vs $C_{5:10} \setminus C_7$	99.61%	100%	99.02%	100%
C_8 vs $C_{5:10} \setminus C_8$	100%	100%	99.46%	100%
C_9 vs $C_{5:10} \setminus C_9$	100%	100%	99.49%	100%
C_{10} vs $C_{5:10} \setminus C_{10}$	100%	100%	98.36%	100%

them resulted in very accurate and promising classifiers, which confirmed that the methodology we defined in Section III is reliable.

V. CONCLUSION AND FUTURE WORK

In this research we analyse e-Learning and e-Assessment data from the Computer Architecture and Organization course at our faculty. The data is collected from various years between 2002 (when the online learning system was implemented) and 2012.

The purpose of the analysis is creating an intelligent agent which based on the previous students knowledge will be able to determine the students' final grades. Therefore, we proposed a methodology that included modelling students' Profiles based on their activities in the e-Learning and e-Assessment processes during the course. The Profiles are then used in a two types of SVM training processes with multiple kernels, and each of them produced 6 classification machines.

The experiments and the results proved the reliability of the methodology we defined by showing that both training cases are capable of producing very accurate classifiers which is very important for our future research.

We believe that this model is essential and of great advantage for both the students, since they will stay highly motivated to successfully finish their exams until the end of the course, and for the teachers, since it can be a good indicator whether there is a need of additional testing of the student's knowledge in order to derive an overall conclusion on the most appropriate grade.

In our future work we will aim to develop different strategies for determining the final grades which will be based on the students' individual manners for passing the given learning objectives.

ACKNOWLEDGEMENT

This work was partially financed by the CodEval (Evaluating performances of program codes and educational processes using ML methods) project at Faculty of Computer Science and Engineering at the "Ss. Cyril and Methodius University", Skopje, Macedonia.

REFERENCES

[1] S. Vonderwell, X. Liang, and K. Alderman, "Asynchronous discussions and assessment in online learning," *Journal of Research on Technology in Education*, vol. 39, no. 3, p. 309, 2007.

[2] M. Gusev and G. Armenski, "E-Assessment Systems and Online Learning with Adaptive Testing," in *e-Learning Paradigms and Applications - Agent-based Approach*, M. Ivanovic and L. Jain, Eds. Springer Verlag, 2013, pp. 229–249.

[3] A. P. Rovai, "Online and traditional assessments: what is the difference?" *The Internet and Higher Education*, vol. 3, no. 3, pp. 141–151, 2000.

[4] M. Hannay and T. Newvine, "Perceptions of distance learning: A comparison of online and traditional learning," *Journal of Online Learning and Teaching*, vol. 2, no. 1, pp. 1–11, 2006.

[5] P.-C. Sun, R. J. Tsai, G. Finger, Y.-Y. Chen, and D. Yeh, "What drives a successful e-learning? an empirical investigation of the critical factors influencing learner satisfaction," *Computers & Education*, vol. 50, no. 4, pp. 1183–1202, 2008.

[6] A. C. Caminero, A. Robles-Gómez, S. Ros, R. Hernández, R. Pastor, N. Oliva, and M. Castro, "Harnessing clouds for e-learning: New directions followed by uned," in *Global Engineering Education Conference (EDUCON), IEEE*. IEEE, 2011, pp. 412–416.

[7] M. Mircea and A. I. Andreescu, "Using cloud computing in higher education: A strategy to improve agility in the current financial crisis," *Communications of the IBIMA*, vol. 2011, pp. 1–15, 2011.

[8] D. G. Chandra and M. D. Borah, "Cost benefit analysis of cloud computing in education," in *Computing, Communication and Applications (ICCCA), 2012 International Conference on*. IEEE, 2012, pp. 1–6.

[9] S. Ristov, M. Gusev, G. Armenski, K. Bozinoski, and G. Velkoski, "Architecture and organization of e-assessment cloud solution," *Global Engineering Education Conference (EDUCON), IEEE*, pp. 736–743, 2013.

[10] M. Simjanoska, M. Gusev, S. Ristov, and A. M. Bogdanova, "Intelligent student profiling for predicting e-assessment outcomes," in *Global Engineering Education Conference (EDUCON), IEEE*, 2014.

[11] F. Castro, A. Vellido, À. Nebot, and F. Mugica, "Applying data mining techniques to e-learning problems," in *Evolution of teaching and learning paradigms in intelligent environment*. Springer, 2007, pp. 183–221.

[12] O. R. Zaíane, "Building a recommender agent for e-learning systems," in *Computers in Education, 2002. Proceedings. International Conference on*. IEEE, 2002, pp. 55–59.

[13] T. Fei, W. J. Heng, K. C. Toh, and T. Qi, "Question classification for e-learning by artificial neural network," in *Information, Communications and Signal Processing, 2003 and Fourth Pacific Rim Conference on Multimedia. Proceedings of the 2003 Joint Conference of the Fourth International Conference on*, vol. 3. IEEE, 2003, pp. 1757–1761.

[14] Y.-C. Chang, W.-Y. Kao, C.-P. Chu, and C.-H. Chiu, "A learning style classification mechanism for e-learning," *Computers & Education*, vol. 53, no. 2, pp. 273–285, 2009.

[15] A. Rashad, A. A. Youssif, R. Abdel-Ghafar, and A. E. Labib, "E-assessment tool: A course assessment tool integrated into knowledge assessment," in *Innovative Techniques in Instruction Technology, E-learning, E-assessment, and Education*. Springer, 2008, pp. 7–12.

[16] Z. Harley and E. Harley, "E-learning and e-assessment for a computer programming course," *EDULEARN11 Proceedings*, pp. 2074–2080, 2011.

[17] M. Uday Kumar, J. Mamatha, S. Jain, and D. Jain, "Intelligent online assessment methodology," in *Next Generation Web Services Practices (NWeSP), 2011 7th International Conference on*. IEEE, 2011, pp. 215–220.

[18] P. Refaeilzadeh, L. Tang, and H. Liu, "Cross-validation," in *Encyclopedia of Database Systems*. Springer, 2009, pp. 532–538.

[19] T. Joachims, "Making large scale svm learning practical," Universität Dortmund, Tech. Rep. LS-8 Report 24, 1999.