BULGARIAN ACADEMY OF SCIENCES

CYBERNETICS AND INFORMATION TECHNOLOGIES • Volume **20**, No 3
Sofia • 2020 Print ISSN: 1311-9702; Online ISSN: 1314-4081
DOI: 10.2478/cait-2020-0032

Towards a Multistep Method for Assessment in e-Learning of Emerging Technologies

Ivan P. Popchev[1], *Daniela A. Orozova*[1,2]

[1]Institute of Information and Communication Technologies, Bulgarian Academy of Sciences, 1113 Sofia, Bulgaria
[2]Faculty of Computer Science and Engineering, Burgas Free University, 8001 Burgas, Bulgaria
E-mails: ipopchev@iit.bas.bg orozova@bfu.bg

***Abstract*:** *In the Fourth Industrial Revolution some important leading technologies are identified as emerging technologies with unknown in advance potential risks. Emphasized is the need for new approaches and solutions for forming of increased information awareness, knowledge and competencies in the present and future generations to use the possibilities of Industry 4.0 for technological breakthroughs. A method for evaluation and prognosis of the knowledge, skills and competencies of the students in the virtual education space is proposed in the form of a five-step process. The method can be adapted to new technologies and applications. Research and analysis of the method are carried out in the academic course 'Artificial Intelligence' at the Burgas Free University with the application of the instruments of the Orange system for experimentation and inference.*

***Keywords*:** *Industry 4.0, Emerging mechnologies, virtual education space, e-Learning, risks, artificial intelligence, machine learning, Orange system.*

1. Introduction

The Fourth Industrial Revolution [17] as a new paradigm leads to digitalization of all assets and economic agents in a uniform chain of the value and the integration in a general digital ecosystem. Unlimited possibilities are created for mutual upbuilding, increasing and merging of technologies and societies, as well as for new technological breakthroughs in many areas. In Germany, there are discussions about "Industry 4.0", a term coined at Hannover Fair in 2011 to describe how this will revolutionize the organization of global value chains.

The integration of the physical and virtual world, as well as of social communities, fundamentally distinguishes the Fourth Industrial Revolution from the previous ones. Prerequisites for its building are the transformations of leading technologies, which are referred to today as *emerging technologies*.

In many studies the *emerging technologies* are identified which form the future digital and hyperconnected world. It is well known that the summary report World

Economics Forum's Global Agenda Council on the Future of Software and Society, publishet in September 2015 [21] indentified 21 technology shifts in which over 800 executives and experts participated.

21 technology shifts are presented in the study and two additional ones, including the tipping points for these technologies and dates of their expected arrival to market: Implantable technologies, Our digital presence, Vision as the new interface, Wearable Internet, Ubiquitous Computing, A Supercomputer in Your Pocket, Storage for All; The Internet of and for Things, The Connected Home, Smart Cities, Big Data for Decisions, Driverless Cars, Artificial Intelligence and Decision Making, Artificial Intelligence and White-Collar Jobs, Robotics and Services, Bitcoin and Blockchain, The Sharing Economy, Goverments and the Blockchain, 3D Printing and Manufacturing, 3D Printing and Human Health, 3D Printing and Consumer Products, Designer Beings, Neurotechnologies. Every Technology Shift must be presented with both positive and negative impacts [15, 17, 20-22].

It is important to mention that each of the *emerging technologies* is in a logically, scientifically proven and practically justified dependence on many different *scientific areas*. The complexity and mutual involvement of these *emerging technologies* is accompanied by a significant increase of the risk factors due to the all-embracing and in some aspects spontaneous digitalization which is a reason for anxiety to the formed disruption in the relation "human-digital environment" [22].

Potential risks in *emerging technologies* can be systematized in the following separate eight groups [15]: R_d – privacy and data security; R_L – change in labor market; R_p – mental distraction; R_M – manipulation and echo camera; R_F – fragmentation; R_A – responsibility and accountability; R_E – ecology, ecosystems and ethics; R_s – change in income/cost structure and ownership of assets.

Each risk has negative, often unknown, undefined in advance impacts. This requires investigation and decision-making about the risk, which can be in the following **scheme of five phases**:

Phase 1. Identification of the risk.

Phase 2. Quantitative/Qualitative evalutation of the risk and its characteristics.

Phase 3. Choice of instrument and/or instruments for risk impact (standards, norms, rules, models, methods, algorithms).

Phase 4. Risk management – direct impact on the environment or the object through the selected instruments;

Phase 5. Monitoring, control and evaluation of the risk management which can be a sufficient reason for going back to previous phases.

Special additional research should also focus on new problem situations in "interactions" between different risks in the digital environment, such as: conflict (collision) or amplification (resonance) to varying degrees between risks. Thus, unknown new systemic risks are formed, which can manifest themselves in cascading, hierarchical or complex multi-connected behavior in cyberspace.

Industry 4.0 emphasizes on the need for new approaches and solutions for forming of increased information awareness, knowledge and competencies in the present and future generations to use the possibilities for technological breakthroughs. This raises new problems and therefore a necessity for the search of non-standard solutions in e-learning. The possibilities for carrying out and management of various processes become huge especially in a connected environment. In order for the concept of virtual learning space to continue its development in this dynamical and heterogeneous environment, it must be able to adapt to the new characteristics and requirements, which the environment imposes.

In the contemporary circumstances, for the academic community of students and teachers the virtual learning space is a real learning space.

The aim of the paper is to propose a method in the form of a multi-step process for evaluation and prognosis of the knowledge, skills and competencies of the students in the virtual learning space of the *emerging technologies,* with possibilities for adaptation to new technologies and applications.

The paper is organized in five sections. The motivation for the conducted work is described in the introduction. The second section is focused on the releated works. The third section presents a five-step method for evaluation and prognosis in e-learning, based on the large amount of accumulated data and experience in the long-term practice of the authors. The application of the evaluation and prognosis method is in the fourth section. The multitude of data accumulated during the training allows to find a connection between them and to derive a model, which is then applied to predict the level of the student at the end of the course. Applying the techniques of the Orange Data Mining System application, models are successively created with the tools: Neural Network, Random Forest, Logistic Regression and Naïve Bayes. The conclusion is in the fifth section.

2. Related works

Various forms and methods can be used for the evaluation of the knowledge, skills and competencies of the students. The final evaluation is complex and includes components, which take part in its forming with different weights. Depending on the studied problems, the criteria for evaluation can be expressed through quantitative, qualitative, fuzzy and mixed model. Various techniques for Multiple Criteria Decision Making (MCDM) and their applications are given in [6, 12]. In addition to the criteria for evaluation, the weights of the criteria can be also expressed through fuzzy numbers or fuzzy relations in order to express the meaning of the criteria [11, 13]. Various approaches exist for evaluation of the knowledge of the students, depending on the aim. Problems related to the prognostic modeling for selection of students are discussed in [2, 4]. Experiments connected to the development of models for prognosis based on data about student admission are presented in [7]. The idea of MCDM in clusters could be applied to the classification and distribution of the students in groups with the aim of their evaluation [16]. Additional coefficients can be taken into account for the significance of the experts' opinions in the determination of the final decision.

The determination of the objective evaluation of students requires the study of different aspects of the acquired knowledge and competencies. Let us take N subcriteria $C_1 = \{C_{11}, C_{12}, ..., C_{1N}\}$, connected to different evaluation components. The teacher can determine the respective weight cofficients, which express the relative significance between the subcriteria. The evaluation of the theoretical knowledge can be made for example through automatic generation and evaluation of various tests.

In [18] the generation of test is supported by specialized ontologies of two intelligent agents – Operative assistant and Evaluating assistant. The first generates test by forming questions selected randomly according to a given theme and using a database. The second one examines the answers of the users using UML ontology. A specific approach is proposed in [19] for presentation of knowledge for learning in three levels: Domain, Extractor and Generator level.

In order for a generalized evaluation of the students to be obtained, taking into account the obtained results from the separate components, a model can be used in which the utility function is

(1) $$\max \sum_{j=1}^{N} w_j e_{ij}, \text{ for } i = 1, 2, \ldots, S,$$

where w_j is a coefficient of relative significance between the subcriteria for evaluation, e_{ij} represents evaluation result of the i-th student with respect to the j-th criterion.

For the evaluation of the work of a student, taking into account the separately acquired knowledge, competencies and skills, the model from [1, 5] can be applied in which the utility function has the form:

(2) $$\max(\alpha \sum_{j=1}^{N} w_j e_{ij} + \beta \sum_{k=1}^{M} w_k e_{ik}) \quad \text{for } i = 1, 2, \ldots, S, \alpha + \beta = 1,$$

(3) $$\sum_{j=1}^{N} w_j = 1, \quad \sum_{k=1}^{M} w_k = 1,$$

where w_j is coefficient of relative significance between the subcriteria of evaluation with respect to the theoretical knowledge, w_k is a coefficient of relative significance of the subcriteria with respect to the practical skills, e_{ij} represents the evaluation result of the i-th student with respect to the j-th criterion about the theory and e_{ik} is evaluation of the i-th student with respect to the j-th criterion about the practice. Weight coefficients α and β are introduced in [1] which show how the theoretical knowledge and practical skills take part in the generalized final evaluation. Additional restriction about these coefficients is expressed through $\alpha + \beta = 1$. Normalization is required in order to provide comparable sizes between the coefficients of relative significance of the subcriteria (w_j and w_k), the additional weight coefficients (α and β) and the subcriteria evaluations e_{ij} and e_{ik}. By contrast with [1], where the focus point is to propose ranking of the students, the model in this paper aims at predicting the students at critical zone during their education on the basis of the obtained evaluations.

Another approach to the dynamical evaluation of the students, using intuitionistic fuzzy sets is presented in [9]. The evaluations of a level of acquiring $\mu(x, t)$ and non-acquiring $\nu(x, t)$ of a unit of knowledge by a student x at time t are real numbers from the set [0, 1]. The degree of uncertainty $\pi = 1 - \mu - \nu$ represents the cases in which the answers cannot be defined exactly, or a technical error has been made. The ordered pairs are defined everywhere in the sense of the theory of the temporal intuitionistic fuzzy sets.

At the beginning, when information about the studied object x has not been obtained yet, all evaluations obtain zero values. The current $(k+1)$-th evaluation is obtained on the basis of the preceding evaluations through the formula:

(4) $$\langle \mu_{k+1}, \nu_{k+1} \rangle = \left\langle \frac{\mu_k k + m}{k+1}, \frac{\nu_k k + n}{k+1} \right\rangle,$$

where $\langle \mu_k, \nu_k \rangle$ is the preceding evaluation, while $\langle m, n \rangle$ is the evaluation from the current problem, for m, n in the interval [0, 1] and $m + n \leq 1$. In this way, information from the preceding events, as well as information from the last solved problem, is included in the evaluation of each skill. The considered model of electronic evaluation proposes not only tracking of the changes of the parameters of the object being taught, but also taking into account the state of the already acquired knowledge and the possibility for application, as well as possibilities for evaluation and change of the educational themes and the criteria for evaluation.

One approach to the automatic evaluation of test units with short answer is presented in [3]. The results of the students are evaluated using a subject ontology. The system scans the obtained words and carries out a search in the vocabulary of the ontology using q-gram metrics. The q-gram metric is a measure based on symbols which evaluates the degree of similarity between two strings and can be described in the following way: Let Σ be a finite alphabet, Σ^* is the set of all strings over Σ and Σq is the set of all strings with length q over Σ for $q = 1, 2, \ldots$ One q-gram is a string $v=a_1a_2\ldots a_q$ in Σq. The q-gram distance between two strings x and y is defined as

(5) $$D_q(x, y) = \sum_{v \in \Sigma^q} \left| G(x)[v] - G(y)[v] \right|,$$

where $G(x)[v]$ and $G(y)[v]$ are the numbers of occurrencies of v in x, and v in y, respectively. The proposed approach evaluates the degree of proximity of the notions in the vocabulary of the used ontology to the words of the student's solution.

3. A method for evaluation and prognosis in e-Learning

Emerging technologies are characterized by a radical novelty, fast grow and impact on the other technologies. In other words, such technologies emerge and evolve in time and have the potential to cause significant impact on the connected to them processes of knowledge production. Despite their indeterminacy, a method can be proposed in the form of a multistep process for evaluation and prognosis of knowledge, skills and competencies of the students in e-Learning.

The method is a generalization of the approaches for the evaluation and prognosis considered in the previous section and can be adapted to various courses of Emerging technologies in the virtual learning space. The method consists five steps.

Step 1. Selection of a way of evaluation. Determination of key knowledge and competencies about the emerging technology being studied. Determination of a degree of weight (importance) for every analysed competency.

- The main **theoretical knowledge** is evaluated with the help of components such as: intermediate tests, problem solving and case solving, exams, generalized discussion, etc. These components evaluate the acquired knowledge and the capabilities to understand the theoretical material being studied. Fill-in test questions, multiple choice questions, enumeration and comparison of objects, giving examples of notions, discussion and use of algorithms are used. Questions related to explanation, interpretation and visualization of solutions are also used.

- The main **competencies** are evaluated by the students' possibilities to apply the acquired knowledge for non-standard decision making in: tests and homework, course problems and projects in which new problems are solved, critical analysis of the solutions is performed, potential risks are determined, self-dependent conclusions are drawn.

Step 2. Conducting of the evaluation process. During the overall learning process, the procedures for evaluation are performed. The results are collected and stored with the aim of specializing the students, follow-up analyses and prognosis about the evaluation of new students.

In the obtained results, dependencies are sought in separate evaluation components between the theoretical and practical knowledge, skills and competencies. In some cases, one student gets excellent score in the test and poor score in test paper, or vice versa. The creation of automatic algorithm for evaluation given specific values of the evaluated components requires non-standard solution.

Step 3. Analysis of the stored data from the training in real environment carried out. Various algorithms for **machine learning** exist. The main input data are the points obtained through the components, which are being evaluated. The output is the corresponding final evaluations of the students on specific subject. In the methods for learning, a part of the sample data is used in the learning of the algorithm. Another part of the sample data is used for testing. In case of poor results from the testing, the learning process can be repeated, or it can be decided that the selected approach is not good at solving of the particular problem.

Step 4. Creating of a prediction model. Based on the collected data, a classifier is created which makes prognosis about current evaluations on the subject. The system for data analysis such as SPSS, Orange, Weka, etc., offer tools based on algorithms for: tree of solutions, logistic regression, Bayes theorem, neural networks and others. After the completion of the algorithm learning, the realized models can be applied to new input data.

Step 5. Evaluation of the prediction accuracy of the algorithms. After the learning of the model, testing of accuracy and precision of its work is carried out:

Accuracy – measure for effectiveness as a ratio of the correctly predicted observation to the general observations,

(6) $$\text{Accuracy} = (TP+TN) / (TP+FP+FN+TN).$$

Precision – ratio of the correctly predicted positive observations to the general predicted positive observations,

(7) $$\text{Precision} = TP/(TP+FP).$$

Recall (Sensitivity) – ratio of the correctly predicted positive observations to all actual positive observations,

(8) $$\text{Recall} = TP/(TP+FN).$$

F1 Score is the mean weight value of Precision and Recall,

(9) $$\text{F1 Score} = 2*(\text{Recall} * \text{Precision}) / (\text{Recall} + \text{Precision}).$$

4. Application of the evaluation and prognosis method

Studies and analyses are performed during the teaching of bachelor students of specialties in the professional field of software engineering. The main subject which is considered is academic course Artificial Intelligence in the e-learning environment, an area in which the authors have many years of experience.

Undoubtedly, **the artificial intelligence is at the heart of the emerging technologies**, because the connected to it scientific breakthroughs form directions whose functioning depends to greatest degree on the knowledge representation and imitation of the capabilities of the human reasoning. With the development of the notions about natural intelligence, as well as the sciences connected to it, appear new directions and applications of the systems with artificial intelligence. The research and analysis in [21] regarding **technology deep shift "Artificial Intelligence"** include the following **positive impacts**: rational, data-driven decision, less bias, removal of "irrational exuberance", reorganization of outdated bureaucracies, job gains and innovation advances in medical science, disease eradication. Respectively there are **negative impacts**: accountability (who is responsible, fiduciary rights, legal), job losses, hacking/cybercrime, liability and accountability, governance, becoming incomprehensible, increased inequality, "falling foul of the algorithm", existential threat to humanity [17]. In artificial intelligence, considerable attention deserves the potential risks which according in [15] are R_d, R_L, R_M, R_F, R_A and R_E.

The course "Artificial Intelligence" was introduced in the Burgas Free University with the establishment of the specialty Informatics. In the period 1994-1996 Assoc. Prof. Danail Dochev, PhD taught the academic course "Knowledge Processing and Expert Systems" and then between 1998-2004 – the course "Artificial Intelligence". At present, lecturers of the course are Acad. Ivan Popchev and Prof. Daniela Orozova, PhD. The course is part of the virtual learning space. It is structured in four modules: Artifical Intelligence – characteristics and problems; Searching for a solution in the state space; Knowledge representation; Intelligent decision-making. Learning materials of size 587 MB in text description, examples and 14 links with useful sources are offered in the e-learning environment. In the teaching process, the Moofle and Microsoft Teams environments are used. As a basic course it is at the base of the courses: "Analysis and projection of data and knowledge bases", "Knowledge management in computer systems", as well as "Business intelligence" with lecturer Acad. Ivan Popchev, as a part of the distance learning master program "Business information technologies". The program is developed under the European Operational Porgram, project BG051R0001-4.3.04, 2007 – 2011.

Step 1. Evaluation of the students in the course "Artificial intelligence". Evaluation components are determined each of which inspects theoretical knowledge, practical skills and competencies with different cognitive level. Four tests are defined ($4\times5 = 20$ points), Project (60 points) and Generalized discussion (20 points).

The project has two parts. In the **First Part** every student selects a theme from the *basic directions* of artificial intelligence and the European policies [25, 26] such as: Ethics guidelines for trustworthy Artificial Intelligence, Ontologies engineering, Semantic Web, Knowledge representation, Computational intelligence, Robotics, Natural language processing, Machine Learning, Deep learning, Soft computing, Pattern recognition, Multi-agent systems, Artificial neural networks, Genetic algorithms, Knowledge based systems, Decision support systems, Business intelligence, Disraptive innovation, Data Science, Fuzzy sets and systems, E-learning, Ethics and Emerging Sciences, Deep fake news, Policy and investment recommendations for trustworthy Artificial Intelligence, Cyber-Physical-Social Space (CPSS), etc. The first part of the project includes state, tendencies and development, unsolved problems, conclusion and bibliography.

The **Second Part** (mandatary) in the project contains the risks. This includes identification, analysis and evaluation of the selected set of instruments for impact on the potential risks. Additional research is presented on *systemic risks* with cascading, hierarchical or complex multiconnected behavior. In conclusion, a *summary assessment* of risk management is given.

Generalized discussion with the student is on the project theme, risk management, monitoring, control and evaluation of risk management of potential risks and possibilities for solving of new problems with non-standard solutions.

The model for evaluation can be dynamically modified and adapted to a specific course. For instance, in the course "Knowledge management in computer systems" an **alternative** model is applied in which the components forming the evaluation are respectively: Test 1 (10 points), Test 2 (10 points), Test 3 (10 points), Test 4 (10 points), Homework (15 points), Project (20 points), Final Exam (25 points).

Step 2. Conducting of teaching and process for evaluation. The problem is the finding of a general approach to the automatized evaluation and prognosis of the students' results.

In view of reducing the subjectiveness of the evaluation of the project and in the summary discussion, outside evaluator from companies in the field of information technologies such as Tehnologika, Scale Focus, etc., is allowed. In every academic course, students can obtain up to 100 points and their final evaluation is formed using the following scale: from 54 to 60 points – Average (3); from 61 to 70 points – Good (4); from 71 to 80 points – Very good (5); from 81 to 100 points – Excellent (6).

Step 3. Analysis of the collected data from the conducted teachings in real environment. Many experiments have been carried out in the work process. The main goal is to solve a qualification problem by determining whether it is possible to predict the evaluation (output variable) with the help of the input variables (points of the separate evaluation components) which are preserved in the model. For the solving of the qualification problem several different techniques are applied.

As an example, Orange Data Mining System, which is an open source based on the language Python [23]. The components of Orange offer wide spectrum of possibilities – from elementary visualization of data, preliminary processing and validation to evaluation of algorithms for learning and construction of models for prediction. Initially, a work process is created; the work panel and the set of instruments are loaded. The data about the students' evaluations on the evaluation components (in number of points) are loaded through the File instrument. They can be entered from Excel (.xlsx), from text file with separators (.txt), file with data separated by comma (.csv) or URL address.

Step 4. Creation of a prognostic model. Models are created consecutively, using the **instruments of the system Orange**: Tree, Random Forest, Logistic Regression, Naïve Bayes, SVM (Support Vector Machines) and Neural Network. Through the learned algorithm, it is expected that upon setting a new combination of values of the selected evaluation components, the current and final evaluation can be determined automatically. The results of the predictions, based on the created models are related to the Summative evaluation [24] for the work of the student. The workflow for the creation of models and prognosis through the intruments of the Orange environment is shown in Fig. 1.

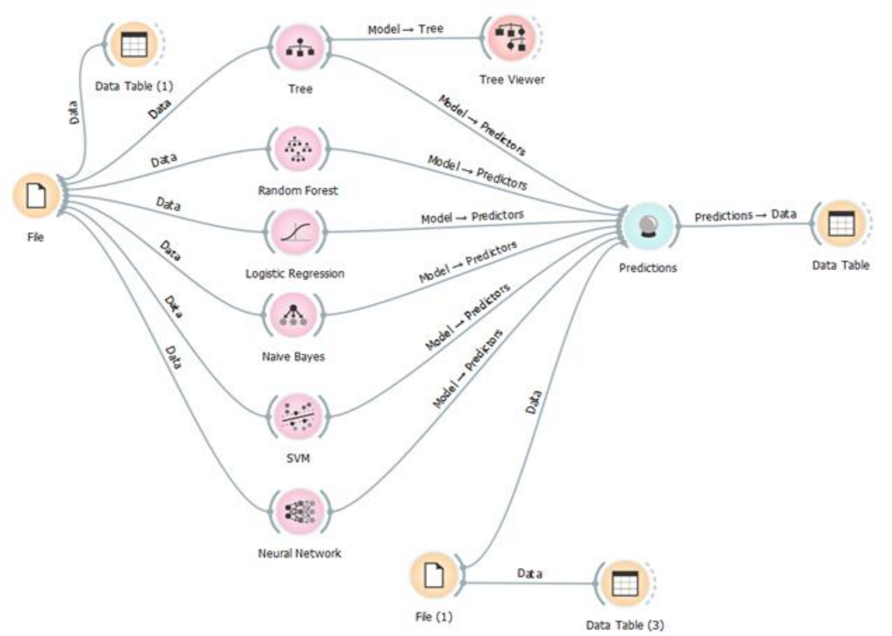


Fig. 1. Workflow for creation of the prognostic models

Table Test_eval.xlsx is prepared with the same structure as the initial table with data but the column of the final evaluation is not assigned values. From the menu Evaluate the instrument Predictions is selected, which performs the prognosis about the data in the file Test_eval.xlsx and determines a value for the unassigned column – evaluation predicted by the created model. The form of the obtained result from the prediction is shown in Fig. 2.

Graphical user interface of the system Orange presents the users with the possibility to concentrate on research analysis of data and not on the encoding of algorithms. Workflow connected to the evaluation of the models through the instruments Test & Score и Confusion Matrix is shown in Fig. 3.

Data & Predictions

	Tree	Logistic Regression	Random Forest	SVM	Naive Bayes
1	0.00 : 0.67 : 0.33 : 0.00 : 0.00 → Excellent	0.00 : 0.49 : 0.02 : 0.06 : 0.43 → Excellent	0.00 : 0.57 : 0.11 : 0.00 : 0.32 → Excellent	0.01 : 0.18 : 0.03 : 0.01 : 0.77 → Very Good	0.00 : 0.85 : 0.01 :
2	0.25 : 0.00 : 0.50 : 0.25 : 0.00 → Good	0.01 : 0.00 : 0.45 : 0.22 : 0.32 → Good	0.20 : 0.10 : 0.20 : 0.50 : 0.00 → Middle	0.22 : 0.02 : 0.23 : 0.51 : 0.03 → Middle	0.39 : 0.00 : 0.04 :
3	0.25 : 0.00 : 0.50 : 0.25 : 0.00 → Good	0.00 : 0.00 : 0.20 : 0.30 : 0.50 → Very Good	0.17 : 0.00 : 0.76 : 0.05 : 0.02 → Good	0.03 : 0.05 : 0.42 : 0.08 : 0.42 → Very Good	0.57 : 0.00 : 0.22 :
4	0.00 : 0.67 : 0.33 : 0.00 : 0.00 → Excellent	0.00 : 0.72 : 0.03 : 0.04 : 0.22 → Excellent	0.00 : 0.60 : 0.20 : 0.00 : 0.19 → Excellent	0.01 : 0.43 : 0.03 : 0.01 : 0.52 → Very Good	0.00 : 0.25 : 0.04 :
5	0.00 : 0.67 : 0.33 : 0.00 : 0.00 → Excellent	0.00 : 0.03 : 0.31 : 0.37 : 0.29 → Middle	0.00 : 0.20 : 0.44 : 0.02 : 0.33 → Good	0.03 : 0.06 : 0.58 : 0.08 : 0.24 → Good	0.01 : 0.04 : 0.86 :
6	1.00 : 0.00 : 0.00 : 0.00 : 0.00 → Bad	0.64 : 0.00 : 0.07 : 0.28 : 0.01 → Bad	0.92 : 0.00 : 0.00 : 0.08 : 0.00 → Bad	0.87 : 0.03 : 0.02 : 0.06 : 0.02 → Bad	0.86 : 0.00 : 0.00 :
7	0.00 : 0.00 : 0.00 : 0.00 : 1.00 → Very Good	0.00 : 0.00 : 0.00 : 0.20 : 0.79 → Very Good	0.10 : 0.15 : 0.00 : 0.00 : 0.75 → Very Good	0.02 : 0.20 : 0.04 : 0.03 : 0.71 → Very Good	0.01 : 0.60 : 0.00 :
8	0.00 : 1.00 : 0.00 : 0.00 : 0.00 → Excellent	0.00 : 0.80 : 0.01 : 0.01 : 0.18 → Excellent	0.00 : 1.00 : 0.00 : 0.00 : 0.00 → Excellent	0.02 : 0.87 : 0.03 : 0.03 : 0.06 → Excellent	0.00 : 1.00 : 0.00 :
9	1.00 : 0.00 : 0.00 : 0.00 : 0.00 → Bad	0.11 : 0.00 : 0.01 : 0.08 : 0.79 → Very Good	0.47 : 0.25 : 0.14 : 0.13 : 0.00 → Bad	0.29 : 0.09 : 0.27 : 0.25 : 0.10 → Good	0.46 : 0.16 : 0.33 :

Fig. 2. Results from the prognosis through different models

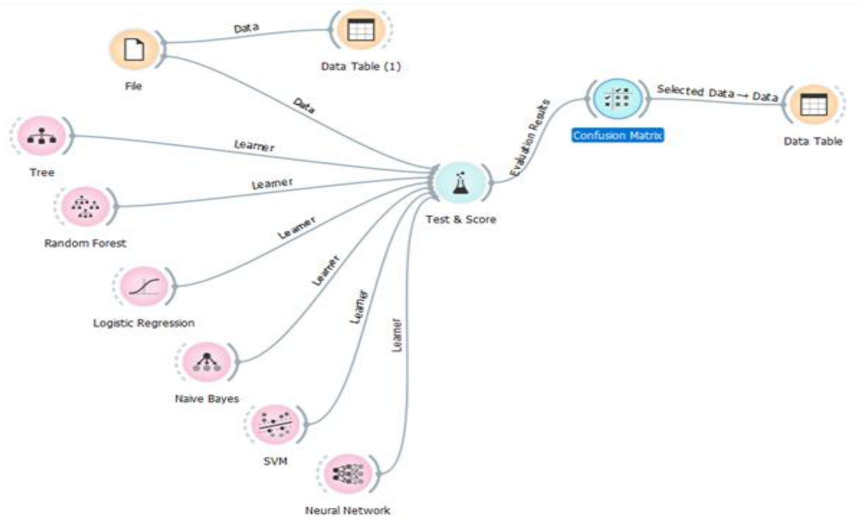


Fig. 3. Workflow process for evaluation of the models

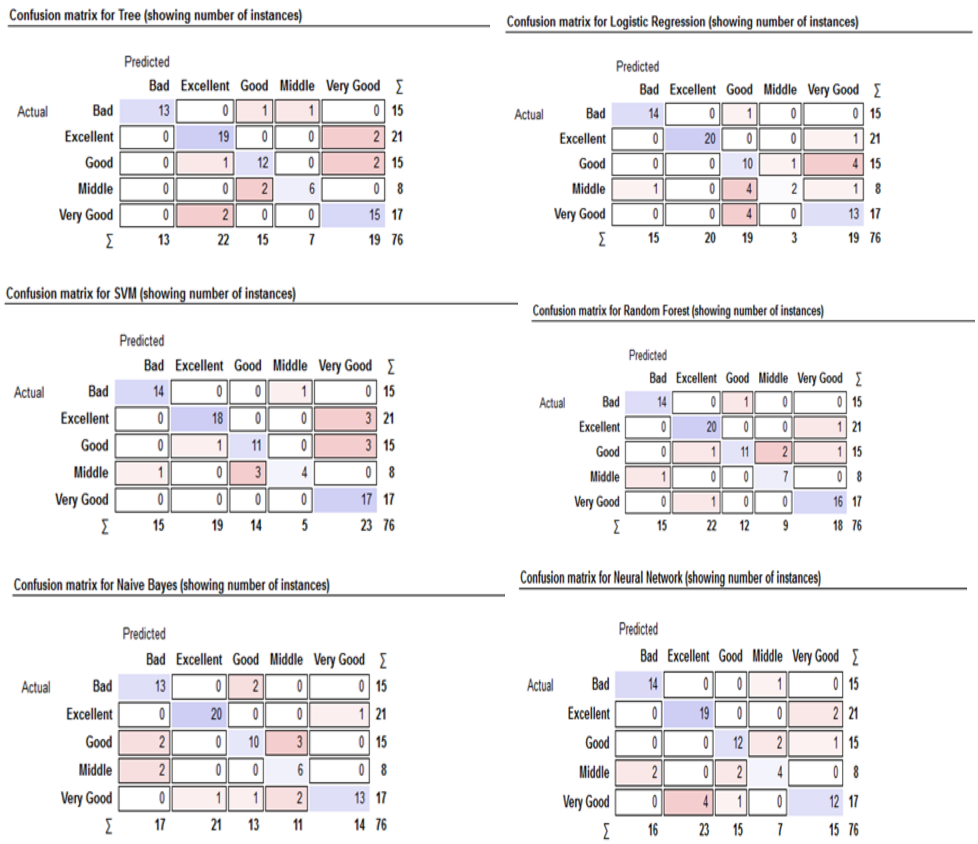
Step 5. Evaluation of the applied qualification models. The evaluation of the functioning of the models with the students' points for different evaluation components is generated with the help of the instrument Test&Score of the system Orange. The result of the work of the instrument Test&Score is a table with evaluations for Accuracy, Precision, Recall (sensitivity) and F_1 Score for the created models. Concrete evaluations of the created models are shown in Fig. 4.

Scores

Method	AUC	CA	F1	Precision	Recall
Tree	0.965	0.855	0.884	0.864	0.905
Logistic Regression	1.000	0.776	0.976	1.000	0.952
Random Forest	1.000	0.895	0.930	0.909	0.952
SVM	0.975	0.842	0.900	0.947	0.857
Naive Bayes	0.994	0.816	0.952	0.952	0.952
Neural Network	0.941	0.803	0.864	0.826	0.905

Fig. 4. Result of the work of the instrument Test&Score over the created models

Considering the accuracy of the prognosis for each of the evaluation classes, it can be summarized that it is worst for the evaluation class Middle. The highest accuracy is achieved for the classes Bad and Excellent. The prognosis for the evaluations Good and Very Good in all considered models is represented with accuracy of approximately 60-75%. The model of Random Forest is the most promising because it is represented with the highest accuracy for all evaluation classes. The *Naive Bayes* model has the lowest evaluation of accuracy with respect to the other models given the data considered. The results of the functioning of the Confusion matrix are presented in Fig. 5 for comparison of the effectiveness of the created models.



Confusion matrix for Tree (showing number of instances)

Actual \ Predicted	Bad	Excellent	Good	Middle	Very Good	∑
Bad	13	0	1	1	0	15
Excellent	0	19	0	0	2	21
Good	0	1	12	0	2	15
Middle	0	0	2	6	0	8
Very Good	0	2	0	0	15	17
∑	13	22	15	7	19	76

Confusion matrix for Logistic Regression (showing number of instances)

Actual \ Predicted	Bad	Excellent	Good	Middle	Very Good	∑
Bad	14	0	1	0	0	15
Excellent	0	20	0	0	1	21
Good	0	0	10	1	4	15
Middle	1	0	4	2	1	8
Very Good	0	0	4	0	13	17
∑	15	20	19	3	19	76

Confusion matrix for SVM (showing number of instances)

Actual \ Predicted	Bad	Excellent	Good	Middle	Very Good	∑
Bad	14	0	0	1	0	15
Excellent	0	18	0	0	3	21
Good	0	1	11	0	3	15
Middle	1	0	3	4	0	8
Very Good	0	0	0	0	17	17
∑	15	19	14	5	23	76

Confusion matrix for Random Forest (showing number of instances)

Actual \ Predicted	Bad	Excellent	Good	Middle	Very Good	∑
Bad	14	0	1	0	0	15
Excellent	0	20	0	0	1	21
Good	0	1	11	2	1	15
Middle	1	0	0	7	0	8
Very Good	0	1	0	0	16	17
∑	15	22	12	9	18	76

Confusion matrix for Naive Bayes (showing number of instances)

Actual \ Predicted	Bad	Excellent	Good	Middle	Very Good	∑
Bad	13	0	2	0	0	15
Excellent	0	20	0	0	1	21
Good	2	0	10	3	0	15
Middle	2	0	0	6	0	8
Very Good	0	1	1	2	13	17
∑	17	21	13	11	14	76

Confusion matrix for Neural Network (showing number of instances)

Actual \ Predicted	Bad	Excellent	Good	Middle	Very Good	∑
Bad	14	0	0	1	0	15
Excellent	0	19	0	0	2	21
Good	0	0	12	2	1	15
Middle	2	0	2	4	0	8
Very Good	0	4	1	0	12	17
∑	16	23	15	7	15	76

Fig. 5. Results of the work of the the instrument Confusion matrix over the created models

After generalization of the results about the mean evaluation of the prognostic accuracy over the classes of grades (Poor, Average, Good, Very Good, Excellent), the following results are obtained: Decision Tree – 85.5%; Logistic Regression – 77.6%; Random Forest – 89.5%; Support Vector Machine – 84.2%; Naive Bayes – 81.6%; Neural Network 80.3.1%. The classification alforithms Decision Tree and Random Forest, which predict with high degree of accuracy the elements of the class

Bad, are suitable for use in similar evaluation problems. These are students with poor grades who are considered as "students at critical zone".

5. Conclusion

The experiments have been carried out using available real data about students in the Artificial Intelligence course. In recent years, the collected data in the e-learning courses generate large volumes of data, which can be analyzed using Map/Reduce processing [10, 27]. Data about students failing the online courses or the distance learning form is of special interest. Here, based on the collected data about the various forms of education, connection can be sought between statistical methods, machine learning, discovery of behavioral models and analysis of large volume data.

The analysis of the data collected in the e-learning gives possibility for modification of the model and for designation of modules, which correspond, to individual necessities of the students for search of the needed information. The evaluation and prognosis allow diagnosing the learner, which could be used as feedback for the personalized learning and this is a successful form of assessing the quality of the learning process. The students' interest, their active participation in the process of knowledge acquiring and the acquiring of skills and competencies can be to a great degree influenced by the quality of the used learning environment.

At this stage of the work, the results from the software experiments are more likely helping the tutors rather than being undisputed final evaluations. In the methods proposed, there are no limitations about the number of evaluation units, which take part in the evaluation process. Strategic decisions about the learning process can be taken, using the data collected by the system for e-learning, through the application of various evaluation and prognosis models [8, 14]. The future research will be aimed towards combining of the results from different algorithms for machine learning for the drawing of a final conclusion as well as the optimization of the access to the data through the application of agents for knowledge extraction and new techniques for data analysis.

It is mandatory to include universities and research centers in European artificial intelligence networks CLAIRE, CLARIN and DARIAN as well as cooperation with European Institute of Innovation and Technology (EIT) and international organizations such as RDA. An equal partnership is a guarantee of success.

***Acknoledgements*:** This research was supported by: SRF of the Burgas Free University as a part of Project "Data Science in the Learning Space for a Blue Career", and SRF "BG PLANT NET establishment of national information network genebank – plant genetic resources", Project KP-06-N36.

References

1. B o r i s s o v a, D., D. K e r e m e d c h i e v. Group Decision Making in Evaluation and Ranking of Students by Extended Simple Multi-Attribute Rating Technique. – Cybernetics and Information Technologies, Vol. **19**, 2019, No 3, pp.45-56.
2. H u s s a i n, S., N. A. D a h a n, F. M. B a-A l w i, N. R i b a t a. Educational Data Mining and Analysis of Students' Academic Performance Using WEKA. – Indonesian Journal of Electrical Engineering and Computer Science, Vol. **9**, 2018, No 2, pp. 447-459.
3. J e c h e v a, V., D. O r o z o v a. Ontology-Based Electronic Test Result Evaluation, Advances in Intelligent and Soft Computing. – In: Proc. of 3rd International Conference of Software, Services and Semantic Technologies S3T, Springer, 2011, pp. 213-214. ISSN: 1867-5662.
4. K a b a k c h i e v a, D. Predicting Student Performance by Using Data Mining Methods for Classification. – Cybernetics and Information Technologies, Vol. **13**, 2013, No 1, pp. 61-72.
5. J a n g, L. C., T. K i m, D.-W. P a r k, D. L a n g o v a-O r o z o v a. Modelling of an Intelligent Training System by a Generalized Net, Issues in Intuitionistic Fuzzy Sets and Generalized Nets. – K. Atanassov, J. Kacprzyk, M. Krawczak, Eds. Wyzsza Szkola Informatyki Stosowanej i Zarzadvania, Warszawa, Vol. **2**, 2004, pp. 17-29.
6. M a r d a n i, A., A. J u s o h, K. M. D. N o r, Z. K h a l i f a h, N. Z a k w a n, A. V a l i p o u r. Multiple Criteria Decision-Making Techniques and Their Applications – A Review of the Literature from 2000 to 2014. – Economic Research-Ekonomska Istraživanja, Vol. **28**, 2015, No 1, pp. 516-571.
7. N a n d e s h w a r, A., S. C h a u d h a r i. Enrollment Prediction Models Using Data Mining. 2009. **http://nandeshwar.info/wp-content/uploads/2008/11/DMWVU_Project.pdf**
8. O r o z o v a, D. Appropriate e-Test System Selection Model. – Compt. Rend Acad. bulg. Sci., Vol. **72**, 2019, No 6, pp. 811-820.
9. O r o z o v a, D. Generalized Net Model of Tutoring System, Issues in Intuitionistic. – Fuzzy Sets and Generalized Nets, Vol. **5**, 2007, pp. 25-34. ISBN 978-83-88311-90-1.
10. O r o z o v a, D., K. A t a n a s s o v. Model of Big Data Map/Reduce Processing. – Compt. Rend. Acad. bulg. Sci., Vol. **72**, 2019, No 11, pp.1537-1545. ISSN 1310-133.
11. P e n e v a, V., I. P o p c h e v. Multicriteria Decision Making by Fuzzy Relations and Weighting Functions for the Criteria. – Cybernetics and Information Technologies, Vol. **9**, 2009, No 4, pp. 58-71.
12. P e n e v a, V., I. P o p c h e v. Fuzzy Multicriteria Decision-Making. – Cybernetics and Information Technologies, Vol. **2**, 2002, No 1, pp. 16-26.
13. P o p c h e v, I., V. P e n e v a. A Fuzzy Multicriteria Decision Making Algorithm. – In: Proc. of 10th International Conference on Multiple Criteria Decision Making, 19-24 July 1992, Taipei, Vol. **II**, 1992, pp. 11-16.
14. P o p c h e v, I., D. O r o z o v a. Towards Big Data Analytics in the e-Learning Space. – Cybernetics and Information Technologies, Vol. **19**, 2019, No 3, pp. 16-25.
15. P o p c h e v, I., I. R a d e v a. Risk Analysis – An Instrument for Technology Selection. – Engineering Sciences, Vol. **4**, 2019, pp. 5-20. ISSN:1312-5702 (Print).
16. R a d e v a, I. Multi-Criteria Models for Clusters Design. – Cybernetics and Information Technologies, Vol. **13**, 2013, No 1, pp. 18-33.
17. S c h w a b, K. The Fourth Industrial Revolution. New York, Crown Publishing Group, 2017. ISBN: 978-5247-5886-8, eBook ISBN: 978-1-5247-5887-5.
18. S t a n c h e v a, N., A. S t o y a n o v a-D o y c h e v a, S. S t o y a n o v, I. P o p c h e v, V. I v a n o v a. An Environment for Automatic Test Generation. – Cybernetics and Information Technologies, Vol. **17**, 2017, No 2, pp. 183-196.
19. S t a n c h e v a, N., A. S t o y a n o v a-D o y c h e v a, S. S t o y a n o v, I. P o p c h e v, V. I v a n o v a. A Model for Generation of Test Questions. – Compt. Rend. Acad. bulg. Sci., Vol. **70**, 2017, No 5, pp. 619-630.
20. R o s s, A. The Industries of the Future. Simon & Schuster. Reprint Edition (7 February, 2017). ISBN-10: 1476753660, ISBN-13: 978-1476753669.

21. Deep Shift Technology Tipping Points and Societal Impact. Survey Report, World Economic Forum, September 2015, p. 44.
http://www3.weforum.org/docs/WEF_GAC15_Technological_ Tipping_Points_report_2015.pdf
22. World Economic Forum. The Global Risks Report. 2019. 14th Edition, p. 107.
www.weforum.org.
23. Orange System [Online]:
https://orange.biolab.si/training/ introduction-to-data-mining/
24. Summative and Formative Evaluation.
https://ieeexplore ieee.org/ document/7462444
25. Policy and Investment Recommendations for Trustworthy Artificial Intelligence [Online].
https://ec.europa.eu/digital-single-market/en/news/policy-and-investment-recommendations-trustworthy-artificial-intelligence
26. Ethics Guidelines for Trustworthy AI [Online].
https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai
27. O r o z o v a, D., I. P o p c h e v. Cyber-Physical-Social Systems for Big Data. – In: Proc. of 21st International Symposium on Electrical Apparatus and Technologies SIELA 2020, 3-6 June 2020, Bourgas, Bulgaria (in print).

Received: 27.05.2020; Second Version: 29.06.2020; Accepted: 15.07.2020