Methodology for Creating Adaptive Online Courses Using Business Intelligence

Despotović, S., Marijana; Bogdanović, M., Zorica; and Barać, M., Dušan

Abstract — *This paper describes methodology for creating adaptive e-learning courses based on learning styles using business intelligence techniques and tools. Building effective e-learning framework depends on finding adequate means for discovering users' interests, preferences, motivations and needs. Data mining and its techniques are discussed as the most appropriate and sophisticated tools for fast discovering students learning styles and classification into groups. Each group attended adapted online course. Evaluation of the system showed that students achieved better results with higher level of satisfaction when attending courses adapted upon learning styles. Experiment was conducted in distance education system of Laboratory for E-Business, Faculty of Organizational Sciences in Belgrade.*

Index Terms — *adaptive e-learning, personalization, clustering students, data mining*

1. INTRODUCTION

The lack of adaptive learning environments or an environment with adaptive features is partly due to the concepts "one-size-fits-all". Very often e-learning courses have a problem of "universal size" as the same static content is presented to all students and objective is getting the learner online and 'into' the technology. A few researches proved that this type of e-learning system organization resulted in failure. Currently, the emphasis is moving toward learner oriented platforms and putting student's expectations, motivation, habits, learning styles, needs, etc. in centre of interest 0.

Manuscript received November 24, 2008.

M.D. is with the Faculty of organizational sciences, University of Belgrade

Z.B. is with the Faculty of organizational sciences, University of Belgrade

D.B. is with the Faculty of organizational sciences, University of Belgrade

According to 0, an e-learning system is considered to be adaptive if it is capable of: monitoring the activities of its users; interpreting these on the basis of domain-specific models; inferring user requirements and preferences out of the interpreted activities, appropriately representing these in associated models; and, finally, acting upon the available knowledge on its users and the subject matter at hand, to dynamically facilitate the learning process. Since the system behavior adapts to a person, this kind of adaptation is also called personalization. Thus, adaptive e-learning system can be described as personalized system, which beside contents discovery and assembly, is able to provide adaptive course delivery, adaptive interaction, and adaptive collaboration support 0. Personalized e-learning uses proactive learning strategy, which enable learner to control learning content, pace and scope.

1.1. Business intelligence and data mining

Global trends, dynamic environment and complexity of issue obligate on high degree of effectiveness, adaptability, integration and coordination of all relevant processes in e-learning. In that context, business intelligence can be recognized as fulfillment of demands for additional, undiscovered knowledge that will be used for improving e-learning process. The term Business Intelligence (BI) presents a wide area of applications and technologies for collecting, storing and analyzing data to help making better business decisions. More details about BI could be found in 0.

Data mining is one of the BI techniques and can be defined as the nontrivial extraction of implicit, previously unknown and potentially useful information from large data sets or databases. Therefore, role of data mining as adaptive mechanism in e-learning systems is obvious 0.

By using data mining tools and techniques it is possible to execute intelligent analysis of large quantities of data stored in database. Data mining can be used both as a mean for predicting unknown or future values of the attributes of interest and for describing embedded patterns that will contribute to generating the best possible personalized e-learning models. 0

Although personalized recommendation approaches that use data mining techniques are first proposed and applied in e-commerce for product purchase, there are also several works 0 about the application of different data mining techniques within recommender systems in E-learning.

1.2. Adaptive hypermedia systems

For a system to be classified as an adaptive hypermedia system - AHS 0, it must essentially meet three criteria: it should be a hypertext or hypermedia based system, it should have a valid user model, and it should be able to adapt the hypermedia using this model (i.e. the system may appear differently to each user depending on each consecutive user model. It is necessary to define basic aspects of adaptation and adequate models 0:

- What should be adapted – adaptation model
- What are parameters and environment for adaptation – user model and context model
- What is the best way for performing adaptation – instruction model and adaptation model

AHS build model of the goals, preferences and knowledge of each learner and uses this model throughout the interaction with the user in order to adapt to the needs of the learner. Popular adaptive techniques applied are adaptive content selection, adaptive navigation and adaptive presentation.

By analogy to 0, suggested solution for architecture of an adaptive e-learning hypermedia system is shown in figure 1.

Proposed model consists of three layers. First layer includes data storages located in *domain* and *context model*. These are networks of connected objects related to e-learning mission, objectives hierarchy, metadata, conceptual design. Whole AHS relied on learning materials repository and *user (learner model)*. Learner model includes data bases about students' preferences and characteristics, behavior and learning knowledge space.

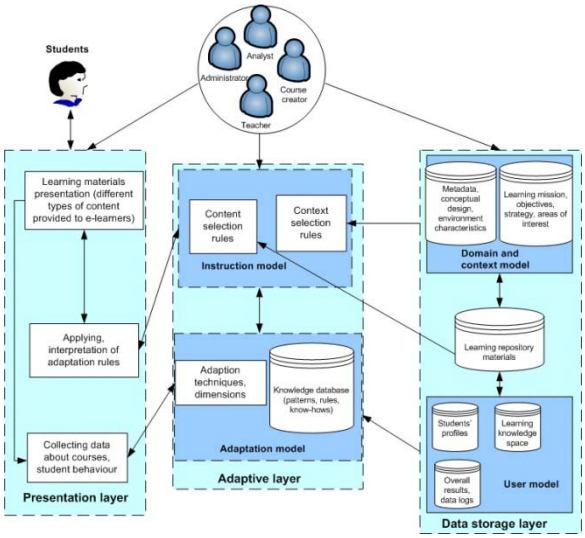

Figure 1. Three-layer architecture of AHS

The instruction (pedagogical) and adaptation models specify the navigational design for an adaptive hypermedia application. Together with the presentation specification they tell *how* the adaptation should be performed, so they describe the dynamics ("flow") of the system. Knowledge base is in the center of middle layer. It possesses variety of applied patterns, rules and "know-how", which in combination with adaptive mechanisms should develop models. Inputs in the adaptation process are sets of students' preferences and profiles. Output of the adaptation is a sequence of content objects personalized to a learner.

Naturally, on the top of the architecture are users. In order to generate effective e-learning systems, students have to be the most important part of the whole paradigm 0. Personalized e-learning implies an active cooperative learning strategy that empowers the learner to be in control of the context and scope of learning experience [9].

Obviously, the best performance in personalization would be achieved if we had information about learner's pre-knowledge, experience, usage of course content, but this issue becomes more complex if that type of data isn't available. Moreover, what if e-learning course has being created before students know nothing about the areas it deals with 0? That problem is examined in the practical example of this paper.

Student modeling is the process whereby an adaptive learning system creates and updates a student model by collecting data from several sources implicitly (observing user's behavior) or explicitly (requesting directly from the user). Traditionally, most of student modeling systems have been limited to maintain assumptions related with student's knowledge (acquired during evaluation activities) not paying too much attention to student's preferences.

1.3. Personalization based on learning styles

Learning is a cognitive activity that differs from student to student. Analyzing adaptability in e-learning system has explicitly pointed out the importance of the modeling learners' cognitive characteristics, particularly, learning styles as the most explored cognitive features.

There are several different learning style models presented in literature; however, Felder-Silverman Learning Styles Model (FSLSM) 0 is often used for providing adaptability regarding learning styles in e-learning environments. Felder-Silverman model describes single student in accordance to four dimensions 0:

- Active and reflexive learning style
- Sensitive and intuitive learning style
- Visual and verbal learning style
- Sequential and global learning style

To build the initial model, the system's authors must firstly establish the rules 0 to match learning styles with the resource's characteristics in order to determine which resources are more appropriate to a particular learning style. In table1 relations between different learning styles and activities in Moodle learning management system are shown. Later in the paper, adaptation of courses is performed in accordance to presented rules

Table 1. Moodle suitability for adaptation

	Active	Reflexive	Visual	Verbal	Sequen-tial	Global	Sensitive	Intuitive
Forum	Concrete problems	Topics for thinking	No	Yes	Yes	Global topics	Facts, ex-amples	Abstract topics
Chat	Yes	No	No	Yes	Frequent	No	Yes	No
Glossary	Many terms	Concepts	No	Yes	Yes	No	Yes	No
Work-shop	Experi-ment	Unexplored topics	Yes	Yes	Yes	Yes	Practical examples	Unexplored topics
Survey	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Choice	Yes	Yes	Yes	Yes	Yes	Rarely	Yes	No
Lesson	Problems examples	Provided topics	Illustra-tion	Written, multimedia	Yes	Rarely	Facts, algorithm	Rarely
Comm.	Face-to-face	E-mail	Com-bined	Combined	Com-bined	Com-bined	Combined	Combined

1.4. Researching goals

Primary goal of the research was to project and implement adaptive, personalized distance education system based on students' characteristics, particularly learning styles. Courses, as key part in architecture of an e-learning system, were organized and fully adapted to students' expectations. Whole process of personalization was conducted through several phases. Concrete aims of the research include:

- Developing generic model and architecture of adaptive e-learning system
- Defining main phases and requirements in developing personalized e-education systems
- Performing clustering students upon their learning styles
- Personalization and adaptation of e-learning courses

2. METHODOLOGY FOR COURSE ADAPTATION

In order to improve process of personalization in e-learning systems and make it more effective, it would be very useful to identify main phases and requirements. Steps shown in figure 2 shouldn't be realized separately, but as integrated components of iterative and dynamic process of developing adaptive e-education system.

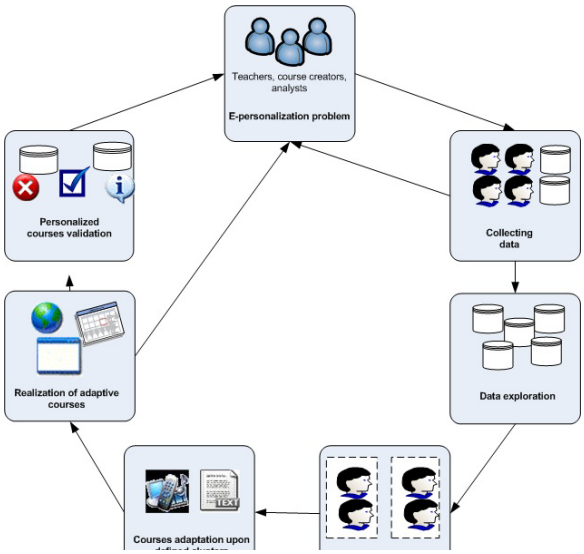

Figure 2. Phases in developing personalized e-education systems

2.1. Participants

For this analysis, data were provided by testing sample of 200 students of undergraduate studies on Faculty of Organizational Sciences, Belgrade. E-learning courses include following areas: Internet technologies, E-business and Computer simulation. In addition, we use a concept of blended learning to carry out whole process of teaching. Moodle[6] LMS has being used for creating and organizing these courses.

2.2. Methods and procedure

Defining e-personalization goals and models

Main goal of this research was to implement an adaptive e-learning system. The substantial part of the research was realized through separating students, who attend our e-learning courses, into different clusters according to their learning styles.

In the beginning it is necessary to define key segments and models[7] in our adaptive hypermedia system 0:

- *Learner model* - collecting personal data (age, sex, etc.), data related to learning styles, as well as relations with domain and context model (year of study, average mark, level of knowledge from similar areas, using Moodle, etc.)
- Domain model - three exams (internet technologies, computer simulation and e-business) on the fourth year of undergraduate studies on our faculty are completely realized through these courses. In addition, we use a concept of blended learning to carry out whole process of teaching 0
- *Content model* - refers to all content objects in scope of the courses
- *Adaptation model* - this is the position where data mining has key role. Namely, by using intelligent analyses, it is feasible to connect particular concepts and content with student characteristics (i.e. learning styles)
- *Instruction model* – set of activities that should be performed based on information realized from developed data mining model

Collecting data about students

I phase - Questionnaire

Due to restraints, which appeared in the process of personalization, not only within the scope of the courses, but also within distance education system we created a questionnaire according to FSLSM. In purpose of coordinating and supporting this uneasiness with personalization problem, it was decided to create questions in such way that they represent four dimensions of learning styles.

This was the most adequate model for adapting personalization to specific conditions. Survey consisted of 30 questions that dealt with some general topics (average grade, year of study) and the others were about motivation for learning, preferred style of communication, way of presenting content, organizing available time, team working.

II phase - Testing

In order to obtain more objective data about students' learning styles, we organized introductory course with all types of activities, resources and materials. Course was organized without any kind of adaptation. In order to complete the course, students had to pass a test that was constructed to evaluate knowledge acquired from different types of materials and activities. Test results were to be used for discovering learning styles of students.

2.3. Exploring collected data about students

All data collected in two described phases were integrated in single Excel table in such way that each question from questionnaire and from test represented one column. Rows were represented through sets of single student's answers. Although denormalized, this table was suitable for further analysis and data mining. After integration, data were transformed and reduced. Number of options in answers were transformed and reduced to three, and missing data were changed with mean values.

[6] Moodle is one of the most used web-oriented LMS. More information on http://www.moodle.org

[7] Components are defined according to the proposed architecture in figure 1

Clustering students using Data Mining Client for Excel

Clustering, as a data mining technique, was applied for building data mining model on the integrated data set prepared in previous phase. Clustering algorithm finds natural groupings among data related to sets of input attributes, so that attributes inside one group (cluster) have fairly the same values, but among groups (clusters) notable differences exist 0. It could be asserted that essential aim of clustering is discovering hidden values and variables, which precisely arrange data. Clustering was carried out using Data mining client for Excel 2007 and SQL Server Analyses services[8].

Some questions were made in order to lead process of creating data mining structures in right way:

- *What number of clusters is the most appropriate?*
- *What are main characteristics within clusters and differences between them?*
- *Which input variable, i.e. learning style, has dominant influence in grouping of the students?*
- *Is created mining model convenient for further forecast?*
- *Which content should be delivered to students from single clusters?*

Experiment was carried outfor two and three clusters. By processing and mining available data it was identified that results were almost of the same accuracy, regardless whether students were divided into two or three clusters, what is explained later in the paper. However, in second case, outcomes were more consistent, logical and of higher quality. Therefore, outputs and conclusions related to the case of three groups will be presented here.

Courses adaptation according to defined groups (clusters)

Using results of this research, changes and adaptations have been made in scope of our e-learning courses. Students have been separated into three groups. The ways adaptation was performed are described in the next chapter.

Adaptive courses realization

This phase was realized through supervising students' behavior during the courses. Courses were carried upon activities, scope and schedule proposed in previous phase.

Personalized courses validation

Final phase in developing adaptive e-learning course includes testing courses effectiveness, how created personalized models work and comparison with non-adaptive course. Further analysis and validation is discussed in next chapter.

[8] Microsoft clustering algorithm based on K-mean

3. RESULTS AND DISCUSION

In this chapter gained results from the phases were explained. First part of results is related to clustering students. For each cluster the most important characteristics and related learning styles are presented in table 2.

Table 2. Clusters and related learning styles

Cluster	Characteristics	Identified learning style
Cluster1	Multimedia materials	Visual
	Going through obligations sequentially	Sequential
	Team work	Active
Cluster2	Students choose topics	Intuitive
	Practical work	Active
	No strict deadlines	Global
Cluster3	Written materials	Verbal
	Going through obligations sequentially	Sequential
	Team work	Active

Distribution of students by clusters is shown in Figure 3. Most of the students were classified into cluster 2, which represents combination of intuitive, active and global learning style.

Figure 4 shows that aptitude about the presentation of teaching materials is of the highest importance for students' classification. Also, it is important who determines topics for essays and deadline for finishing exam obligations, professors or students themselves.

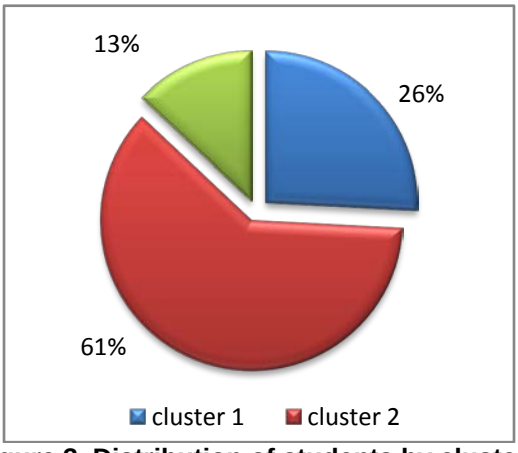

Figure 3. Distribution of students by clusters

After data mining model is created, it is necessary to validate it. The method of comparison used here is called mining accuracy lift chart 0. Results are sorted and plotted in the graph together with ideal model, which presents theoretical model with accuracy of 100%.

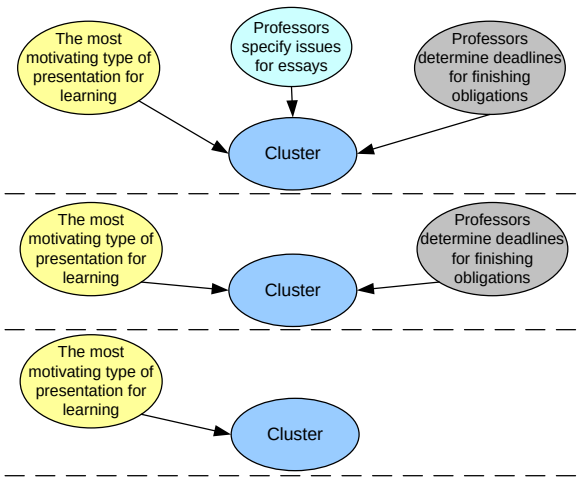

Figure 4. Dependency network

If the state of the predictable column, i.e. forecasted class in this case, is specified, than the quality of model could be analyzed. The X-axis of the chart represents percentage of the test dataset that is used to compare the predictions. The Y-axis of the chart represents percentage of values that are predicted to be in specified state, i.e. Cluster 1. The most important is that it is significantly above the blue line in the chart, which represents "random guess line". Green line, which indicates the "ideal model", shows that it would "catch" 100% of predicted population (Cluster 1) by using about 50% percentage of available data in the model with three classes. Classifying model would catch 100% of students arranged in Cluster 1 by using approximately 80% of total data.

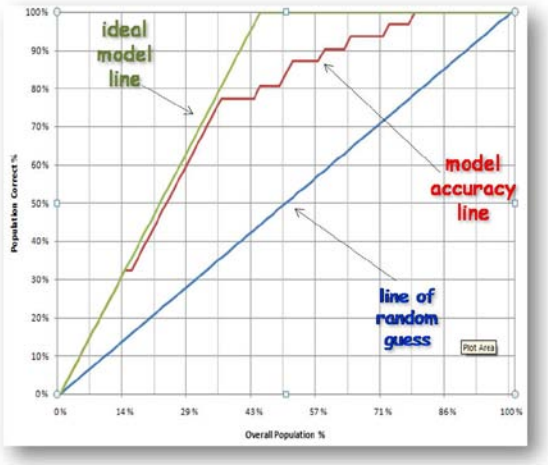

Figure 5. Model validation - lift chart

Adaptation was performed upon gained results in clustering. There were some similar characteristics for all students, like working in teams or passing exams sequentially. Indicated demands have been fulfilled for all of them by making some possibilities globally available.

Although percentages of these characteristics were fairly high in every cluster, this decision could raise question about what happened with students who didn't have same expectation. They were given possibility to choose whether they wanted to use set of global characteristics or to adjust mentioned options by themselves.

Due to the fact that the main goal was to do some fine-tuning of the courses, final adaptations based on proposed generic model and discussed relationships between learning styles and different types of presentation were reflected in:

- *Course level adaptation* - students in Group 1 and 2 were let to determine time-limits for finishing exam obligation, but students in Group 3 have been provided with the certain terms for exams.
- *Teaching materials* - Picture, video, graphs, animation and hypertext materials were delivered to Group 1, but text and audio materials were given to Group 3. Group 2 has been provided with combination of multimedia and written materials.
- *Examination* - Students in Groups 2 and 3 had projects instead of essays. In Group 1 learners could choose between practical and theoretical tasks. Homework, quizzes and oral exams were obliged for all students.
- *Activities* - Generally, almost all activities were available for students. The only difference was the way how these activities were organized (Table 3).

In the figures 6 and 7 examples of different type of materials (text and multimedia) are presented.

Table 3. Course organization and clusters

			Cluster 1	Cluster 2	Cluster 3
Type of presentation			Multimedia materials	Combined	Text materials
Course activities	Assessment	Choosing topics	Teacher	Student	Teacher
		Type of activity	Project / Case study / Survey	All types	Essay
		Final exam	Test with multimedia questions	No	Oral examination
		Deadline	Strictly specified by teacher	Not defined	Strictly specified by teacher
	Moodle activity		Video lessons, workshops	Wiki, Glossary, Lessons	Text lessons
Communication			Face-to face, Video conference	Forum, chat, Face-to face	Forum, chat, Face-to face

Figure 6. An example of text materials

Figure 7. An example of multimedia materials

Figures 8 and 9 show percentage of students who passed or failed exam in e-business organized through non-adaptive and adaptive online course. Percentage of students who passed exam is 11% higher in case of adaptive e-learning environment (figure 10).

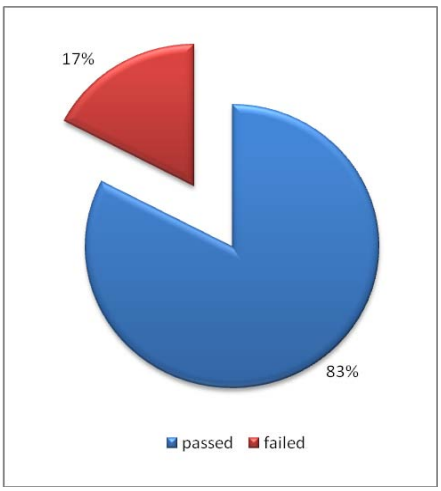

Figure 8. Non adaptive course results

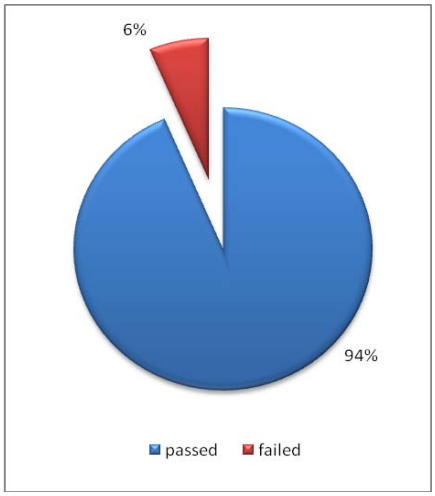

Figure 9. Adaptive course results

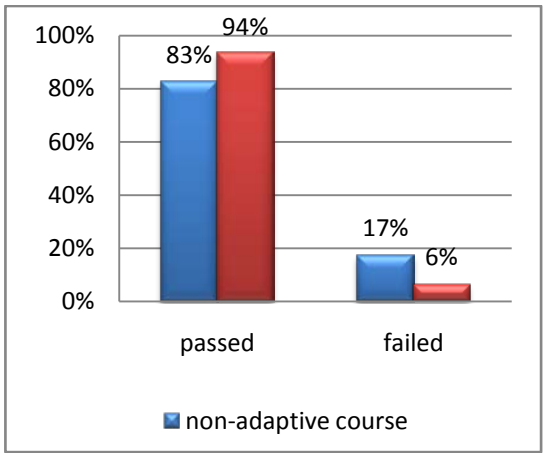

Figure 10. Comparative analysis of students' passing exams in non-adaptive and adaptive course

Comparing students' marks on non-adaptive and adaptive course, it could be realized that when applying adaptive way of organizing courses and learning, higher number of students got higher marks. In the case of non-adaptive course, students' grades are distributed near to normal distribution. On the contrary, in adaptive course there are quite a big number of students with the highest mark (10).

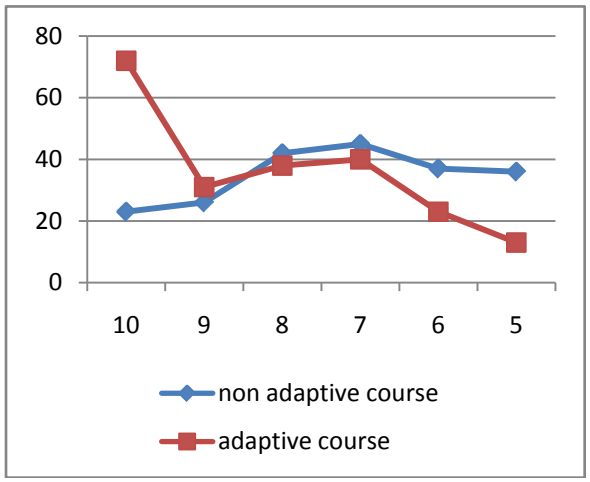

Figure 11. Grade comparison in non-adaptive and adaptive course

Since the most important for classification of students into clusters was type of presentation, we asked students if presentation of teaching materials within adapted courses was appropriate to their learning style. Only 4% of students replied that the type of presentation was not adjusted to their learning style (Figure 12.)

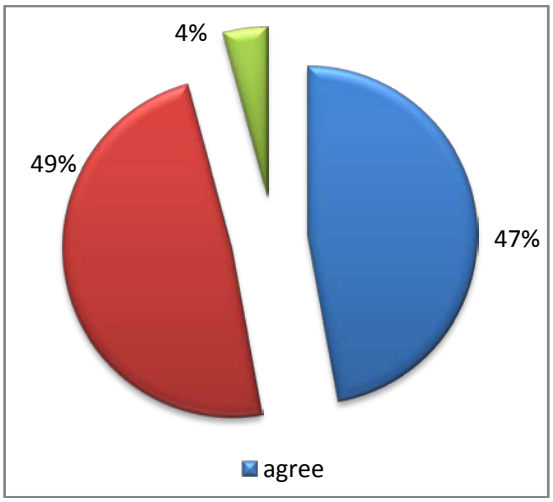

Figure 12. Suitability of teaching materials and type of presentation

When asked if adaptive course suited their learning style most students agreed that is does, while only 5% of students said that adaptive course did not fit their learning style.

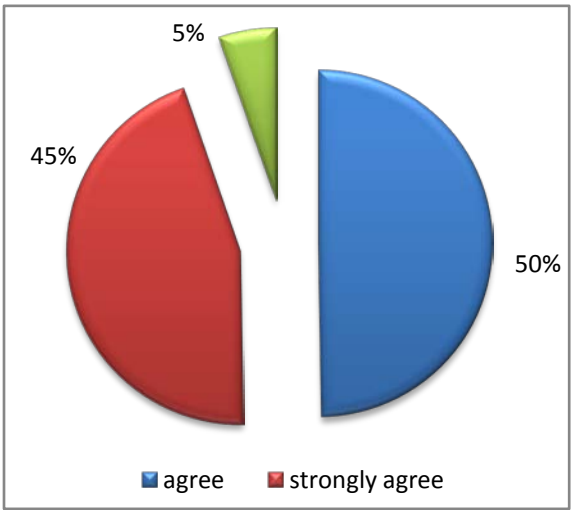

Figure 13. Suitability of adaptive course to students learning style

4. CONCLUSION

E-learning systems generate an exponentially increasing amount of data, and much of this information has the potential to become new knowledge to improve all instances of e-learning. New concepts in academic analytics imply a higher degree of standardization and uniformity in adaptation process and request real time analyzing of collected data. Simultaneously, by following the idea of "eLearning 2.0" 0, e-education systems should fulfill demands for blended learning, open access, connectivity, putting students in the centre.

In this paper the following has been done:

- Generic model and architecture of adaptive e-learning system were proposed
- Main phases and requirements in developing personalized e-education systems were identified
- Data mining model was created upon students' learning styles
- E-learning courses have been adapted and evaluated

Research will be expanded by providing some more data about students' characteristics and by improving courses according to new data. Finally, it would be very useful to develop a real-time feedback loop between intelligent analysis and the adaptive e-learning system.

REFERENCES

[1] Aroyo, L., Dolog, P., Houben, G-J., Kravcik, M., Naeve, A., Nilsson, M., Wild, F., "Interoperability in Personalized Adaptive Learning", Educational Technology & Society, 2006, 9 (2), pp. 4-18.

[2] Brusilovsky, P., "Adaptive hypermedia", User Modeling and User Adapted Interaction, Ten Year Anniversary Issue (A. Kobsa, ed.), 2001, 11 (1/2), pp. 87-110.

[3] Brusilovsky, P., "Methods and Techniques of Adaptive Hypermedia", User Modeling and User-Adapted Interaction, 1996., 6, pp. 87-129

[4] Derry, S.J., Potts, M.K., "How Tutors Model Students: A Study of Personal Constructs in Adaptive Tutoring", American Educational Research Journal, 1998., 35(1), pp. 65-99.

[5] Ebner, M., "E-Learning 2.0 = e-Learning 1.0 + Web 2.0?", The Second International Conference on Availability, Reliability and Security ARES'07, 2007., pp. 1235-1239.

[6] Felder, R.M., Silverman, L.K., "Learning and Teaching Styles in Engineering Education", Engineering Education, 1988, 78(7), pp. 674-681.

[7] Graf, S., "Fostering Adaptivity in E-Learning Platforms: A Meta-Model Supporting Adaptive Courses", CELDA 2005, 2005, pp. 440-443.

[8] Karampiperis, P., Sampson, D., "Adaptive Learning Resources Sequencing in Educational Hypermedia Systems", Educational Technology & Society, 2005, 8 (4), pp. 128-147.

[9] Burgos, D.; Specht, M., "Adaptive e-Learning Methods and IMS Learning Design: An Integrated Approach", Advanced Learning Technologies, 2006, pp.1192 - 1193

[10] Paramythis, A., Loidl-Reisinger, S., "Adaptive Learning Environments and e-Learning Standards", Electronic Journal of e-Learning, 2004 2(1), pp.181-194.

[11] Rashid, A.M., Albert, I., Cosley, D., Lam, S.K., McNee, S.M., Konstan, J.A., Riedl, J., "Getting to know you: learning new user preferences in recommender systems", ACM IUI, 2002, 127-134.

[12] Romero, C., Ventura S., (editors), "Data mining in e-learning", *WIT Press*, 2006.

[13] Tang, Z., Mac Lennan, J., "Data mining with SQL server 2005", *Wiley*, 2005.

[14] Vassileva, J., "A task-centred approach for user modelling in a hypermedia office documentation system", User Modelling and User-Adapted Interaction 1996, 6(2-3), pp. 185-223.

[15] Watson, H., Wixom, B., "The Current State of Business Intelligence", Computer, 2007, 40(9), pp. 96-99.