KNOWLEDGE DATABASES FOR E-LEARNING SYSTEMS

Mihai ANDRONIE
Spiru Haret University, 13 Ion Ghica Street, Bucharest, Romania
E-mail: mihai_a380@yahoo.com

Abstract: *Technological developments in areas such as databases and advanced data analysis techniques, but not only are directly reflected in a variety of activities, among which worth mentioning are the information and communication technology assisted learning processes, generally known as e-learning.*

With the development of knowledge databases, people interested in their own education can more easily access information they are interested in, in a shorter period of time with a higher efficiency. Knowledge data bases combined with advanced data analysis techniques can be a very important factor to improving the technology assisted learning processes carried on by learning institutions such as universities. It is possible to design and implement algorithms that, by evaluating a person's knowledge in a particular field of study can select the exact themes one needs to study, optimizing in this way the educational processes that are carried on and reducing the time needed for tuition programs. Using powerful tools such as advanced data analysis techniques it is possible to better correlate the content of a course to each student, customizing it in order to increase the efficiency of the learning processes. Universities and other educational institutions can now offer students the opportunity to quickly access exactly the information they need, when they need it.

In this paper are explored the possibilities to apply advanced data analysis techniques, also known as data mining techniques, to improve e-learning platforms in order to offer the students customized courses and to increase the efficiency of e-learning systems, offering the students more flexibility.

Key Words: *databases, knowledge, learning resources, e-learning.*

I. INTRODUCTION

The greatest advantage that is obtained through e-learning and open distance learning is the possibility of any student to attend lifelong learning programs in order to enhance their knowledge in a certain field, without attending in class courses offered by different learning institutions.

The quality and performances that are obtained by using e-learning techniques depend on both the quality of the content available and the quality of adapting that content to the needs of the students.

In order to improve the quality of e-learning techniques, one idea would be to develop and use advanced algorithms to adapt the knowledge to the needs of the students. These algorithms can be based on advanced data analysis techniques that are currently used to extract knowledge from large volumes of data. In the following sections there will be identified some areas where advanced data analysis techniques can be applied to improve the learning content and to optimize the e-learning processes.

II. KNOWLEDGE DATABASES FOR E-LEARNING

Modern e-learning IT and multimedia based technologies offer an interactive and arguably more efficient way to transmit the information from the teacher to the student. However, for the information to be possible to be transmitted to the student only by means of technology, it is necessary for it to be structured in a specific format, in other words in specialized databases, especially designed for e-learning purposes. Such databases contain all the necessary knowledge needed for the instruction processes.

Information present in e-learning databases is practically knowledge structured by the teachers in the required format.

In Figure 1 it is presented the way knowledge is transmitted from the teachers to the students, passing through knowledge databases and being then transferred to the final beneficiaries by using e-learning technologies.
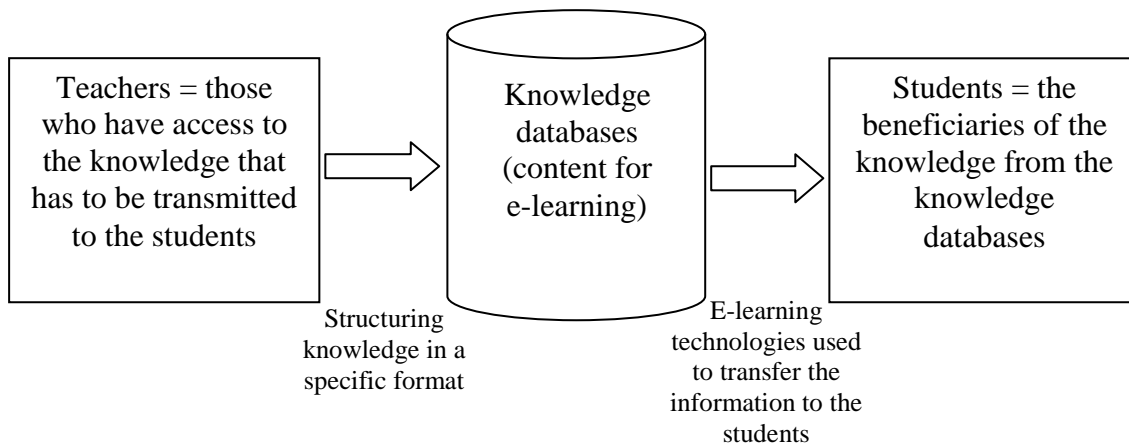

Figure 1. The knowledge on its way from the teacher to the students through the use of e-learning technologies and knowledge databases

As it can be observed from the model presented previously, all the information and knowledge that forms the content of the e-learning courses must be first known and structured by the teacher. This also happens in traditional learning.

For general purpose courses, where the content of courses changes infrequently, the approach described previously can be a satisfactory one, but for specialized courses, where the information that has to be transmitted to the students must be as much as possible up to date, there can be designed systems that automatically take some of the information contained in a knowledge database from the surrounding environment. In this way part of the content can be updated automatically.

In order to be possible to take information from the surrounding environment (like for example from different technological processes) and use it in the learning processes, this information must be extracted from large volumes of data. To accomplish such a task there must be used advanced data analysis techniques, also known as data mining techniques that will be described in the next section.

III. ADVANCED DATA ANALYSIS TECHNIQUES APPLIED ON LARGE DATABASES

Following the significant advances in information technology there appeared increased volumes of data produced and stored in various fields of activity of human society, these data containing valuable information that can be used to train all those that are interested. To extract the

useful information from data there were developed various advanced techniques suggestively called data mining techniques.

Is inconceivable to make an advanced data analysis without developing effective methods to store data, retrieve and process it, this work being facilitated by database technologies.

Advanced data analysis techniques intensively use the information technology (IT) to analyze huge volumes of data that could not otherwise be used because of the large amount of computation required for this operation.

Databases are one of the most common forms of data storage where the application of appropriate data mining techniques could lead to the extraction of valuable knowledge for the field of activity where they are used.

Databases and data warehouses were the main support for the application of advanced data analysis techniques from the beginning, while other data sources such as streams of data, graphs, multimedia data, text and web started to be analyzed much more recently.

Databases, of which the most used, are the relational databases, from the beginning were an important source of data from witch there could be extracted knowledge. They are structured collections of data organized in tables, each table having a number of columns (attributes) and a number of lines.

Databases are created and managed using database management systems (DBMS), users not having direct access to files containing data. These systems are designed to allow easy access, satisfying multiple users in a timely manner.

The relational databases permit the use of query languages such as SQL that allow retrieving data. In addition to relational databases, data mining techniques can be applied on other types of databases such as object databases or object-relational databases etc.

It is important not to confuse the application of specialized techniques for advanced data analysis of data from a database with common queries. While queries only retrieve data from database, data mining techniques can find hidden patterns, identify trends, make predictions or provide information not found anywhere in the analyzed database explicitly. Such data mining techniques are a more advanced data analysis.

Databases have the advantage that they are structured clearly and this makes them easy to be analyzed. Databases contain accurate and precise, unambiguous information, and are among the most common sources of data available. Virtually any modern organization has such a database. For this reason databases remain some of the most valuable sources of information.

Besides relational databases, data warehouses are a valuable source of data that can be exploited.

Data warehouses are homogeneous collections of data gathered from multiple sources and integrated. The result does not necessarily contain all information from source databases, but can summarize and retain only those data which are considered important. The physical structure of a relational data warehouse can be or as a multidimensional cube [2].

Besides the application of advanced data analysis techniques on databases and data warehouses, they can be applied on data formats less conventional, such as data flows, time series, sequences and even unstructured data such as images, text, etc.

Applying advanced data mining techniques on data flows has huge potential for tasks such as Internet traffic analysis, analyzing the data generated by surveillance systems, real-time analysis of data generated from industrial production processes, and analyzing other data generated by activities with significant dynamics [2], making the information thus obtained useful for instruction processes based on e-learning technologies.

A time series database consists of sequences of "values or events resulted from repeated measurements over time" [2]. Data from time series databases can be acquired at regular intervals but this is not compulsory. It can also be obtained at irregular time periods provided they are in chronological order.

As data flows, time series data may be present in large quantities and therefore requires high-performance algorithms for later analysis.

Extracting knowledge from graphs and social networks is another direction of development in advanced data analysis. Graphs are important because they allow modeling of heterogeneous

structures such as text data, web-based data, social networks, and other data showing a high degree of complexity. Data analysis on social network structures is similar to analysis of graph structured data.

New developments in information technology have made possible the extraction of knowledge from new kinds of data less conventional. A data type that can be analyzed is text data in electronic format, becoming increasingly frequent in recent years. Although data mining techniques applied on documents, books, and other electronic texts is intended to extract knowledge, we must not confuse this process with a simple retrieval of information from these sources [4].

With the very rapid development of Internet, it became one of the most important data sources available. Consequently, there began to appear different advanced data analysis techniques to extract knowledge from the web that have great potential uses.

Over time there have been developed a large number of data mining algorithms, more or less complex, usually implementing specific functions.

IV. APPLYING ADVANCED DATA ANALYSIS TECHNIQUES FOR E-LEARNING IMPROVEMENT

It is possible to apply advanced data analysis techniques to obtain useful knowledge from large volumes of data. This knowledge can be integrated into knowledge databases and further used in the e-learning processes. As it can be seen in Figure 2, there are three possible uses of advanced data analysis techniques that can improve the e-learning processes:

- Applying advanced data analysis techniques on the available information, data and knowledge related to the studying programs; in this way it is possible to improve the quality of the knowledge used in the e-learning processes, keeping it up to date;
- Applying advanced data analysis techniques on the data from the knowledge database, in order to select the exact information a certain student needs (to customize the lesson to the student's requirements);
- Applying advanced data analysis techniques on the data selected for a certain student in order to give the student the knowledge he needs in an optimized format.

Figure 2 presents the information flux from gathering data from the real processes to knowledge databases and then to the student. By using advanced data analysis techniques and e-learning as presented above it is possible to offer each student the exact knowledge he needs, making the whole educational process more efficient. A student will focus more on the content he needs most, presented in a format that is more accessible to him.

In order to customize the e-learning content it is important to follow two steps:

- Selecting content from the content database (knowledge database), based on an assessment test specifically designed to cover the particular needs of the student;
- Identifying a form of presentation that is the most suitable to a particular student by using an assessment test of the learning habits of that student.

As it can be observed in Figure 2, based on the feedback regarding the knowledge a student needs, content is selected from the database. At this stage there are necessary algorithms to interpret the feedback (the completed assessment test) and based on the results to select the knowledge necessary. If for example, after the assessment test regarding the knowledge needed it results that a student has problems with at mathematics, then in the content selected for him there will be more mathematics lessons and exercises.

For interpreting the feedback regarding the individual habits of a student, other specialized algorithms are needed. These will take as an input the raw content provided by the previous stage and put it in the most appropriate format. Depending on the algorithms used, there can be needed some templates for different lessons. One simple example of adapting the lesson to the needs of a particular student can be to adapt the length of the lesson to the time he is able to learn. For most adults for example, lessons must not be longer than about half an hour if they are learning in the evening, after work.
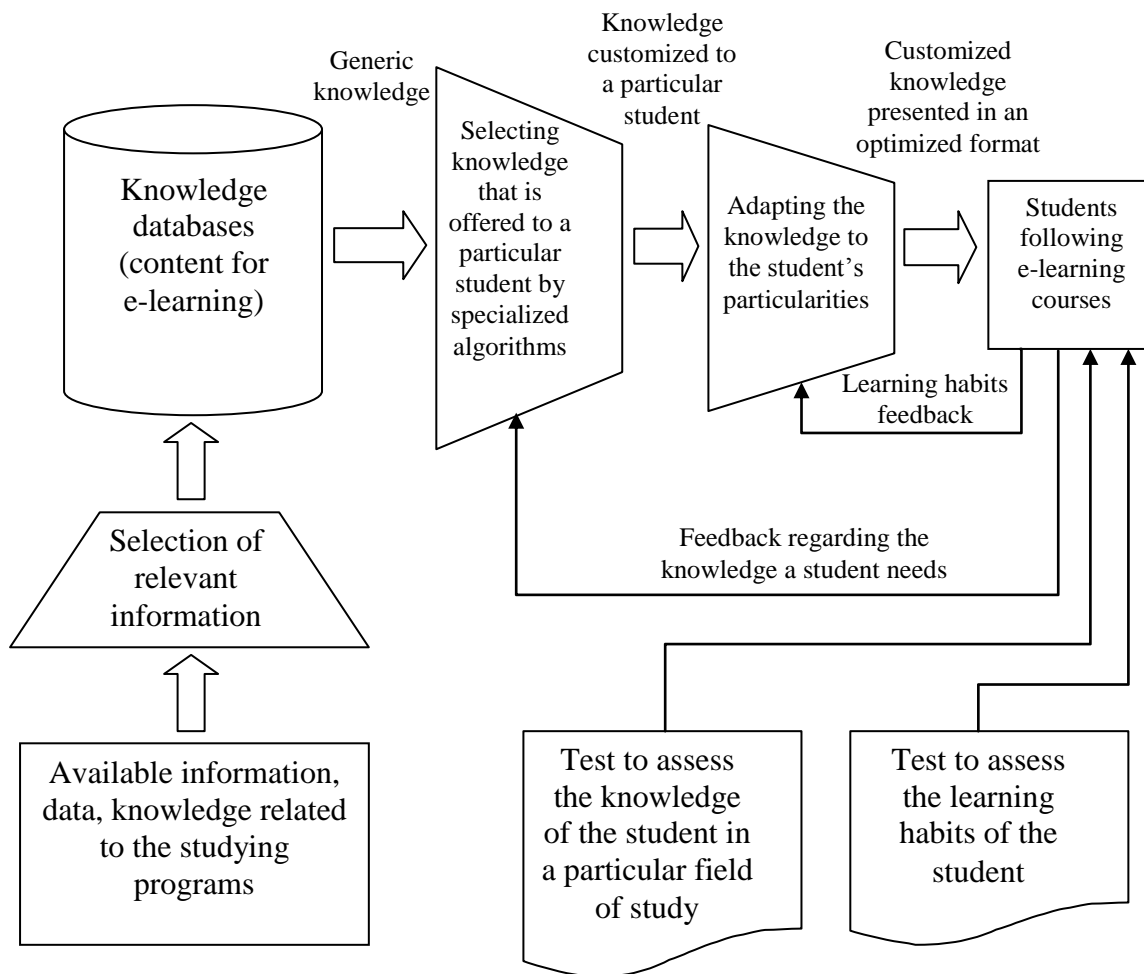
Figure. 2. Offering customized e-learning services by analyzing the student's needs

V. CONCLUSIONS

It is possible to apply advanced data analysis techniques to obtain useful knowledge from large volumes of data to improve the e-learning processes. These techniques can be applied in more directions like: improving the e-learning knowledge databases by updating the content automatically, selecting content that is needed for each sent and adapting the presentation of that content to the individual learning habits of each student.

To accomplish this tasks specific algorithms must be developed, based on the advanced data analysis techniques that are currently used to extract knowledge from large volumes of data.

References

[1] Mihai Andronie, *Modern data mining techniques*, The Ninth International Conference on Informatics in Economy IE, 2009;

[2] Jiawei Han, Micheline Kamber: *Data Mining. Concepts and Techniques*, Second Edition, Elsevier, 2006;

[3] Ingrid Fischer, Thorsten Meinl: *Graph Based Molecular Data Mining - An Overview*, 2005;

[4] Marti A. Hearst: *Untangling Text Data Mining*;

[5] Andronie Maria: *E-Learning Course Guide,* International Scientific Symposium- Quality Management in higher education- Spriru Haret University- The Faculty of Marketing and International Economic Business".