

# Intelligent learning system for High education

Aleksandar Kotevski<sup>1</sup> and Gjorgji Mikarovski<sup>2</sup>

**Abstract** – The purpose of this paper is proposing model for intelligent system for learning in high education, by giving suggestions to users. The main goal is to be found and present the most useful and practice text content to student from a large set of data, which are stored in database. In other words, this model is going to present adaptive learning system, so throughout the interaction with the students it's going to adapt to their needs.

The most appropriate content that will be presented to students will be found in using phrase searching, under the following criteria: the most visited page from student, student's keyword in his user profile, pattern relations and the words and phrases that are often used from users while searching content.

Student interaction is the most important part for this model. The user's responses are the most crucial factor that affecting the increasing intelligence of the learning system.

**Keywords** – Data mining, intelligent, learning system, patterns, phrases

## I. WHAT IS DATA MINING

Data mining is an important data analysis methodology that has been successfully employed in many domains, and which has become especially popular after the World Wide Web made large volumes of data on many topics widely available. Data mining is an important paradigm for educational assessment. The usual assumption is that mining is performed after educational activity with that activity having been designed without regard for the mining process [1].

Gartner Inc.'s definition of data mining is the most comprehensive: "...the process of discovering meaningful new correlations, patterns, and trends by sifting through large amounts of data stored in repositories and by using pattern recognition technologies, as well as statistical and mathematical techniques".

It uses a combination of an explicit knowledge base, sophisticated analytical skills, and domain knowledge to uncover hidden trends and patterns. These trends and patterns form the basis of predictive models that enable analysts to produce new observations from existing data. Speaking about intelligent systems we stress that these systems apply techniques from the field of Artificial Intelligence (AI) to provide broader and better support for the users of Web-based educational systems. [2]

From educational aspect, data mining can be used as tools for better allocating resources, content and useful links that are related with student needs, and also improving the effectiveness of learning systems. Because almost every

learning system has large volumes of data, data mining can be used to discover hidden patterns and relation that are very helpful when system has to make decision for next content that should be present to the student. By collecting information on a particular student's performance as well as other cognitive and no cognitive variables, the software can make inferences about strengths and weaknesses, and can suggest additional work [3]. Since web-based educational systems are capable of collecting vast amounts of student profile data, data mining and knowledge discovery techniques can be applied to find interesting relationships between attributes of students, assessments, and the solution strategies adopted by students [4].

It means that the main goal of intelligent learning system is to recognize uncovered hidden patterns and the relation between learning content.

## II. RELATED WORK

Using system for online learning has more advantages, but also and some disadvantages. Because each learning system has number of courses and learning content, large bodies of text, images and multimedia materials, user will waste the time while they are searching in large set of information and found something useful for them. Some recent works addressed this problem. One approach supports the analysis of large bodies of texts by interaction techniques together with a meaningful visualization of the text annotations. For example Compus [5] supports the process of finding patterns and exceptions in a corpus of historical document by visualizing the XML tag annotations. Another approach is to use data-mining algorithms integrated with visual interfaces so that non-specialists can derive benefit from these algorithms [6]. Data mining is an emerging methodology used in educational field to enhance our understanding of learning process to focus on identifying, extracting and evaluating variables related to the learning process of students [7]. Galit has completed a case study that uses student's data to analyze their learning behavior to predict the results and to warn students at risk before their final exams [8]. Han and Kamber explained that k-means is a well-known clustering algorithm tends to uncover relations among variables already presented in dataset. Erdogan and Timor used educational data mining to identify and enhance educational process which can improve their decision-making process [9].

## III. PROPOSED MODEL

The goal of this paper is to propose model for intelligent system for learning in high education, by giving suggestions to users. It means that system will learn from students visiting history and will adapt to user's needs. In that content, learning system is going to propose learning content to registered

<sup>1</sup>Aleksandar Kotevski is with the Faculty of Law, Partizanska B.B. 7000, Bitola, Macedonia E-mail: aleksandar.kotevski@uklo.edu.mk.

<sup>2</sup>Gjorgji Mikarovski is with the Faculty of Technical Scientist of Bitola, Ivo Lola Ribar B.B.,7000,Bitola,Macedonia, Email:gjorgji.mikarovski@tfb.edu.mk.

students by following criteria: history visits, keyword in student's profile, searching words, pattern relations and student response.

On the other site, the learning content author manually fills-up the phrases in field's title, META tags and keyword. Phrases included within the keyword and description META tags are valuable precursors to the contents of the document.

The titles, which are enclosed within the TITLE tag of the Web documents' headers, were identified as another useful source of phrase sets whenever the contents of the TITLE tag are delimited either by commas, semi-colons or vertical bars[10]. Each phrase within the TITLE tag can be considered as a potential phrase that could be used to define the target phrase.

```
<html>
<title>title page</title>
<meta name="keywords" content="word1, word2, word3,
word4, ..." >
<body>
learning content
learning content
learning content
...
learning content
</body>
```

For more efficient searching, the system is going to search for matching phrases in title, META tags and learning content.

*A. User history*

The learning system will bookmark all pages that the user had been visiting, and also order of visiting. Visiting history may be getting as parameter while proposing next content. To enable this feature, we are going to use database table with following fields: preview-page (preview visit page), active-page (page that user is viewing), next-page (next page that user will view) and repeat (parameter that indicate how many times this order of page visiting was matched).

Let say user is visiting page B. Preview visit page was A. In that situation, system will search in table and find out each record that fulfills the condition: preview-page = A and active-page = B. Result set is two-dimensional array: next-page and indicator repeat. After query execution, user will get list for recommendation page for visiting, ordered by repeat indicator. It means that result with the highest repeat indicator will be to the top of list of suggestions.

```
$prev_page = $_SESSION['prev_page'];
$current_page = $_SESSION['current_page'];

$result = mysql_query("SELECT next_page FROM
tbluser_histiry where preview_page = '". $prev_page.'" And
active_page='". $current_page.'" order by repeat desc");
```

*B. Keyword*

Each user has ability to insert keywords which are important for searching content. User can insert keyword while registering to the system, or later in his profile settings page. When user has inserted some keywords, system will search in database and will show all learning content related to the keywords. Keyword repeating in the content will determine the order of proposed list. Let say we have keyword K, and there are few records A,B,C and D.

Keyword K is found in A 5 times, in B 3 times, in C 6 times and in D 10 times. So, proposed list will have following order: D, C, A, B.

```
$result_kewywords = mysql_query("SELECT keywords
FROM tbluser_profile where user_id='".
$_SESSION['user_id']."'");

while($row_kewywords =
mysql_fetch_array($result_kewywords)){
$list_keyword[]=$row_kewywords['keywords'];}

$result = mysql_query("SELECT * FROM tblcontent where
title like '%" . $row_kewywords['keywords'] . "%' and keywords
like '%" . $row_kewywords['keywords'] . "%' and content like
'" . $row_kewywords['keywords'] . "'");
```

*C. Searching keywords*

The proposed model will have the ability to store searching history of each user. In addition, each word and phrase that user was using in search form will be saved in database. The table has the following fields: word (unique), repeat. Let say table already has record with value intelligent (word field). If user inserts word intelligent in search form, then system will check in database if that word already exists in table. If yes, system will update only repeat value (increment for 1). Otherwise, system will insert new row in the table. While user searching through learning content, system updates list permanently. The list is sorted by the repeat field and is using while system decision for proposing learning content.

```
$result_kewywords = mysql_query("select keywords from
tbluser_search_history where user_id ='" .
$_session['user_id']."'");
while($row_kewywords =
mysql_fetch_array($result_kewywords)){
$list_keyword[]=$row_kewywords['keywords'];
}
$result = mysql_query("select * from tblcontent where title
like '%" . $row_kewywords['keywords'] . "%' and keywords like
 '%" . $row_kewywords['keywords'] . "%' and content like
'" . $row_kewywords['keywords'] . "'");
```

On the other words, the system will give suggestion to user through searching learning content by words and phrases from the list.

*D. Related patterns*

Each searching pattern has own rating value. Patterns with the highest rating value are often in use in searching process instead of patterns with lowest rating value. On the other site, each pattern has relation with other pattern. That's why this model proposes using kind of relationship mining. In that content, this model proposes using relationship mining for discovering relationships between patterns. This may take the form of attempting to find out which patterns are most strongly associated with a single pattern of particular interest, or may take the form of attempting to discover which relationships between any two patterns are strongest.

For this reason, the proposed model is going to use table for relation between searching patters plus field with synonyms for selected searching pattern.

*E. User response*

User response is the most important part of this model. Response will be the best direction for making learning system more intelligent. In that content,we'll require from user to fillquestioner with following questions:

- Does suggested page was adequate for user
- Specify next page that is related with selected page, but it isn't into the suggested list

User interaction is the most important part for this model. The user is factor that influence in system learning for user purposes. It means that after content showing, user has to rate if the content has helped to them. That rating tell to system does selected text pattern is useful for user or not.

Each user has own setting profile, so he can set up what factor has more priority while searching learning content. Let say user has following priority settings:

- 1) Searching keywords
- 2) User history
- 3) Keywords
- 4) User response

It means that system will recognize following array:  
 \$sort\_array = ('searching\_kewywords', 'user\_history', 'keywords', 'user\_response')

Result set will be order by relevance:  
 select \* from tblcontent where title like '%"\$.keyword.%'  
 and keywords like '%"\$.keyword.%' and content like '%"\$.keyword.%'  
 order by (  
 ( case when '".sort\_array[0].'" like '".keyword.'" then 1 else 0 end ) + (  
 ( case when '".sort\_array[1].'" like '".keyword.'" then 1 else 0 end ) + (  
 ( case when '".sort\_array[2].'" like '".keyword.'" then 1 else 0 end ) + (  
 ( case when '".sort\_array[3].'" like '".keyword.'" then 1 else 0 end ) ) desc

```

case when '".sort_array[1].'" like '".keyword.'"
then 1
else 0
end
) + (
case when '".sort_array[2].'" like '".keyword.'"
then 1
else 0
end
) + (
case when '".sort_array[3].'" like '".keyword.'"
then 1
else 0
end
)
) desc
    
```

IV. CONCLUSION

For universities, data mining techniques could help to provide more personalized education, maximize educational system efficiency, and reduce the cost of education processes. It may guide us to increase student's retention rate, increase educational improvement ratio, and increase student's learning outcome.

The data collected from different applications require proper method of extracting knowledge from large repositories for better decision making. Knowledge discovery in databases (KDD), often called data mining, aims at the discovery of useful information from large collections of data [11].

REFERENCES

- [1] Steven L. Tanimoto, University of Washington, Dept. of Computer Science and Engineering: Improving the Prospects for Educational Data Mining
- [2] Adaptive and Intelligent Web-based Educational Systems, Peter Brusilovsky, School of Information Sciences, University of Pittsburgh, 135 North Bellefield Avenue, Pittsburgh, PA 15260, USA, ChristophPeylo, Software LogistikimArtland, Friedrichstr. 30, 49610 Quakenbruck, Germany
- [3] Ducational measurement and intelligent systems, Valerie J. Shute, Diego Zapata-Rivera
- [4] Data mining for a web-based educational system, Behrouz Minaei-Bidgoli
- [5] Telecommunications, 2007. ConTel 2007. 9th International Conference, Vranic, M.; Pintar, D.; Skocir, Z.:The use of data mining in education environment
- [6] Mining "Hidden Phrase" Definitions from the Web, Hung. V. Nguyen, P. Velamuru, D. Kolippakkam, H. Davulcu, H. Liu, Department of Computer Science and Engineering Arizona State University, Tempe, AZ, 85287, USA
- [7] Mining Educational Data to Analyze Student's Performance, Brijesh Kumar Baradwa, Saurabh Pal, (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 2, No. 6, 2011
- [8] Galit.et.al, "Examining online learning processes based on log files analysis: a case study". Research, Reflection and Innovations in Integrating ICT in Education 2007

- [9] A data mining application in a student database. *Journal of Aeronautic and Space Technologies* July 2005 Volume 2 Number 2 (53-57)
- [10] Heikki, Mannila, *Data mining: machine learning, statistics, and databases*, IEEE, 1999
- [11] Church, K.W., and Helfman, J.I., *Dotplot: A Program for Exploring Self-Similarity in Millions of Lines of Text and Code*, In *Proc. of the 24th Symposium on the Interface, Computing Science and Statistics V24*, 58-67. 1992
- [12] Fekete, J. and Dufournaud, N., *Compus: visualization and analysis of structured documents for understanding social life in the 16th century*. In *Proc. of the Fifth ACM Conference on Digital Libraries*, 47-55. 2000
- [13] Jing Luan, Chief Planning and Research Officer, Cabrillo College Founder, *Knowledge Discovery Laboratories: Data Mining Applications in Higher Education*
- [14] Heikki, Mannila, *Data mining: machine learning, statistics, and databases*, IEEE, 1999