

WCLTA 2010

Using algorithms for evaluation in web based distance education

Tuncay Sevindik ^{a*}, Zafer Cömert ^b

^aAssist.Prof.Dr., Yıldız Technical University, Faculty of Education, Department of Computer and Inst. Tech. Edu, Istanbul, 34220, Turkey

^bRes. Assist., Bitlis Eren University, Faculty of Engineering, Department of Computer Engineering, Bitlis, 13000, Turkey

Abstract

Traditional assessment approaches are still being used in distance education environments. Positive changes have been experienced on dimensions of user, management and teacher in distance education systems at each passing day. In addition to these positive changes, new approaches to be used at the evaluation of distance education are emerging. Each of these approaches is an algorithm. In this study, the algorithms to be used at the evaluation of distance education platforms are analyzed and compared. Distance education algorithms as K-means, Apriori, C45, Support Vector Machines (SVM), KNN and Naive Bayes are created the universe and sample of this research. As a result, it is determined that which algorithms can be effective at analyzing of the student behavior, dimension of management and giving more impressive decision of the teachers.

© 2010 Published by Elsevier Ltd. Open access under [CC BY-NC-ND license](http://creativecommons.org/licenses/by-nc-nd/3.0/).

Keywords: Algorithm, Assessment, Distance Education, Evaluation;

1. Introduction

Evaluation concept in education is described as the last step among the items of developing a programme. Evaluation item accepted as a feedback in deciding whether a concept of education is effective or not, is the most important keystone in modern education system. Also it is an important step in traditional or internet based distance education system at the time of decision to reach the target. Evaluation concept is also divided in to two approaches as traditional evaluation and alternate evaluation. However both types of the evaluation are used in distance education in addition to traditional education. But in distance education environment generally traditional evaluation approach is used. Its most important reason is that data, web and text mining is not generally employed in evaluation perception in distance education environment. Firstly what is distance education?

Distance education is a concept which helps teacher and student get together in dual interaction, different place and time or simultaneous different places via single or dual interaction and saves up time and place. One of the most important questions in modern education systems is there are not enough qualified teachers and the ones who exist just educate a very limited group. And the limited education environment is at schools in certain determined time and place. Distance education approaches eliminate these limits and enable qualified educators reach anywhere in the world. Even if distance education cannot be seen as a significant rival to face to face education, owing to limitations of face to face education, distance education is needed as an alternative way. At this point distant

* Tuncay Sevindik. Tel.: +0-506-894-2277

E-mail address: tsevindik@gmail.com.

education will be the first option to be chosen because of its ability reaching larger groups and being used without limitation of time and place.

The previous distance education systems were one way newspaper, radio, letter or television based ones. We are introduced to computer based distant education with the development of computers and in 1990's with internet becoming widespread we are introduced to totally internet based education environments. Even if it is common idea that distance education concept appeals to students only, it is also for the usage of the adults and the people who have professions like providing in-service training.

There are various distance education models used around the world and a lot of studies to improve these models. Since 1955 the options of distance education have increased with the development of radio and television. The emergence of satellite technology has enabled the distance education in different countries. As an example "Open University" in England founded in 1968 provides distance education in other countries. In 1980's when the computer technology became cheaper, the usage of computers became widespread. With the increase of opportunities provided by the internet, computer becomes very effective in distance education system (Daş & Varol, 2001).

"Distance education is a kind of revolutionary breakthrough and pioneer contemporary education model in education environment which has emerged during 1920's technological developments." (Daş & Varol, 2001).

In distance education, content and education management systems are the second and the third part of developing a programme but there are still shortcomings in the evaluation size. In distance education, only multiple-choice, true-false or matching exams are applied but still it cannot be checked by written examination to check whether the student is on the level of cognitive evaluation.

More than sixty years after the introduction of the first computer system, Internet has become the standard platform for e-learning environments. E-learning, the contemporary version of distance education, is mainly web-based, conducted by means of Internet-connected computers running special programs (learning content management systems, LCMS), which bring learners, teachers, courses and collaborative technologies into contact (Yannis and Mavrommatis, 2009).

It can be seen that programme elements are being used in distance education systems as well as in traditional education systems. When programme elements are examined, it is observed that it consists of target, content, educational status and evaluation stages. Then after the questions of target or why they do distance education are answered it comes to the stage of content and educational status in accordance to the programme elements. At this point according to content, education management systems step in content management systems and in developing or guiding content.

We see a lot of content and education management systems has been developed in present conditions. But there is not a contemporary approach to complete these systems. During the evaluation stage to check whether the individual reaches the goal by written examinations in distance education systems. To be able to evaluate and conclude the individuals' or students' navigations during the distance education environments. To reshape the content by means of students' purposely educational navigations between the units of content. The answer to all these applications emerges the considerations of evaluation of data mining which is taken as a base in internet based distance education approaches.

When applications, data mining in internet based environments, are examined, we can see text mining as a core, web mining as middle layer, data mining as the last part of the applications. Of course for all these applications semantic web will be an important inference. Semantic web applications in parallel to development of mining applications will create the structure of semantic web in distance education platforms.

2. Evaluation in Distance Learning

In any programme development system, evaluation is one of the most important fundamental parts. Because of the fact that the success of distance education depends on the general performance of the institution, evaluation step

has to be planned carefully. As well as evaluation of students, evaluating the program is one of the matters that would influence the success of the institution. The desired goals that is to be reached in program evaluating should be determined well and lacks should be analysed objectively. The situation of reaching the goals, the contentment of students and the teachers should also be evaluated (Monolescu, Schifter, & Greenwood, 2004).

Distance education is going towards a student-centred structure like in the other kinds of education. The role of the teachers is changing into a way that eases learning. And most institutions make changes that would control the students' attendance and the work steps. Although technological means changed, the necessity of teachers have never changed, but the roles have changed. Teachers have become a means between students and information. And using technological means have become a must for teachers. (Beaudoin, 1990).

In general, the success rate of distance education is measured with the success in control of the students. How successful students are in lessons, how well they understand the lessons, how much they developed, whether the content is enough or not can be measured by the means of control. Students should be controlled well and in which areas they have problems should be identified. These deficiencies should be worked on to resolve an new techniques should be developed.

Methods used in distance education ranked as follows (Bayam & Urin, 2002):

1. Quizzes: After each topic, how well the students have learned the subjects is evaluated with quizzes. In this way, how much they understand the lesson can be seen.
2. Assignments: students should sometimes be given assignments and they should be encouraged to make research.
3. Online examinations: after a certain level of courses, with a general quiz, how well the students follow the courses, how much they understand and how well they learned can be measured. The most important point is how well they learned.
4. The control module of the platform of distance education: in this module, there should be reports that can control the students. Some of these reports should be reached by students for self-control. In this module, there should be the reports of attendance of the students, examination and homework grades, general situation in the class. Moreover, there should be form and chat part and students should be able to ask questions, get help about problematic topics and reach an area of discussion to develop themselves.

In addition, new evaluation means are being developed appropriate for the new teaching models. In constructivist learning approach, the evaluation phase is not within the frame of the explanations about what information and learning is, but it is about the process of constructing the information not evaluating it. Constructivist education environment makes it necessary to develop assessment methods that require long-term review instead of evaluation at the final stage in traditional approach. Applying constructivist education and portfolio evaluation method to distance education are examples of this.

3. Algorithms Used in Evaluation

C45: This data processing algorithm is used to provide machine learning to solve classification problems. By using a large set of information a decision tree is created under certain conditions and enables to have inferences accordingly. For example: If the weather is sunny and cloudless golf can be played, if it is cloudy and the temperature is low, golf cannot be played.

K-Means: This algorithm is a classification system which has a wide usage area. It enables the classification of the given records according to needed qualifications so the same characteristics of groups is classified. Before starting to group, it needs to be inserted the same type of data. In addition, via k-means algorithm data can be divided into desired number of parts according to the similar features. Because algorithm is plain the results are fast and because it is one of the old algorithms many suitable software can be found.

SVM: Support vector machines is one the most effective way which is used in data mining. In the system it is counted with Support Vector Classifier (SVC) and Support Vector Regressor (SVR). These systems at the beginning of 1990's developed by Vapnik. With its need less samples is one of the reason which makes it prominent.

KNN: This algorithm classifies objects according to their similarities for the nearest classification (Pilavcılar, 2007).

Naive Bayes: The information provided on the basis of a set of objects belonging to known groups of elements will be used to determine which group would take. Naive bayes is one the easiest and the most effective method, it can be adapted easily and fast to large-sized data. Because it is easy to understand, it is easy to apply and it is easy to learn who dont know much about the system.

Apriori: In the systems of machine learning a lot of structures were developed to find a design and to classify. Among these the easy ones always are adapted easily and give effective results. Apriori is to methods to determine design. All these algorithms mentioned are also used easily in text mining at the same time. (Karabatak and Ince, 2010; Wu & Kumar, 2009)

Text mining: Text mining mostly focuses on defining data and at the stage of inference. Preliminary stages of this instructed data , are composed of more regular documents into the required format of conversion, this kind of decomposition is not usually seen in data mining systems. Furthermore because they obtained information consist of the texts in natural language format, text mining focuses on natural language processing which is very different field according to computer sciences. Maybe the most striking points about text mining , information extraction, data collection and heap-based linguistic calculation process (Feldman & Sanger, 2007).

5. Conclusion

According to the evaluation, evaluation seems to be the most important stage in distance education determining the effectiveness of all programme elements. At the stage of evaluation algorithm C45 is decided to be the effective one in classifying students' characteristics and academic success. This algorithm is determined to be the effective to decide inferences and a decision tree.

Another classifying algorithm K-means classifies students, behaviour and academic success features which are alike when compared to C45 algorithm. According to SVM algorithm it is the best method to evaluate when we lack the number of data. It is used when it is need to decide according to students' different features and when the data is limited. KNN algorithm's the most important feature is to match every single document. This situation helps to use the data effectively during evaluation stage.

Naive Bayes algorithm is superior to other algorithms because it is known to begin with the data's which are given by the known resources and decide in which group new students should stay. Apriori algorithm with more than one substance decides the relationship between behaviour and data.

Text mining is incorporation of all algorithms which have outstanding features to accommodate the art of installing meaning to the texts. Finally algorithms are effective to be used in text mining, have productive features to use in distance education to determine positive and negative features of distance learning. And it also helps to analysis student behaviour, academic success, content analysis at the base of evaluation.

References

- Bayam, Y., & Urin, M. (2002). Uzaktan Eğitimde Öğrenci Takibi ve Değerlendirmesi. *Açık ve Uzaktan Eğitim Sempozyumu*. Eskişehir: Anadolu Üniversitesi.
- Beaudoin, M. (1990). The Instructor's Changing Role in Distance Education. *The American Journal of Distance Education* , 21-29.
- Daş, R., & Varol, N. (2001). Günümüzde Uzaktan Eğitim Uygulamalarına Genel Bakış. 1.
- Feldman, R., & Sanger, J. (2007). *The Text Mining Handbook*. New York: Cambridge University Press.
- Karabatak, M., Ince, M.C.. Apriori Algoritması ile Öğrenci Başarısı Analizi. www.emo.org.tr Retrieved July 2010
- Kayri, M. (2008). Elektronik Portfolyo Değerlendirmeleri için Veri Madenciliği Yaklaşımı. *Yüzüncü Yıl Üniversitesi, Eğitim Fakültesi Dergisi* , 100-101.
- Monolescu, D., Schifter, C., & Greenwood, L. (2004). *The Distance Education Evolution: Issues and Case Studies*. London: Idea Group Inc.
- Pilavcılar, İ.,F. (2007). Metin Madenciliği İle Metin Sınıflandırma. Fenbilimleri Enstitüsü, Yıldız Teknik Üniversitesi,
- Wu, X., & Kumar, V. (2009). *The Top Ten Algorithms in Data Mining*. New York: CRC press.