



## Adaptive Learning using Data Mining Techniques

Dhaval Doshi<sup>\*</sup>, Himey Jain,<sup>3</sup> Nikhil Tiwari, Bobby JasujaDepartment of Information Technology  
Medicaps Institute of Science and Technology, Indore, India

**Abstract** - E-learning is defined as a Virtual Learning Environment (VLE) which provides resources for interaction between students and teachers. The latest advancement in the field of Information Technology and enhanced use of internet has facilitated tremendous proliferation towards usage of E-learning. It provides a new way of distributing information with better efficiency and lesser cost by clustering similar resources into groups of meaningful data using data mining algorithms. One of the recent advancement in e-learning is adaptive e-learning. This type of learning provides the content to the users according to their current level of learning. The learning abstraction system consists of huge amount of data so there is a requirement of some specific way to retrieve useful data with speed and ease. The traditional approach does not take into account the fact that the mental level of the learners varies significantly from person to person. Consequently it would be unfair if every learner gets access to similar kind of data. Adaptive E-learning is a concept which provides content dynamically to users according to their mental level by employing data mining technique. We propose the use of k-means algorithm for implementing adoptive learning.

**Keywords** -e-learning, adaptive learning, data mining, clustering, k-means algorithm, knowledge discovery in database, kdd

### I. INTRODUCTION

#### A. System of e-learning

Over the last few years, Internet has become a wide spread medium that has revolutionized how world treats and exchanges information and education system is one of the major domain that has a profound impact of it. In fact, advancement in Internet and telecommunication has made sharing and manipulation of information in real times a reality.

E-learning (a web based learning method) provides a new way of education where a large amount of data is generated and should be available ubiquitously. It is a blessing to education industry which provides a way to connect and educate people even in the remotest part of the country to learn at just a click of mouse. However the challenge lies in managing the huge amount of data efficiently so as to provide information as and when needed. Data mining is used to tackle such kind of problems.

#### B. Concept of adaptive-learning

The e-learning system accumulates huge amount of information by recording student activities such as online assessment reports, answers, lectures and student questions. Adaptive E-learning is a concept which provides content dynamically to users according to their mental level by employing data mining technique. Therefore adaptive e-learning methodology ensure complete development of an individual by providing user specific content. Data mining can be effectively used along with these adaptive learning technologies so as to customize the learning experience. The data mining technique is used to extract meaningful data for every student. The system also provides essential data which differs from person to person. A personal course design can ensure that the learners are provided the resources as per their mental level as and when required.

### II. APPLICATION OF DATA MINING IN ADAPTIVE LEARNING

Data Mining is a recent advancement in information technology also known as KDD (Knowledge discovery in database). It is a process of extracting meaningful, novel and potentially useful patterns which are implicitly stored in databases, data warehouses and other huge data repositories. Knowledge Discovery in database is in close relationship with Artificial Intelligence. Finding of association rules makes it possible for web based tutors to recognize knowledge patterns and identify virtual learning course based on the recognized patterns. This paper presents an overview of available data mining and AI techniques to overcome the challenges in e-learning. This paper proposes the use of k-means algorithm for implementing adaptive learning.

#### A. Use of Data Mining Technique

Clustering is one of the most prominent techniques in data mining for grouping student's data having similar characteristics to promote group based learning and to provide learning diagnosis. K-means algorithm is often used for this purpose. These algorithms are applied when instance of data are to be divided into natural groups. Data is divided into group prior to the application of algorithm.

### III. CONCEPT OF CLUSTERING

Clustering is defined as the technique of division of data in groups of relevant objects. Clustering is used for searching relevant data as a linear request when the data is not present in the related server. Clustering technique is broadly classified into hierarchical clustering and partitioning relocation. Hierarchical clustering makes clusters gradually whereas partitioning relocation makes clusters directly. They make clusters either by identifying clusters as highly populated area with data or iteratively relocate points between subsets.

Partitioning relocation methods are further classified as probabilistic clustering k-medoids method which concentrates on how well points fit into their clusters in order to build proper convex shapes. Second type of partitioning algorithm is section based density partitioning which are flexible in terms of shape and hence can discover cluster of irregular shapes. They are usually applied for low dimensional data of numerical entry called as spatial data.

### IV. K-MEANS CLUSTERING ALGORITHM

One of the simplest unsupervised algorithm to solve the clustering problem is k-means algorithm. The idea is to define k-centres one for each cluster. Next step is to associate each data point to the nearest centre by calculating the distance from each of the centres chosen. This distance is called as Euclidean distance. After completion of first iteration (when no points are pending) is called as early age group. At this point we calculate k-new centroids as centre of clusters resulting from the previous step and again new binding is done by calculating the distance. The loop ends when centres do not move any more. This algorithm aims at minimizing an objective function known as squared function.

$$J(V) = \sum_{i=1}^c \sum_{j=1}^{c_i} (x_i - v_i)^2$$

Where,

- ' $\|x_i - v_j\|$ ' is the Euclidean distance between  $x_i$  and  $v_j$ .
- ' $c_i$ ' is the number of data points in  $i^{th}$  cluster.
- ' $c$ ' is the number of cluster centres.

#### A. Use of k-means in e-learning

We propose the use of k-means algorithm for effective implementation of adaptive learning. We will like to divide the users of an e-learning program in the following categories on the basis of their performance and knowledge gained:

- 1) Beginner level: These are novice users who have just started the learning process. This level is initially assigned by the system to those people who have just started learning in the virtual environment.
- 2) Intermediate Level: Those users who have mastered the basics of a particular program are promoted to intermediate levels. These levels are designed to provide deeper insight on the topic and with the increase in level of difficulty.
- 3) Master Level: A user is promoted to this level only when he has sufficient knowledge of the topic and wishes to achieve expertise in it.

#### B. Systematic steps for k-means algorithm

- 1) Let  $X = \{x_1, x_2, x_3, \dots, x_n\}$  be the set of data points and  $C = \{c_1, c_2, c_3, \dots, c_n\}$  be the set of centres.
- 2) Randomly select any 'c' centres.
- 3) Calculate the distance between each data point and the centre using Euclidean distance formula

$$\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

where  $x_1, x_2$  and  $y_1, y_2$  are the coordinates of the data point and the centre.

- 4) Assign the data point to the centre whose distance from the centre is minimum of all cluster centre.
- 5) Recalculate cluster centre using

$$v_i = \left(\frac{1}{c_i}\right) \sum_{j=1}^{c_i} x_i$$

Where, ' $c_i$ ' represents the number of data points in  $i^{th}$  cluster.

- 6) Data points are then mapped to the nearest centre, thus grouping students according to their mental level.
- 7) This is an iterative process that constantly maps data points to centre points
- 8) This process is repeated periodically for evaluating individual performances. Thus learners are regrouped into different centre points on the basis of their recent performances.

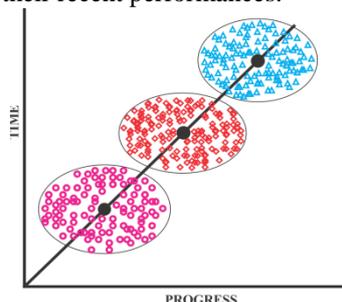


Fig. Error! No sequence specified. A graphical representation of data clusters

In the above figure points on the centre line represent the centre points and points in proximity to centre point represent data points. Points in vicinity to centre points are clubbed together to form clusters using k-means algorithm.

## V. CONCLUSION

This research proposes data mining model for adaptive learning based on K-Means clustering technique. The education data mining analyzed and investigated in this research focuses on better understanding the learning and teaching process by implementing data mining. We believe that the model presented will be able to comprehensively, logically and meaningfully analyze the data and make the learning experience a more personalized one. To further test the effectiveness of this model and increase generality of the proposed research more extensive experiments are required to be conducted.

## ACKNOWLEDGEMENT

We would like to express our sincere gratitude to all those people who were directly or indirectly involved in successful completion of this research paper.

## REFERENCES

- [1] Nichols and Schedule "Theory for E-Learning Educational Technology and SOC, Mvol,6,no,2,pp,1-10,2003
- [2] MCSER-Mediterranean Center of Social and Educational Research Rome, Italy, 2011
- [3] Romero, C., Ventura, S. (2010), Educational Data Mining: A Review of the State-of-the-Art. IEEE Transaction on Systems, Man, and Cybernetics, Part C: Applications and Reviews.
- [4] Uribe-Tirado, A., Melgar-Estrada, L.-M., &Bornacelly-Castro, J.-A. (2007). Moodle learning management
- [5] System as a tool for information, documentation, and knowledge management by research groups. Profesional de la Información.
- [6] Silva, D., & Vieira, M. (2002). Using data warehouse and data mining resources for ongoing assessment of distance learning. In Proceedings of the 2002 IEEE Conference on Advanced Learning Technologies Kazan, Russia.
- [7] Sison, R., & Shimura, M., (1998). Student Modelling and Machine Learning. International Journal of Artificial Intelligence in Education.
- [8] Srikant, R., & Agrawal, R. (1996). Mining Sequential Patterns: Generalizations and Performance Improvements. In P. M. G. Apers, M. Bouzeghoub, & G. Gardarin (Eds.) Proc. 5th Int. Conf. Extending Database Technology,EDBT, vol. 1057, (pp. 3–17). Springer-Verlag.
- [9] Srivastava J., Cooley R., Deshpande M., & Tan P. (2000, January) Web Usage Mining: Discovery and Applications of usage Patterns form Web Data, SIGKDD Explorations.
- [10] Talavera, L., &Gaudioso E. (2004). Mining student data to characterize similar behavior groups in unstructured collaboration spaces. Workshop on Artificial Intelligence in CSCL. ECAI.
- [11] Tang, T.Y., McCalla, G., (2005).Smart Recommendation for an Evolving e-Learning System:Architecture and Experiment. International Journal on e-Learning.
- [12] Tsai, C. J., Tseng, S. S., & Lin, C. Y. (2001). A Two-Phase Fuzzy Mining and Learning Algorithm for Adaptive.
- [13] Learning Environment. In ICCS '01: Proceedings of the H. (2004).
- [14] Effective personalized recommendation based on time-framed navigation clustering and association mining. Expert Systems with Applications 27(3). Wu X., (2004).
- [15] Data Mining: An Artificial Intelligence Perspective. The IEEE Intelligent Informatics Bulletin.Yoo, J., Yoo, S., Lance, C., Hankins, J. (2006).
- [16] Student Progress Monitoring Tool Using Treeview. In The 37th Technical Symposium on Computer Science Education,. Houston, USA: ACM Press.Zaïane, O., & Luo J., (2001).
- [17] Web usage mining for a better web-based learning Environment.Proceedings of conference on advanced technology for education, Banff, Alberta. Xu, R., &Wunsch, D. I. I. (2005).