Volume-3, Issue-11, Nov Special Issue -2014 • ISSN No 2277 - 8160

Research Paper | Systems

Evaluation of E-Learners Behaviour in Distributed E-Learning Environments

F.Nirmala Sherine Assistant Professor, Department of Information Technology, Periyar Maniammai University, Thanjavur, Tamil Nadu.

V.Hamsadhwani Assistant Professor (SS), Department of Information Technology, Periyar Maniammai University, Thanjavur, Tamil Nadu.

ABSTRACT *Learning is a cognitive activity that differs from student to student .In recent years, due to large amounts of network-based teaching and learning data continue to grow inexorably in size and complexity, knowledge clustering becomes more important in e-learning. This paper introduces an evaluation methodologies forth e-learners' behaviour that will be an automatic feedback to the decision makers in e-learning system. Learner's profile plays a imperative role in the evaluation process to improve the e-learning process performance. The work focuses on clustering the e-learners based on their behaviour into specific categories that represent the learner's profiles. This paper presents the use of different fuzzy clustering techniques: fuzzy c-means and kernelized fuzzy c-means to find the learners' categories and predict their profiles. The proposed framework consists of the following phases such as data description, preparation, features selection, and the experiments design using the fuzzy clustering models. Experimental results shows that comparison between FCM and KFCM proved that the KFCM is much better than FCM in predicting the learners' behaviour.*

KEYWORDS : E-Learning, Learner Profile, Clustering, Behaviour Patterns, Fuzzy C-Means Clustering, Kernelized FCM.

1. Introduction

The tremendous growth of web based education in recent past has also helped in the growth of online education. One of the main challenges in web based education is to appreciate and give confidence to student participation. Web based system is based on semantic web technologies which is adaptive learning. Adaptive learning systems employed clustering the e-leaners behaviour into specific categories that represent the leaner's profile by fuzzy c-means and kernelized fuzzy c-means. Many researchers are adopting semantic web technologies to find new ways for designing adaptive learning systems based on describing knowledge using fuzzy.

The open problems in the e-leaners behaviour we are trying to address are:

How to support e-leaner identification and profile in such a distributed environment.

How to integrate e-learners behaviour with other functionalities needed to provide support for leaners.

The aim of this paper is to present e-learner behaviour in distributed environments based on fuzzy clustering models. E-learning system based on the design of semantic content, learner and domain models to tailor the teaching process for individual learner's needs . The proposed new adaptive e-learning has the ability to support personalization based on learner's ability, learning style, preferences and levels of knowledge. In our approach the user profile is updated based on achieved learner's abilities. At the core of online education is the fact that students can have access to instructional resources anytime, anywhere and connect with the material at their convenience. These systems gather a great deal of information which is very precious in analyzing students' behaviour and supporting teachers in the revealing of possible errors, shortcomings and improvements.

2. Related Work

a. Data Clustering

Data clustering is suggested as a means to promote groupbased collaborative learning and to provide incremental student diagnosis [1]. In [2], user actions associated to students' Web usage were gathered and preprocessed as part of a Data Mining process. The Expectation Maximization (EM) algorithm was then used to group the users into clusters according to their behaviors. These results could be used by teachers to provide specific control to students fitting to each cluster. The shortening statement that students belonging to each cluster should share Web usage behaviour makes personalization [5] strategies more scalable. Clustering was proposed in [4] to group similar learning documents based on their topics and rank the available data features according to their relevance for the definition of student clusters.

b. Prediction Techniques

The anticipating of students' behavior and performance when using e-learning systems abides the prospective of levelling the improvement of effective courses as well as e-learning environments in general. A methodology to improve the performance of developed courses through adaptation was presented in [4]. Course log-files stored in databases could be hewed by teachers using evolutionary algorithms to discover important relationships and patterns, with the target of discovering relationships between students' knowledge levels, e-learning system usage times and students' scores. A system for the automatic analysis of user actions in Web-based learning environments, which could be used to make predictions on future uses of the learning environment [4].

3. Proposed System

The proposed approach employs fuzzy logic theory to determine the difficulty levels of test items according to the learning status and personal features of each student, and then applies to cluster the test items into groups, as well as dynamic programming [8] for test sheet construction.

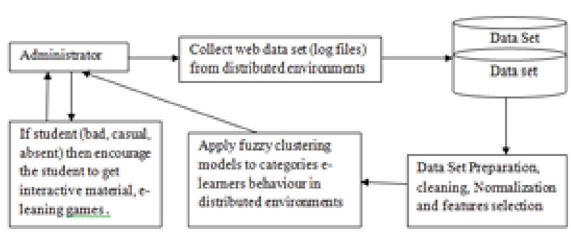

Fig 1: Categories E-learners Behaviour

A. Log Files Description

The data logged in server logs forwards the access of a Web site by multiple users. Web server-side data and client-side data constitute the main sources of data for Web usage mining. Web server access logs establish the most widely used data because it obviously records the browsing behaviour of site visitors. For this reason, the term Web log mining is

Sometimes used. Web log mining should not be confused with Web log analysis. The data sets used in this study were obtained from web access logs for studying a two courses; the first is for teaching "C Programming (CP)". The second course is "Data Structure(DS)", for the first year. Student's behaviour through teaching courses it proposed that, visits from students attending this course could fall into one of the following five categories:

1. Regular students: These e-learners download the current set of material. Since they download a limited/current set of material, they probably study class-notes on a regular basis.
2. Bad students: These learners download a large set of material. This indicates that they have stayed away from the class-notes for a long period of time.
3. Worker students: These visitors are mostly working on class or lab assignments or accessing the discussion board.
4. Casual students: those students who did not interact with the course material and if they visit the web course, they do not download any documents.
5. Absent students: those students who are absent during the teaching course where after many experiments we found that the casual students and the absent students do not affect the study of learner's profiles because the paper focuses on the learners profiles based on number of hits, downloaded documents, time of accessing the web course, and day of accessing the course materials.

B. Data Preparation and Cleaning

Data preparation consists of two phases: data cleaning, and data abstraction and normalization.

Data cleaning process: Data cleaning process consists of two steps Hits cleaning and Visits cleaning as follows.

- *Hits Cleaning:* To remove the hits from search engines and other robots. In the second data set; the cleaning step reduced the log files data set by 3.7%, the number of hits was reduced from 50184 before cleaning to 38004 after cleaning.
- *Visits cleaning:* To clean the data from those visits, which didn't download any class-notes, were eliminated, since these visits correspond to casual visitors. The total visits were 3048; after visits cleaning the visits were reduced to 1102 as shown in Table 1.
- Remove the Casual and absent classes from the data sets: where those two cleaning steps were not interested in studying the learners who did not download any Byte, as well as the casual learners.
- *Data privacy and learners security:* It is required for the identification of web visits; it is done using Linux commands. Certain areas of the web site were protected, and the users could only access them using their IDs and passwords. The activities in the restricted parts of the web site consisted of submitting a user profile, changing a password, submission of assignments, viewing the submissions, accessing the discussion board, and viewing current class marks. The rest of the web site was public. The public portion consisted of viewing course information, a lab manual, class-notes, class assignments, and lab assignments. If the users only accessed the public web site, their IDs would be unknown. Therefore, the web users were identified based on their IP address. This also made sure that the user privacy was protected. A visit from an IP address started when the first request was made from the IP address.The web logs were pre-processed to create an appropriate representation of each user corresponding to a visit.

2. Data Abstraction and Normalization: The abstract representation of a web user is a critical step; that requires a good knowledge of the application domain. Previous personal experience with the students in the course suggested that some of the students print preliminary notes before a class and an updated copy after the class. Some students view the notes online on a regular basis. Some students print all the notes around important dates such as midterm and final examinations. In addition, there are many accesses on Tuesdays and Thursdays, when the in-laboratory assignments are due. On and Off-campus points of access can also provide some indication of a user's objectives for the visit. Based on some of these observations, it was decided to use the following attributes for representing each visitor:

a. On campus/Off campus access (binary values 0 or 1).
b. Day time/Night time access: 8 a.m. to 8 p.m. were considered to be the Daytime (day/night).
c. Access during lab/class days or non-lab/class days: All the labs and classes were held on Wednesday and Thursdays. The visitors on these days are more likely to be Worker Students.
d. Number of hits (decimal values).
e. Number of material downloads (decimal values).

The first three attributes had binary values of 0 or 1. The last two values were normalized. The distribution of the number of hits and the number of class-notes was analysed for determining appropriate weight factors. The numbers of hits were set to be in the range [0, 10]. Since the class-notes were the focus of the clustering, the last variable was assigned higher importance, where the values ranged from 0 to 15.

Table:1 Data Sets Before and After Preprocessing

Data set	Hits	Hits after Cleaning	Visits	Visits after Cleaning
CP	372100	354000	25741	6824
DS	50184	38004	3048	1102

Table 2: FCM Result for 1 st Data Set

Class Name	Behavior of each Class					Size
	Camp	Time	Lab	Hits	Req	
Regular	0.254	0.28	0.2	0.782	.8	125
Workers	0.258	0.25	0.5	0.481	.8	259
Bad	0.583	0.56	0.4	0.567	.4	48
R&W	0.254	0.59	0.5	0.234	.6	15
R&B	0.482	0.45	0.6	0.482	.5	255
W&B	0.356	0.78	0.3	0.235	.6	152
R&W&B	0.782	0.56	0.5	0.554	.8	12

Table 3: KFCM Result for 1 st Data Set

Class Name	Behavior of each Class					Size
	Camp	Time	Lab	Hits	Req	
Regular	0.02	0.25	0.82	0.5	.2	1862
Workers	0.02	0.52	0.25	0.8	.8	125
Bad	0.41	0.21	0.15	0.6	.7	256
R&W	0.35	1.50	0.12	0.4	.6	248
R&B	0.56	0.36	0.48	0.8	.8	325
W&B	0.23	0.55	0.78	0.6	.4	25
R&W&B	0.45	0.54	0.48	0.8	.3	48

Table 4: FCM Result for 2nd Data Set

Class Name	Behavior of each Class					Size
	Camp	Time	Lab	Hits	Req	
Regular	0.001	0.52	0.3	0.32	0.6	180
Workers	0.92	0.80	0.2	0.2	0.8	233
Bad	0.62	0.72	0.5	2.5	0.4	123
R&W	0.22	0.21	0.8	0.5	0.7	185

Class Name	Behavior of each Class					Size
	Camp	Time	Lab	Hits	Req	
R&B	0.6	0.63	0.9	0.8	0.3	158
W&B	0.66	0.71	0.4	0.4	0.4	135
R&W&B	0.50	0.12	0.6	0.2	0.5	15

Table 5: KFCM Result for 2nd Data Set

Class Name	Behavior of each Class					Size
	Camp	Time	Lab	Hits	Req	
Regular	0.285	0.56	0.5	0.593	.5	158
Workers	0.566	0.45	0.4	0.158	.3	142
Bad	0.785	0.56	0.2	0.585	.8	125
R&W	0.568	0.55	0.3	0.878	.3	256
R&B	0.55	0.32	0.7	0.258	.4	256
W&B	0.125	0.45	0.3	0.556	.2	156
R&W&B	0.235	0.55	0.2	0.485	.3	48

Table 6: Comparison Between Results of FCM and KFCM

Data Set	Model	Ratio between Size of Clusters and Real Results		
		Regular/Real	Worker/Real	Bad/Real
1st	FCM	80%	87%	80%
1st	KFCM	79%	85%	79%
2nd	FCM	89%	90%	94%
2nd	KFCM	90%	91%	96%

4. Result analysis

It was possible to classify the learners using the two fuzzy clustering techniques into five clusters as regular students, worker students, bad students, casual students, and absent students using both of fuzzy c-means, kernelized c-means, and KFCM Method table 1,2,3,4,5,6. But the problem here is that the absent students were not found in the data sets as the absent student is characterized by the casual interaction with the web course, they did not download any materials documentation related to the course when they visited the web site. The evaluation achieves that both of the two methods were good enough; moreover the KFCM was better and its performance from the points of matching with the real marks and the speed was high.

5. Conclusion

This paper proposed the two different fuzzy clustering techniques, the FCM and KFCM, where the clustering is one of the most important models in data mining. The work focused on student categories like regular, worker, bad, casual, absent based on e-learners profile. The student get additional timing, interactive e-leaning materials. It supports categories of e-leaners and enforces continuous learning. Finally the paper proved that the KFCM was better in predicting the e-learners behavior.

REFERENCES

[1] M.Farida Begam and Gopinath Ganapathy,Adaptive Learning Management System Using Semantic Web Technologies, International Journal on Soft Computing (IJSC) Vol.4, No.1, February 2013. [2] Mukta Goyal, Divakar Yadav, Alka Choubey, E-learning: Current State of Art and Future Prospects, IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 3, No 2, May 2012 | [3] Alexandros Paramythis and Susanne Loidl-Reisinger ,Adaptive Learning Environments and e-Learning Standards, Electronic Journal on e-Learning Volume 2 Issue 1 (February 2004) 181-194 | [4] Mofreh A.Hogo ,Evaluation of E-Learners Behaviour using Different Fuzzy Clustering Models: A Comparative Study, International Journal of Computer Science and Information Security, Vol.7 No.2 2010. | [5] Peter Dolog, Michael Sintek, Personalization in Distributed eLearning Environments, ACM 2004. | [6] Nabila Bousbia , Issam Rebaï , Jean-Marc Labat , Amar Balla ,Analysing the Relationship between Learning Styles and Navigation Behaviour in Web-Based Educational System , Knowledge Management & E-Learning: An International Journal, Vol.2, No.4. | [7] Fernando Alonso, Genoveva López, Daniel Manrique and José M Viñes,An instructional model for web-based e-learning education with a blended learning process approach, British Journal of Educational Technology Vol 36 No 2 2005 217–235. | [8] Mandeep Kaur, Kewal krishan , Cluster Analysis of Behavior of E-learners , International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-3, Issue-2, May 2013. | [9] C. Bouras, A. Philopoulos and Th. Tsiatsos e-Learning through distributed virtual environments , Journal of Network and Computer Applications (2001) 24, 175–199 |