Discovering Learners’ Characteristics Through Cluster Analysis for Recommendation of Courses in E-Learning Environment

Bhupesh Rawat, Department of Computer Science, Babasaheb Bhimrao Ambedkar University (BBAU), Lucknow, India
Sanjay K. Dwivedi, Department of Computer Science, Babasaheb Bhimrao Ambedkar University (BBAU), Lucknow, India

ABSTRACT

With the emergence of the web, traditional learning has changed significantly. Hence, a huge number of ‘e-learning systems’ with the advantages of time and space have been created. Currently, many e-learning systems are being used by a large number of academic institutions worldwide which allow different users of the system to perform various tasks based on their goals. However, most of these systems follow a ‘one size fits all’ approach where same resources are offered to learners irrespective of their unique learning requirements. Therefore, personalization is required as a part of e-learning systems which offers resources to learners based on their profile. This research aims to perform cluster analyses in order to validate clusters created through a k-means algorithm. The clusters will be used to classify a new learner into its appropriate class and recommend relevant courses. Finally, the accuracy of the recommendation is evaluated using various evaluation metrics. The proposed recommendation system helps learners to improve their academic performance and hence overall learning process as well.

KEYWORDS

Data Pre-Processing, E-Learning, Elbow Method, K-Means, Moodle, Silhouette Method, Weka

1. INTRODUCTION

E-learning is defined as the delivery of content using electronic mediums such as satellite communications, audio & videotape, Interactive Television, and CD-ROM among others. A large number of educational institutions adopted the digital revolution through e-learning programs and applications (Hammouri et al., 2018). There are some popular definitions of e-learning (Ozkon et al., 2009; Huang, 2002; Castro et al., 2007; Romero & Ventura, 2007) found in the existing literature on e-learning systems. It has changed the face of traditional learning in which students were not able to learn according to their time and space. However, with the emergence of the world wide web, sophisticated e-learning environments are being used increasingly by academic institutions worldwide due to their high volume of data, the strong interactivity, the great coverage, and no spacetime

DOI: 10.4018/IJICTE.2019010104

Copyright © 2019, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.
restrictions as compared to the traditional learning system. These e-learning systems are also termed as learning management systems (LMS). While there are a large number of LMS, Moodle is one of the widely used collaborative LMS due to its open-source nature and rich functionalities which make it a natural choice for our purpose. Additionally, it is provided freely and currently, it is the most popular e-Learning platform with over 85,000 registered sites worldwide, over 8 million courses, almost 76 million students, and over one million teachers. The different functionalities offered by Moodle are organized nicely into different modules such as the assignment module, quiz module, chat module and messaging module among others (Rice, 2006). While assignment module allows the instructor to gather, review and grade learners’ assignment, quiz module helps instructors to measure learners’ comprehension of the learning material. Each activity performed by a learner in Moodle is recorded such as time spent on the forum, time is taken to solve quizzes and total time taken in completing assignment etc. In addition to this, Moodle is also used by learners for several purposes such as to communicate with their teachers and peers regarding their learning problems, to post a question on the forum, to read messages on the forum, to create a new topic, to answer the posted question, to visit resources, to access resources among others.

However, one of the major problems with the learning management systems is that they provide the same learning resources to all the learners without taking into account the differences in their level of knowledge, skill, interest, and goals. For instance, if a learner is lacking the basic understanding of a subject, it would serve no purpose to teach them advanced concepts of the same subject. This situation negatively affects learners’ academic performance and the overall learning process. One of the solutions to this problem is to build learner’s profile based on their preferences and then recommend courses based on this profile. A learner’s profile is a set of information representing a user via user related rules, settings, needs, interests, behaviors, and preferences (Araniti et al., 2003). In (Tang & McCalla, 2005; Margo, 2004) the authors conducted a study on how data mining techniques could be used in e-learning environment and how they could improve the overall learning process.

The objective of this research is to recommend different categories of data mining courses such as ‘courses for beginner’, ‘course for intermediate learner’, ‘course for advanced learner’ to different type of learners belonging to one of the three clusters namely ‘not active’, ‘average’ and ‘active’. These clusters form the basis of the learners’ profile and are created by applying the \( k \)-means algorithm to learners’ usage data which is collected from Moodle server. The learners’ usage data is generated by letting the learners interact with the Moodle for performing their learning tasks such as assignment and quizzes through collaborative means such as messaging, forums and chat among others. Our research proposes that a learner’s academic performance can be improved by taking into account the differences in learners’ skill, interest, and level of knowledge while making recommendations of data mining courses to learners. Furthermore, recommending advanced courses to a learner who is falling in ‘not active cluster’ may demotivate a learner to study the recommended course which will further lead to the poor academic performance by the learner.

The following tasks are carried out in order to achieve the above objective:

1. Collect learners’ data from Moodle’s server;
2. Apply data pre-processing to this data in order to make it suitable for data mining algorithm;
3. Apply clustering algorithm particularly \( k \)-means to the pre-processed data obtained in the previous step in order to create the group of learners with similar learning patterns;
4. Perform clusters validation to ensure that the clusters represent the true characteristic of learners;
5. Classify a new learner into its appropriate cluster in order to recommend relevant courses;
6. Create ‘user-item rating matrix’ to be used by collaborative filtering algorithm;
7. Generate a list of recommendation by applying collaborative filtering algorithm to the ‘user item rating matrix’;
8. Evaluate the quality of recommendations through well-known evaluation metrics.