

Personalized Learning Approach In Learning Management System Using Cluster Models

Kirk Alvin S. Awat, Maria Rona L. Perez, Ace C. Lagman, Tim Jamison S. Awat, John Benedict C. Legaspi

Abstract—This paper focuses on the integration of cluster models using K-means algorithm as a main mechanism in providing a personalized learning approach integrated in a learning management system (LMS). The cluster models provide intelligent questions mechanism suited to the individualized learning of students. The LMS will provide necessary examinations to gauge the students' learning capability and based on the result of the examination, the item analysis and clustering algorithm were performed to provide reinforcement questions in helping students to pass a certain examination. The derived cluster models consist of easy, average, and hard question categories. The process repeats until such time that the students will be able to answer the questions correctly. In this way, the students increase their academic performance. The system was also evaluated by the LMS experts using the FURPS model. Overall, the system was rated with a "Very Satisfactory" rating.

Index Terms— data mining, item analysis, item bank, K-Means clustering, reclassification

1 INTRODUCTION

The use of Learning Management System has become prevalent in the higher education. This has been widely used by many universities that support learning and teaching initiatives. It will become vital for the universities to take advantage of it to enhance teaching and learning as technological advances. Technology serves a main tool in providing quality education and research to the faculty, staff, and students [1]. The LMS provides an application that can be customized per specific methods and learning strategies. It includes, peer review, online chat, forums, discussion pages, supporting resources, video conferencing, exchanging of e-mails, and content repositories. By doing this, the LMS allows anytime access from any location. It also provides a method of consistency for the content evaluation and delivery. Learning Management System provides a tool where teachers can create online examination for the students. However, the system does not support and tailor to students' different learning styles. The success of a learning environment platform boils down to how it can support individualized learning. Individualized learning, or individualized instruction, is a method of teaching in which content, instructional technology, and pace of learning are based on the abilities and interest of each learner. With this, the researcher aims to integrate adaptive learning in the LMS environment.

Adaptive Learning is one of the key research areas in the educational setting. This educational technique can also be integrated in the learning management system as most open source learning management systems follow a size fit framework. One size fit framework resulted to no differentiated learning. Addressing this problem is very crucial as the main purpose of LMS is to support student academic performance and one way to solve this problem is the integrating of cluster models using K-means algorithm. Clustering algorithms under supervised learning is one of the popular methods under machine learning. The main purpose of clustering technique is to profile datasets by computing the similarities of each point using Euclidean distance. One of the most popular approaches under this area is the K-means clustering technique. This technique works on identifying K number of centroids and assign each point to the nearest computed clusters. The cluster technique also serves as a statistical method in grouping similar objects into categories, which refers to cluster groups. Cluster analysis has no means to provide ways to differentiate relevant and irrelevant attributes. It is imperative that the choice of attributes to be included in the cluster analysis must be carefully determined [2]. The clustering technique enables the grouping of questions in accordance to its difficulty levels. The cluster models derived from the algorithm can tag questions whether it is easy, average, or difficult. The goal is to sort data points, referring the questions into groups in a manner that the degree of association between two objects is high if they belong to the same group and low if they belong to different groups. Cluster analysis discovers structures in data without explaining why those structures exist [3]. It allows the researchers to discover and define patterns between data elements. Learning Management System provides a tool where teachers can create online examination for the students. The goal of this paper is to provide an implementation of clustering algorithm to the questions in the item bank that can provide a personalized approach suited to each student's learning capability.

- Kirk Alvin S. Awat is a professor at FEU Institute of Technology. He is a Doctor in Information Technology graduate. Philippines. E-mail: kirkalvin_awat@yahoo.com
- Maria Rona L. Perez is a professor at FEU Institute of Technology. She is a Doctor in Information Technology graduate. Philippines. E-mail: ronaloboperez@gmail.com
- Ace C. Lagman is the Program Director in IT at the FEU Institute of Technology. He a Doctor in Information Technology graduate and currently pursuing PhD in Computer Science in Technological University of the Philippines. E-mail: aclagman@feutech.edu.ph
- Tim Jamison S. Awat is a professor in FEU Diliman and currently pursuing master in Information Technology at the AMA University, Philippines. E-mail: tj.awat@feudiliman.edu.ph
- John Benedict C. Legaspi Awat is a professor at FEU Institute of Technology. He is a Master in Information Technology

2 CONTRIBUTION

The information gained by using data mining helps in assisting and facilitating in the enhancement of learning. The main objective of this study is to discover a way of integrating cluster model in a learning management system. This study aims to answer the following research questions:

- What cluster models can be extracted from the dataset using K-Means algorithm.?
- How are cluster models integrated in LMS?
- How acceptable is the system as perceived by the experts using FURPS model?

3 RELATED STUDIES

3.1 Data Mining in the Field of Education

The educational data mining (EDM) focused on the student understanding aspect. EDM is the field of using data mining techniques in educational environments. This field focuses on extracting hidden and useful patterns from the educational databases. Masaleno [4] reveals that learning analytics main goal is to improve students' learning performance. Its main objective is to assess the students' data, such as behaviors and performance in the development of adaptive learning [4].

3.2 Clustering Model

Clustering is the grouping of similar objects based on their distance once plotted. The main purpose of the clustering is to profile and provide concise summaries of the data used. The algorithm plays a very important role in large datasets consisting of many attributes [5].

3.3 Learning Management System

Adaptive Learning System, conform its learning components to build the most suitable adaptive learning experience to students through learning design. The learning design is a series of learning activities as pre-requisites to complete the lesson [6]. Ballera and his colleagues [7] mentioned well-designed course materials and assessment tools must be present in learning management system to make it effective tool to students. The understanding of Bloom's Taxonomy must be employed in the assessment in order to provide holistic approach of learning [7].

4 METHODOLOGY

4.1 Datasets

Students enrolled in Networking Fundamentals were selected as participants to simulate cluster models integrated in LMS platform.

An examination which contains questions from the item bank was developed. The assessment contains ten questions. The result of the assessment underwent item analysis and clustering algorithm procedures. To determine the level of difficulty of question in item analysis, the formula for question difficulty assessment and item discriminations were used.

$$\text{Question Difficulty} = TA/N \quad (1)$$

where, TA: students who get correct answers
S: total number of students

$$\text{Item Discrimination} = (T1/T2) - (B1/B2) \quad (2)$$

where, T1: Top 27% of students with correct answers
T2: Top 27% of students
B1: Bottom 27% of students with correct answers
B2: Bottom 27% of students

4.2 K-means Clustering Algorithm

K-means presents K as cluster. Each cluster model has its own characteristics and profile.

- The level of difficulty of each item will be determined and will be used as observation instance
- Five mean points will be chosen randomly.
- The distance will be calculated using the Euclidean formula
- Distance $[(x), (a)] = \sqrt{(x-a)^2}$ (3)
- The generated clusters will be analyze with reference to the clusters.
- The clusters will be presented
- The mean will be recalculated for the new cluster. Steps 3-6 will be repeated.
- Mean = SOV/C (4)
- where, SOV: sum of all observation values
- C: total number of observations in the cluster
- Stop the process if the generated clusters are repetitive.

4.3 Decision-Based LMS

The decision-based LMS aims to provide an intelligent mechanism to provide reinforcement questions to students who failed to get questions correctly. The system will decide which follow-up questions will be asked to the student if they are having difficulty in a certain question item. These questions are randomly selected on a pool of questions related to the topic. Through this process, students can understand more about a topic by answering related topics that deemed hard for them. The algorithm for the Decision-based LMS is as follows:

Step 1: Answer the Main Questionnaire

Step 2: Clustering Algorithm

Step 3: Identify the question items deemed *Hard*

Step 4: Generate review exam containing topics related to the *Hard* items.

Step 5: Repeat Step 4 until the student will correctly answer all the items.

4.4 Statistical Analysis Tools

To evaluate the system using FURPS model, the following statistical analysis were used: Weighted Mean. Considered the measuring central tendency.

$$\text{Weighted Mean (WM)} = \frac{\sum f(X1+X2+\dots+Xn)/n}{N} \quad (5)$$

where:

$\sum f(X1+X2+\dots+Xn)$ – represents the sum of all mean (M) of each criterion

n – refers to the total number of respondents

N – refers to the total number of criteria for evaluation.

Rating Scale. To interpret the perception of the respondents as to the effectiveness of the motivation factors, the five-point Rating Scale was used with the following equivalent points.

TABLE 1
RATING SCALE

Rate	Scale	Verbal Interpretation
4.51- 5.00	5	Excellent
3.51-4.50	4	Very satisfactory

Rate	Scale	Verbal Interpretation
2.51-3.50	3	Satisfactory
1.51-2.50	2	Fair
1.00-1.50	1	Needs Improvement

5 RESULTS AND DISCUSSION

5.1 Clustering

The difficulty level and the discrimination index per item were determined in this section. The difficulty per item was used as input to the K-Means algorithm. The profile of the cluster models are shown on Figure 1.

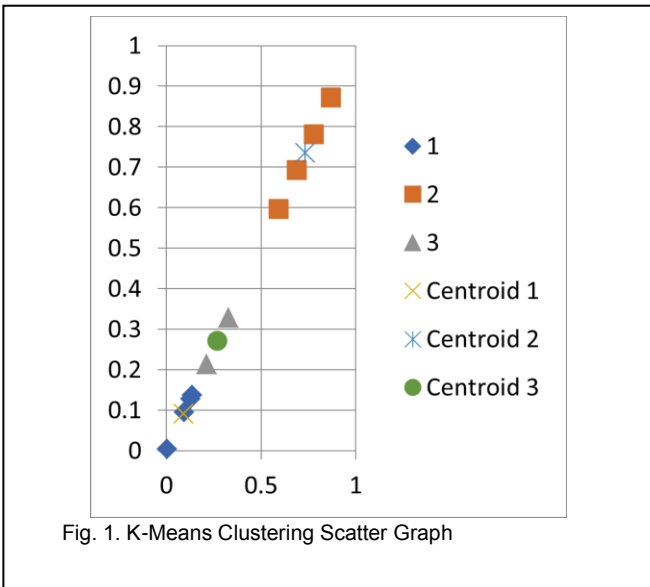


Fig. 1. K-Means Clustering Scatter Graph

Figure 1 shows that three cluster models were generated. Each cluster model has its own characteristics. Based on the given figure, Cluster 1 serves as the cluster with the hardest sets of questions. The profile of the clusters are shown below.

TABLE 2
GENERATED CLUSTERS

CLUSTER	CENTROID (MEAN VALUE)	DIFFICULTY LEVEL	NO. OF ITEMS
1	0.090811	HARD	4
2	0.734685	EASY	4
3	0.27087	AVERAGE	2
TOTAL			10

The generated clusters are presented in Table 2. This includes the centroid or the mean value, the difficulty level determined for each cluster, and the total number of items per cluster. The centroid value was used to determine the difficulty level. If the value is low, this means that there are few students who got the correct answer. If the value is high, the easier the questions are.

5.2 Placement of Questions

TABLE 3
PLACEMENT OF QUESTIONS ACCORDING TO CLUSTER MODELS

CLUSTER	QUESTION ITEM
CLUSTER 1	ITEM 1, ITEM 7, ITEM 8, ITEM 10
CLUSTER 2	ITEM 4, ITEM 5, ITEM 6, ITEM 9
CLUSTER 3	ITEM 2, ITEM 3
TOTAL	10

Table 3 shows the membership of question items to the clusters generated. The questions are grouped together using the Euclidean Distance of the K-means algorithm. Items 1, 7, 8, and 10 belong to Cluster 1. Cluster 2 contains items 4, 5, 6, and 9. And, Cluster 3 has items 2 and 3. With this clustering, items in the Cluster 1 are deemed hard (as seen in Table 1). These items have prepared reinforcement questions in case they will be grouped in the Hard items cluster.

TABLE 4
DECISION-BASED LEARNING

QUESTION ITEM	REINFORCEMENT QUESTIONS
ITEM 1	A1, A2, A3, A4, A5, A6...
ITEM 7	G1, G2, G3, G4, G5, G6...
ITEM 8	H1, H2, H3, H4, H5, H6...
ITEM 10	J1, J2, J3, J4, J5, J6...
TOTAL	105

Table 4 shows the structure of the reinforcement questions per item that are deemed Hard. A randomized algorithm was executed to get reinforcement questions and be able to create a review examination. In this decision-based learning, the prototype gets to decide if the student will have to take a review exam or if the student can proceed with the next activity. With this process, a personalized learning can be implemented suited the level of understanding of the students taking the examinations.

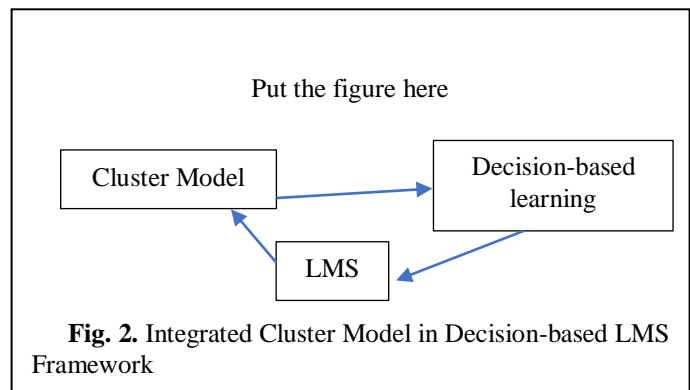


Fig. 2. Integrated Cluster Model in Decision-based LMS Framework

Figure 2 shows the framework of the research. The questionnaire was created on the Learning Management System. An item analysis was conducted to initially determine the difficulty level of each item. The clustering method was executed to cluster these questions. This determines a possible new set of groups for the Easy,

Average, and Hard items on the cluster model. Based on the cluster model, a decision was made by the system to determine the items that need reinforcements and the reinforcement questions that will be asked to the students. These reinforcement questions were selected in a pool of questions linked to a certain item.

TABLE 5: SIMULATION OF DECISION-BASED LEARNING

QUESTION ITEM	REINFORCEMENT QUESTION	STUDENT CORRECTNESS OF RESPONSE
ITEM 1	A1	WRONG
	A3	WRONG
	A4	WRONG
	A6	WRONG
	A9	CORRECT

In the sample data sets, items 1, 7, 8, and 10 are deemed Hard. The simulation of the decision-based learning from the clustered model is presented in Table 5 for Item #1. Assuming each row is from a different review exam. Reinforcement questions are pulled from the question pool through the use of random algorithm which is connected to item #1. As long as the student answers wrong for each reinforcement questions, more reinforcement questions will be asked until such time that the student answers correctly. This ensures that a student will have a better understanding on a certain topic.

TABLE 6
SUMMARY OF FURPS EVALUATION

CRITERIA	WEIGHTED MEAN	VERBAL INTERPRETATION
FUNCTIONALITY	4.6	EXCELLENT
USABILITY	4.4	VERY SATISFACTORY
RELIABILITY	4.2	VERY SATISFACTORY
PERFORMANCE	4.6	EXCELLENT
SUPPORTABILITY	4.4	VERY SATISFACTORY
OVERALL WEIGHTED MEAN	4.44	VERY SATISFACTORY

Table 6 shows the summary of the survey conducted by the researchers following the FURPS model. The system was evaluated by five LMS experts. Based on the table, *Functionality* and *Performance* got the highest weighted mean of 4.6. This shows that the prototype developed includes feature sets, many capabilities, and good security. It was also able to respond in a short time and consumes few computing resources. *Usability* and *Supportability* received a weighted mean of 4.4. Overall aesthetics, consistency, user-friendliness, and proper documentation were considered. The system also shows adaptability, extensibility, testability, serviceability, and compatibility. *Reliability* received the lowest mean of 4.2. The prototype showed few errors and bugs that can be solved easily. Overall, the experts evaluated the prototype *Very Satisfactory* with an overall mean of 4.44.

6 CONCLUSION

The general objective of the study is to discover a way of integrating cluster model in a learning management system. Specifically, it sought answers to the following research questions: (1) What cluster models can be extracted from the dataset?; (2) How are cluster models integrated in LMS?; (3) How acceptable is the system as perceived by the experts using FURPS model? The prototype was able to generate cluster models for items in the question bank. These models were determined as *Easy*, *Average*, and *Hard* questions. These cluster models were integrated in an examination. With the application clustering algorithm, the researchers found a new way of categorizing questions. With this process, the researchers were able to build a framework on how to integrate cluster models in a Learning Management System. The system was also evaluated by the LMS experts using FURPS model. Functionality and Performance gained a 4.6 weighted mean that commensurate to Excellent. While Usability and Supportability gained a weighted mean of 4.4 that interprets to Very Satisfactory. And lastly, Reliability has a 4.2 weighted with a verbal interpretation of Very Satisfactory. Overall, the experts rated the prototype as Very Satisfactory.

REFERENCES

- [1] D. Walker, J. Linder, T. P. Murphey and K. Dooley. Learning Management Usage Perspective from University Instructors, *The Quarterly Review of Distance Education*, vol. 17, pp. 41-50, 2016.
- [2] R. Cornish. Cluster Analysis, 2007. [Online]. Available: <http://www.statstutor.ac.uk/resources/uploaded/clusternalysis.pdf>.
- [3] B. Foley. An Introduction to Cluster Analysis, 15 February 2018. [Online]. Available: <https://www.surveymzmo.com/resources/blog/cluster-analysis/>. [Accessed July 2019].
- [4] Maselena, M. Huda, N. Sabani and B. Basiron. Demystifying Learning Analytics in Personalised Learning, *International Journal of Engineering and Technology*, 2018.
- [5] Berkhin. A Survey of Clustering Data Mining Techniques in Grouping Multidimensional Data, Berlin, Heidelberg, Springer, 2006.
- [6] Barbosa H., Garica F. and Rodriguez, M. Construction of Assessments with double adaptation processes. *Innovations in E-Learning Instructional Technology Assessment and Engineering*, Iskander (ed), pp. 133-13
- [7] M. Ballera, I. A. Lukandu and A. Radwan. Improving Learning Throughput in E-learning using Interactive-Cognitive Based Assessment, *The International Journal of E-Learning and Educational Technologies in the Digital Media*, vol. 1, no. 1, pp. 32-49, January 2015.