International Journal of Psychological Rehabilitation, Vol.24, Issue 2, 2020
ISSN: 1475-7192

Reliability Measurement of Learning Outcome using K-Means Cluster Technique

Muhammad Rozahi Istambul[1], Egi Abinowi [2)]

Abstract: *E-Learning has begun to be widely used in universities in Indonesia. However, various obstacles are commonly encountered in its implementation such as infrastructure and learning design. Usually, the problem of providing infrastructure is adjusted to the ability of a university to procure needs according to the conditions in the field. This is very different from the design of E-Learning which demands a change in pedagogical paradigm in the interaction between lecturers and students. Learning outcomes of a single course at the university under study are often not optimal/reliable and tend to indicate passive student participation in face-to-face activities in class. Therefore, the design of E-Learning is very important to be made in such a way that the level of student participation becomes higher and results in reliable learning outcomes. Reliability is measured using the K-means Cluster technique by monitoring the extent to which students interact in using E-Learning facilities.*

Keywords: *E-Learning, Pedagogic, Learning Outcome, Learning Design, K-Means Cluster*

I. INTRODUCTION

The use of E-Learning in several universities in Indonesia is indeed still not massive considering the need for technological infrastructure and learning designs that can change the habit of interacting in conventional learning activities.

The hope of those involved in the use of E-Learning is that these findings do not merely utilize technology, but must show advantages compared to the usual face-to-face method in class as well. The fundamental difference that occurs is that during class meetings, students are not actively participating in given the role of the lecturer as the main source of information for students.

Various assignments given to students cannot be completed directly because they are still dependent on meetings with lecturers the following week. This results in the unreliable material that must be mastered by students because the discussion of the material cannot be conveyed briefly, and students are still waiting for the lecturer's presentation.

On the other hand, students also have the habit of not completing assignments quickly because they rely on explanations at meetings that will take place a week later, and if done quickly even the results of the discussion cannot be immediately obtained. These various conditions cause learning outcomes to become unreliable which demands motivation and mastery of students in understanding topics more deeply based on lecturer explanations.

Learning design is one approach to make students actively participate in learning. This means that students must do many things that are not monotonous and get a variety of support or assistance from various sources. Convenience and ease of obtaining information should be a priority for students because what they will do is available on the internet, so students' insights and understanding will increase and most importantly students are given the confidence to obtain information from various sources with the time and way of learning according to their convenience.

Evaluation of student assignments is also important because students believe the assignment gets a good rating from the lecturer. This is based on the ability of learning outcomes obtained by students from various sources of information. As research data conducted at the engineering faculty of Widyatama University related to learning using E-Learning, further analysis will be carried out on the extent of student activity in the learning process. The courses that are used as a reference in this study are courses that lead to qualitative data where every student can apply it even in a small organizational environment.

Based on the various designs of learning activities and learning outcomes obtained by students, this research will measure the extent of active participation of students in the learning process so that this can become one of the reference models in the further development of E-Learning.

II. METHODOLOGY

[1] Information System Department, Widyatama University, Indonesia
Email : rozahi.istambul@widyatama.ac.id [1)], egi.abinowi@widyatama.ac.id [2)]

Learning outcomes are the expected results of student learning. Various learning designs need to be planned to get the right learning outcomes. As explained by Mrunal Mahajan and M. K. Sarjit Singh that learning outcomes are guidelines that guide students to achieve the desired results from learning planned subjects [5]. Learning outcomes are also an indicator of the success of a study program and provide clear ideas about what can be achieved by engaging in a particular program [5].

Meanwhile, the use of E-Learning to obtain meaningful learning outcomes to be utilized as optimal as possible, especially the various facilities available in the learning process. As explained by Valentina A. & Nelly A. that adopting E-Learning in education, especially for universities, will provide many benefits and advantages because E-Learning is considered as one of the best educational methods [11].

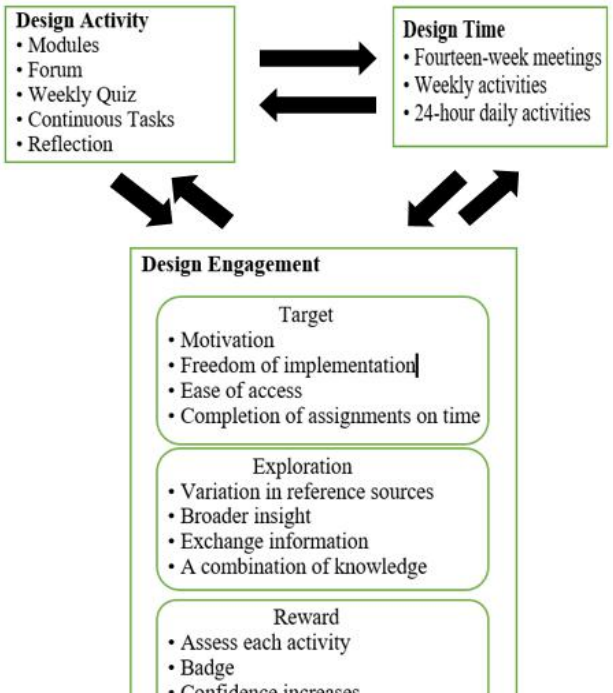
At present, the design of learning is very important to be a concern that demands the role of information technology as a reference for various sources and directions in learning through E-Learning. According to Richard and Haya in the quote Valentina A. & Nelly A. that the internet is one of the vital things in providing shared research and learning resources [11]. Whereas Olojo, Adewumi, and Ajisola said that E-Learning can improve knowledge and performance, where E-Learning offers students to control the content, a sequence of learning, time, and adjust their own learning experiences [7]. Likewise, Mohammed M. Alhawiti and Yasser Abdelhamid explained that E-Learning and information technology now have the potential to support various instructional designs for students that are different and well-integrated [4].

There were 45 students involved in this study and one subject was chosen to be given instructions on each weekly activity. Students, of course, have been given directions on freedom that refer to various sources [7] [11] and can carry out various learning activities according to the period provided. This course is in semester V where students are expected to be able to determine the stages and the best way to understand a learning topic already.

According to Gamze Tezcan & Gürsoy Meriç that a process evaluation can be used together with performance evaluation compared to traditional methods. Performance evaluation is more beneficial, but only at the initial stage. When evaluating performance, it is not possible to obtain general results, but evaluating items one by one is more meaningful [1]. All performance is optimized to meet the needs of students through facilitation of various interactions of student activeness in carrying out each instruction, as explained by Rozahi Istambul that the capabilities contained in the framework require identification of goals, needs, and individual learning styles and designs, as well as program management and evaluation, to fulfill these individual requirements [3].

While the learning framework that will be used in this E-Learning is by the limits of the various supporting components, as written by Samnan A., Amaad U., Stephen R. [10] which needs to be provided to support research activities. However, the framework in this study tends to lead to instructional designs conducted by lecturers and students. Besides, planning various types of interactions to optimize the function of E-Learning is also very important, as described by Rozahi Istambul that efforts are needed to prepare and distribute teaching materials together with forms of interaction between lecturers and students as a guide for lecturers and students in interacting online and or offline [2].

The model that has been designed to support the creation of a learning outcome in E-Learning activities involves three main component groups, namely; design activity, design time, and design engagement as shown below:



III. Analysis and discussion

In measuring the reliability of learning outcomes, data clustering is used. Data Clustering according to Saima Bano and M. N. A. Khan is a procedure in which a group of entities is created based on similar features [9]. Clustering is one of the widely used knowledge discovery techniques to reveal structures in a data set that can be extremely useful to the analyst. Besides that, according to N. Valarmathy, S. Krishnaveni, data clustering or grouping is a set of categories that are identified as groups based on the existing class similarity [6]. Clustering has a variety of applications in different domains viz data mining and knowledge discovery.

As data, references that will be used in clustering are all activities carried out by students for 14 weeks, namely; downloading modules, activity on discussion forums, taking weekly quizzes to discuss related topics, making ongoing assignments and doing reflection. Furthermore, the K-Means cluster method will be applied, i.e. the data that has been obtained is grouped into several clusters based on the similarity of the data so that the data that has the same characteristics will be grouped into one cluster and those that have different characteristics will be grouped in other clusters that have similar characteristics. According to Oyelade, O. J, Oladipupo, O. O, and Obagbuwa, I. C, K-means always meets with the local minimum. The specific local minimum found depends on the start of the centroid cluster [8].

The data used in this study comes from monitoring the activity log on one of the E-Learning system platforms. The following table illustrates the number of logs of activities that students do every day from Monday to Sunday. In this measurement, data processing uses Minitab 16 software based on the number of activities logs (Table 1) which are then processed using the K-Means Cluster method.

Table 1. The number of activities logs

	Weekly					
	2 & 3	**4 & 5**	**6 & 7**	**8 & 9**	**10 & 11**	**12 & 13**
Monday	280	359	242	46	190	106
Tuesday	576	233	231	231	206	290
Wednesday	332	206	111	116	109	254
Thursday	399	318	118	154	269	182
Friday	485	396	195	181	283	114
Saturday	503	386	1036	435	370	281
Sunday	1954	1953	1023	1445	1517	1333

The grouping is divided into three clusters using random centroids where initial clusters are chosen randomly with data between the lowest value and the highest value. The results can be seen in the following table 2.

Table 2. Results of K-Means Cluster Data Processing

Cluster	Number of observation	average distance from centroid	Maximum distance from centroid
1	1	0.000	0.000
2	1	0.000	0.000
3	5	179.925	248.151

In cluster 1, there was one observation with an average distance from observation to the centroid of 0,000 and the furthest distance from the centroid of 0,000. Whereas in cluster 2 there was one observation with an average distance from observation to the centroid of 0,000 and the furthest distance from the centroid of 0,000.

Finally, in cluster 3 there were five observations with an average distance from observations to the centroid of 179,925 and the farthest distance from the centroid of 248,151. Clusters that have a smaller average distance tend to be more compact than a cluster that has a greater average distance. Clusters that have a higher value indicate greater observational variability within the cluster.

Table 3. Centroid Clusters

Variable	Cluster1	Cluster2	Cluster3	Grand centroid
Weeks 2 & 3	503	1954	414.4	647
Weeks 4 & 5	389	1353	302.4	550.5714
Weaks 6 & 7	1036	1023	179.4	422.2857
Weaks 8 & 9	435	1445	157.6	381.1429
Weeks 10 & 11	370	1517	211.4	420.5714
Weeks 12 & 13	281	1333	189.2	365.7143

The picture in Table 3 is a centroid cluster with activity as a variable in weeks 2 through 13. Each variable has a centroid value for each cluster, so the grand centroid is obtained from several observations.

Table 4. Observations on centroids

Cluster	1	2	3
1	0	2827.3802	927.1704
2	2827.3802	0	3236.7851
3	927.1704	3236.7851	0

Table 4 is a table of distances between cluster centroids which measures how far the centroid cluster is in the final partition between each other.

Table 5. Results of Cluster Processing 1

	Weekly					
	2 & 3	4 & 5	6 & 7	8 & 9	10 & 11	12 & 13
Saturday	503	386	1036	435	370	281

Table 5 above shows the results of the K-Means cluster which produced cluster 1 on Saturday. The table shows the total number of activity logs in cluster 1 totaling 3011 log activities carried out by students for 12 weeks.

Table 6. Results of Cluster Processing 2

	Weekly					
	2 & 3	4 & 5	6 & 7	8 & 9	10 & 11	12 & 13
Sunday	1954	1953	1023	1445	1517	1333

Table 6 above shows the results of K-Means Clusters which produced cluster 2 on Sunday. The table shows the total number of activity logs in cluster 2 as many as 9225 student activity logs.

Table 7. Results of Cluster Processing 3

	Weekly					
	2 & 3	4 & 5	6 & 7	8 & 9	10 & 11	12 & 13
Monday	280	359	242	46	190	106
Tuesday	576	233	231	231	206	290
Wednesday	332	206	111	116	109	254
Thursday	399	318	118	154	269	182
Friday	485	396	195	181	283	114

Table 7 shows the results of the K-Means Cluster that produced cluster 3 on Monday, Tuesday, Wednesday, Thursday, and Friday with the sum of the total log activity in cluster 3 totaling 7212 log activities conducted by students.

Based on the results of measurements on the overall active participation of students related to the learning achievements obtained, the creation of three clusters marks three groups of time activities that are often done by students. The resulting measurement value shows the lowest activity log of the total number of students in cluster 1 with an average number of 66.91 activity logs per student on Saturday and the highest activity log being in cluster 2 with an average number of 205 activity logs per student on Sunday. While cluster 3 is between clusters 1 and 2 with an average of 160.27 logs of activity per student from Monday to Friday.

IV. CONCLUSIONS

1. Based on the measurement results above, daily activities show very high student participation every week as evidenced by the activity log.
2. There is a trend in the similarity in the implementation of each student's activities based on the cluster.
3. The day with the highest active participation of students on Sunday and Saturday has the lowest active participation. While participation from Monday to Friday is quite low so that it is grouped in the same cluster. Thus, the most active participation is between Saturdays and Sundays.

REFERENCES

[1] Gamze Tezcan, Gürsoy Meriç, “Developing a Science Process Skills Test Regarding the 6th Graders”, The International Journal of Assessment and Evaluation, First published in 2013 in Champaign, Illinois, USA, Volume 19

[2] Mh. Rozahi Istambul, “Optimizing the Implementation of e-Learning Strategies in Higher Education”, International Journal of e-Education, e-Business, e-Management and e-Learning, Vol. 2, No. 2, April 2012

[3] Muhammad Rozahi Istambul, Hari Supriadi, “Evaluation of Blended Learning Implementation which is Conditioned to Optimize the Mastery of Student Knowledge and Skills”, International Journal of Engineering & Technology, 7 (4.33), 2018, pp.195-200

[4] Mohammed M. Alhawiti, Yasser Abdelhamid, “A Personalized e-Learning Framework”, Journal of Education and e-Learning Research, Vol. 4, No. 1, 2017, pp.15-21

[5] Mrunal Mahajan and Manvinder Kaur Sarjit Singh, “Importance and Benefits of Learning Outcomes”, IOSR Journal Of Humanities And Social Science, Volume 22, Issue 3, Ver. V, March 2017, pp. 65-67

[6] N.Valarmathy, S.Krishnaveni, “Performance Evaluation and Comparison of Clustering Algorithms used in Educational Data Mining”, International Journal of Recent Technology and Engineering, Volume-7, Issue-6S5, April 2019

[7] Olojo Oludare Jethro, Adewumi Moradeke Grace, Ajisola Kolawole Thomas, “E-Learning and Its Effects on Teaching and Learning in a Global Age”, International Journal of Academic Research in Business and Social Sciences, January 2012, Vol. 2

[8] Oyelade, O. J, Oladipupo, O. O, Obagbuwa, I. C, “Application of k-Means Clustering algorithm for prediction of Students’” Academic Performance”, International Journal of Computer Science and Information Security, Vol. 7, 1, 2010

[9] Saima Bano and M. N. A. Khan, “A Survey of Data Clustering Methods”, International Journal of Advanced Science and Technology Vol.113, 2018, pp.133-142

[10] Samnan Ali, M. Amaad Uppal, Stephen R. Gulliver, “A conceptual framework highlighting e-learning implementation barriers”, Information Technology & People Vol. 31 No. 1, 2018, pp. 156-180

[11] Valentina Arkorful, Nelly Abaidoo, “The role of e-learning, the advantages and disadvantages of its adoption in Higher Education”, International Journal of Education and Research, Vol. 2 No. 12 December 2014