Journal of Advanced Technology and Multidiscipline (JATM)
Vol. 03, No. 01, 2024, pp. 13-20
e-ISSN: 2964-6162

Student's Behavior Clustering based on Ubiquitous Learning Log Data using Unsupervised Machine Learning

Ika Qutsiati Utami[1], Wu-Yuin Hwang[2], Ratih Ardiati Ningrum[3]

[1,3]*Data Science Technology, Engineering Department, Faculty of Advanced Technology and Multidiscipline, Universitas Airlangga, Surabaya, Indonesia*
[1,2]*Graduate Institute of Network Learning Technology, National Central University, Taoyuan, Taiwan*
[2]*Department of Computer Science and Information Engineering, National Dong Hwa University, Hualien, Taiwan*

Abstract*—Online learning is the source of data generation related to learner's learning behaviors, which is valuable for knowledge discovery. Existing research emphasized more on an understanding of student's performance and achievement from learning log data. In this study, we presented data-driven learning behavior clustering in authentic learning context to understand students' behavior while participating in the learning process. The objective of the study is to distinguish students according to their learning behavior characteristics and identify clusters of students at risk of unsuccessful learning achievement. Learning log data were collected from ubiquitous learning applications before conducting Exploratory Data Analysis (EDA) and cluster analysis. We used partitional clustering using K-means algorithm and hierarchical clustering based on the agglomerative method to improve clustering strategies. The result of this study revealed three different clusters of students supported by data visualization techniques. Cluster 1 comprised more students with active learning behavior based on the total logs, total problems posed, and the total attempts in fraction operation and simplification. Students in clusters 2 and 3 had a higher attempt at problem-solving instead of problem-posing. Both clusters also focused on fraction's conceptual understanding. Knowledge discovery of this study used real data generated from ubiquitous learning application namely U-Fraction. We combined two different types of clustering method for delivering more accurate portrait of a student's hidden learning behaviors. The outcome of this study can be a basis for educational stakeholders to provide preventive learning strategies tailored to a different cluster of students.

***Keywords*—Learning analytics, behavior clustering, unsupervised learning, learning log-data, education research, educational policies.**

I. INTRODUCTION

Over ten years, starting from 2011, learning analytics (LA) with data-driven analysis has arisen by exploiting machine learning in the educational field [1], [2]. Several research studies in educational data mining and artificial intelligence have tempted to distinguish the LA movement in an educational context [3], [4], [5], [6]. LA used educational data for knowledge discovery and transform data into meaningful insights. It is used for leveraging educational data to support the teaching and learning process. The main purpose of LA is to utilize models in improving learning and evaluating the process through instrumental investigation. According to the first definition of LA, it is an approach to collecting, analyzing, and reporting educational data related to learning, learners, and its related context [7]. There are several techniques used for the analytical process of learning related data such as supervised and unsupervised learning methods. The main difference between the two approaches is the use of labeled data (Nafis and Biswas, 2022; Shakarami, Shahidinejad and Ghobaei-Arani, 2021). In the unsupervised learning method, the process of data analysis to learn the patterns from data does not require labeled input and output data. It is being used generally for clustering and segmentation-related tasks. The algorithm performs natural clustering over the dataset to identify similar patterns and characteristics. The process of learning about user behavior from log data typically involves partitioning the data into meaningful subsets, called partitions, and comparing the different partitions.

In an educational context, cluster analysis can be used to gain insight into structured data such as student behavior grouping, finding similar learning patterns, and student performance clustering [10], [11]. However, despite the potential of unsupervised learning or cluster analysis for LA, it is seldom utilized for supporting teaching and learning analysis in ubiquitous learning contexts based on students' learning log data [12]. Log data is automatically produced files and timestamps relevant to the system or software application [13]. Log data can provide a portrait of a student's hidden learning behavior and give a more complete or accurate picture of all behaviors. Yet, log data generated by the learning application server had left the characteristic prone to data noise. The process of mining and reducing noise in log data is considered as challenging task. In addition to that fact, this study tries to perform an unsupervised learning method on student behavior based on learning log data generated from ubiquitous learning applications. In this study, log data refers to all students' activity while using the learning system namely ubiquitous fraction (U-Fraction) [14], [15]. This learning application is installed on a tablet device with an Android operating system. By analyzing learning log data produced from the application,

Journal of Advanced Technology and Multidiscipline (JATM)
Vol. 03, No. 01, 2024, pp. 13-20
e-ISSN: 2964-6162

the educational stakeholder can obtain learning problems at the earliest possible time. Additionally, it can enable them to resolve learning issues in a timelier fashion. Most importantly, a lot of data from learning systems and applications can be analyzed using machine-learning techniques to support decision-making in the educational field. We structured this paper as follows. In Section 1, an overview of LA especially for cluster analysis in an educational context was presented. Section 2 presented a literature review of related studies. Section 3 described the methodological part of the research. Furthermore, section 4 explained the result of the study followed by a research discussion. Last, section 5 provided the conclusion of the study. .

II. Literature review

A. *Learning log data in educational context*

Learning log data is defined as important source to provide powerful portrait of students learning patterns and their hidden behaviors during participation on learning process. Log data are commonly collected from online learning platforms such as virtual learning environment, e-learning, or mobile learning applications. Accessing and analyzing learning log data is challenging due to privacy issue and proper storage management. Effective learning log data management requires more time to be processed because the huge amount of information collected from online server need complex treatment like understanding of application usage, preprocessing task, data engineering, and data architecture provision. In the past research, some studies focused on the direction how to interpret learning log data in understanding student learning process in flipped classroom [16], [17], [18]. Commonly, researchers on learning analytics used learning log data from Learning Management System (e.g., Moodle, Canvas, etc.) or Massive Open Online Course (e.g., Coursera, Udemy, etc.). The learning analytics goals emphasized teaching and learning processes in asynchronous learning networks. For example, data collection related to the number of posts, the number of posts read, the number of posts replied, and content viewed.

A limitation of previous studies is that they focus on student performance and student satisfaction which typically rely on self-reporting and may be inaccurate [19], [20]. Therefore, more studies based on log-file data are needed in order to add an additional level of research validity to the understanding of students' behavior in relation to the authentic learning approach. It has been argued that log-file data may be more genuine and authentic than survey data, which are prone bias into students' interpretations [21]. Instead, learning analytics can reflect real and uninterrupted user behavior [17]. Therefore, rather than relying on student perceptions, this study examines ubiquitous application learning log data on how student interact in authentic learning context and how students accessed learning material. However, the use of learning log data in mobile application particularly for authentic learning context had not yet fully exploited to unveil students learning behaviors. Whereas, the adoption and acceptance of mobile learning has led to a dramatic increase in available learner data. Furthermore, students learning and social interaction with real-world situation is critical to be learned and analyzed. Currently, there have been few studies that examine students' actions like their interaction behaviors rather than their perceptions and performance. The present study takes a further step toward the direction to propose an approach in interpreting students learning log data to understand how students learn in the authentic situation over time. These findings are hinting that log data could be an important source to identify behavioral interaction in authentic learning context.

B. *Cluster analysis used in educational purpose*

To support the call for LA in education, several cluster analyses have been researched in the literature. While some research studies have collected log-file data from virtual learning environments, the data was frequently evaluated using more conventional statistical techniques like regression, correlation, and t-tests rather than analytics algorithms [22]. Instead, cluster analysis serves as an exploratory method that aims to identify naturally occurring homogeneous groups that were either unclear or previously unknown [23]. With a rapid increase in available learner data, cluster analysis becomes the potential in understanding and unveiling hidden information about students in educational settings [24]. Studies by Yadav [25], [26] proposed a new approach known as hybrid clustering to assess students’ academic performance. The clusters are formed based on the intelligence level of students. Walsh and Risquez [22] used cluster analysis to explore the engagement of native and non-native English-speaking management students in a flipped classroom. They used log file data to identify hidden patterns in student behavior, paying particular attention to the institution's native language proficiency.

Research shows the exploratory potential of cluster analysis on log file data in other contexts such as peer tutoring [27], [28]. However, despite its potential, cluster analysis is still underutilized in the context of education. Moreover, the rare previous application of cluster analysis to study student learning behavior in ubiquitous learning contexts remained unclear. The present study is adapted from the work of Jovanovic et al. [29]. However, in the present study, we applied cluster analysis to log-file data to identify patterns in how students access online resources over time while engaging with a ubiquitous learning application. This paper attempts to address the lack of research using learning analytics in the ubiquitous learning context, using students' learning log data from a mobile application, and the cluster analysis algorithm using hierarchical and partitional methods.

III. Method

In the present study, we employed EDA as an initial technique for understanding the dataset. Investigation of data using EDA is used to discover unseen patterns, data anomalies, and a summary of the data [30]. Two important practices in EDA i.e., descriptive statistics and data visualization were used to gather insight from the data [31]. Before conducting EDA, we accessed the data from the online repository and organized it using Structured Query Language (SQL) operations such as data selection, data join, and data aggregation. We used

Journal of Advanced Technology and Multidiscipline (JATM)
Vol. 03, No. 01, 2024, pp. 13-20
e-ISSN: 2964-6162

learning log data generated from a ubiquitous learning application namely U-Fraction. The dataset is related to student learning activity while using the application such as problem-solving activities and peer assessment. It was adapted from an experimental study conducted by Hwang et. al. in 2018 [32]. The data log structure before data preprocessing is represented by the database design (Figure 2). There are 10 variables selected for cluster analysis after the data preprocessing stage and feature selection stage. The attributes of the dataset are presented in Table 1.

TABLE 1
THE ATTRIBUTE OF THE DATASET

No	Attribute name	Description
1	Operation	Total attempts of fraction operation
2	Success_oper	Total of successful fraction operation
3	Simplification	Total attempts at fraction simplification
4	Success_simp	Total of successful fraction's simplification
5	Asking	Problem posing
6	Answer	Problem-solving
7	Comment	Peer assessment
8	Understanding	Fraction understanding
9	Log1	Total data logging 1
10	Log2	Total data logging 2

After the data preprocessing step with EDA, we followed a two-step cluster analysis using a K-means algorithm and agglomerative method. The K-means algorithm is a partition-based clustering method, while agglomerative is a hierarchical clustering method [27], [28]. K-means is best suited for a small-to-medium number of clusters, as is the case for student behavior clustering of this study [29]. The clustering process in K-means started by defining the number of clusters k [30], [31]. In addition, each of k is represented by a cluster center and each data point is assigned to the nearest cluster center namely the centroid. The algorithm group data that has similar characteristics into the sample cluster, while data with different characteristics are grouped into other clusters [32], [33]. Typically, the Euclidean distance is used as a distance measure. The calculation using the Euclidian Distance formula (equation 1) with the description of the formula in Table 2 is as follows:

$$d\,(\mathrm{x}, \mathrm{y}) = \sqrt{\sum_{i=1}^{n}(y_i - x_i)^2} \tag{1}$$

TABLE 2
THE FORMULA DESCRIPTION

Symbol	Description
d	Calculation of the distance to the center of the cluster
x	Point coordinates of the object
y	Centroid coordinate
$\sum_{i=1}^{n}(y_i - x_i)^2$	The amount of data to be measured, while i = 1 is the clustering process starting from the first iteration
x_i	Coordinate the point of the i object
y_i	i centroid coordinate point

In the next step, new cluster centers are defined as the center of mass of each cluster candidate. Unless the following termination criterion is met, this process is repeated. The algorithm terminates if the last iteration did not lead to changes in the assignment of each data point to the current cluster centers [25]. The pseudocode is given in Table 3. Beside a K-means algorithm, we also performed the agglomerative method as a bottom-up approach to hierarchical clustering. Recursively, each observation starts in its cluster, and pairs of clusters are merged as one moves up the hierarchy. This method works from the dissimilarities between the objects to be grouped. A type of dissimilarity can be suited to the subject studied and the nature of the data. Overall, the process in research methodology is presented in Figure 1.

TABLE 3
K-MEANS PSEUDOCODE

```
Algorithm 1. K-Means Algorithm
  Data: number of clusters k, dataset X
  Result: cluster centres C = {c1, ..., ck}
  Start
    Randomly select k data points as initial cluster centres;
    Repeat
      Reinitialize all partition S subsets as empty:
      S1 = S2 = ··· = Sk = {};
      Compute the distance of each data point to each cluster centre;
      Assign each data point to the closest cluster centre:
      for i ∈ {1, ..., N} do
        respective label l = argmin_{j ∈ {1, ..., k}} ‖ xi − cj ‖²;
        Sl = Sl ∪ {xi};
      End
      Define new cluster centres based on the current partition:
      for j ∈ {1, ..., k} do
        cj = ∑_{i ∈ {1, ..., N} xi ∈ Sj} xi / |Sj|
      End
    until the cluster assignment converges;
  End
```

IV. RESULT AND DISCUSSION

In this section, we explained the results of the present study. The results of the study are categorized into two sub-sections as follows:

A. *Exploratory data analysis*

In this step, we performed data pre-processing using EDA such as data cleaning (i.e., missing value computation and data noise treatment), data transformation, and data reduction. EDA is important step in data analytic task because it performs initial investigation on data to discover patterns, to spot some anomalies, to test hypothesis, and to check assumptions using summary statistics and graphical representations. In present study, we employed several Python libraries such as Pandas, NumPy, Matplotlib, and ScikitLearn to perform the EDA's operation and cluster analysis. The learning log dataset comprises 4202 observations and 11 characteristics. We used data.head(10) function to show the dataset with only ten rows available (see Table 4). Furthermore, dataset information including summary and missing value checking results is presented in Figure 3. From the dataset summary, we can identify the total of the column and the data type of each column. Data has only non-null and integer values. In addition, missing value analysis is used to check whether the dataset contains a null value after the data pre-processing step. From the result, we concluded that all columns have no missing values.

Journal of Advanced Technology and Multidiscipline (JATM)
Vol. 03, No. 01, 2024, pp. 13-20
e-ISSN: 2964-6162

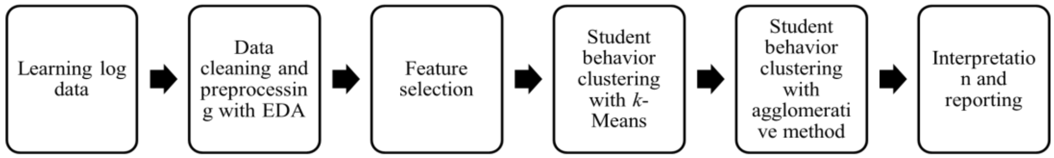

Fig. 1 Research design

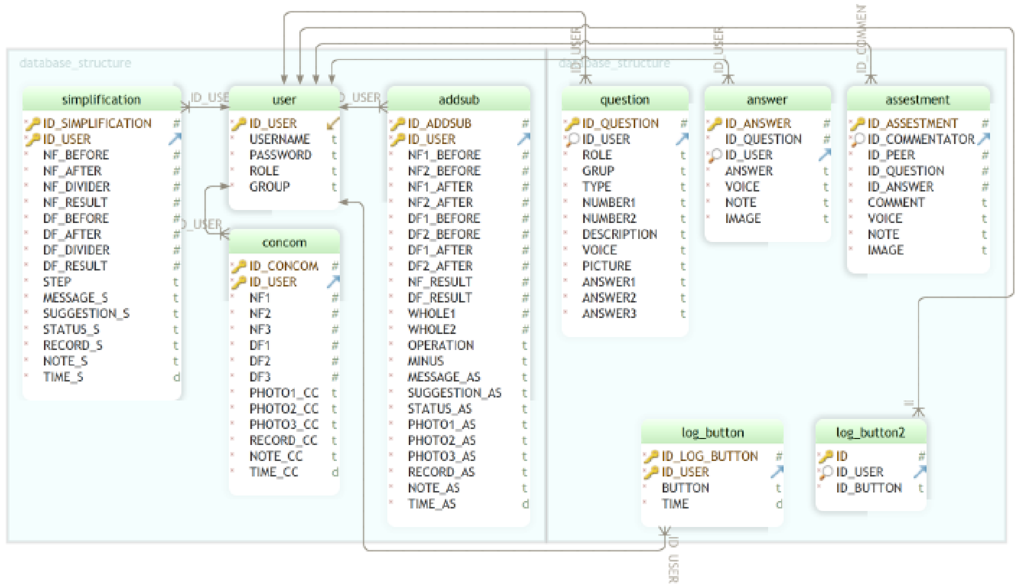

Fig. 2 The data log structures

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 25 entries, 0 to 24
Data columns (total 11 columns):
 #   Column         Non-Null Count  Dtype
---  ------         --------------  -----
 0   USER           25 non-null     int64
 1   OPERATION      25 non-null     int64
 2   SUKSES_OPER    25 non-null     int64
 3   SIMPLIFY       25 non-null     int64
 4   SUKSES_SIMP    25 non-null     int64
 5   ASKING         25 non-null     int64
 6   ANSWER         25 non-null     int64
 7   COMMENT        25 non-null     int64
 8   UNDERSTANDING  25 non-null     int64
 9   LOG1           25 non-null     int64
 10  LOG2           25 non-null     int64
dtypes: int64(11)
memory usage: 2.3 KB
```

(a)

```
USER             0
OPERATION        0
SUKSES_OPER      0
SIMPLIFY         0
SUKSES_SIMP      0
ASKING           0
ANSWER           0
COMMENT          0
UNDERSTANDING    0
LOG1             0
LOG2             0
dtype: int64
```

(b)

Fig. 3 Dataset information. (a) data summary, (b) missing-value check

TABLE 4
THE DATASET WITH THE TOP 10 ROWS

User	Operation	Success_oper	Simplify	Success_simp	Asking	Answer	Comment	Understanding	Log1	Log2
1	24	22	162	16	243	69	20	292	2621	518
2	45	44	128	20	242	74	14	287	2928	512
3	148	18	177	18	74	70	21	290	2601	541
4	9	38	138	44	97	73	12	258	2300	560
5	62	29	137	35	98	81	12	264	2382	1429
6	109	33	381	12	203	87	19	241	2290	3401
7	33	22	181	15	192	85	12	255	2892	506
8	26	18	239	12	107	93	12	291	2334	943
9	39	49	144	53	161	89	12	254	2284	567
10	48	20	155	16	153	135	12	244	2212	1418

Journal of Advanced Technology and Multidiscipline (JATM)
Vol. 03, No. 01, 2024, pp. 13-20
e-ISSN: 2964-6162

B. Cluster analysis with K-means algorithm and agglomerative method

Cluster analysis in step 1, the student's clusters of learning behavior are identified using a K-means algorithm. We used two methods to select the optimum number of clusters k. The first method is based on Elbow Method, an empirical method to obtain the best value of k. This method calculates the sum of the square of the points and the average distance. Figure 4 shows the result of the elbow method. We concluded that the optimal value of the cluster is 3 as presented by the last elbow point (Figure 4). The second method is called the Silhouette method. It calculates the silhouette coefficient of every point. The value of the Silhouette score varies from -1 to 1. Silhouette score 1 means the cluster is dense and well-separated than other clusters. A value near 0 represents overlapping clusters with samples very close to the decision boundary of the neighboring clusters. A negative score indicates that the samples might have got assigned to the wrong clusters. In this cluster analysis, we obtained 3 clusters as the optimum value because it has a higher score of the Silhouette method (Figure 5). Furthermore, the characteristic of each cluster is shown in Table 5. The result revealed different clusters of students based on their learning behavior variable related to a ubiquitous learning activity. We also presented the comparison of each cluster of students' learning behavior using parallel coordinates plots (Figure 6).

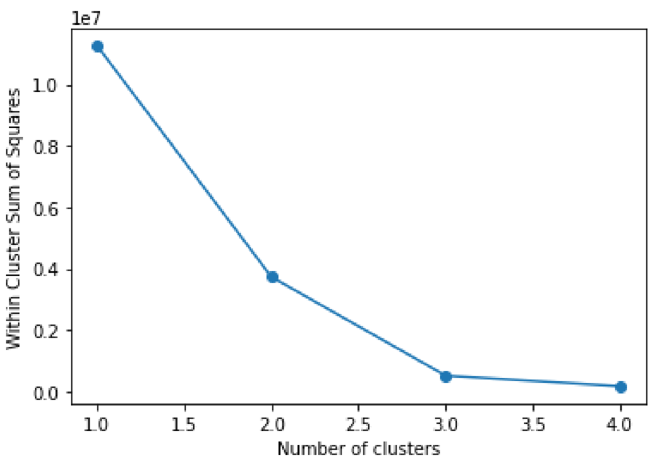

Fig. 4. Elbow method

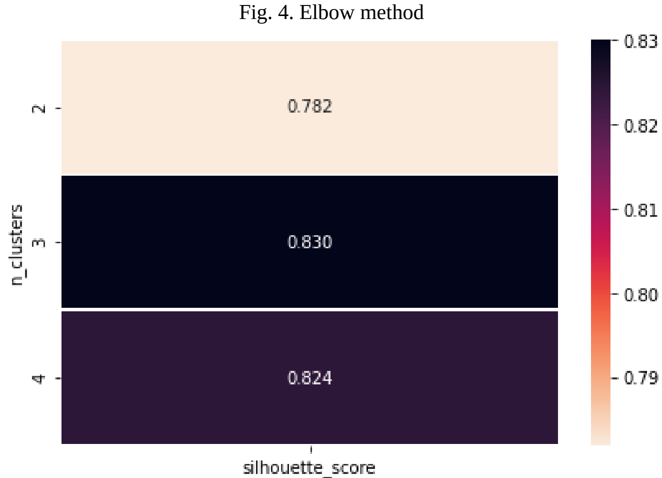

Fig. 5 Silhouette score

TABLE 5
THE CLUSTER CHARACTERISTIC

Characteristic	Cluster 1 (*N*=16)	Cluster 2 (*N*=8)	Cluster 3 (*N*=1)
Total attempts of fraction operation	109.0	42.4	36.9
Total of successful fraction operation	33.0	22.9	26.0
Total attempts at fraction simplification	381.0	173.1	186.6
Total of successful fraction's simplification	12	22.7	22.8
Problem-posing	203.0	158.0	125.9
Problem-solving	87.0	99.1	96.6
Peer assessment	19.0	19.5	15.4
Fraction understanding	241.0	263.5	264.0
Total data logging 1	2290.0	2426.4	2445.2
Total data logging 2	3401.0	569.9	1501.1

Based on the visualization, three clusters of students were identified. Cluster 1 comprised more students with active learning behavior based on the total logs, total problems posed, and the total attempts in fraction operation and simplification. Students in clusters 2 and 3 had a higher attempt at problem-solving instead of problem-posing. Both clusters also focused on fraction understanding. Additionally, students in cluster 1 were similar to those in cluster 2 in terms of peer assessment activity. In step 2 of cluster analysis, we performed hierarchical clustering based on the agglomerative method. Figure 7 shows a dendrogram, a diagram of the hierarchical relationship between the students. The clades that are close to the same height are similar to each other. The result revealed all students of each cluster that has similar learning behavior characteristics. According to the result, there are three clusters performed. Cluster 1 consists of 16 students, followed by cluster 2 with 8 students and cluster 3 with 1 student. This result of hierarchical clustering using the agglomerative method is similar to the previous method using a K-means algorithm as partitional clustering.

V. CONCLUSIONS

This research used the unsupervised learning method of machine learning to discover a similar pattern of students' learning log data and perform cluster analysis in order to obtain students' behavior clustering. The dataset is collected from students' learning activity while using the ubiquitous fraction app called U-Fraction. Data are processed in the initial step using EDA for data cleaning and transformation. Partition-based clustering methods using the K-means algorithm and hierarchical clustering methods using an agglomerative approach are used to create a cluster of students. The result showed three different clusters of students with different learning behavior characteristics. Cluster 1 comprised more students with active learning behavior based on the total logs, total problems posed, and the total attempts in fraction operation and simplification. Students in clusters 2 and 3 had a higher attempt at problem-solving instead of problem-posing. Both clusters also focused on fraction understanding. However, no significant difference in peer assessment activity among the groups. The outcome of this study can help educational

Journal of Advanced Technology and Multidiscipline (JATM)
Vol. 03, No. 01, 2024, pp. 13-20
e-ISSN: 2964-6162

stakeholders to provide preventive learning strategies tailored to different clusters of students.

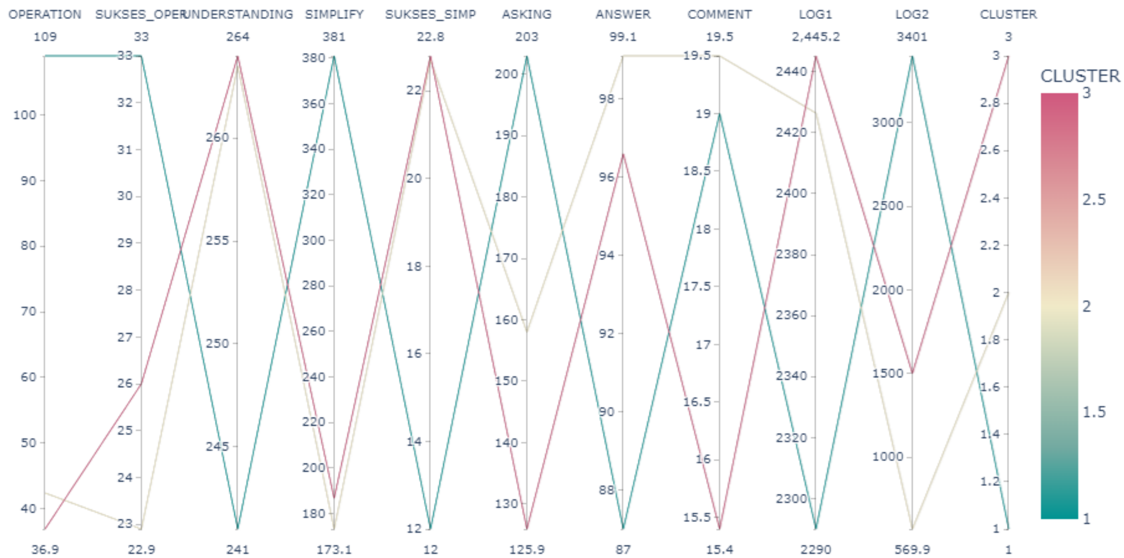

Fig. 6 The comparison of each cluster

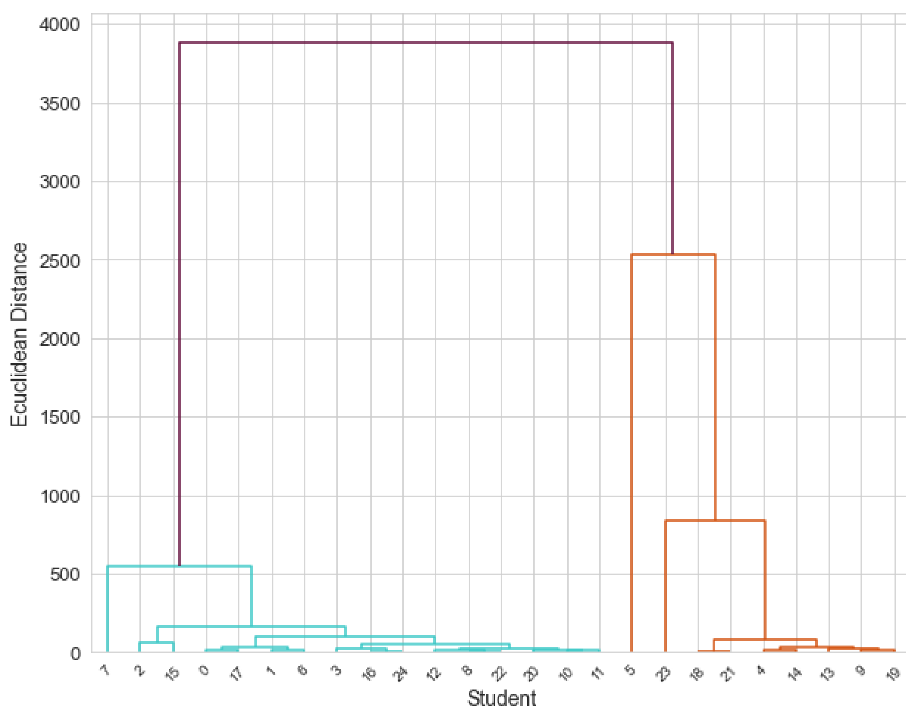

Fig. 7 Clustering results using hierarchical clustering

ACKNOWLEDGMENT

All authors read and approved the final manuscript.

REFERENCES

[1] B. Quadir, M. Chang, and J. C. Yang, "Categorizing learning analytics models according to their goals and identifying their relevant components: A review of the learning analytics literature from 2011 to 2019," Computers and Education: Artificial Intelligence, vol. 2, p. 100034, 2021, doi: https://doi.org/10.1016/j.caeai.2021.100034.

[2] D. J. Lemay, C. Baek, and T. Doleck, "Comparison of learning analytics and educational data mining: A topic modeling approach," Computers and Education: Artificial Intelligence, vol. 2, p. 100016, 2021, doi: https://doi.org/10.1016/j.caeai.2021.100016.

[3] A. A. Mubarak, H. Cao, and S. A. M. Ahmed, "Predictive learning analytics using deep learning model in MOOCs' courses videos," Educ Inf Technol (Dordr), vol. 26, no. 1, pp. 371–392, 2021, doi: 10.1007/s10639-020-10273-6.

[4] M. Shorfuzzaman, M. S. Hossain, A. Nazir, G. Muhammad, and A. Alamri, "Harnessing the power of big data analytics in the cloud to support learning analytics in mobile learning environment," Comput Human Behav, vol. 92, pp. 578–588, 2019, doi: https://doi.org/10.1016/j.chb.2018.07.002.

[5] R. Bodily, R. Nyland, and D. Wiley, "The RISE Framework: Using Learning Analytics to Automatically Identify Open Educational Resources for Continuous Improvement The RISE Framework: Using Learning Analytics to Automatically Identify Open Educational Resources for Continuous Improvement Bodily, Nyland, and Wiley," 2017.

[6] K. M. L. Jones, "Learning analytics and higher education: a proposed model for establishing informed consent mechanisms to promote student privacy and autonomy," International Journal of Educational Technology in Higher Education, vol. 16, no. 1, p. 24, 2019, doi: 10.1186/s41239-019-0155-0.

[7] J. Han, K. H. Kim, W. Rhee, and Y. H. Cho, "Learning analytics dashboards for adaptive support in face-to-face collaborative argumentation," Comput Educ, vol. 163, p. 104041, 2021, doi: https://doi.org/10.1016/j.compedu.2020.104041.

[8] M. T. Nafis and R. Biswas, "A secure technique for unstructured big data using clustering method," International Journal of Information Technology, vol. 14, no. 3, pp. 1187–1198, 2022, doi: 10.1007/s41870-019-00278-x.

[9] A. Shakarami, A. Shahidinejad, and M. Ghobaei-Arani, "An autonomous computation offloading strategy in Mobile Edge Computing: A deep learning-based hybrid approach," Journal of Network and Computer Applications, vol. 178, p. 102974, 2021, doi: https://doi.org/10.1016/j.jnca.2021.102974.

[10] S. S. Hamidi, E. Akbari, and H. Motameni, "Consensus clustering algorithm based on the automatic partitioning similarity graph," Data Knowl Eng, vol. 124, p. 101754, 2019, doi: https://doi.org/10.1016/j.datak.2019.101754.

[11] T. Li, A. Rezaeipanah, and E. M. Tag El Din, "An ensemble agglomerative hierarchical clustering algorithm based on clusters clustering technique and the novel similarity measurement," Journal of King Saud University - Computer and Information Sciences, vol. 34, no. 6, Part B, pp. 3828–3842, 2022, doi: https://doi.org/10.1016/j.jksuci.2022.04.010.

[12] A. Mathrani, T. Susnjak, G. Ramaswami, and A. Barczak, "Perspectives on the challenges of generalizability, transparency and ethics in predictive learning analytics," Computers and Education Open, vol. 2, p. 100060, 2021, doi: https://doi.org/10.1016/j.caeo.2021.100060.

[13] S. Dumais, R. Jeffries, D. M. Russell, D. Tang, and J. Teevan, "Understanding User Behavior Through Log Data and Analysis," in Ways of Knowing in HCI, J. S. Olson and W. A. Kellogg, Eds., New York, NY: Springer New York, 2014, pp. 349–372. doi: 10.1007/978-1-4939-0378-8_14.

[14] W.-Y. Hwang, I. Q. Utami, and H. Chen, "An Evaluation Study of Learning Behaviors and Achievements with Ubiquitous Fraction (u-Fraction) for Elementary School Student," in 2018 IEEE 18th International Conference on Advanced Learning Technologies (ICALT), 2018, pp. 350–354. doi: 10.1109/ICALT.2018.00087.

[15] I. Q. Utami and W.-Y. Hwang, "The impact of collaborative problem posing and solving with ubiquitous-decimal app in authentic contexts on math learning," Journal of Computers in Education, vol. 9, no. 3, pp. 427–454, 2022, doi: 10.1007/s40692-021-00209-5.

[16] J. N. Walsh and A. Rísquez, "Using cluster analysis to explore the engagement with a flipped classroom of native and non-native English-speaking management students," International Journal of Management Education, vol. 18, no. 2, Jul. 2020, doi: 10.1016/j.ijme.2020.100381.

[17] W. Greller and H. Drachsler, "Translating Learning into Numbers: A Generic Framework for Learning Analytics," 2012. [Online]. Available: http://groups.google.com/group/learninganalytics

[18] D. Gašević, S. Dawson, T. Rogers, and D. Gasevic, "Learning analytics should not promote one size fits all: The effects of instructional conditions in predicting academic success," Internet and Higher Education, vol. 28, pp. 68–84, Jan. 2016, doi: 10.1016/j.iheduc.2015.10.002.

[19] M. B. Gilboy, S. Heinerichs, and G. Pazzaglia, "Enhancing student engagement using the flipped classroom," J Nutr Educ Behav, vol. 47, no. 1, pp. 109–114, Jan. 2015, doi: 10.1016/j.jneb.2014.08.008.

[20] J. Jovanović, D. Gašević, S. Dawson, A. Pardo, and N. Mirriahi, "Learning analytics to unveil learning strategies in a flipped classroom," Internet and Higher Education, vol. 33, pp. 74–85, Apr. 2017, doi: 10.1016/j.iheduc.2017.02.001.

[21] I.-H. Jo, D. Kim, and M. Yoon, "International Forum of Educational Technology & Society Constructing Proxy Variables to Measure Adult Learners' Time Management Strategies in LMS," Source: Journal of Educational Technology & Society, vol. 18, no. 3, pp. 214–225, 2015, doi: 10.2307/jeductechsoci.18.3.214.

[22] J. N. Walsh and A. Rísquez, "Using cluster analysis to explore the engagement with a flipped classroom of native and non-native English-speaking management students," International Journal of Management Education, vol. 18, no. 2, Jul. 2020, doi: 10.1016/j.ijme.2020.100381.

[23] S. Brown, S. White, and N. Power, "Cluster Analysis of Assessment in Anatomy and Physiology for Health Science Undergraduates," International Journal of Teaching and Learning in Higher Education, vol. 28, no. 1, pp. 102–109, 2016, [Online]. Available: http://www.isetl.org/ijtlhe/

[24] W. Greller and H. Drachsler, "Translating Learning into Numbers: A Generic Framework for Learning Analytics," 2012. [Online]. Available: http://groups.google.com/group/learninganalytics

[25] J. Patel and R. S. Yadav, "Applications of Clustering Algorithms in Academic Performance Evaluation," OAlib, vol. 02, no. 08, pp. 1–14, 2015, doi: 10.4236/oalib.1101623.

[26] R. S. Yadav, "Application of hybrid clustering methods for student performance evaluation," International Journal of Information Technology, vol. 12, no. 3, pp. 749–756, 2020, doi: 10.1007/s41870-018-0192-2.

[27] M. De Smet, H. Van Keer, and M. Valcke, "Blending asynchronous discussion groups and peer tutoring in higher education: An exploratory study of online peer tutoring behaviour," Comput Educ, vol. 50, no. 1, pp. 207–223, Jan. 2008, doi: 10.1016/j.compedu.2006.05.001.

[28] G. Lust, M. Vandewaetere, E. Ceulemans, J. Elen, and G. Clarebout, "Tool-use in a blended undergraduate course: In Search of user profiles," Comput Educ, vol. 57, no. 3, pp. 2135–2144, Nov. 2011, doi: 10.1016/j.compedu.2011.05.010.

[29] J. Jovanović, D. Gašević, S. Dawson, A. Pardo, and N. Mirriahi, "Learning analytics to unveil learning strategies in a flipped classroom," Internet and Higher Education, vol. 33, pp. 74–85, Apr. 2017, doi: 10.1016/j.iheduc.2017.02.001.

[30] A. K. Jain, R. P. W. Duin, and J. Mao, "Statistical pattern recognition: a review," IEEE Trans Pattern Anal Mach Intell, vol. 22, no. 1, pp. 4–37, 2000, doi: 10.1109/34.824819.

[31] C.-M. Liu, Z. Niu, and K.-T. Liao, "Mechanisms to improve clustering uncertain data with UKmeans," Data Knowl Eng, vol. 116, pp. 61–79, 2018, doi: https://doi.org/10.1016/j.datak.2018.05.004.

[32] W.-Y. Hwang, I. Q. Utami, S. W. D. Purba, and H. S. L. Chen, "Effect of Ubiquitous Fraction App on Mathematics Learning Achievements and Learning Behaviors of Taiwanese Students in Authentic Contexts," IEEE Transactions on Learning Technologies, vol. 13, no. 3, pp. 530–539, 2020, doi: 10.1109/TLT.2019.2930045.

[33] N. Nidheesh, K. A. Abdul Nazeer, and P. M. Ameer, "An enhanced deterministic K-Means clustering algorithm for cancer subtype prediction from gene expression data," Comput Biol Med, vol. 91, pp. 213–221, 2017, doi: https://doi.org/10.1016/j.compbiomed.2017.10.014.

Ika Qutsiati Utami is a Ph.D. student at the Graduate Institute of Network Learning Technology, College of Electrical Engineering & Computer Science, National Central University, ROC, Taiwan. She received B.S. Degree in Information Systems from the Faculty of Computer Science, Brawijaya University, Indonesia & obtained M.S. Degree in Network Learning Technology from the College of Electrical Engineering & Computer Science, National Central University, Taiwan. She is also working as a lecturer at the Faculty of Advanced Technology and Multidiscipline, Universitas Airlangga, Indonesia. Her research interests are in the areas of Human-Computer Interaction, AI/ML for Future Education, and Educational Data Science.

Wu-Yuin Hwang is a professor affiliated with both the Department of Computer Science and Information Engineering, College of Science and Engineering, National Dong Hwa University, Taiwan, and the Institute of Network Learning Technology, National Central University, Taiwan. His current research interests are related to the integration of IOT, AI, and multimedia sensors of mobile devices for learning and interactions among humans and all things in AR contexts like smart buildings and campuses. Dr. Hwang received the Outstanding Research Award, from the Ministry of Science and Technology, Taiwan in 2021. He is also ranked in the top 7 scholars of the world in terms of high-quality journal publication performance of instructional design and technology.

Journal of Advanced Technology and Multidiscipline (JATM)
Vol. 03, No. 01, 2024, pp. 13-20
e-ISSN: 2964-6162

Ratih Ardiati Ningrum is a lecturer and researcher in Engineering Faculty at Airlangga University where she has been a faculty member since 2020. She graduated with a Statistics master's joint degree program from National Chiao Tung University, Taiwan, and Institut Teknologi Sepuluh Nopember, Surabaya. Currently, her research focuses on survival analysis and structural equation modeling.