



Clustering of Learning Media User Data During Covid-19 Pandemic Using K-Means Method Based on Multicultural Culture in Indonesia

Evi Fatimatur Rusydiyah^{(1)*}, Hikmah Rossyta Virgiannada⁽²⁾, Mujib Ridwan⁽³⁾, Bayu Adhi Nugroho⁽⁴⁾, Moh Rifqi Rahman⁽⁵⁾

(1) UIN Sunan Ampel Surabaya, Indonesia

(2) UIN Sunan Ampel Surabaya, Indonesia

(3) UIN Sunan Ampel Surabaya, Indonesia

(4) Curtin University, Australia

(5) UIN Sunan Ampel Surabaya, Indonesia

*Correspondence to: evifatimatur@uinsby.ac.id

Abstract: Learning applied in Indonesia from elementary school to undergraduate level generally uses face-to-face or direct learning methods. However, after the Covid-19 disease outbreak, all methods change to online learning methods at all levels and cultures. This affects all aspects of learning such as comfort, understanding and learning outcomes. To find out the difference in the results using face-to-face methods and online methods, a study was carried out on classifying learning media users from elementary school to university level using the k-means method. This study aimed to determine the differences in learning outcomes between the semester before the Covid-19 pandemic and the semester during the Covid-19 pandemic which applies online learning. The results of grouping data on online learning media users showed that all levels of education were considered sufficiently ready to implement online learning. As well as cultural differences in Indonesia have not had an impact on the commitment of schools to implement online learning during the Covid-19 disease outbreak.

Keywords: k-means; clusters and learning media

Recommended citation: Rusydiyah, E. F., Virgiannada, H. R., Ridwan, M., Nugroho, B. A., & Rahman, M. R. (2021). Clustering of Learning Media User Data During Covid-19 Pandemic Using K-Means Method Based on Multicultural Culture in Indonesia. *Journal of Innovation in Educational and Cultural Research*, 2(2), 67-76.

INTRODUCTION

Based on Law No.2 of 2003, education is a conscious and planned effort to create a learning process and atmosphere for developing the potential of students in order to have control, intelligence, noble character and formation of spiritual strength. There are 3 education pathways according to Law No.20 of 2003, namely formal education, non-formal education and education. The learning process from various levels of education uses more face-to-face methods, only in certain situations using online learning methods.

The learning method that usually uses face-to-face methods changes to online methods due to the Covid-19 disease outbreak. The Covid-19 outbreak first appeared in Indonesia on February 14, 2020 (detik.com, 2020) and quickly spread so that starting March 2020 there will be comprehensive online learning at all levels of education, both formal, non-formal and informal. The transfer of learning methods that is carried out suddenly and in the long term of course has an impact on several aspects such as the suitability of the type of learning media for educators and students, the learning results obtained by students, the readiness of educators and students in learning online and so on. Likewise, the culture of the Indonesian people is so diverse, where habits and readiness to carry out learning in this network must also vary.

The fundamental problem with the application of online learning is the condition of Indonesia itself which is a very diverse country, meaning that the readiness of one region to another in Indonesia does not necessarily have the same readiness to hold online learning. Different demographic and psychological conditions in Indonesia are important to note ([Rahmawati & Sujono, 2021](#)). Regarding the readiness of different technological infrastructure between regions, the readiness of remote communities who are not necessarily familiar with technology, not to mention the diverse social and economic statuses. [Rasmitadila et al \(2020\)](#) revealed how the readiness of schools varies from one region to another to implement this online learning, even though the government has issued a regulation to provide operational funding for schools related to internet costs, this has not fully guaranteed the implementation of online learning. [Phan & Dang](#)

(2017) said that teachers also need e-readiness, which means teacher attitudes, teacher training, and technical skills to carry out online learning, but once again the readiness of teachers also varies when referring to the conditions of Indonesia's diversity. But on the other hand Indonesia has the slogan *bhinneka Tunggal Ika*, that even though it is different, it is still one. Thus, whether this stretching difference also affects the choice of learning media during online learning during the Covid-19 outbreak is an interesting question to answer. In other words, the importance of doing this research is to find out whether there are differences in the use and selection of online learning media during the Covid-19 outbreak with Indonesia's conditions, which are basically diverse and different.

From the results of the exposure to the problem, a research topic can be raised, namely the Clustering of learning media users online during the Covid-19 pandemic using the k-means method. Many previous studies have used the k-means method, one of which is entitled "Clustering of Student Data Using the K-means Method to Support the Selection of Marketing Strategies" from the results of these studies yielding information on groups of national examination scores that are directly proportional with a GPA (Suprawoto, 2016). Previous research can be used as a reference in resolving the research topics raised. The purpose of this research was to obtain the data on the Clustering of learning media users that provide information related to the readiness of educators and students in online learning, knowing the differences in learning outcomes during the Covid-19 pandemic and before the Covid 19 pandemic of various kinds, of course, the character of students and the style of teaching teachers will also be different, and this has more or less influence on how teachers choose learning media during COVID-19. Likewise, to find out how different the use of learning media is based on the multicultural Indonesian cultural conditions. This grouping is important because the culture in Indonesia is so diverse, of course, the characters of students and the styles of teaching teachers will also differ, and this has more or less influence on how teachers choose learning media during COVID-19. The results are expected to be able to determine the comparison of learning outcomes using online learning and face-to-face learning, then it can be used to determine the appropriate learning method for educators and students. Meanwhile, the purpose of this study was to cluster the use of learning media during the covid-19 outbreak using the k-means method based on multicultural culture in Indonesia.

METHODS

The research flow carried out in this study is in accordance with the framework in Figure 1 below:

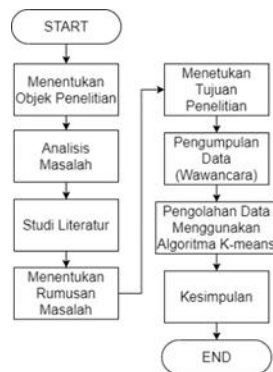


Figure 1. Research Flow

The research started from a problem that will be solved or given a solution. The problem to be examined in this study focused on how the feasibility and quality of learning using learning media during the online Covid-19 pandemic period. The data used to be processed in this study were the average semester scores before the Covid-19 pandemic, the average semester scores during the Covid-19 pandemic, teaching and learning activities conducted by teacher and the kinds of learning media used.

Partisipant

Participants are consisted of 172 respondents at university level, 71 respondents at Senior High School/Vocational High School level and 23 respondents from Junior High School and Elementary School/equivalent.

Instrument

Data mining is an activity to find new information from large data by looking for hidden patterns or rules that were previously unknown (Beynon-Davies, 2004). There are six functions in data mining (Larose, 2005). They are 1) Description function (Description). The purpose of this description function is to better understand the observed data. This description function will lead to the discovery of hidden patterns in the data being studied, 2) Estimation function (Estimation). The estimation function is an estimate for guessing unknown values, but the variables for this estimation function are more numerical rather than categorical. The model is built through a complete record by providing the value of the target variable as a predictive value, 3) Prediction function (Prediction). The prediction function is a process of estimating values based on patterns in a data set, using several data attributes to predict the values of the currently unknown variables, 4) Clustering Function (Clustering). Clustering or grouping is done without a specific target or label into the same object class. In other words, a cluster is a collection of records that have similarities as well as dissimilarities between one another. The goal is to group objects that are similar to each other in certain groups, 5) Grouping function (clustering). This function can be said as supervised induction, with the aim of analyzing historical data that can produce a model that can predict future behavior. This induction model helps to generalize, detect classes from which have not been classified, and 6) Association function (association). This function aims to find a relationship contained in the attribute value of a data set that often appears simultaneously at the same time.

The data mining stage or extracting information is one of the stages in the Knowledge Discovery in Database (KDD). Before entering the data mining stage, data processing must be carried out coherently according to the stages in KDD. It is intended so that the data processed during the data mining stage becomes better and produces accurate information (Sadewo et al., 2017). The following are the stages of data processing in KDD (Fayyad et al., 1996):

1. Data Selection

At the data selection stage, data selection activities are carried out to be used in data mining.

2. Preposition / Cleaning

Cleaning or data cleaning is intended to select quality data. It can be done for data cleaning such as deleting duplicate data, repairing data and checking inconsistent data.

3. Transformation

The transformation stage is the stage for transforming data so that it is ready for use at the data mining stage. Data transformation is carried out by adjusting the techniques or methods to be used in the data mining process.

4. Data Mining

The data mining stage is the core of the KDD stage that is an activity to find information from interesting patterns that exist in data. In this stage, certain related techniques and methods are needed.

5. Interpretation

Patterns or information that has been generated from data mining are translated into simple information that is easy to understand, it can be compared with existing facts or hypotheses.

Data collection

The data collection technique used in this study was interviews, where the object of research was all students at the elementary school to university levels who used learning media for online learning during the Covid-19 pandemic. After data collection was complete, the collected data will enter the data processing stage to obtain useful information.

Data Analysis

Data mining techniques are adjusted to the objectives of a study. In this study, the objective of this research focuses on grouping learning media user data at all levels from elementary school to university. The purpose of this research is in accordance with the clustering function of data mining. Data clustering can use several methods, one of which is k-means. K-means is a grouping method that partitions data into 2 or more groups by grouping similar and similar data (Wardhani, 2016). The goal is to minimize variation within a group but maximize variation between groups (Prasetyo, 2012). The following are the steps of the k-means method in clustering data (Agusta, 2007):

1. Determine the number of clusters k.

Clusters (groups) are determined according to the demands and needs.

2. Determine the center of the cluster (centroid)

The selection of cluster centers is carried out randomly at the initial stage, if it has entered the second iteration or repetition stage, the determination of the centroid must use the applicable formula. The formula for determining the centroid in the second iteration stage:

$$\bar{v}_{ij} = \frac{1}{N_i} \sum_{k=0}^{N_i} X_{kj}$$

Description:

- \bar{v}_{ij} : centroid average of the I-th cluster for the j-th variable
- N_i : the number of members of the I-th cluster
- i, k : index of the cluster
- j : index of the variable
- x_{kj} : the k-th data value of the j-th variable in the cluster

3. Calculate the distance of the data from the center of the centroid.

Calculation of data distances with centroid use euclidean distance theory (Venkateswarlu & Raju, 2013). The equation is:

$$C_{ij} = \sqrt{(X_{1i} - X_{1j})^2 + (X_{2i} - X_{2j})^2 + \dots + (X_{ki} - X_{kj})^2}$$

Description:

- C_{ij} : Distance from I-th data to j cluster center
- X_{ki} : I-th Data to k-th attribute
- X_{kj} : Central point to I-th in k-th attribute

4. Clustering data based on the distance of the closest centroid.

Data that has been calculated for the distance to the center of the centroid will be entered into the cluster that has the smallest centroid distance value.

5. Repeat step 2 by performing the second iteration. The process will be terminated when there is no difference between the old cluster results and the latest cluster results.

RESULT AND DISCUSSION

The results of the interview found 172 respondents at university level, 71 respondents at Senior High School/Vocational School level and 23 respondents at Junior High School and Elementary School equivalent. Interviews were conducted randomly from various regions using google form media. and from all respondents are respondents who come from three different regions, namely Madura, Java and Sunda. The purpose of selecting these three different regions is to find out the differences in the use of multicultural learning media in Indonesia. The following is the preliminary interview data:

1. University Level:
 - a. Name
 - b. University
 - c. Education level
 - d. Location
 - e. Study program
 - f. Media used during the Covid-19 pandemic
 - g. Learning intensity
 - h. GPA for the semester of the Covid-19 pandemic
 - i. GPA of semester before the covid-19 pandemic
 - j. Average age of lecturers
2. High school level and equivalent:
 - a. Name
 - b. School name
 - c. Location (city)

- d. Department
 - e. Media used online during the Covid-19 pandemic
 - f. Learning intensity
 - g. Average semester covid-19 pandemic scores
 - h. Average score of the semester before the Covid-19 pandemic
 - i. Average age of teachers
3. Junior High School and Elementary School equivalent:
- a. Name
 - b. Education level
 - c. School name
 - d. Location (city)
 - e. Online Media used during the Covid-19 pandemic
 - f. Learning intensity
 - g. Average semester covid-19 pandemic scores
 - h. Average score of the semester before the Covid-19 pandemic
 - i. Average age of teachers

Initial data become the source of interview data that will be processed. Data processing follows the stages in the KKD as follows :

Data Selection

Data collection using interviews contained 10 questions for university level, 9 questions for senior high school, junior high school and elementary school levels. Of the 10-9 data that were asked, only 4 were selected to be processed using the k-means method, namely:

- a. Learning media data
- b. Learning intensity data
- c. Data on the semester scores for the covid-19 pandemic
- d. Data on semester scores before the Covid-19 pandemic
- e. Average age of educators

Preposition/Cleaning

The source of data or initial data consisted of 172 respondents at university level, 71 respondents at Senior High School/Vocational High School level and 23 respondents from Junior High School and Elementary School/equivalent. Not all data in the initial data can enter the next stage at the KDD stage, unless all initial data were confirmed to be of quality data. To select quality data, the preposition/cleaning stage performs several data cleaning including duplicate data cleaning, data checking and data correction. The cleaning results obtained data that were ready for the next stage as many as 128 data at the university level, 70 data at the high school/vocational school level and 22 data at the junior and elementary school level and equivalent.

Transformation

The data transformation in this study was adjusted to the terms and conditions in the k-means method. The following are the regulatory requirements for the k-means method (A. Yadav and S. Dhingra 2016):

- a. *Determine the number of clusters*
- b. *Numeric type Data*

In the transformation stage, non-numeric data will be converted into numeric types for processing. There were four data to be processed from the data selection stage are all numeric, except for 1 data, namely learning media data. For that it is necessary to transform the data in this study using the excel function to convert it into numeric type. The following is an example of the results of transforming learning media data:

Table 1. Learning Media Data

| Learning Media | The Number of Learning Media |
|---|------------------------------|
| Whatsapp, Zoom, Google Meet, Google Form, Google Drive, Webex | 6 |
| Whatsapp, Google Classroom | 2 |
| Google Classroom, Zoom, Google Meet, Google Form | 4 |
| Whatsapp, Google Classroom, Zoom, Youtube, Google Form | 5 |
| Whatsapp, Google Classroom, Zoom, Google Meet, Google Form | 5 |
| Whatsapp, Google Classroom, Zoom | 3 |

Data Mining

Data clustering was divided into 2 such as data clustering at university level and data clustering for Senior High School Junior High School and Elementary School level. Due to the difference in the format for the semester average scores the data will be processed separately. The following data that is ready for use at the data mining stage:

Table 2. Processed Data for University Level

| Data | The Number of Learning Media | Learning Intensity | Average Score during Covid-19 | Average Score Before Covid-19 |
|------|------------------------------|--------------------|-------------------------------|-------------------------------|
| 1 | 6 | 4 | 3,31 | 2,78 |
| 2 | 2 | 5 | 3,58 | 3,05 |
| 3 | 4 | 4 | 2,92 | 2,45 |
| ... | | | | |
| 126 | 1 | 5 | 3,50 | 3,86 |
| 127 | 6 | 4 | 3,20 | 3,60 |
| 128 | 3 | 3 | 2,86 | 3,50 |

Table 3. Processed Data for Senior High School, Junior High School, Elementary School/Equal

| Data Ke- | The Number of Learning Media | Learning Intensity | Average Score during Covid-19 | Average Score Before Covid-19 |
|----------|------------------------------|--------------------|-------------------------------|-------------------------------|
| 1 | 1 | 4 | 76,5 | 66,5 |
| 2 | 2 | 3 | 90 | 80 |
| 3 | 1 | 3 | 90 | 80 |
| ... | | | | |
| 126 | 1 | 4 | 76 | 78 |
| 127 | 3 | 5 | 85 | 89 |
| 128 | 1 | 2 | 65 | 75 |

Data mining processing can be done manually or using tools such as excel or special data processing tools. In this research, Rapidminer is used for processing. The use of tools aims to reduce human error in data processing. In addition to processing the processed data, the level of accuracy will be calculated using the Davies Bouldin Index (DBI) technique. DBI is a matrix used to evaluate a clustering algorithm (Davies & Bouldin, 1979). Setting tools for data processing and validation using DBI can be seen in the Figure 2:

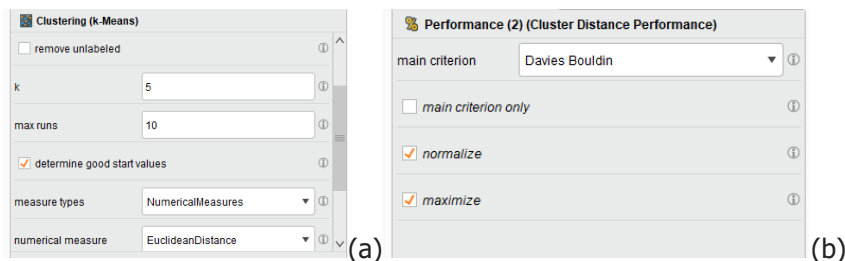


Figure 2. setting clustering (a) and setting DBI (b)

Interpretation

The KDD interpretation stage is useful for conveying information in language that is easy to understand. The following are the results of the final clustering of learning media user data that are presented so that they are easy to understand:

University Level

The results of clustering at the university level with Davies Bouldin validation of 0.242 or the accuracy reached 75.8% after averaging the cluster results data obtained as below:

Table 4. The Results of Final Score at University Level Cluster

| Cluster | \bar{x} The Number of Learning Media | \bar{x} Learning Intensity | \bar{x} Semester GPA during Covid-19 | \bar{x} Semester GPA before Covid-19 | \bar{x} Score Difference | Total |
|---------|--|------------------------------|--|--|----------------------------|-------|
| 0 | 6,36 | 3,36 | 3,54 | 3,52 | 0,03% | 25 |
| 1 | 3,52 | 4,45 | 3,51 | 3,50 | 0,01% | 33 |
| 2 | 5,61 | 4,75 | 3,56 | 3,54 | 0,02% | 28 |
| 3 | 1,42 | 4,31 | 3,55 | 3,53 | 0,02% | 26 |
| 4 | 3,25 | 2,69 | 3,52 | 3,53 | -0,01% | 16 |

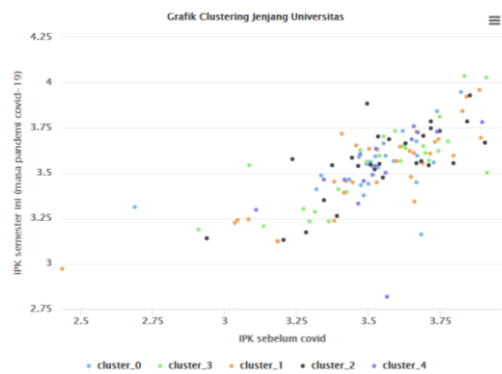


Figure 3. Clustering Graph for University Level

From the results of clustering, other information was extracted from the results of the data to form new information as below:

Table 5. Results of Additional Information Processing Clusters at University Level

| Cluster | Percentage of Teacher Age | | | | Level | | |
|---------|---------------------------|-------------|-------------|-----------|--------|-------|---------|
| | <35 years | 35-45 years | 45-55 years | >55 years | D3 | D4 | S1 |
| 1 | 8,00% | 60,00% | 28,00% | 4,00% | - | - | 100,00% |
| 2 | 9,09% | 72,73% | 9,09% | 9,09% | 9,09% | 3,03% | 87,88% |
| 3 | 10,71% | 71,43% | 14,29% | 3,57% | 10,71% | 3,57% | 85,71% |
| 4 | 3,85% | 73,08% | 23,08% | - | 11,54% | 3,85% | 84,62% |
| 5 | 6,25% | 56,25% | 31,25% | 6,25% | 12,50% | - | 87,50% |

Learning media used by more than 50% of university level respondents were WhatsApp, Google Classroom and Zoom.

Senior High School, Junior High School and Elementary School/Equal

The results of clustering at the senior high school, Junior High School and Elementary School levels are equivalent with the Davies Bouldin validation of 0.230 or the accuracy reaches 77% after averaging the cluster data results obtained as below:

Table 6. Final Average Cluster Score of Senior High School, Junior High School and Elementary School Levels.Equal

| Cluster | \bar{x} The Number of Learning Media | \bar{x} Learning Intensity | \bar{x} Semester GPA during Covid-19 | \bar{x} Semester GPA before Covid-19 | \bar{x} Score Difference | Total |
|---------|--|------------------------------|--|--|----------------------------|-------|
| 0 | 3,18 | 4,21 | 80,92 | 79,36 | 1,93% | 28 |
| 1 | 3,45 | 4,30 | 88,09 | 87,39 | 0,79% | 20 |
| 2 | 3,41 | 4,28 | 84,91 | 82,21 | 3,19% | 29 |
| 3 | 2,00 | 3,83 | 77,27 | 74,09 | 4,11% | 12 |
| 4 | 1,33 | 4,33 | 94,67 | 94,00 | 0,70% | 4 |

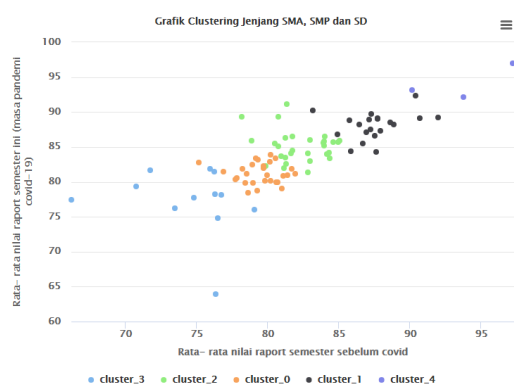


Figure 4. Clustering Graphic of Senior High School, Junior High School and Elementary School/Equal

From the results of clustering, other information was extracted from the results of the data to form new information as below:

Table 7. Results of Additional Information Processing Clusters at Senior High School, Junior High School and Elementary School/Equal

| Cluster | Percentage of Teacher Age | | | | Level | | |
|---------|---------------------------|-------------|-------------|-----------|-------------------|--------------------|--------------------|
| | <35 years | 35-45 years | 45-55 years | >55 years | Elementary School | Junior High School | Senior High School |
| 1 | 7,14% | 85,71% | 7,14% | - | 10,71% | 3,57% | 85,71% |
| 2 | 10,00% | 75,00% | 10,00% | 5,00% | 20,00% | 20,00% | 60,00% |
| 3 | 13,79% | 75,86% | 10,34% | - | 6,90% | 13,79% | 79,31% |
| 4 | 33,33% | 58,33% | 8,33% | - | 16,67% | 25,00% | 58,33% |
| 5 | 33,33% | 33,33% | 33,33% | - | 33,33% | - | 66,67% |

Learning media used by more than 50% of respondents at Senior high school, Junior High School and Elementary School levels were WhatsApp and google classroom while google form was used by 49% of respondents.

Based on the results of the calculation of data on learning media users, information is obtained including: The age of teacher does not determine The Number of Learning Media used.;

1. The Number of Learning Media used slightly affects the learning outcomes of students.
2. The intensity of teaching and learning activities affects the learning outcomes of students.
3. The use of online methods for learning has been shown to improve student learning outcomes.
4. The most popular learning media at all levels of education are WhatsApp and Google Classroom.
5. Judging from the difference in value, high school, Junior High School and Elementary School levels are considered to be more ready to apply online methods in learning compared to university level.
6. Cultural differences, multiculturalism that are unique to Indonesia, do not really influence the use of learning media. This means that different cultures do not have an impact on the commitment of schools in different cultural areas to conduct online learning (using almost the same learning media) during the Covid-19 outbreak.

Information resulting from data analysis is not necessarily true in accordance with the facts in society. As is found in point three, namely the use of online methods for learning, it is proven to be able to improve student learning outcomes. Student learning outcomes are not only fixated on increasing grades but also how the level of understanding of students. Although the results of the data showed that 92.7% of students from

all levels have increased the value of learning outcomes, it does not mean that students understand better the learning material that is conveyed online than face-to-face. This is supported by a study entitled "Identification of Student Barriers to Online Biology Learning During the Covid-19 Pandemic in Jember Regency". The study explains that one of the internal barriers to online biology students is that students were not able to absorb material well with a cognitive presentation of 42.8% (Hariyanti et al., 2020). Learning material can be absorbed well if a comfort is built in teaching and learning activities between educators and students. Students feel pressured by online learning which is only assigned to educators. Evidenced by the research entitled "The Effect of Online College Learning Technology in the Covid-19 Era and Its Impact on Student Mental". There were 47% of Telkom University student respondents agreed that lecturers give more assignments than face-to-face learning. This task is of course a psychological burden on students so that 70% of respondents experience difficulty sleeping due to this (Hifzul & Sumarni, 2020). From the explanation above, a solution can be drawn from the awareness of both parties between educators and students where they must grow enthusiastically in carrying out teaching and learning activities in order to develop effective and comfortable teaching and learning activities during the Covid-19 pandemic.

The previous concern is about Indonesia's diverse conditions which in turn cause readiness to carry out online learning as well, such as the readiness of technology infrastructure, social and economic status between regions. (Rahmawati & Sujono, 2021), or the readiness of the school itself to support the implementation of online learning (Rasmitadila et al., 2020), Likewise, the readiness of the teacher itself is not necessarily able to carry out online learning itself (Phan & Dang, 2017). These readiness are different from these studies. In the context of conditions in Indonesia, it turns out that this study does not have an impact on the enthusiasm of schools in three different regions, namely Java, Sunda and Madura to carry out online learning. In general, schools in three different regions can use WhatsApp or google classroom as online learning media.

Thus it can be seen that Indonesia's diverse, multicultural conditions that are unique to Indonesian culture, demographic and psychographic conditions which are also so different from one school to another, have absolutely no impact on the use of learning media during the COVID-19 outbreak. All schools face the same conditions, namely the demand for online learning, while the currently developing media is only limited to Whatsapp, Zoom, Google Meet, Google Form, Google Drive and Webex which can be used to carry out learning, and this is the reason why the conditions of Indonesia's diverse culture do not really make a difference to the use of online learning media itself.

CONCLUSION

Clustering learning media users at elementary school to university level using the k-means method shows some information including: the age of teachers does not determine the amount of The Number of Learning Media used, The Number of Learning Media used slightly affects student learning outcomes, the intensity of learning activities teaching affects student learning outcomes, the use of online methods for learning is proven to improve student learning outcomes, the most popular learning media at all levels of education are WhatsApp and google classroom, seen from the difference in value, high school level, Junior High School and Elementary Schools and the equivalent are considered to be more prepared to apply online methods in learning than at the university level. Cultural differences, multiculturalism that are unique to Indonesia, apparently do not have much influence on the use of learning media. This means that different cultures do not have an impact on the commitment of schools in different cultural areas to conduct online learning (using almost the same learning media) during the covid-19 outbreak. In general, all levels of education have so far been considered ready for online learning because only 7.3% of all respondents experienced a decline in value. However, the quality of learning is not only determined based on the final score but also seen from the level of understanding of the participants, the comfort of educators and students. For this reason, further research is needed related to differences learning outcomes with understanding level of students, as well as the level of comfort in the continuity of teaching and learning activities.

REFERENCES

- Agusta, Y. (2007). K-Means–Penerapan, Permasalahan dan Metode Terkait. *Jurnal Sistem Dan Informatika*, 3(1), 47–60.
- Beynon-Davies, P. (2004). *Database systems*. Springer.

- Davies, D. L., & Bouldin, D. W. (1979). A cluster separation measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2, 224–227.
- detik.com. (2020, April). Kapan Sebenarnya Corona Pertama Kali Masuk RI? 26 April 2020 07:38 WIB.
- Fayyad, U. M., Piatetsky-Shapiro, G., & Smyth, P. (1996). Knowledge Discovery and Data Mining: Towards a Unifying Framework. *KDD*, 96, 82–88.
- Hariyanti, D., Mun'im, A. H., & Hidayat, N. (2020). Identifikasi Hambatan Mahasiswa dalam Pelaksanaan Pembelajaran Biologi Secara Daring Selama Pandemi Covid-19 di Kabupaten Jember. *ALVEOLI: Jurnal Pendidikan Biologi*, 1(1), 11–21.
- Hifzul, M. M., & Sumarni, N. (2020). Pengaruh Teknologi Pembelajaran Kuliah Online Di Era Covid-19 Dan Dampaknya Terhadap Mental Mahasiswa. *EduTeach: Jurnal Edukasi Dan Teknologi Pembelajaran*, 1(2), 153–165.
- Larose, D. T. (2005). An introduction to data mining. *Traduction et Adaptation de Thierry Vallaud*.
- Phan, T. T. N., & Dang, L. T. T. (2017). Teacher Readiness for Online Teaching: A Critical Review. *Int. J. Open Distance E-Learn (IJODEL)*, 3(1), 1–16.
- Prasetyo, E. (2012). Data mining konsep dan aplikasi menggunakan matlab. *Yogyakarta: Andi*.
- Rahmawati, A., & Sujono, F. K. (2021). Digital Communication through Online Learning in Indonesia: Challenges and Opportunities. *Jurnal ASPIKOM*, 6(1), 61–76.
- Rasmitadila, A. R. R., Rachmadtullah, R., Samsudin, A., Syaodih, E., Nurtanto, M., & Tambunan, A. R. S. (2020). The perceptions of primary school teachers of online learning during the COVID-19 pandemic period : A Case study in Indonesia. *Journal of Ethnic and Cultural Studies*, 7(2), 90–109.
- Sadewo, M. G., Windarto, A. P., & Hartama, D. (2017). Penerapan datamining pada populasi daging ayam ras pedaging di indonesia berdasarkan provinsi menggunakan k-means clustering. *InfoTekJar (Jurnal Nas. Inform. Dan Teknol. Jaringan)*, 2(1), 60–67.
- Suprawoto, T. (2016). Klasifikasi data mahasiswa menggunakan metode k-means untuk menunjang pemilihan strategi pemasaran. *JIKO (Jurnal Informatika Dan Komputer)*, 1(1).
- Venkateswarlu, B., & Raju, P. G. S. V. (2013). Mine Blood Donors Information through Improved K-Means Clustering. *ArXiv Preprint ArXiv:1309.2597*.
- Wardhani, A. K. (2016). K-means algorithm implementation for clustering of patients disease in kaje clinic of pekalongan. *Jurnal Transformatika*, 14(1), 30–37.