The K-Medoids Clustering Method for Learning Applications during the COVID-19 Pandemic

Samudi[1]*, Slamet Widodo[2], Herlambang Brawijaya[3]
[1]STMIK Nusa Mandiri,Indonesia, [2,3]Bina Sarana Informatika University, Indonesia
[1]samudi.net@gmail.com, [2]slamet.smd@bsi.ac.id, [3]herlambang.braw@gmail.com

Abstract: The central government took a policy to carry out social distancing with this social distancing impact on all activities such as impacting the learning process that usually takes place in class turning into online learning which uses several supporting applications in the learning process during the Covid 19 pandemic. Such as using WhatsApp, Moodle, and Zoom for the learning process. Of the three applications commonly used by students and lecturers for the learning process, they can be grouped using the K-Medoids method which will get the preferred and disliked application clusters. Researchers took data from questionnaires made with Google Form which were distributed to 100 students who often did the learning process with these three applications. Can produce groupings of applications that are liked and disliked by students in the learning process. By looking at the level curve in fig 7. There are two clustering in red and blue. If the curve line to the right is increasing, many things are not liked in the application used. Conversely, if the line from the curve towards the right decreases, the less preferred online applications. The results of this study can be seen from the curve image that learning is the most preferred in the learning process during the Covid 19 pandemic using the Moodle and WhatsApp applications, and for the use of zoom is not preferred in the learning process.

Keywords: Data Mining, K-Medoid Algorithm, Clustering, Learning, Covid 19

INTRODUCTION

All countries are currently experiencing a phenomenon that is very dangerous to health, namely the attack of COVID 19 or commonly called the Corona Virus, one of which is the State of Indonesia which consists of several islands that cannot be separated from the attack of COVID 19. COVID 19 is one of the diseases that first appeared in Wuhan City, Hubei Province, China at the end of December 2019 (Etikasari, Puspitasari, Kurniasari, & Perdanasari, 2020). Until now, there has not been found a drug or vaccine that can be used to overcome the spread of the COVID 19 pandemic (Arianto & P, 2020). In Indonesia, the addition of new cases As of August 20, 2020, an increase of 2,266 cases, while the number of patients recovering from COVID 19 has exceeded the number 100,674 cases or 68.3% (Tim Komunikasi Komite Penanganan Corona Virus Disease 2019 (Covid-19) dan Pemulihan Ekonomi Nasional, 2020).

For this reason, the central government continues to urge everyone to implement social distancing to minimize the transmission of COVID 19. Social distancing itself is an action where everyone is required not to be close to one another (Haqien & Rahman, 2020). With social distancing, all activities hurt all fields (Dwitri, Tampubolon, Prayoga, Zer, & Hartaman, 2020), one of which is the learning process at the school to college-level must be done at home. Therefore, the learning process carried out at home requires learning media that is affordable online (Chesilia, Oktaviany, & Dewi, 2012), the challenge is determining the right platform (Yulianto, Cahyani, & Silvianita, 2020) to support the learning process during the COVID pandemic. 19.

The learning process that takes place in tertiary institutions that researchers have done Online, is what researchers often do for the learning process with three applications such as WhatsApp, zoom, and Moodle. With the three applications online, researchers feel.

Based on the constraints in determining applications to support learning during a pandemic, it can be grouped using the K-Medoids algorithm, grouping or what is commonly called clustering based on previous researchers is a technique used for statistical data analysis in various fields, one of which is data mining (Madhulatha TS, 2012)

*Corresponding Author

and aims to share data in clusters (Mustakim, 2017). In making clusters, K-Medoids can use because they have a lot of information on the cluster creation process (Tamura & Miyamoto, 2014).

In this study, the researchers applied the K-Medoids algorithm in grouping the use of applications for learning during the COVID 19 pandemic from various applications that could support the learning process, so researchers wanted to find out information about the media used from various platforms such as zoom, WhatsApp, and Moodle applications which is often used for learning media is it acceptable during the learning process.

LITERATURE REVIEW

Medoid is better than K-Means because in K-Medoids we find k as a representative object to minimize the number of inequalities of data objects, whereas in K-Means we use the total Euclidean distance for data objects (Marlina, Putri, Fernando, & Ramadhan, 2018). Clustering is a technique used for statistical data analysis, in various fields, one of which is data mining (Madhulatha TS, 2012) and aims to divide into clusters (Sabzi A, Farjami Y, 2011) with a similarity level (Irwansyah & Faisal, 2015). Grouping using the K-Medoids method is widely proposed by several researchers and can be applied such as clustering the spread of COVID 19 in Indonesia (Sindi et al., 2020), which results in grouping the spread of COVID 19 for the community. And there is another researcher for the K-Medoids algorithm for disease grouping in Pekanbaru Riau (Juninda, Mustakim, & Andri, 2019) which produces 4 clusters. The steps for the K-Medoids algorithm are as follows (Pramesti, Lahan, Tanzil Furqon, & Dewi, 2017)

1. Initialize k cluster centers (number of clusters)
2. Allocate each data (object) to the nearest cluster using the Euclidian Distance measurement equation with the equation:

$$d(x,y) = ||x - y||$$
$$= \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \quad ; 1,2,3,\dots n$$

3. Randomly select objects in each cluster as candidates for the new medoid.
4. Calculate the distance of each object in each cluster with the new medoid candidate.
5. Calculate the total deviation (S) by calculating the value of the new total distance - the old total distance. If S < 0, then swap objects with cluster data to form a set of k new objects as medoids.
6. Repeat steps 3 to 5 until there is no change in medoids so that you get the cluster and the members of each cluster.

METHOD

The method used in this study is from identifying problems that exist at this time, then by collecting data, after the data is collected, the initial data processing is then divided into several categories, then the application of the K-Medoids algorithm by entering data into the rapid miner application the last is the evaluation of the results of the data that has been inputted to the rapid miner, then a diagram will appear that can provide information related to data grouping using K-Medoids. You can also see the sequence based on the previous researcher diagram in Fig 1 below (Juninda et al., 2019).
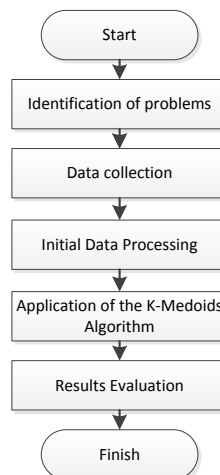
Source : (Juninda et al., 2019)

Fig. 1. Research methods

RESULT

Identification Problem

Researchers took three applications that would be grouped based on the learning process that was liked and disliked in the learning process during the COVID 19 pandemic. The three applications that the researchers used in this study were Moodle, Zoom, and Whatsapp.

Data Collection

In collecting data the researchers used an online questionnaire using a google form which was addressed to 100 students who used to carry out the teaching and learning process online with WhatsApp, Zoom, and Moodle.

Initial Data Processing

The data processing that the researchers did was divided into two categories of likes and dislikes from the three Zoom, Moodle, and Whatapp applications used in the learning process during the COVID 19 pandemic as follows:

Table 1.
Application Users

Application	Like	Dislike
Moodle	11	16
Zoom	36	68
WhatsApp	53	16

Source : Research Result (2020)

Application of the K-Medoids Algorithm

The data that has been obtained, the researcher enters the Rapid Miner software by selecting New Process as follows.
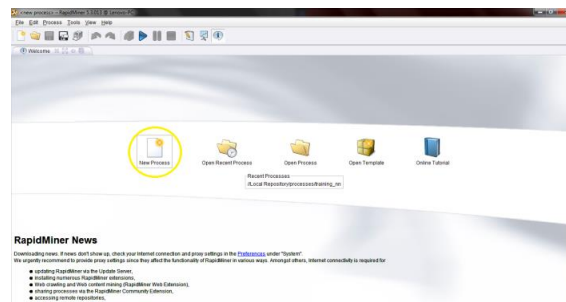

Fig. 2. Rapid Miner

After that, a worksheet consisting of the left middle, and right columns will appear. To enter data from Excel, select the lower left part where there is a sign that the import data into an existing repository is as follows:
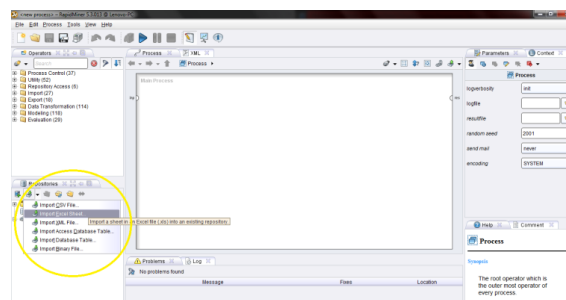

Fig. 3. Entry Data

After entering data from Excel then click next so that it will display the type attribute that will be used in the grouping process in Rapid Miner. Several things must be changed in determining attributes, including changing the application attribute to nominal and changing it to id as follows:
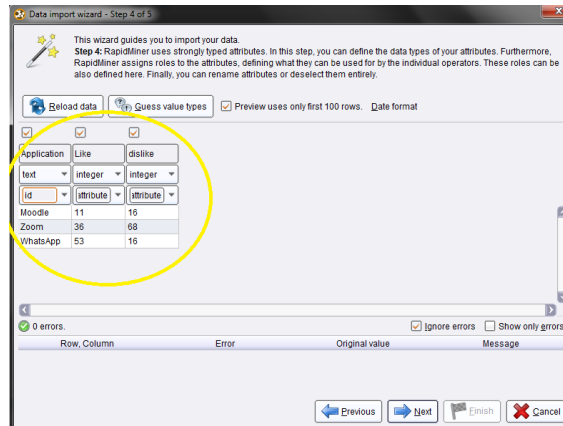
Fig 4. Attribute Selection

Click next, then save it with the name Learning Data With Applications in the COVID Period. After saving, the next step is to enter the K-Medoids into the worksheet on the top left in the modeling, clustering, and segmentation section. Then move the worksheet, then the Excel data set is also moved to the worksheet, then draw a line so that it forms as below:
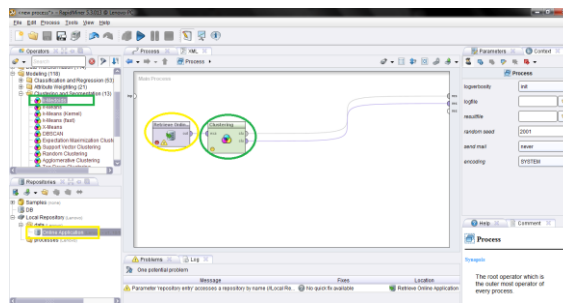

Fig. 5. Inserting K-Medoids

After that, the next step is the kiln run which has the blue triangle symbol above, a data view, a metadata view, a plot view, advanced charts, and annotations will appear. So that the formation of cluster 0 and cluster 1 will appear, for cluster 0 consists of moodle and WhatsApp, for cluster 1 consists of zoom. The display is as follows:

Fig 6. Cluster formation

To be clearer in grouping likes or dislikes in the learning process using K-Medoids, it can be seen in Figure 7 to show it, you can choose the Cluster Model (Clustering) so that it will be clear that the ascending and descending curves will be seen, such as cluster 0 which consists of Moodle and WhatsApp with a blue stripe showing that the more to the right, the smaller is not liked, unlike cluster 1 which consists of zoom, the higher is not liked.
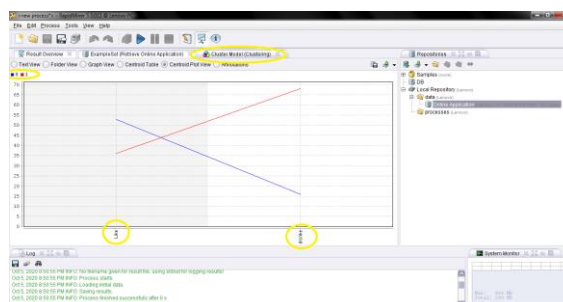
Fig 7. Cluster Model (Clustering)

5. Evaluation of Results

In the evaluation and the results in this fifth stage, it can be seen the results of the clustering model that is formed in the rapid miner which has been described above that grouping with K-Medoids can form 2 clusters.

DISCUSSION

By clustering online applications such as WhatsApp, zoom, and moodle which are often used in the learning process for 100 students, it can produce a grouping of applications that students like and dislike in the learning process. By looking at the level curve in fig 7. There are 2 clusterings in red and blue. If the curve line to the right is increasing, many things are not liked in the application used. Conversely, if the line from the curve tow

CONCLUSION

Based on the results of the research, the authors can conclude that during the COVID-19 pandemic that spreads throughout the world, wherein the learning process that is usually delivered directly meeting one another must be replaced by a social distancing system The lessons that are delivered must be presented online with a variety of applications used, such as zoom, WhatsApp or moodle. In this study, the results show that the clustering or grouping that has been carried out on the rapid miner results that learning is the most preferred in the learning process during the current COVID 19 pandemic using the Moodle and WhatsApp applications, for the use of zoom is not preferred in the learning process. For future research, we can examine the factors that influence the application in learning during a pandemic which are the most liked and disliked.

REFERENCES

Arianto, F. S. D., & P, N. (2020). PREDIKSI KASUS COVID-19 DI INDONESIA MENGGUNAKAN METODE BACKPROPAGATION DAN FUZZY TSUKAMOTO. *(Jurnal Teknologi Informasi*, *4*(1), 120–127. Retrieved from http://jurnal.una.ac.id/index.php/jurti/article/view/1265%0A

Chesilia, S., Oktaviany, D., & Dewi, D. (2012). Sistem Informasi Manajemen Penjualan dan Persediaan Barang Berbasis Web pada CV. Matrik Jaya. *Jurnal Sistem Informasi STMIK GI MDP*, (x), 1–10. Retrieved from http://eprints.mdp.ac.id/1738/1/Jurnal-2012240058.pdf

Dwitri, N., Tampubolon, J. A., Prayoga, S., Zer, P. P. P. A. N. W. F. I. R. H., & Hartaman, D. (2020). PENERAPAN ALGORITMA K-MEANS DALAM MENENTUKAN TINGKAT PENYEBARAN PANDEMI COVID-19 DI INDONESIA. *Jurnal Teknologi Informasi*, *4*(1), 128–132. Retrieved from http://jurnal.una.ac.id/index.php/jurti/article/download/1266/1104%0A

Etikasari, B., Puspitasari, T. D., Kurniasari, A. A., & Perdanasari, L. (2020). Sistem Informasi Deteksi Dini Covid-19. *Jurnal Teknik Elektro Dan KOmputer*, *9*(2). Retrieved from https://ejournal.unsrat.ac.id/index.php/elekdankom/article/view/28278%0A

Haqien, D., & Rahman, A. A. (2020). PEMANFAATAN ZOOM MEETING UNTUK PROSES PEMBELAJARAN PADA MASA PANDEMI COVID-19. *Susuan Artikel Pendidikan*, *5*(1). Retrieved from https://journal.lppmunindra.ac.id/index.php/SAP/article/view/6511%0A

Irwansyah, E., & Faisal, M. (2015). *Advanced Clustering: Teori dan Aplikasi*. Deepublish.

Juninda, T., Mustakim, & Andri, E. (2019). Penerapan Algoritma K-Medoids untuk Pengelompokan Penyakit di Pekanbaru Riau. *Seminar Nasional Teknologi Informasi, Komunikasi Dan Industri (SNTIKI) 11*, (November), 42–49. Retrieved from http://ejournal.uin-suska.ac.id/index.php/SNTIKI/article/view/8002/4465

Madhulatha TS. (2012). An Overview on Clustering Methods. *IOSR Journal of Engineer*, *2*(4), 719–725.

Marlina, D., Putri, N. F., Fernando, A., & Ramadhan, A. (2018). Implementasi Algoritma K-Medoids dan K-Means untuk Pengelompokkan Wilayah Sebaran Cacat pada Anak, *4*(2), 64–71.

Mustakim. (2017). Effectiveness of K-means clustering to distribute training data and testing data on K-nearest neighbor classification. *Journal of Theoretical and Applied Information Technology*, *95*(21), 5693–5700.

Pramesti, D. F., Lahan, Tanzil Furqon, M., & Dewi, C. (2017). Implementasi Metode K-Medoids Clustering Untuk Pengelompokan Data. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, *1*(9), 723–732. https://doi.org/10.1109/EUMC.2008.4751704

Sabzi A, Farjami Y, Z. M. (2011). An Improved Fuzzy K-Medoids Clustering Algorithm with Optimized Number of Clusters. *IEEE International Conference on Hybrid Intelligent System*, 206–210.

Sindi, S., Ratnasari, W., Ningse, O., Sihombing, I. A., Zer, F. I. R. H., & Hartama, D. (2020). Analisis algoritma k-medoids clustering dalam pengelompokan penyebaran covid-19 di indonesia, *4*(1), 166–173.

Tamura, Y., & Miyamoto, S. (2014). Two-Stage Clustering Using One-Pass K-Medoids and Medoid-Based Agglomerative Hierarchical Algorithms. In *IEEE International Conference on Soft Computing and Intelligent Systems and 15th International Symposium on Advanced Intelligent System* (pp. 484–488). Kitakyushu.

Tim Komunikasi Komite Penanganan Corona Virus Disease 2019 (Covid-19) dan Pemulihan Ekonomi Nasional.

(2020). Kesembuhan Covid-19 di Indonesia Tembus 100.000. Retrieved from https://covid19.go.id/p/berita/kesembuhan-covid-19-di-indonesia-tembus-100000

Yulianto, E., Cahyani, P. D., & Silvianita, S. (2020). Perbandingan Kehadiran Sosial dalam Pembelajaran Daring Menggunakan Whatsapp group dan Webinar Zoom Berdasarkan Sudut Pandang Pembelajar Pada Masa Pandemic COVID-19. *Jurnal Riset Teknologi Dan Inovasi Pendidikan*, *3*(2), 331–341. Retrieved from https://journal-litbang-rekarta.co.id/index.php/jartika/article/view/277%0A