Jurnal Teknologi Pendidikan, August 2022, 24 (2), 246-257
DOI: http://dx.doi.org/10.21009/JTP2001.6
p-ISSN: 1411-2744 e-ISSN: 2620-3081
Accredited by Directorate General of Strengthening for Research and Development
Available online at http://journal.unj.ac.id/unj/index.php/jtp

Clustering Of Student Learning Styles in the Industry 4.0 Using KMeans Algorithm

Medika Risnasari[1], Nabila Aulia[2], Laili Cahyani[3]
Universitas Trunojoyo Madura, Jawa Timur, Indonesia

Received: : July 4, 2022
Revised: : August 22, 2022
Accepted: : August 23, 2022

Abstract

Clustering is a technique for grouping homogeneous data so that the points in each cluster are as similar as possible according to convenience measures such as Euclidean-based distance or correlation-based distance. In the industrial era 4.0, learning media, the environment, the way teachers teach will affect student learning styles. From research on learning styles, many researchers agree on the importance of identifying learning styles to accelerate their learning performance. The purpose of this study is to clustering student learning styles in the industrial era 4.0 using the Kmeans algorithm. Knowing the clustering of student learning styles in the classroom can make it easier for teachers to determine the type of ICT-based learning media in the industrial era 4.0. The accuracy of learning media is a factor in improving students' understanding. The research method used is the development of the waterfall model. The number of research subjects was 108 students. the results of the research on the number of clusters (K), namely 6, obtained cluster 1 as many as 27 students, cluster 2 as many as 24 students, cluster 3 as many as 21 students, cluster 4 as many as 17 students, cluster 5 as many as 11 students and cluster 6 as many as 8 students. The performance of the grouping results based on the silhouette coefficient is 0.302, which means the grouping structure is weak. In cluster 1, the highest number has auditory elements, followed by kinesthetic and visual elements. The development of ICT-based media is one of the factors of student learning styles in the industrial era 4.0

Keywords: clustering, learning style, Kmeans, silhoutte coeficient

(*) Corresponding Author: medika.risnasari@trunojoyo.ac.id, 081357302354

How to Cite: Risnasari, M., Nabila Aulia, & Laili Cahyani. (2022). Clustering Of Student Learning Styles in the industri 4.0 Using KMeans Algorithm. JTP - Jurnal Teknologi Pendidikan, 24(2), 246-257. https://doi.org/10.21009/jtp.v24i2.28029

INTRODUCTION

A literature review on learning styles states that learning styles are very contextual. Learning styles are also influenced by external factors such as learning facilities, teaching models, technological sophistication (Asbari, 2020). In learning style theory, factors that influence learning styles include physical, emotional, sociology and the environment (Dunn,1990). various fields, including education. In the industrial era 4.0, Information and communication technologies (ICT) have been used in various fields including education. Student learning media in the industrial era 4.0 have many in the form of ICT media such as media with Augmented reality technology, virtual reality, educational games, electronic simulations and others. In addition, the distance learning period that students have experienced during the Covid-19 pandemic has made students use ICT-based learning media. The learning model used by the teacher also adapts to distance

learning. The use of learning media based on Learning Management System such as Moodle has a relationship with learning styles and students' cognitive abilities.

In general, there are two main categories of how a person learns, namely how the easiest way to absorb information (modalities) and how to organize and process that information (brain dominance). Learning style is a combination of how to absorb information and organize and process that information (Deporter, 2006). In identifying student learning styles, the relationship between student and teacher learning styles (Awla, 2014).

Learning style as a combination of how individuals absorb information in an easy way (modality) and how individuals manage and organize that information (Asbari, 2020). In the book "Belajar dan Pembelajaran", there are three main types of learning modalities, namely visual, auditory, and kinesthetic. But it usually happens that a person has a combined learning modality. Some combinations of learning modalities that usually occur consist of a combination of two learning modalities, resulting in variations in audio-visual, audio-kinesthetic, and visual-kinesthetic learning modalities (Suyono, 2014).

In the learning process, the ability of students to understand the subject matter varies, some are fast and some are slow. Of the 108 students of class XI Software Engineering, 65 students were able to absorb the material well, 32 students were able to absorb the material sufficiently, and 11 students were less able to absorb the material. Some of the factors that influence this include student interest in certain fields, the facilities owned by students are not the same, and each student's learning effort is different. Of the three things, the most dominant factor is interest in the subject matter and student learning efforts. From the student factor as many as 25 respondents, it was found that students pay attention to the teacher and take notes on what the teacher has said, prefer to learn by direct practice, do not like to listen to explanations from the teacher and only take the core material, ask directly to the teacher, look for material by reading on the internet. textbooks, and only listen to the teacher's explanation without re-recording.

The data mining approach in the grouping process in the field of education has been carried out by many researchers. Correlation between variables that affect learning styles can also be done through technology-based data analysis. The correlation between learning styles and student characteristics can be done by the system in data mining in a machine learning (Han, 2011; Glazunova, 2020; Rasheed, 2021).

In information technology, a method that is able to clustering students into a learning style can be done by the clustering process. In this research, the KMeans algorithm is used in the clustering process of student learning styles. Machine learning is a branch of science from artificial intelligence, namely a system that can adapt human abilities to learn. K-Means is a simple clustering method that has the ability to process large amounts of data and partition the dataset into several clusters. The K-Means algorithm is the simplest algorithms and most common clustering method. K-means has the ability to group large amounts of data which have efficient time for computation and fast relatively (Khotimah, 2016; Indriyani, 2019; Purba, 2018; Sinaga, 2020; Tamba, 2019). However, K-Means has a disadvantage depending on the initial cluster center determination. K-Means cluster test results in the form of solutions that are local optimal. From the trial process is

expected to have similarities or closeness between data so that can be grouped into several clusters, where among cluster members have a high level of similarity (Bain, 2015). The K-Means algorithm is also versatile, which is easy to modify and simple in the distance calculation function with Euclidean distance, and depends on iteration termination criteria. The advantages of K-Means are that the algorithm is quite easy to implement and operate, relatively fast, and efficient. On the other hand, K-Means also has shortcomings, namely in determining the number of clusters and determining the cluster center (Nainggolan,2019; Hartanti, 2020; Hidayati, 2021).

Clustering is the process of grouping a set of data into a certain number of clusters or groups where the data in the cluster has high similarities but has high dissimilarities with data in other clusters. Similarities and dissimilarities are calculated based on attribute values and involve a measure of the distance between data (Han, 2011; Suyanto, 2019).

The basic stages of the K-Means algorithm work in four steps as follows(priyatman, 2019): a) Choose a number of k objects in the data set randomly as the initial centroid; b) Assigning every object that is not a centroid to the nearest cluster based on a certain distance measure; c) Updates the centroid based on the average of the objects contained in each cluster; d) Repeat the second and third steps (iterations) until all centroids are stable and convergent

In previous studies, the K Means algorithm was widely used in the clustering process in various fields. In this research, the K-Means algorithm is focused on grouping student learning styles in the industrial era 4.0. In the industrial era 4.0, students have utilized information technology facilities as a learning resource. And these results will be part of one of the factors of student learning styles. Knowing the grouping of student learning styles in the classroom can make it easier for teachers to determine the type of ICT-based learning media in the industrial era 4.0. The accuracy of learning media is one of the factors in improving students' understanding.

METHODS

The research subjects were 108 students of XI RPL SMK 4 Bojonegoro. This research will produce the percentage of data collection for clustering student learning styles with the Kmeans algorithm starting with filling out questionnaires related to student learning styles. This student learning style questionnaire was first validated by a psychologist in their field.

This research method uses the waterfall model. The waterfall development model has several stages that are systematically arranged in developing a system, these stages are requirements analysis and definition (needs analysis), system and software design (design), implementation and unit testing (implementation), integration and system testing (integration and testing), and operation and maintenance (operations and maintenance). The stages of the waterfall model could be seen in figure 1.

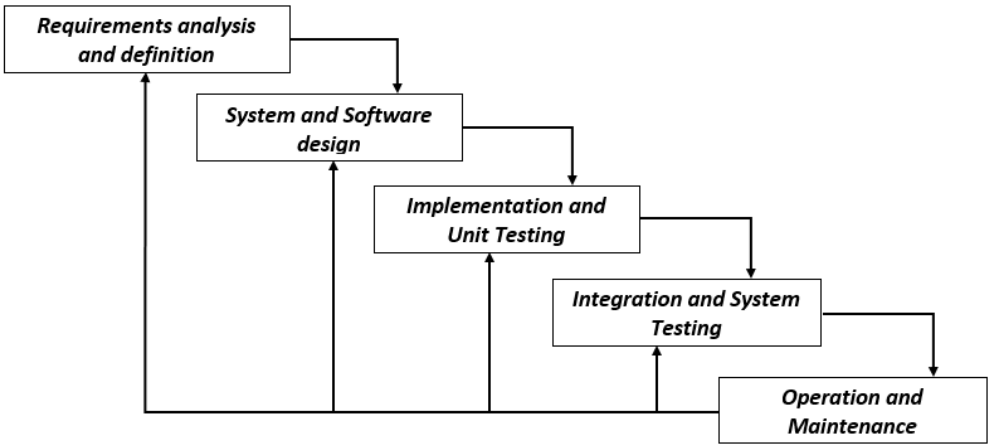

Figure 1. the stages of waterfall model for method research (Sasmito, 2017)

The explanation of the waterfall model are:

1. The requirements analysis and definition stage are the stage where system services, constraints, and objectives are determined by the results of consultations with system users which are then explained in detail and functioned as system specifications. This stage is also known as the stage for analyzing and defining requirements that will be used in development. Steps that can be taken at this stage are consulting with users in order to obtain information related to the needs desired by users. Data collection techniques for needs analysis using interview and questionnaire techniques.
2. System and software design is a system design stage that determines the hardware and software system requirements by forming an overall system architecture. The designs that will be discussed at this stage are UML (Unified Modeling Language) design, database design, algorithm design, and interface design.
3. Implementation and unit testing is a software design realized as a series of programs or program units. This stage implements the design that has been made into a programming language.
4. Integration and testing system, a clustering algorithm scenario trial was conducted and a trial subject was tested. The test scenario of the clustering algorithm is that the program will be tested as an application of the clustering algorithm with K-Means on clustering student learning styles. The test subjects in this study were system experts, web design experts, and users of SMKN 4 Bojonegoro students. Software trials can use WAPT software version 5.0.
5. Operation, which is a system that has been tested and revised, will be submitted to the user so that it can be used properly as needed. Maintenance includes repairs if there are problems with the system.

While the stages of the cluster data processing process can be seen in figure 2.

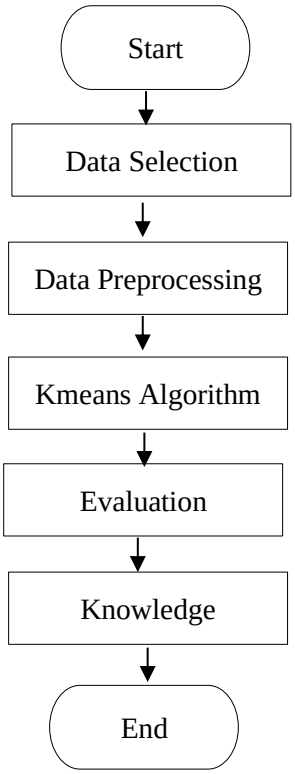

Figure 2. The stages of the cluster data processing process

The explanation of the stages thats data processing process are:

1. Data Selection, The data used in this study are student data and data from student learning style questionnaires. The student data that will be tested in this study amounted to 108 data, namely data from students of class XI Software Engineering at SMK Negeri 4 Bojonegoro.
2. Data Preprocessing, In this research, the preprocessing used is data cleaning and data transformation. Data cleaning is carried out so that there is no missing value in the data. Transformation data is used to change the learning style questionnaire answer data in the form of nominal data into numerical data, this is done so that the data can be processed using the K-Means algorithm which can only process numerical data.
3. K-Means Algorithm, Data mining which in this step includes the choice of a suitable data mining strategy such as classification, clustering, regression, association, and the choice of the data mining algorithm itself. In this study, the data mining strategy chosen was clustering with the K-Means algorithm.
   The steps in the K-Means algorithm are:
   a. Determine the number of K clusters. At this stage the number of K clusters has been determined based on the type of learning style, namely visual, auditory, kinesthetic, visual auditory, kinesthetic auditory, and visual kinesthetic.

b. Determine the centroid or cluster center. In this study, the initial centroid was determined by randomly selecting one object representing all objects in the cluster (Purba, 2018).
c. Calculate the distance of each data with the centroid by calculating the distance Euclidean distance. Euclidean distance can be calculated by the formula:

$$d(p,q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + ... + (p_n - q_n)^2} \quad (1)$$

d. Grouping of data based on the closest Euclidean. The results of the Euclidean calculations can be made in a table and the cluster that has the minimum distance from the centroid to the data will be a cluster of the data.
e. Calculate the new cluster centroid that has been formed in the previous iteration. The new centroid is the average of all objects contained in each cluster that has been formed.
f. If the data or centroid changes then repeat back to the third step. The iteration stops if the data is fixed or the centroid has been stable, in other words, there is no change in value.

4. Evaluation

Evaluation can be done by measuring the quality of clustering. In this study the results of clustering are calculated using the silhouette coefficient with the formula:

$$s(o) = \frac{b(o) - a(o)}{\{a(o), b(o)\}} \quad (2)$$

5. Knowledge

Knowledge representation and visualization serve as the delivery of knowledge to users. In this study, knowledge or clustering results will be conveyed to users through a diagram and table

RESULTS & DISCUSSION

RESULT

This learning style cluster research is web-based. The results of clustering in this system are visualized in the form of diagrams and tables. The results of grouping the learning styles of 108 students are 11 students visual learning style, 27 students auditory, 17 students kinesthetic, visual auditory a total of 24 students, visual kinesthetic a total of 8 students, and auditory kinesthetic as many as 21 students. In Table 1 is a display of centroid value for 11^{th} iteration of the cluster process that occurs with the Kmeans algorithm and Figure 3 is result of clustering learning style.

Table 1. Centorid Centroid Value For 11th Iteration

Cluster	Centroid value for 11th iteration		
	Visual	**Auditory**	**Kinestetic**
K1	40,3636363	34, 3636363	31,1818
K2	36,037037	36,88888	36,5925
K3	38,82352	39,88823	43,4706
K4	45,51666	39,5	38,5
K5	41,75	32,5	45,125
K6	43,33333	47,14285	45,3333

In table 1 we know the value of the centroid in the 11th iteration. in the Kmeans algorithm, this centroid value performs an iterative process and will stop when it reaches a stable value.

Table 2. Centorid value and Results of the cluster learning style

Number	NIS	Name	Result Learning Style			Learning Stye
			Visual	**Auditory**	**Kinestetic**	
1	4696	ABDUL HARIS PRATAMA ARDIYANTO	56	35	35	K2
2	4697	ACHMAD FAIZAL MILADI	36	40	35	K1
3	4698	ADITIYA GUSTINIAWAN	36	45	34	K1
4	4699	AGINA MEY RINTAN MAHRANY	44	44	30	K2
5	4700	ALFIAMITA	44	30	44	K6

In table 2, the results of the development of the web-based student learning style cluster were successfully developed by knowing the learning styles of each student.

The final result of the learning style cluster process with the Kmeans algorithm can be seen in table 3. Visualisation diagram in figure 3.

Table 3. Percentage of each cluster

Cluster	Sum	Percentage (%)
K1	27	25
K2	24	22,22
K3	21	19,44
K4	17	15,74
K5	11	10,19
K6	8	7,4

In the figure 3 shows of presentage learning style result. From table 3 and figure 3, we know that of the 108 students as test subjects from SMKN 4 Bojonegoro that as many as 25 students clustered in cluster 1 (K1). So that we can also know that as many as 25 students have similar characteristics from the results of the learning style questionnaire.

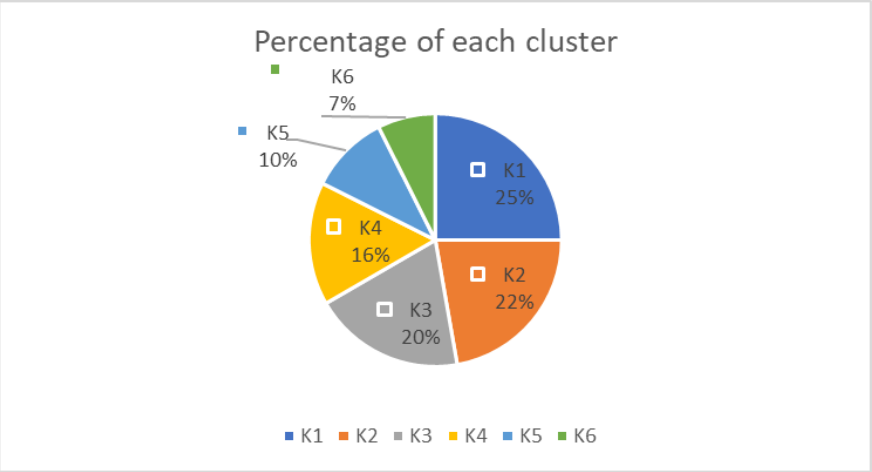

Figure 3. Diagram of presentage learning style result

While in cluster k6 there are 8 students who cluster themselves as having similar characteristics based on the results of the learning style questionnaire. So this cluster is a process where students will cluster learning styles based on the similarity of the characteristics possessed by these students. The statistical value as part of the cluster characteristics can be seen in table 4.

Table 4. Statistical value as part of the cluster characteristics

Cluster	Parameter	Visual	Auditory	Kinestetic
K1	Average	36,04	36,89	36,59
	Max	40	45	40
	Min	30	31	34
K2	Average	45,42	39,5	38,5
	Max	56	44	46
	Min	41	35	28
K3	Average	43,33	47,14	45,33
	Max	52	52	53
	Min	38	44	38
K4	Average	38,82	39,88	43,47
	Max	42	43	48
	Min	37	37	40
K5	Average	40,36	34,36	31,18
	Max	46	41	36
	Min	29	26	24
K6	Average	41,75	32,5	45,13
	Max	45	35	48
	Min	33	30	40

From the results of the learning style cluster with 6 clusters, we can see that the determination of the cluster with the Kmeans algorithm obtained clusters with the characteristics as in table 2. Random generation of initial centroid values in the Kmeans algorithm and the calculation of the Euclidean distance then the data will group itself according to the similarity of the characteristics of the data.

In cluster 1, the highest number has auditory elements, followed by kinesthetic and visual elements.leraning style in era industri 4.0 with based on ICT made students like listening to learning. But a kinestetic and visual also a part of style learning. The quality test of the Kmeans algorithm in the learning style cluster is shown using the coefficient silhoute. The results of the silhoute coefficients in table 5.

Table 5. Silhouette Coefficient result

Cluster	Silhouette Coefficient
K1	0,387
K2	0,205
K3	0,291
K4	0,466
K5	0,043
K6	0,341
Average	0,302

The overall clustering result has a value of 0.302 which means that the clustering quality has a weak structure. This indicates that there is some data that is in the middle of two or more clusters.

DISCUSSION

The Kmeans algorithm is able to carry out the clustering process from the results of the student learning style questionnaire. However, the performance obtained with the silhoutee coefficient is weak. This indicates that the clustering quality of the K-Means algorithm has a weak structure where there is data in the middle of two or more clusters. To improve the clustering performance, optimization using the elbow method (Hartanti, 2020). The results of this study in helping teachers determine learning media based on student learning styles are sufficient to provide information for teachers. In these 108 students have different learning styles with a blend of visual, auditory, kinesthetic. In general, students still use visuals in their learning style. By identifying the results of the questionnaire from the learning styles of students, the Kmeans algorithm is able to determine the type of student learning styles. Meanwhile, the performance of the KMeans algorithm in clustering student learning styles is measured by the silhoute coefficient. We know that a ICT with many smartphone, fast of internet, many media kind of media learning like augmented reality, virtual reality, simulation electronic become a student must a adapted learning. So style of learning is affected with a environment like a era industri 4.0 with technology.

The K-Means algorithm can identify students' learning styles and have a significant influence on student learning outcomes (Widayanti, 2013). Each

individual has their own abilities and skills. Individual abilities and skills can develop well if they are trained continuously. The ability to understand and absorb information for each student is different, some have fast, medium and slow abilities. Therefore, by knowing the learning styles that students have, they can provide learning activities with various methods, strategies and media in the learning process according to the students' learning styles. Identification of student learning styles can be done in various ways, one of which is by using a simple K means algorithm to be able to map student learning styles so that the selection of methods, strategies, media and teaching materials as well as learning resources according to student needs can be done. achieve maximum learning outcomes.

CONCLUSION

Kmeans algorithm is used in the process of clustering learning styles on 108 students at SMKN 4 Bojonegoro. This learning style clustering is used to help teachers understand students' learning styles so that teachers are able to determine the right learning media for students. Students in the industrial era 4.0 with the development of information technology are very user friendly in using the internet as a learning resource.

The results of clustering student learning styles on 108 students into six learning style clusters obtained are 11 students in K5, there are 27 students in K1, there are 17 students in K4, there are 24 students in K2, there are 21 students in K3, and there are 8 students in K6. While the results of the clustering quality of the K-Means algorithm is 0.302, which means that the clustering quality has a weak structure.

This could be due to the close resemblance of traits between individuals. From these results it can be concluded that the Kmeans algorithm is able to cluster learning styles with a weak structure. So, to get a better clustering of learning styles, you can use other cluster algorithms and optimize the number of clusters such as the elbow method.

CONFLICT OF INTEREST

The research we do is for the education sector. By knowing the types of learning styles in students, it becomes a parameter for teachers to teach better.

ACKNOWLEDGEMENT

We would like to thank all those who have assisted in the completion of this research, including SMK 4 Bojonegoro who have been willing to serve as a research site.

REFERENCES

Asbari, M., Tukiran, M., PURWANTO, A., Santoso, P. B., Wijayanti, L. M., & Hyun, C. C. (2020). Masih Relevankah Pengukuran Gaya Belajar Pada Pembelajaran Online?(Sebuah Kajian Literatur Sistematis). *Journal of Industrial Engineering & Management Research*, *1*(3), 267-275.

Awla, H. A. (2014). Learning styles and their relation to teaching styles. *International journal of language and linguistics*, *2*(3), 241-245.

Bain, K. K. (2015). Customer segmentation of SMEs using K-Means clustering method and modeling LRFM. In *International Conference on Vocational Education and Electrical Engineering, Universitas Negeri Surabaya*.

DePorter, B. (2006). Quantum learning: Membiasakan belajar nyaman dan menyenangkan.

Dunn, R. (1990). Rita Dunn answers questions on learning styles. *Educational leadership*, *48*(2), 15-19.

Glazunova, O., Morze, N., Golub, B., Burov, O. Y., Voloshyna, T., & Parhomenko, O. (2020). Learning style identification system: Design and data analysis.

Hartanti, N. T. (2020). Metode Elbow dan K-Means Guna Mengukur Kesiapan Siswa SMK Dalam Ujian Nasional. *J. Nas. Teknol. dan Sist. Inf*, *6*(2), 82-89.

Han, J., Pei, J., & Kamber, M. (2011). *Data mining: concepts and techniques*. Elsevier. Suyanto, D. (2017). Data Mining untuk klasifikasi dan klasterisasi data. *Bandung: Informatika Bandung*.

Hartanti, N. T. (2020). Metode Elbow dan K-Means Guna Mengukur Kesiapan Siswa SMK Dalam Ujian Nasional. *J. Nas. Teknol. dan Sist. Inf*, *6*(2), 82-89.

Hidayati, R., Zubair, A., Pratama, A. H., & Indana, L. (2021). Analisis Silhouette Coefficient pada 6 Perhitungan Jarak K-Means Clustering. *Techno. Com*, *20*(2), 186-197.

Indriyani, F. dan Irfiani, E. 2019. *Clustering Data Penjualan pada Toko Perlengkapan Outdoor menggunakan Metode K-Means*. JUITA : Jurnal Informatika. Vol.7, No.2, November 2019, pp.109-113.

Khotimah, B. K., Irhamni, F. I. R. L. I., & Sundarwati, T. R. I. (2016). A Genetic algorithm for optimized initial centers K-means clustering in SMEs. *Journal of Theoretical and Applied Information Technology*, *90*(1), 23.

Lwande, C., Muchemi, L., & Oboko, R. (2021). Identifying learning styles and cognitive traits in a learning management system. *Heliyon*, *7*(8), e07701.

Nainggolan, R., Perangin-angin, R., Simarmata, E., & Tarigan, A. F. (2019, November). Improved the performance of the K-means cluster using the sum of squared error (SSE) optimized by using the Elbow method. In *Journal of Physics: Conference Series* (Vol. 1361, No. 1, p. 012015). IOP Publishing.

Priyatman, H., Sajid, F., & Haldivany, D. (2019). Klasterisasi Menggunakan Algoritma K-Means Clustering untuk Memprediksi Waktu Kelulusan Mahasiswa. *JEPIN (Jurnal Edukasi dan Penelitian Informatika)*, *5*(1), 62-66.

Purba, Windania., dkk. 2018. *The Effcet of Mining Data K-Means Clustering Toward Students Profile Model Drop Out Potential*. Journal of Physics : Conference Series. 1007(2018) 012049 pp.1-6.

Rasheed, F., & Wahid, A. (2021). Learning style detection in E-learning systems using machine learning techniques. *Expert Systems with Applications*, *174*, 114774.

Sasmito, G. W. (2017). Penerapan metode Waterfall pada desain sistem informasi geografis industri kabupaten Tegal. *Jurnal Informatika: Jurnal Pengembangan IT*, *2*(1), 6-12.

Sinaga, K. P., & Yang, M. S. (2020). Unsupervised K-means clustering algorithm. *IEEE access*, *8*, 80716-80727.

Tamba, S. P., Batubara, M. D., Purba, W., Sihombing, M., Siregar, V. M. M., & Banjarnahor, J. (2019, July). Book data grouping in libraries using the k-means clustering method. In *Journal of Physics: Conference Series* (Vol. 1230, No. 1, p. 012074). IOP Publishing.

Widayanti, F. D. (2013). Pentingnya Mengetahui Gaya Belajar Siswa Dalam Kegiatan Pembelajaran Di Kelas. *Erudio Journal of Educational Innovation*, *2*(1). https://doi.org/10.18551/erudio.2-1.2