K-MEANS ANALYSIS IN GROUPING ABILITIES OF BATTUTA UNIVERSITY INFORMATICS STUDY PROGRAM STUDENTS

Baginda Harahap
Battuta University
professionalbaginda@gmail.com

Abstract

This study aims to identify and classify students based on the learning outcomes of visualization and animation subjects measured based on the variables of attendance, assignments, midterm exams (UTS), and final semester exams (UAS) to be further used to evaluate learning for subjects that require proficiency. good quantitative analysis. This study uses k-means cluster analysis in classifying students into three groups based on their learning outcomes. After being grouped, there are 3 categories of cluster 0 with 14 competent conditions, 9 in incompetent condition cluster 1, and 9 students in very competent condition cluster 2.

Keywords: score value, k-means cluster and informatics study program

1. Introduction

Value as a concrete noun. Value here is a value or values that are often used to refer to something of value, such as its value, its value, and the value system. Then it is used for anything that has value or value as opposed to what is not considered good or valuable. Value is also used as a verb in the expression assess, rate and be judged. Judging is generally synonymous with evaluation when it is actively used to judge actions. Dewey distinguishes two things about judging, it can mean appreciating and evaluating.

In simple terms, it can be said that value is an idea or concept about something that is important in a person's life and becomes his concern. As a standard of behavior, of course the value according to someone to do so. There are some differences of opinion in interpreting value. Differences in perspective in understanding the meaning or understanding of value are a treasure for experts in interpreting the value itself, because each perception is based on theoretical, empirical, and analytical points of view.

According to Mulyana, value is a reference and belief in making choices. Value is something that is desired so that it gives birth to action in a person (Mulyana, 2004: 11). According to Frankel, values are standards of behavior, beauty, justice, truth, and efficiency that bind humans and should be carried out and maintained (Kartawisastra, 1980: 32-35). In addition to the two classifications of values as mentioned above, the values that are often used as human references in their lives are the six values contained in Spranger's theory, namely theoretical values, economic values, aesthetic values, social values, political values, and religious values.

Theoretical value involves logical and rational considerations in thinking and proving the truth of something. Economic value, related to the balance of values that contain profit and loss, which means prioritizing the use of something for humans. Aesthetic value, also known as the value of beauty, is very dependent on a person's subjective opinion. Social value accumulates at the highest value, namely affection between humans.

According to Scheler, value is a quality that does not depend on objects. An object is something of value. This independence includes every empirical form, value is an a priori quality. Addiction refers not only to objects that exist in the world such as paintings, sculptures, actions, people, and so on, but also to our reactions to objects and values.

From some of the definitions above, value is something inherent in humans that deserves to be carried out and maintained, as creatures created by God who have a distinctive character from other creatures. Humans have reason, feelings, conscience, affection, morals, character, and ethics are the distinctive characteristics of humans compared to other creatures, and this character is inherent in humans as a form of value itself.

Assessment is an important factor in the learning process. Therefore, respondents think that judging skills are an important element in shaping the professionalism of lecturers. There are several attributes attached by respondents related to this element, namely: fair and objective, not stingy and understanding. However, when viewed from the attributes attached by these students, it can be seen that the respondents are not consistent in expecting the professionalism of the lecturers. On the one hand, respondents said that lecturers must be fair and objective in giving grades. This means, lecturers are expected to give grades as they are. But on the other hand, respondents expect that lecturers are also not stingy and understanding in giving grades. This means that students expect lecturers to give subjective grades.

Ability comes from the word capable which means power (able, able) to do something, while ability means ability, skill, strength (Team Compilation of the Big Indonesian Dictionary, 1989: 552-553). Ability means the capacity of an individual to perform various tasks in a job. (Stephen P. Robbins & Timanthy A. Judge, 2009: 57). From these understandings, it can be concluded that ability is the ability or skill of an individual to master a skill and be used to do various tasks in a job.Ability is the most important thing from a person's job to be issued in a job, from that ability people will judge us whether we are competent or not. According to Hasibuan (2005, p94) work ability is a work achieved by a person in carrying out the tasks assigned to him based on skills, experience, and sincerity and time. Meanwhile, according to Robert Kreitner (2005, p185) what is meant by ability is a stable characteristic related to a person's maximum physical and mental ability.

In general, the definition of ability according to the KBBI (Big Indonesian Dictionary) is an ability, a person's ability to do something. A person is said to have the ability or ability if he can and is able to do something that he must do. Ability is the power (power) to do an action. Ability can also be an innate ability from birth, or is the result of practice or practice. According to Yusdi "ability is the ability to do something. A person is said to be capable if he can do something he must do.According to the Practical Indonesian Dictionary, students are those who are studying in higher education (Taufik, 2010). Salim and Salim (in Spica, 2008) say that students are people who are registered and undergoing education at universities. Susantoro (in Siregar, 2006) states that the figure of a student is also thick with nuances of dynamism and scientific attitude in seeing things based on objective, systematic and rational reality. A student is literally a person who studies in a college, be it a university, institute or academy. Those who are registered as students at universities can automatically be referred to as students (Takwin, 2008). According to Budiman (2006), students are people who study at tertiary level schools to prepare themselves for an undergraduate level expertise.

Meanwhile, according to Daldiyono (2009) a student is someone who has graduated from Senior High School (SLTA) and is currently pursuing higher education. Based on some of the expert opinions above, the researcher concludes that students are people who are undergoing higher education at a university or college.According to the General Indonesian Dictionary, competence means (authority) the power to determine or decide something. "Competence is the ability, skill and skill possessed by a person with regard to the duties of his position or profession" (Trianto, 2006). According to Law no. 14 of 2005 concerning Teachers and Lecturers. Competence is a set of knowledge, skills, and behaviors that must be possessed,

internalized and controlled by a teacher or lecturer in carrying out professional duties. According to Sabri in Puspitarini (2011), to be able to carry out teaching tasks well, lecturers must have professional abilities.

Lecturers are professional educators and scientists with the main task of transforming, developing and disseminating science, technology, and art through education, research and community service. Currently the government is trying to improve the condition of education in Indonesia, one of the steps it takes is in the Law of the Republic of Indonesia Number 14 of 2005 concerning Teachers and Lecturers. Lecturers are one of the important components that determine the quality of higher education, because the role of lecturers in learning is very central as a facilitator. as well as a motivator for the development of student thinking, as future community leaders.

2. Method

This study uses the K-Means algorithm to determine the best cluster. This best cluster is used for grouping student abilities in visualization and animation courses. So that it will be an evaluation in the learning process both for students and also for lecturers who are effective in the subject.

2.1 Visualization and Animation

In life we are often faced with terms such as visualization, animation and simulation. We need to know that there are similarities and differences when we look at their functions and definitions.

a. Visualization

Visualization is a form of conveying information that is used to explain something with animated images or diagrams that can be explored, calculated and analyzed the data. According to McCormick (et al, 1987), visualization provides a way to see the unseen. Some of the things that make up the formation of visualization:

i.   Use of signs
ii.   Image (drawing)
iii. Coat of arms and symbols
iv.   Science in writing letters (typography)
v.   Illustration and color

Visualization is a human effort to describe certain intentions into a form of information that is easier to understand. Usually nowadays people use computers.

Visualization develops with technological developments, including engineering, product design visualization, education, interactive multimedia, medicine, etc. Basically, visualization is used to diagnose and analyze the displayed data in order to predict conclusions.

b. Animation

Animation is a collection of images, either 2-dimensional or 3-dimensional composed of a set of objects/images arranged according to the storyline to produce moving images. At first, animation was only in the form of pieces of illustration or photography, which were then moved so that it seemed as if they were alive. Animation can be regarded as a movement simulation created by displaying sequential images or frames (Prakosa, 2010). Animated films are not only for entertainment for children but can also be used in various fields, because animated films contain information that is conveyed to the audience (Henry, 2005).

Animation comes from the Latin "anima" which means soul, life, spirit. While characters are people, animals or other real objects that are expressed in the form of 2D and 3D images. so that animated characters can be interpreted as images containing objects that seem to be alive, caused by the collection of images changing regularly and alternately being displayed. Objects in images can be in the form of text, object shapes, colors and special effects.

2.2 K-Means Cluster

The K-Means algorithm is a clustering algorithm that groups data based on the cluster center point (centroid) closest to the data. The purpose of K-Means is to group data by maximizing the similarity of data in one cluster and minimizing the similarity of data between clusters. The measure of similarity used in the cluster is a function of distance. So that maximizing the similarity of the data is obtained based on the shortest distance between the data and the centroid point. The initial step in the data clustering process using the K-Means algorithm is the formation of the starting point of the centroid cj. In general, the formation of the starting point of the centroid is generated randomly. The number of cj centroids generated corresponds to the number of clusters that were determined at the beginning. After k centroids are formed, the distance between each xi data and the jth to kth centroid is calculated, denoted by d (xi, cj). There are several distance measures that are used as a measure of the similarity of a data instance, one of which is the Euclidean distance. Euclidean distance calculation as in Eq.

K-Means is a non-hierarchical data collection method or data partition method into two or more groups. This method groups data into several partitions by entering data with the same characteristics into the same group and other data with different characters will be grouped into groups according to their respective characters (Putri et al., 2015). K-Means is a method of testing population components of data and grouping the data into a defined cluster depending on the minimum distance between the population components and each cluster center (Agustina et al., 2013). Kmeans Clustering is a constant of the desired number of clusters, while Means or can be defined as a cluster is an average value of a set of population data. In other words, K-Means Clustering can be defined as a data mining model that groups data into a partition system and performs a mapping process without using supervision (Malik Rio Andika, 2018).

The K-Means Clustering method groups and maps a population of data into several cluster groups where each data from the cluster group has the same characteristics as its group and is different from other groups. We can determine the basic algorithm of K-Means Clustering with the following steps:
1. Determine the desired number of clusters.
2. Choose a cluster randomly and group the other data into these clusters based on their closest distance.
3. Calculate the centroid / average of the existing data generated from each cluster.
4. Re-allocate each data into the nearest centroid/cluster average.
5. Repeat step 3, if there is still data that moves clusters, causing a change in the value of the cluster centroid.

This study uses primary data that comes from the record grades of fourth semester students in the visualization and animation course, the Informatics Engineering study program at Battuta University in the 2020/2021 academic year. The number of students who became the object of research were 32 students. Several independent variables were used in this study.

Table 1. Independent Variables

VARIABLE	VARIABLE NAME	DATA TYPE
X1	Attendance Value	Continuous
X2	Task Average	Continuous
X3	UTS value	Continuous
X4	UAS Score	Continuous

2.3 Clustering

One of the clustering techniques in data mining is the clustering method. Understanding scientific clustering in data mining is the grouping of a number of data or objects into clusters (groups) so that each in the cluster will contain data that is as similar as possible and different from objects in other clusters (NASARI & SIANTURI, 2016). One of the clustering methods that have efficient and fast properties that can be used is the k-means method, this method aims to make object clusters based on attributes into k

partitions (Wardhani, 2016). Hierarchical clustering is a data grouping method that starts by grouping two or more objects that have the closest similarities. Then the process is passed to another object that has a second affinity (Bastian et al., 2018).

Clustering is called grouping a number of data or objects into clusters (groups) so that each cluster will contain data that is as similar as possible and different from objects in other clusters (Putri, Kom, & Kom, 2015). There are two clustering methods that we know, namely Hierarchy and Non Hierarchy (Putri et al., 2015).

Clustering is grouping records, observing, or paying attention and forming classes of objects that have similarities. Cluster is a collection of records that have similarities with one another and have dissimilarities with records in other clusters. Clustering differs from classification in that there is no target variable in clustering. Clustering does not attempt to classify, estimate, or predict the value of the target variable. However, the clustering algorithm tries to divide the entire data into groups that have similarities (homogeneous), where the similarity of records in one group will be of maximum value, while similarity of records in other groups will be of minimal value.

2.4   RapidMiner

RapidMiner is a software that is open (open source). RapidMiner is a solution for analyzing data mining, text mining and predictive analysis. RapidMiner uses a variety of descriptive and predictive techniques to provide users with insights so they can make the best decisions. RapidMiner has approximately 500 data mining operators, including operators for input, output, data preprocessing and visualization. RapidMiner is a stand-alone software for data analysis and as a data mining engine that can be integrated into its own product. RapidMiner is written using the Java language so that it can work on all operating systems.

2.5 Implementation

The definition of implementation put forward by Hanifah Harsono, it can be said that implementation is a policy in the completion of decisions in order to achieve good goals depending on how the implementation goes well in carrying out the final refinement process. Therefore, a good implementation is expected in each program to create the expected goals. In simple terms, implementation can be interpreted as implementation or application. Majone and Wildavsky (2004) in (Nurdin and Usman, (2004:70) suggest implementation as evaluation. Majone and Wildavsky argue that implementation is an expansion of activities that adjust to each other.

In this study, the RapidMiner software will be the decision-making system in processing data on the ability of students in the informatics study program based on attendance scores, average assignments, UTS and UAS in visualization and animation courses.
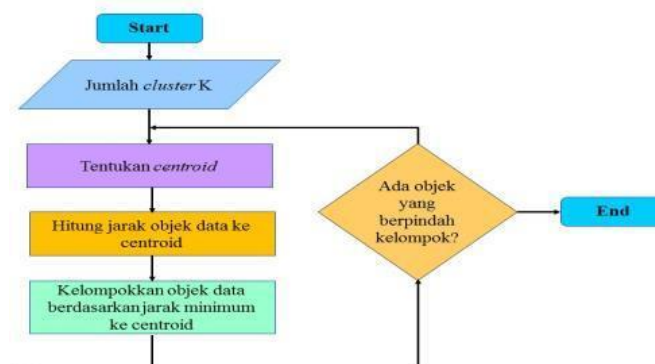

Figure 1. RapidMiner framework

3 Results and Discussion

Figure 2. The Form of RapidMiner's Implementation of the Results of Visualization and Animation Courses in the Informatics Study Program Based on Figure 2. the operator needs that are used to manage the student's ability scores in the visualization and animation courses, namely the Retrieve operator functions to input data, while the data we use is in excel format. Then the Clustering operator functions to group, in this case we group them into 3 groups, namely competent, incompetent and very competent.
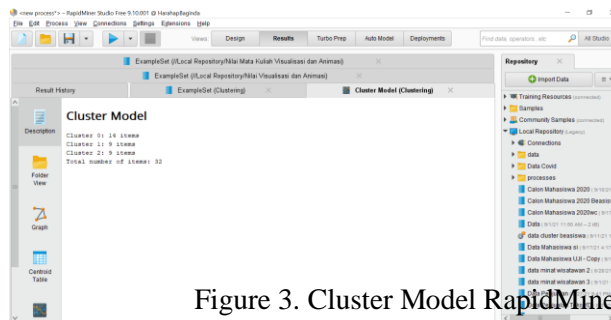
Cluster Model

Figure 3. Cluster Model RapidMiner

In Figure 3. we can see that the amount of data entered is 32 items, while cluster 0 is 14 items competent students, cluster 1 is 9 items incompetent students, and cluster 2 is 9 students who are very competent in visualization and animation courses. items.
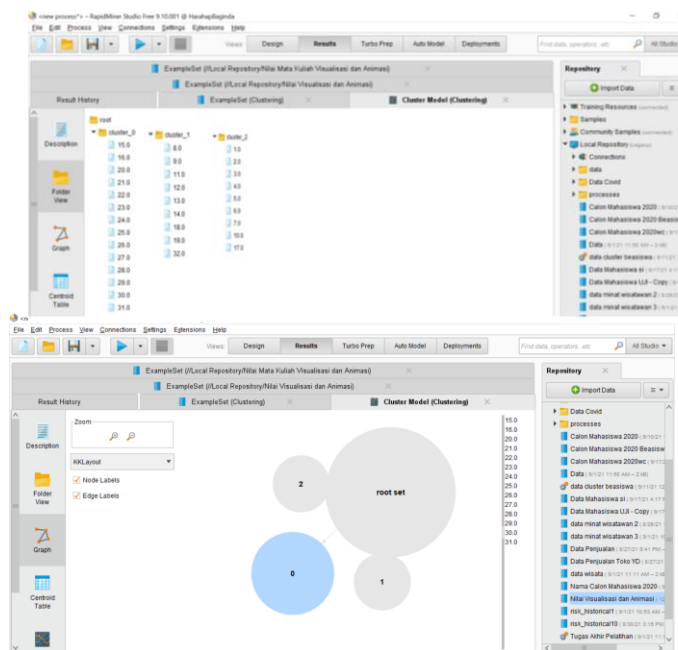

Figure 5. Appearance of Graph

Figure 6. Appearance of Centroid Table

Figure 7. Appearance of Plot
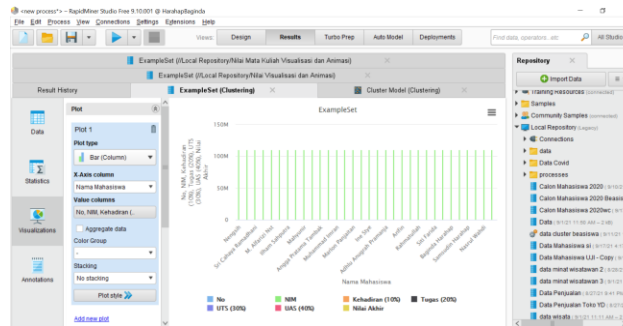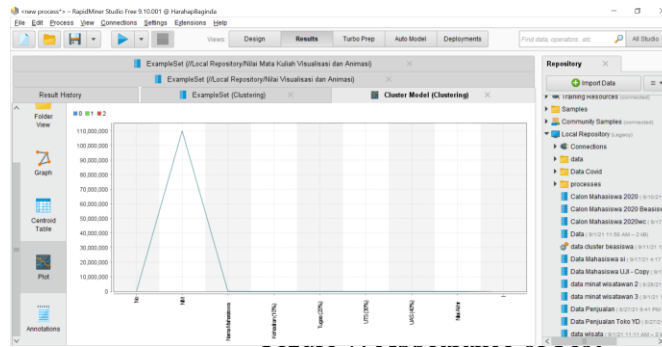
Figure 8.   Appearance of Visualizations
Table 2.   Incompetent Student

NIM	Student Name
110119008	Sanny Vikas
110119009	Mahyunir
110119011	Angga Pratama Tambak
110119012	Ilham Abdillah Siregar
110119013	Muhammad Imran
110119014	Muhammad Darmiansyah Kurniawan
110119018	Erman Efendi
110119019	Adhlu Anugrah Pramanja
110119032	Arbinto Boyke Ginting

Table 3. competent student

NIM	Student Name
110119015	Marlon Panjaitan
110119016	Muhammad Willy Nazhmi
110119020	Teuku Muhammad Ravyza Husein
110119021	Arifin
110119022	Muhammad Yusuf Nasution
110119023	Rahmatullah
110119024	Yusuf Dalimunthe
110119025	Siti Farida
110119026	Anwarul Khair

110119027	His Majesty Harahap
110119028	Arief Budirahman Lubis
110119029	Samsudin Harahap
110119030	Ari Pratama Yudha Tj
110119031	Nasrul Wahdi

Table 4. Very Competent Student

NIM	Student name
110119001	sad
110119002	Muhammad Fikri Alhafidz
110119003	Sri Cahaya Ramadhani
110119004	Java Paimen
110119005	M. Alfarizi Nst
110119006	Akhmad Arfansyah
110119007	Ilham Sahputra
110119010	Ari Swanda
110119017	Ine Stye

4 Conclusion

Based on the results obtained from this study with the K-Means clustering method. The results obtained 3 categories of student abilities in achieving visualization and animation subjects, namely incompetent, competent and very competent. This is evidenced by the calculation of the closest distance based on the determination of the centroid value randomly using the Euclidean Distance formula based on the attendance value, average assignment, UTS and UAS. So there are 9 students in the incompetent category, 14 competent students, and 9 very competent students.

Reference

[1]. Harahap, Baginda. Rambe, Aripin. (2021). Implementation of K-Means Clustering for Students who Receive Scholarships from the Battuta Education Foundation at Battuta University 2020/2021 Case Study of Informatics Study Program. INFORMATICS Information Management Labuhanbatu University Vol. 9 No. 3 / September /2021 2615-1855 (E-ISSN) 2303-2863 (P-ISSN). Medan.

[2]. Sri, M. (2015). Application of Data Mining With Clustering Methods for Grouping Bird Shipping Data. Proceedings of the National Scientific Seminar on Computer Technology (SENATKOM 2015) Putra Indonesia University YPTK Padang. Vol. 1, ISSN : 2460 – 4690. Padang.

[3]. Harahap. Baginda. (2018). Application of the K-Means Algorithm to Determine Best Selling Building Materials (Case Study at UD. YD Indarung Building Store). Ready Star-2 ISSN(Print) : 2620-6048 ISSN(Online) : 2686-6641. Medan.

[4]. Sugiono, Dkk. (2019). Grouping Student Behavior in E-Learning Lectures with K-Means Clustering. Journal of Scientific Studies Bhayangkara University, Greater Jakarta Volume 19, No. 2, May 2019 p-ISSN 1410-9794 e-ISSN 2597-792X. Jakarta.

[5]. Defit, S. (2013). The Use of Apriori Algorithm in Analyzing Student Behavior in Choosing Courses (Case Study: FKIP UPI "YPTK"). Journal of Media Processor Vol.8, No.3, October 2013. Padang.

[6]. Du, W, dkk. (2016). A New Projection-based K-Means Initialization Algorithm. Proceedings of 2016 IEE Chinese Guidance, Navigation and Control Conference. 978-1-4673-8318-9/16/S31.00©2016 IEEE.

[7]. Gornitz, N, dkk. (2017). Support Vector Data Descriptions and k-Means Clustering : One Class?. 2162237X© 2017.

[8]. Hu, J, dkk. (2017). Improved K-Means Algorithm Based on Hybrid Fruit Fly Optimization and Differential Evolution. 978-1-5090-2508-4/17/$31.00©2017 IEEE.

[9]. Gu, L. (2016). A Novel Locality Sensitive k-Means Clustering Algorithm based on Subtractive Clustering. IEEE. 978-1-4673-9904-3/16/S31.00 ©2016. Ikhwan.A.dkk., 20015. Application of Data Mining with Fp-Growth Algorithm for Support.

[10]. Bastian, A, dkk. (2018). APPLICATION OF K-MEANS CLUSTERING ANALYSIS ALGORITHM IN HUMAN TRANSMITTED DISEASES (CASE STUDY OF MAJALENGKA DISTRICT) Ade, (1), 26–32.

[11]. Fitriyani, F. (2018). Bagging Method For Balance Class In Thoracic Surgery Using Naive Bayes. Journal of Scientific Studies, 18(3), 278-282.

[12]. Kaur, M. (2013). Cluster Analysis of Behavior of E-learners. (2), 344–346.

[13]. Berkhin, Paul. (2006). A survey of clustering data mining techniques. In: Grouping multidimensional data. Springer Berlin Heidelberg. p. 25-71.

[14]. DU, K.-L. (2010). Clustering: A neural network approach. Neural Networks. 23.1: 89-107.

[15]. Garcia, dkk. (2002). Application of k Means Clustering algorithm for prediction of Students Academic Performance." arXiv preprint arXiv:1002. London.

[16]. Asroni, Ronald Adrian. (2015). Application of the K-Means Method for Clustering Students Based on Academic Values with the Weka Interface Case Study at the Department of Informatics, UMM Magelang. JOURNAL OF SCIENTIFIC SEMESTA TEKNIKA Vol. 18, No. 1, 76-82, May 2015. Magelang.