Clustering K Means for Criteria Weighting With Improvement Result of Alternative Decisions Using SAW and TOPSIS

[1]Erna Daniati
Department Of Information System
Universitas Nusantara PGRI Kediri
Kediri, Indonesia
ernadaniati@unpkediri.ac.id

[2]Hastari Utama
Department Of Informatic Engineering
Universitas Amikom Yogyakarta
Yogyakarta, Indonesia
utama@amikom.ac.id

***Abstract*— In today's world, decision making can be helped by using decision support systems. This system is an approach to support decision making. Decision makers use relative weights for each attribute. The obtained total score is the score of each alternative decision. The process of making alternative decision uses a combination of SAW and TOPSIS methods. In addition, the determination of criteria weight is also influenced by K Means method. This clustering method serves to provide an alternative weighting value so that decision makers no longer need to give the initialization value. The combination of these methods generate a more absolute alternative value than only using SAW. This research is a development of previous research by adding TOPSIS method. The resulting alternative decision has a more absolute and significant alternative value. This is indicated by comparison of TOPSIS usage results and using only SAW. This research uses data student of information system in Universitas Nusantara PGRI Kediri. It aims to proof the significant result for comparing with further method.**

Keywords—K Means, SAW, TOPSIS, DSS, Thesis

I. INTRODUCTION

Thesis is a scientific work done by students [1]. This thesis is used as a graduation requirement for students and become a report of research activities. This research covers the field and science that understood by students. Students will explore and add science related to thesis. Arranged thesis uses the basic concept of research methods so the results can be accounted for.

Student Research in form of thesis begins in deciding topics that match with the fields that students want. This topic can be reinforced by some concentration courses [1]. Students will study subject deeply in accordance with the thesis topic that has been determined. The courses have been adapted and conducted to the lecture achievement curriculum. So, students who choose the topic according to the courses that follow it will be easier process of research.

In today's world, decision making can be helped by using decision support systems. This system is an approach to support decision making [2], [3]. This system is also computer-based supportive. The decisions resulting from these decision-makers strongly influence the organization's policy. Decision-makers still often use intuition to determine their decisions. However, decisions based on this intuition are more subjective and less satisfactory. It uses an interactive and flexible view to support solutions on unstructured management addresses.

In the decision support system, there are also several methods for modeling. These methods such as Simple Additive Weighting (SAW), TOPSIS, and AHP [4]. SAW is one of the techniques which used to accomplish problem of decision analysis. Decision makers use relative weights for each attribute. The obtained total score is the score of each alternative decision. In addition, SAW method can also be combined with other methods [5]. Furthermore, TOPSIS is often used in evaluating various criteria in object evaluation [6]. This method tries to calculate the distance between object evaluation to ideal solution of optimal and worst. If it is near the optimal solutions and far from worst solution then it is a good solution. Then, the clustering method of K Means is used to data input of the TOPSIS method. This clustering method works to generate weight value as input for TOPSIS method so decision maker is not defined from beginning.

Decision support systems can also be used in higher education environments. On this issue, the decision support system serves to help students in selecting a appropriate curriculum to the background of their ability than campus requirement [7]. It is applied to the process of new student registration. In addition, decision support systems can also be used in assisting new students in determining faculty according to ability [8]. Determination of this faculty using scoring model, financial, and mixed. These three models will help students decide when they want to find the faculty and even the appropriate curriculum. In this study also make use of the decision to help students determine the topic of the thesis.

Decision making of various criteria or decision making in several criteria (MCDM) is a difficult task because of the same alternative. This study carries out the MCDM hybrid method with simple additive weighing (SAW), a technique for order preference by similarity with ideal solutions (TOPSIS) and gray relational analysis (GRA) [9]. A feature of this method is that it uses experimental design techniques to determine the weight of attributes and methods used to make hybrid models. This model can guide decision makers in making valuable ones without professional reasons or extensive experience. The ranking results agreed upon by more reliable MCDM methods generated by a single MCDM method.

Other studies also use the SAW and TOPSIS methods as the main structure for managing information from fuzzy intervals. Fuzzy sets of RATs are more uncertain intervals than ordinary fuzzy sets and can be used for uncertain or uncertain information in fields that complete multi-criteria decision analysis (MCDA) [10]. The use of fuzzy set method is available in SAW-based and TOPSIS-based MCDA methods and conducts comparative studies. Discussions that have been carried out jointly between functions and scores, where the score function will have positive and negative effects from conditions that are not biased, positive biases, and negative biases. The correlations and levels of contradictions obtained in the experiment indicate that there is something clear between SAW fuzzy intervals and TOPSIS ratings.

This study is expected to generate benefits from the combination of K Means cluster and the SAW and TOPSIS modeling methods. This clustering method is expected to help the decision maker in order to be not initialize criteria weight. It will prune decision modelling phase. The expected final result of the calculation is more absolute and significant so the difference of weight score is more specific and clear. It helps the decision maker to decide based on alternative decisions comfortably. In addition, implementation to higher education is also expected to able contributes students in determining the decision in deciding topic of the thesis.

II. METHODOLOGY

This study uses an experimental approach that attempts to experiment changes on the variables. Experimental research is a study that manipulates or controls the natural situation by creating artificial conditions [11]. Formulation of this condition done by the researcher. Thus, experimental research is a research which is conducted by manipulating the object of research, and there is the existence of deliberate control of the object. In addition, there are three important elements in experimental research that must be considered in doing this research, namely control, manipulation, and observation. The control variable here is the essence of the experimental method, because it is the control variable that will be the standard to see if there are any changes, as well as the differences that occur due to the different treatments. While the manipulation is a deliberate operation in experimental research. Changes in variables is conducted in this study involves the variable value of the course and the grade point average (GPA) of each semester.

The stages of this study are shown in Figure 1. This study begins with a literature study. This stage aims to find references of several papers and books relating to the theme of decision support systems. The second step is data collection. Data collection is done by observation approaches about the process that occurs in the determination of the topic of the thesis and look for the data in the archive in the form of file report of each student. Problems of the background is analyzed then made a decision support model formulation. This formulation requires a case example as a simulation. The results of this simulation are tested by comparing with the results of previous research so that the obtained results can be explained in this study.

This research uses several methods in building decision support system. This method consists of K Means method, Simple Additive Weighting (SAW) and Technique for Order Performance by Similarity to Ideal Solution (TOPSIS). The workflow of this system is shown in Figure 2. The process begins with obtaining a sample of student grades in semesters 1 through 7. The selected student is randomized in several different year. Some of the collected student score is calculated the grade point average (GPA) per semester so every student has GPA score.

Each student with an GPA score from semesters 1 through 7 is used as a category in the clustering to generate multiple groups or clusters. These established groups are used as the basis for topic groups. The average calculation result from each topic group is used input value as the weight value of each topic in SAW method calculation.

Calculation using K Means method in Figure 2 is an algorithm that uses distance as evaluation of equality index. This method has advantages in terms of speed, simplicity and high efficiency for large amounts of data as well as high scalability. In addition, K Means method is also widely used in various areas of research such as social networking, image processing, attack detection on android, and so on [6]. The K Means method can also be upgraded to Seeded-K Means where the centroid is calculated based on the random selection of k documents in the X dataset [12]. Once the dataset is determined then the next step is the same as the previous step.

The K Means method has a formulation pattern as shown in (1) where the process tries to find the shortest distance. In Equation (1), the result of the 1st c group where the shortest distance is obtained from the distance of the category xi value to the centroid of the test. This searching distance uses the euclidean method. Central point of the test is obtained by using equation (2). This center is searched by doing the average of each group member. This average value is calculated from the sum of category xi values in each group ci divided by the number of members m in each ci group. This group search is always repeated until it gets the composition of members of each group in each loop. Groups with the same members on each iteration are the result of the K Means clustering.

$$C^i := \arg\min \left\| x^i - \mu_j \right\|^2 \quad (1)$$

$$\mu_j := \frac{\sum_{i=1}^{m} \{c^i = j\} x^i}{\sum_{i=1}^{m} \{c^i = j\}} \quad (2)$$

The SAW method used in this study is a method often used in problem analysis for decision making [1], [4]. The decision maker assigns a relative weight value for each of its attributes. The total value is obtained by multiplying the weight value of each category weight and its alternate value. Calculations using SAW are performed only on the input of alternative values and the normalization of the matrix containing the alternate values. The normalization phase of the matrix is shown in (3). In the equation Z_i is the result of the sum of all data n. The value of Z_i can also be referred to as an alternative value of multiplication of the normalization of matrix r_{ij} with w_j weight.

$$Z_i = \sum_{j=1}^{n} r_{ij} . w_j \quad (3)$$

The alternative value of SAW calculation is the input for TOPSIS method. This TOPSIS method was first developed by Hwang and Yoon in 1981 [4]. This method was developed to solve the problem of multiple criteria decision making (MCDM). MCDM seeks to focus on the structuring and resolution of decision issues involving multiple criteria and can be solved by the existence of methods. The resolution of the MCDM problem corresponds to the alternative selection of a number of alternative decisions provided [13]. In accordance with the technique, the best alternative is one of the closest values to a positive ideal solution and far with the value of a negative ideal solution. In Figure 2 the TOPSIS method is only depicted in an empty box. This means that the description of the TOPSIS method path has a detailed description. This detail description is shown in Figure 3.

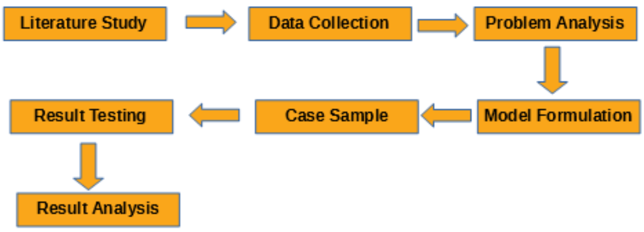

Fig. 1. Research Phases

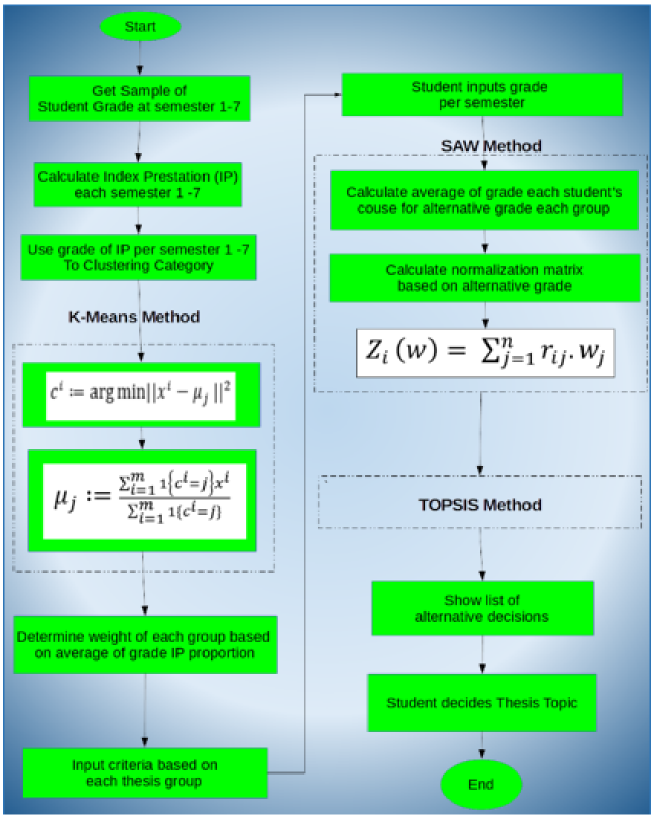

Fig. 2. Overview of Decision Support System Design

MCDM can be broadly divided into multi-objective decision-making (MODM) and multi-attribute decision-making (MADM) [9]. There are several methods in each of these categories. Different methods use different normalization procedures and have different treatments for benefit and cost criteria. In this study, we focus on hybrid methods involving SAW, TOPSIS and GRA techniques. The simplicity of SAW makes it very popular in practical applications. This is often used as a benchmark for comparing results obtained from other MCDM methods. Meanwhile, TOPSIS is a distance-based approach.

The most widely used method in multi-criteria decision making model (MCDM) is SAW [14]. In this method, the score for each option is obtained by combining the values of the option in different criteria, taking into account the weight of each criterion, so that the relative weight is given directly by the decision maker. The SAW method for rating options can be described as follows:

1) Normalization of the initial matrix to the decision matrix counts as

$$Rij = \frac{x_{ij}}{\sum_1^M x_{ij}} \qquad (4)$$

where, R_{ij} is the normalized weight of the criteria for j, m is the number of criteria, and X_{ij} is the initial weight.
2) Determining the value of $Rij = Rij^{min}$ or $Rij = Rij^{max}$.
$Rij = Xij^{min}$ if the efficiency index is minimized, $Rij = Rij^{max}$ if the efficiency index is maximized.
3) Nilai normal dari reliabilitas indeks sebagai Persamaan. (5) dan (6), masing-masing:

$$Rij = \frac{x_{ij}}{x_{ij}^{max}} \qquad (5)$$

if the efficiency index is maximized.

$$Rij = \frac{x_{ij}^{min}}{x_{ij}} \qquad (6)$$

if the efficiency index is minimized.
4) The calculation of the weighted decision matrix is weighted as:

$$V_{ij} = R_{ij} \times W_j \qquad (7)$$

where, V_{ij} is a normalized decision matrix element, R_{ij} is the score of the i-th alternative with respect to the j-th criteria, and w_j is the criteria weight using the AHP model.
5) The final step in the SAW method is the final score data integration each option will obtain according to:

$$A_i = \sum R_{ij} \times W_j \qquad (8)$$

where, A_i is the final weight of each option, R_{ij} is the score of the alternative i in relation to criterion-j, and w_j is the criteria weight using the AHP model.

TOPSIS, one of the most famous classical MCDM methods, was originally developed by Hwang and Yoon in 1981, with further development by Chen and Hwang in 1992 [9]. The TOPSIS method introduces two 'reference' points: positive ideal solutions and negative ideal solutions. Positive ideal solutions are solutions that maximize benefit criteria and minimize cost criteria, while negative ideal solutions maximize cost criteria and minimize 12 benefit criteria. TOPSIS determines the best alternative by minimizing the distance to the ideal solution and maximizing the distance to the negative ideal solution. This method assumes that each attribute monotonically increases or decreases. TOPSIS utilizes Euclidean distance to measure alternatives with positive ideal solutions and negative ideal solutions. The order of alternative preferences is generated by comparing Euclidean distances. The order of alternative preferences is produced by comparing Euclidean distances. The TOPSIS process is carried out in Figure 3.

In Figure 3 a detailed step is shown on the TOPSIS method. The first step in this method is to build an alternative matrix derived from the calculation using the SAW method. The matrix is normalized from each element so that once the matrix has been prepared the elements are changed and different from before. This normalization process involves dividing each element by the square of the element. This normalization is shown in equation 9 where r_{ij} shows the normalization value of the j-criterion for the alternative A_i. Next, calculate the weight of the normalized decision matrix as in equation 9 [9]. This normalized matrix is described again to find the value of positive and negative ideal solutions.

The equations for calculating the positive and negative ideal solutions are shown in expression (10) and (11). Each ideal solution produces value. Expression (10) shows about the calculation of positive ideal solution whom the result is represented in A+. Besides that, negative ideal solution is represented in A- at expression (11). TOPSIS attempt to determine the positive ideal solution and the negative ideal solution. The positive ideal solution maximizes the benefit criteria and minimizes the cost criteria, while the negative ideal solution maximizes the cost criteria and minimizes the benefit criteria. The benefit criterion is a criterion where the value of the criterion gets bigger then it is more feasible to choose. While the cost criterion is the opposite of the benefit criterion, the smaller value of the criterion then it is more feasible to be chosen. In the TOPSIS method, the optimal alternative is the closest to the positive ideal solution and the farthest from the negative ideal solution.

The value of this calculation is calculated by the rule of approaching a positive ideal solution and knowing the ideal solution. The result of this calculation is the preference value of each alternative decision. Thus, decision makers can choose preferences with a certain value from each alternative decision.

$$r_{ij} = \frac{x_{ij}}{\sqrt{\sum_{k=1}^{m} x_{kj}^2}}, i = 1, \dots, m; j = 1, \dots, n \quad (9)$$

$$A+ = \left\{ \begin{array}{l} (\max V_{ij} | j \in J), (\min V_{ij} | j \in J'), \\ i = 1,2,3, \dots, m\} = V_1 + \ V_2 + , \dots, V_n + \end{array} \right\} \quad (10)$$

$$A- = \left\{ \begin{array}{l} (\max V_{ij} | j \in J), (\min V_{ij} | j \in J'), \\ i = 1,2,3, \dots, m\} = V_1 - \ V_2 - , \dots, V_n - \end{array} \right\} \quad (11)$$

III. Result and Discussion

Formulation of calculation flow of decision support system with K Means, SAW, and TOPSIS method should be applied in case examples to show the advantages and weaknesses. The data as a case example is data of 12 students who have Grade Point Average (GPA). The data are shown in Table I. The data is taken on the Information System Programme of Universitas Nusantara PGRI Kediri. Each student is calculated GPA from semester 1 through 7. This data is used as input data for grouping topic list.

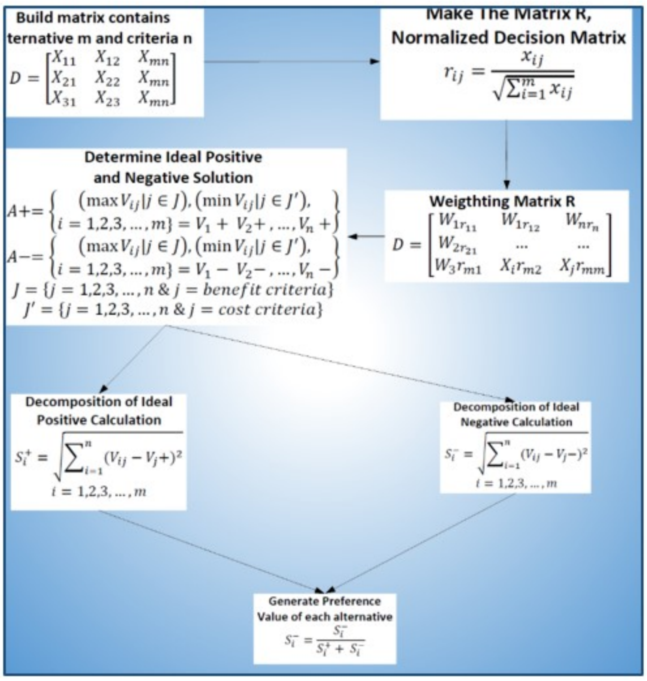

Fig. 3. Flow of TOPSIS Calculation

The data presented in Table I contains some values of GPA per semester. This data is selected a random number of 5 rows. Selection of this data aims to determine the centroid on clustering using K Means method. The data that has been selected is shown in Table II. Each data has a group label with the group sample names being A, B, C, D, and E. This centroid is used as a reference for calculating other data distance to the centroid this. If it produces the shortest distance to the centroid of the group then the data belongs to that group member. This is done in several iterations. Repetitive iterations are done to look for group members that do not change again every iteration change. Groups generated from the iteration process where the members are not, then this iteration will stop and produce groups with fixed members. The clustering results are shown in Table III.

TABLE IV. Weight of Criteria

	C1	C2	C3	C4	C5
Average:	3.300	3.240	3.442	2.969	2.259
Proportion:	0.217	0.272	0.397	0.568	1.000
Weight:	0.088	0.122	0.202	0.362	1.000

TABLE V. Entriying Course Scores by Student

1. Business Intellegence:	**Credit**	**Letter Score**	**Number Score**	**Total**
Information System Concept	4	B+	3.5	14
Statistic	2	C+	2.5	5
Management Information System	2	B	3	6
Database	3	C+	2.5	7.5
Analysis of Business Process	2	C	2	4
Operation Research	2	C	2	4
Information Systen Analysis and Design	4	C+	2.5	10
Object Oriented Programming	3	C+	2.5	7.5
Information System Development	4	B+	3.5	14
Distributed System	3	E	0	0
DSS	4	D	1	4
Data Mining	3	E	0	0
Expert System	2	D	1	2
	Grade	**2.05**		**0.41**

TABLE VI. Total Alternative Scores Using SAW Method

Topic Group	C1	C2	C3	C4	C5
Business Intellegence	0.41	0.41	0.41	0.41	0.41
System Analyst and Design	0.43	0.43	0.43	0.43	0.43
Audit and Control	0.51	0.51	0.51	0.51	0.51
Software Development	0.26	0.26	0.26	0.26	0.26
Networking	0.35	0.35	0.35	0.35	0.35
Square Of Each Criteria	0.8032	0.8032	0.8032	0.8032	0.8032

Matrix of Normalized Decision

Topic Group	**C1**	**C2**	**C3**	**C4**	**C5**
Business Intellegence	0.510458167	0.510458167	0.51045817	0.510458	0.510458

Topic Group	C1	C2	C3	C4	C5
System Analyst and Design	0.535358566	0.535358566	0.53535857	0.535359	0.535359
Audit and Control	0.634960159	0.634960159	0.63496016	0.63496	0.63496
Software Development	0.323705179	0.323705179	0.32370518	0.323705	0.323705
Networking	0.435756972	0.435756972	0.43575697	0.435757	0.435757

TABLE VII. MATRIX OF WEIGHTED NORMALIZED DECISION

Topic Group	C1	C2	C3	C4	C5
Business Intellegence	0.045132291	0.06207784	0.10313899	0.184882	0.510458
System Analyst and Design	0.047333866	0.065106027	0.10817016	0.193901	0.535359
Audit and Control	0.056140167	0.077218777	0.12829484	0.229975	0.63496
Software Development	0.028620477	0.039366435	0.06540521	0.117242	0.323705
Networking	0.038527566	0.052993278	0.08804548	0.157826	0.435757
MAX	0.056140167	0.077218777	0.12829484	0.229975	0.63496
MIN	0.028620477	0.039366435	0.06540521	0.117242	0.323705

TABLE VIII. RESULT OF POSITIVE AND NEGATIVE IDEAL SOLUTION

Ideal Solution	Score
D1+	0.13607857
D2+	0.108862856
D3+	0
D4+	0.340196424
D5+	0.217725712
D1-	0.204117855
D2-	0.231333569
D3-	0.340196424
D4-	0
D5-	0.122470713

TABLE IX. GENERATED ALTERNATIVE DECISIONS

Topic	Score
Business Intellegence:	0.6
Audit and Control:	0.68
System Analyst and Design:	1
Networking:	0
Software Development:	0.36

TABLE X. ALTERNATIVE DECISIONS USING SAW

Topic	Score
Business Intellegence:	1.648039216
Audit and Control:	1.812745098
System Analyst and Design:	2.55
Networking:	0.662745098
Software Development:	1.200980392

The number of groups formed is 5. Furthermore, this is done by calculating the weight of criteria for each group. Each group is represented in the average GPA score. This average calculation result is represented in C_i. In general, the value of the weight of this criterion is inputted by the decision maker. However, this step is less simple and decision makers should know the intent of the score that have been entered. Therefore, the criteria weight is searched by calculating the average of each group values. Weight value is obtained by calculating the value of proportion. The value of the proportion is obtained by dividing each C_i by the total value of C. It is also similar when calculating the weight value where each C_i value is divided by the mean of the value of C.

Furthermore, the students enter the value of each course that has been followed. The value entered is shown in Table V. This value is used to obtain an alternative weight using the SAW method. Alternative value calculation results are shown in Table VI. These alternative values as input values for matrix normalization where the results are shown in Table 7. Normalization of this matrix is the first step in the use of TOPSIS method. This is followed by performing a calculation to find a weighted normalization matrix whose results are shown in Table VIII. Then, the calculation of positive and negative ideal solutions is performed on the basis of a weighted weighing matrix. The results obtained from the positive and negative solutions are shown in Table IX. The final step, in the TOPSIS method is to generate an alternative decision. This final step is shown by the calculation of the preference value of each alternative decision. This is done by finding a value that is close to a positive ideal solution but to know the ideal negative solution. The results obtained are shown in Table X. The results in Table X are compared with Table XI which is an alternative decision produced only by the SAW method without combined with the TOPSIS method. The comparison result shows the alternative decision value that is generated using TOPSIS more absolute and simple there is alternative value that produced by SAW.

IV. CONCLUSION

In this study contains several stages and discussion. Some of these stages and discussions produce some conclusions. The use of the K Means method seeks to derive the criterion weight derived from the average value of each topic group. The resulting weights need not be determined by decision makers and as input values for SAW calculations. Then, the calculation using the SAW method aims to obtain the total alternative value as the input of the decision matrix to calculate the calculation using TOPSIS method. The final result is an alternative decision where each alternative decision has a weight. The highest weight of the alternative decision is expected to be a decision supporter. The result of comparison indicates that the priority weight sequence has the TOPSIS method is more absolute and simple than just using SAW.

ACKNOWLEDGMENT

I thank all to the people of Universitas Nusantara PGRI Kediri for all their support in completing this research. In addition, I would also like to thank for the 2019 ICITISEE committee who prepared the conference and provided the publication media for this scientific work.

REFERENCES

[1] E. Daniati and A. Nugroho, “K-Means Clustering With Decision Support System using SAW,” IEEE Int. Conf. Control Syst. Comput. Eng., vol. 6, no. November, pp. 25–27, 2016.

[2] V. L. Sauter, Decision Support Systems for Business Intelligence. Hoboken, NJ, USA: John Wiley & Sons, Inc., 2011.

[3] E. Turban, J. E. Aronson, and T.-P. Liang, “Decision Support Systems and Intelligent Systems,” Decis. Support Syst. Intell. Syst., vol. 7, p. 867, 2007.

[4] J. Seyedmohammadi, F. Sarmadian, A. A. Jafarzadeh, M. A. Ghorbani, and F. Shahbazi, "Application of SAW, TOPSIS and fuzzy TOPSIS models in cultivation priority planning for maize, rapeseed and soybean crops," Geoderma, vol. 310, no. November 2016, pp. 178–190, 2018.

[5] Y. J. Wang, "A fuzzy multi-criteria decision-making model based on simple additive weighting method and relative preference relation," Appl. Soft Comput. J., vol. 30, pp. 412–420, 2015.

[6] X. Lu, W. Chang, and S. Zhou, "Quality Evaluation of Diesel Engine Using an Improved TOPSIS Method Based on K-means Algorithm," in Proceedings of the 35th Chinese Control Conference, 2016, pp. 9628–9633.

[7] L. Ploywattanawong, "Decision Support Systems Model for Admission to Higher Education," pp. 1359–1363, 2016.

[8] J. Chanwijit, W. Lomwongpaiboon, O. Dowjam, and P. Tangworakitthaworn, "Decision Support System for Targeting Higher Education," pp. 4–7, 2016.

[9] P. Wang, Z. Zhu, and Y. Wang, "A novel hybrid MCDM model combining the SAW, TOPSIS and GRA methods based on experimental design," Inf. Sci. (Ny)., vol. 345, pp. 27–45, 2016.

[10] T. Y. Chen, "Comparative analysis of SAW and TOPSIS based on interval-valued fuzzy sets: Discussions on score functions and weight constraints," Expert Syst. Appl., vol. 39, no. 2, pp. 1848–1861, 2012.

[11] C. Kothari, Research methodology: methods and techniques. 2004.

[12] R. K. Roul, "Semi-supervised Clustering using Seeded- k Means in the Feature Space of ELM," pp. 0–5, 2016.

[13] B. Dong, P. Gao, H. Wang, and S. Liao, "Clustering Human Wrist Pulse Signals via Multiple Criteria Decision Making," Proc. - Int. Conf. Tools with Artif. Intell. ICTAI, vol. 2014–Decem, pp. 243–250, 2014.

[14] [14] A. A. Ameri, H. R. Pourghasemi, and A. Cerda, "Erodibility prioritization of sub-watersheds using morphometric parameters analysis and its mapping: A comparison among TOPSIS, VIKOR, SAW, and CF multi-criteria decision making models," Sci. Total Environ., vol. 613–614, pp. 1385–1400, 2018.

TABLE I. GRADE POINT AVERAGE PER SEMESTER

ID	Student Name	GPA Per Semester						
		GPA-1	GPA-2	GPA-3	GPA-4	GPA-5	GPA-6	GPA-7
13.11.5473	AFLAKHATIS RATNA C.	3.35	3.25	3.42	3.35	3.16	3.26	3.31
13.11.5474	ANIS YULIASIH	3.4	3.25	3.18	3.3	3.29	3.21	3.31
13.11.5475	AHMAD ARIF SETYO UTOMO	2.65	2.5	2.93	3.4	3.16	2.79	3
13.11.5476	APRIZAL EFENDI	2.55	1.44	1.16	0.86	2.68	2.47	2.5
13.11.5477	AHMAD EFENDI	3.25	3.55	3.18	3.5	3.39	3.11	3.44
13.11.5478	ANI BINTI FITRIAH	3.3	3.28	2.53	3.2	3.45	3	3.13
13.11.5479	APRILIA DITA HENING SUCI	3.1	3.05	2.89	2.9	3.37	3.13	2.81
13.11.5480	ELFA MIFTAKHUL JANNAH	3.25	3.28	3.29	3	3.26	3.16	3.18
13.11.5481	ADITYA RAMADHAN TRY P.	2.85	3.1	2.92	2.9	3.24	3	3.21
13.11.5482	AHMAD SETYO UTOMO	3.4	3.48	3.39	3.53	3.61	3.42	3.94
13.11.5483	AGUNG FEBRIANSAH	2.7	2.9	2.89	1.7	3.39	2.97	2.57
13.11.5484	AHMAD PRASETYA PUTRA	2.55	2.9	2.05	1.69	2.91	2.45	3.42

TABLE II. INITIALIZING CENTROID OF CLUSTER

ID	Student Name	GPA-1	GPA -2	GPA -3	GPA -4	GPA -5	GPA -6	GPA-7	Cluster
13.11.5473	AFLAKHATIS RATNA C.	3.35	3.25	3.42	3.35	3.16	3.26	3.31	**A**
13.11.5474	ANIS YULIASIH	3.4	3.25	3.18	3.3	3.29	3.21	3.31	**B**
13.11.5477	AHMAD EFENDI	3.25	3.55	3.18	3.5	3.39	3.11	3.44	**C**
13.11.5481	ADITYA RAMADHAN TRY P.	2.85	3.1	2.92	2.9	3.24	3	3.21	**D**
13.11.5484	AHMAD PRASETYA PUTRA	2.55	2.9	2.05	1.69	2.91	2.45	3.42	**E**

TABLE III. CLUSTERING RESULT IN 4TH ITERATION

ID	Student Name	GPA Per Semester							Dist. to centroid A	Dist. to centroid B	Dist. to centroid C	Dist. to centroid D	Dist to. centroid E	Cluster Candidat
		GPA-1	GPA-2	GPA-3	GPA-4	GPA-5	GPA-6	GPA-7						
13.11.5473	AFLAKHATIS RATNA C.	3.35	3.25	3.42	3.35	3.16	3.26	3.31	0.000	0.313	0.613	1.064	3.210	A
13.11.5474	ANIS YULIASIH	3.4	3.25	3.18	3.3	3.29	3.21	3.31	0.286	0.191	0.570	0.923	3.066	B
13.11.5480	ELFA MIFTAKHUL JANNAH	3.25	3.28	3.29	3	3.26	3.16	3.18	0.433	0.191	0.809	0.736	2.880	B
13.11.5477	AHMAD EFENDI	3.25	3.55	3.18	3.5	3.39	3.11	3.44	0.522	0.519	0.342	1.141	3.291	C
13.11.5482	AHMAD SETYO UTOMO	3.4	3.48	3.39	3.53	3.61	3.42	3.94	0.845	0.933	0.342	1.606	3.641	C
13.11.5475	AHMAD ARIF SETYO UTOMO	2.65	2.5	2.93	3.4	3.16	2.79	3	1.270	1.194	1.563	0.837	2.576	D
13.11.5478	ANI BINTI FITRIAH	3.3	3.28	2.53	3.2	3.45	3	3.13	1.001	0.761	1.054	0.728	2.666	D
13.11.5479	APRILIA DITA HENING SUCI	3.1	3.05	2.89	2.9	3.37	3.13	2.81	0.947	0.693	1.269	0.304	2.484	D
13.11.5481	ADITYA RAMADHAN TRY P.	2.85	3.1	2.92	2.9	3.24	3	3.21	0.899	0.672	1.130	0.339	2.424	D
13.11.5483	AGUNG FEBRIANSAH	2.7	2.9	2.89	1.7	3.39	2.97	2.57	2.057	1.806	2.361	1.206	1.776	D
13.11.5476	APRIZAL EFENDI	2.55	1.44	1.16	0.86	2.68	2.47	2.5	4.091	3.860	4.381	3.158	1.062	E
13.11.5484	AHMAD PRASETYA PUTRA	2.55	2.9	2.05	1.69	2.91	2.45	3.42	2.475	2.230	2.631	1.645	1.062	E