Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control
Journal homepage: http://kinetik.umm.ac.id
ISSN: 2503-2267
Vol. 8, No. 2, May, Pp. 525-534

Integration of fuzzy C-Means and SAW methods on education fee assistance recipients

Abdul Fadlil[1], Imam Riadi [2], Yana Mulyana*[3]
Department of Electrical Engineering, Universitas Ahmad Dahlan, Yogyakarta, Indonesia[1]
Department of Information System, Universitas Ahmad Dahlan, Yogyakarta, Indonesia[2]
Master Program of Informatics, Universitas Ahmad Dahlan, Yogyakarta, Indonesia[3]

Article Info

Keywords:
Fuzzy C-Means, SAW, KEMDIKBUD, KIP-Kuliah, Clustering

Article history:
Received: December 21, 2022
Accepted: March 08, 2023
Published: May 31, 2023

Cite:
A. Fadlil, I. Riadi, and Y. Mulyana, "Integration of Fuzzy C-Means and SAW Methods on Education Fee Assistance Recipients", KINETIK, vol. 8, no. 2, May 2023.
https://doi.org/10.22219/kinetik.v8i2.1636

*Corresponding author.
Yana Mulyana
E-mail address:
yana2107048012@webmail.uad.ac.id

Abstract

Every year, UMTAS gets a quota for KIP tuition fee assistance provided by KEMDIKBUD. This program is intended for high school / vocational/equivalent graduates from poor and vulnerable families. The evaluation results of its implementation have problems, including the number of applicants exceeding the quota given by KEMDIKBUD and some applicants coming from well-off families. This research uses the Fuzzy C-Means method for data clustering and the SAW method for ranking. The results of data grouping using the Fuzzy C-Means method obtained the first cluster (C1) of 72 data and the second cluster (C2) of 119 data. Group C1 is closer to the provisions of aid recipients (eligible) compared to data group C2 (ineligible) because Data C1 consists of 100% DTKS recipients, 50% KIP and KKS card owners, 100% parental income <750,000, 40.28% parental dependents >=2 people and 29.17% applicants with achievements. 72 registrant data included in Data C1 are then ranked using the SAW technique to get weights, and 30 data with the highest weight will be used as a decision on recipients of KIP-Kuliah Education fee assistance according to the quota provided. The optimization of Fuzzy C-Means with SAW methods in selecting recipients of education fee assistance is objective and right on target.

1. Introduction

Education has a significant role for a country, be it a developed country or a developing country because it can improve the welfare and prosperity of its people. Education development programs in each country are different, including Indonesia. The budget allocation of 20% of the total State Budget *(APBN)* is allocated for education [1]. One of the programs carried out by the government in equalizing education is the *KIP-Kuliah* program, where all elements of society classified as poor and vulnerable to poverty have the same right to be able to continue their education to a higher level. The *KIP-Kuliah* program is listed in regulation of the ministry of education and culture (*Permendikbud*) No. 10 of 2020 concerning the Smart Indonesia Program, stating that students who are eligible to be admitted to universities consist of persons with disabilities, students who come from poor/vulnerable families, *KIP* cardholder students, or with special considerations, affirmation students (Papua and West Papua and 3T and TKI) as well as students affected by disasters, social conflicts or special conditions [2].

Universitas Muhammadiyah Tasikmalaya (UMTAS) is one of the private universities that has received a quota for this program from 2018 until now, with a different number of quotas each year. In 2022 the number of quotas received by UMTAS was 30 prospective *KIP-Kuliah* student recipients while the number of applicants is 191 registrants, making it difficult for universities to determine which prospective students are entitled to receive this program following the guidelines provided by financing and education service center. Determining admission under the *KIP-Kuliah* guidelines is a sign that a student is from a disadvantaged background by looking at the ownership of the KIP or KKS card, which makes the admission decision inaccurate. The problem that exists in determining the acceptance of *KIP-Kuliah* is that not all registrants come from poor/vulnerable families so data must be grouped into eligible and ineligible categories. In addition, the number of quotas received by UMTAS is less than the number of registrants therefore ranking must be done. If there is a recipient cancellation, the ranking will facilitate the replacement process. The combination of Fuzzy C-Means algorithm and SAW method in this research is expected to determine the recipients of tuition assistance more objectively, and right on target. Fuzzy C-Means is one of the clustering algorithms that determine results based on the proximity distance in each cluster to the data center point so that it provides optimal results and does not require many variables in its assessment [3]. The basic concept of the Fuzzy C-Means algorithm is to determine the cluster center to calculate the average membership value in each cluster based on the membership degree value. In the initial stages, the centroid assessment is considered less precise and must be repeated from each membership value until the centroid moves closer to the correct membership value [4]. The SAW method is made by finding the sum of the performance values of each characteristic, using the weighting results as a guide for decision-making. The criteria contained in the SAW method are the criteria needed and determined by each person or company [5].

In previous research conducted by Trisudarmo et.al, related to the combination of Fuzzy C-Means and SAW methods in determining scholarship recipients at SMK Auto Matsuda, the Fuzzy method is used for clustering where the most significant value is declared feasible, and the smallest value is declared not feasible, for ranking using the SAW method based on weighting the value of each alternative that is declared feasible. Determining using these two methods in decision-making is effective, efficient, and easy [3]. Furthermore, research conducted by Nunik Destria Arianti and Sardjoeni Moedjiono regarding scholarship acceptance using Fuzzy C-Means and SAW methods states that decision-making can be more objective, more precise, and more efficient than manual decision-making with an accuracy rate of 73.3% [6]. In this study, the author will also use the same method, namely Fuzzy C-Means and SAW, where the Fuzzy C-Means method was used for grouping eligibility data as a recipient of education fee assistance and SAW for ranking. The assessed variables in this method are *DTKS* status, *KIP* and *KKS* card recipients, average parent income per month, a number of parent dependents and student achievement. *DTKS* is information that serves as a reference source for implementing social assistance, including information on income support and social needs [7]. *KIP* is a government program created to overcome problems related to education due to financial difficulties [8]. *KKS* is a conditional social assistance program from the government for poor families [9]. With the existence of this research, the author hopes to help advise UMTAS college leaders in determining prospective recipients of education fees so that they can be more objective and right on target.

2. Research Method

2.1 Planning

An essential overview in making the initial steps of research by planning is examining the problems to be studied and determining the methods to overcome them. The data to be processed is the data of *KIP-Kuliah* registrants at UMTAS in 2022. The data will be grouped into two categories, namely eligible and ineligible. All data members who are considered eligible are then ranked to get results. The initial steps of research conducted by the author can be seen in Figure 1.

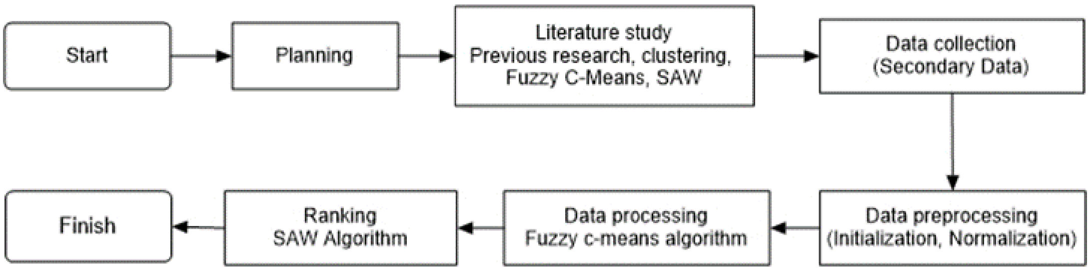

Figure 1. Research Workflow

At this stage, the author identifies problems in the data selection of tuition assistance recipients, where the determination of acceptance still needs to be more targeted. The data clustering method in this study uses the Fuzzy C-Means algorithm. The research to be carried out is to group all registrant data into 2 clusters and perform ranking.

2.2 Literature Research

Literature research is a scientific activity carried out to find answers to a problem. The aim is to provide theoretical or practical contributions to developing relevant disciplines [10]. Literature studies in this study include data processing of education fee assistance receipts at *UMTAS* using the Fuzzy C-Means and SAW methods. The literature study is one of the essential elements in research to examine the problems that will be discussed theoretically through the study of various journals and other references [4].

2.2.1 Metode Fuzzy C-Means

Fuzzy C-Means is one of the fuzzy logic clustering algorithms that determine results based on the proximity distance in each cluster to the data center point to provide optimal results [11]. The basic concept of the Fuzzy C-Means algorithm is to determine the cluster center to calculate the average membership value in each cluster based on its membership degree value [12]. The initial stage of assessing the cluster center is considered less precise. It must be repeated from each membership value until the cluster center moves closer to the correct membership value [13]. Fuzzy C-Means is an unsupervised clustering method in which each cluster data is determined by its membership degree [14]. This method is very good at pattern recognition when the number of clusters is known [15]. Fuzzy C-Means was proposed by Dunn in 1973. Then in 1981, it was further developed by J. C. Bezdek [16],[11]. A Fuzzy C-Means method is a form of Euclidian for the distance between vectors or points in each cluster because it is determined based on the value of the degree of membership [17]. The calculation steps of the Fuzzy C-Means algorithm Equation 1, Equation 2, and Equation 3 are as follows [18]:

1. Define cluster data X, in the form of an n x m matrix (n = number of data samples, m = attributes of each data). X_ij = sample data to i (i=1,2,...,n), attribute to j (j=1,2,...,m).
2. Specify :
   Number of clusters (c) = 2; Rank (w) = 2; Maximum iteration = MaxIter = 100; Smallest expected error (ξ) = 0.01; Initial objective functions (F0) = 0; Initial Interation = t = 1.
3. Distribute a random number μ_{ik}, i=1,2,3 …,n; k=1,2,3…c; as the elements of the initial partition matrix μ.
4. Calculate cluster centers up to k: V_{kj}, with k = 1,2,...,c; and j = 1,2,...,m

$$V_{kj} = \frac{\sum_{i-1}^{n}\left((\mu_{ik})^{w} \text{x } X_{ij}\right)}{\sum_{i-1}^{n}(\mu_{ik})^{w}} \tag{1}$$

5. Calculate the objective function at iterations up to t.

$$P_t = \sum_{i=1}^{n}\sum_{k=1}^{c}\left(\left[\sqrt{\sum_{j=1}^{m}\left(X_{ij} - V_{kj}\right)^{2}}\right](\mu_{ik})^{2}\right) \tag{2}$$

6. Calculating the change of partition matrix.

$$\mu_{ik} = \frac{\left[\sum_{j=1}^{m}\left(X_{ij} - V_{kj}\right)^{2}\right]^{\frac{-1}{w-1}}}{\sum_{k=1}^{c}\left[\sum_{j-1}^{m}\left(X_{ij} - V_{kj}\right)^{2}\right]^{\frac{-1}{w-1}}} \tag{3}$$

7. Checking the stop condition.
   if (|Pt – Pt-1|< ξ) or (t > MaxIter) then it stops.
   if no t = t + 1, repeats to step four.

2.2.2 Simple Additive Weighting (SAW)

The SAW method is one of the multi-attribute decision methods, wherein determining the decision is determined by the weight value in each criterion [5]. This method is used to find the optimal alternative value from many alternatives. The fundamental concept of this method is to find the weighted sum of each performance rating in each alternative attribute [19],[20]. This method requires the normalization process of the decision matrix first to a scale so that it can be compared with all alternative ratings [21],[22]. The Equation 4 is the SAW method equation matrix.

$$r_{ij} = \left\{\begin{matrix} \dfrac{x_{ij}}{max_i x_{ij}} & If\ J\ is\ the\ benefit\ attribute \\ \\ \dfrac{min_i x_{ij}}{x_{ij}} & If\ J\ is\ the\ cost\ attribute \end{matrix}\right\} \tag{4}$$

where:
r_{ij} is a normalized performance rating from the alternative A_I on the attribute C_j
i is 1,2,…, m
j is a sequence 1,2,….,n

The prefence value for each alternative V_i calculated using the following Equation 5.

$$v_i = \sum_{j=1}^{n} w_j r_{ij} \tag{5}$$

where:
v_i is the preference value for each alternative
w_j is the value of weights
r_{ij} is a normalized performance rating value

2.3 Data Collection

The data collection method is the method for obtaining the information needed to achieve research objectives [18]. At this stage, the data collection process is carried out by interviewing the student affairs manager, the head of the academic bureau of student affairs, and the alums of *UMTAS*. The data collection stages are as follows :

1. The author conducts interviews with the head of academic, student and alumni bureau.
2. The author requests data on *KIP-Kuliah* applicants at *UMTAS*.
3. The author categorizes the collected data based on *DTKS* status, *KIP* and *KKS* card recipients, parents' income, number of dependents and achievements.

This study involves all data of registrant at *Universitas Muhammadiyah Tasikmalaya* in 2022 as much as 191 data, the data includes *DTKS* status, ownership of *KIP* and *KKS* cards, parental income, number of parental dependents and achievement levels. All data prepared in .xslx format. can be seen in Table 1.

Table 1. KIP-Kuliah Registrant Data

Registrant Data	Status of DTKS	No. KIP	No. KKS	Parents' Income	Total Dependents	Student Achievement
Data 1	recorded	-	-	-	2	As a Participant of the 2018 Scout Movement Kwartir at the District / City Level
Data 2	unrecorded	F43P6H	-	-	2	
Data 3	recorded	-	-	-	1	
…...	…..	….	….	….	….	….
Data 189	unrecorded	-	-	Rp. 2,500,000	4	
Data 190	unrecorded	-	-	Rp. 2,000,000	4	
Data 191	recorded	P57KC1	-	-	6	

2.4 Preprocessing Data

2.4.1 Data Initialization

Data pre-processing is an important and necessary step in the data mining process because it is closely related to the preparation and creation of the original data set [23]. Data quality determines the data mining process, therefore the data should not be empty or have missing data values, so it must be ensured in advance that the data entered is correct and there is no anomalous data because it can affect the clustering results [24][25][26]. The initialization stage identifies data according to its group and determines the variables used [27]. This stage is carried out to facilitate clustering by converting nominal data into numerical data [28]. The variables or attributes initialized in this study are the *DTKS* status attribute, ownership of *KIP* and *KKS* cards, total parent income, and achievement attributes.

1. Status of *DTKS*

For the *DTKS* status attribute, a change was made from nominal to numeric, the parameters for this assessment are Not detected *DTKS* (1); Not detected *DTKS* but parents' income <Rp.750,000 (2); Not detected *DTKS*, has a *KIP*/*KKS* card (3); and Detected *DTKS* (4).

2. *KIP* and *KKS* card ownership

For the *KIP / KKS* card attribute, changes are made from nominal to numeric, the parameters in this assessment are not having a card (1); having a *KKS* card (2); having a *KIP* card (3); and having a *KIPK* and *KKS* card (4).

3. Parent's income

The attribute of parental income is changed from nominal to numeric, the parameter on parental income is assessed from the amount of average monthly income. Income >Rp.1,750,001 (1); Rp.1,250,001-Rp.1,750,000 (2); Rp.750,001-Rp.1,250,000 (3); and <=Rp.750,000 (4).

4. Achievements

The achievement attribute is changed from nominal to numeric, the parameters for achievement are international level (4); national (3); provincial (2); city / district level achievement (1); and no achievement (0).

The results of initialization on the overall education fee assistance registrant data can be seen in the Table 2.

Table 2. Initialization of Registrant Data

Registrant Data	Status of DTKS	KIP&KKS	Parents' Income	Total Dependents	Student Achievement
Data 1	4	1	4	2	0
Data 2	3	3	4	2	0

Data 3	4	1	4	1	0
…...	…..	….	….	…..	…..
Data 189	1	1	1	4	1
Data 190	1	1	1	4	1
Data 191	2	3	6	6	1

2.4.2 Normalization

Data normalization is a change in the value of each data so that no data is superior in value. Data normalization in this study uses the min-max algorithm where in this calculation the smallest and largest values are sought in the education fee assistance registrant data with a value range of 0-1 so that it will produce a new dataset [29]. The Equation 6 is the algorithm equation *min-max.*

$$X_i = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{6}$$

where:

$\boldsymbol{X_i}$ is normalization
X_{max} is the value of the largest variable
X_{min} is the smallest variable value

The results of the calculation of tuition fee assistance registrant data that has been normalized can be seen in Table 3.

Table 3. Normalization of Registrant Data

Registrant Data	Status of DTKS	KIP&KKS	Parents' Income	Total Dependents	Student Achievement
Data 1	1.00	0.00	1.00	0.17	0.00
Data 2	0.67	0.67	1.00	0.17	0.00
Data 3	1.00	0.00	1.00	0.00	0.00
…...	…..	….	….	…..	…..
Data 189	0.00	0.00	0.00	0.50	0.00
Data 190	0.00	0.00	0.00	0.50	0.00
Data 191	1.00	0.67	1.00	0.83	0.00

3. Results and Discussion

This study uses data on registrants receiving tuition fee assistance *(KIP-Kuliah)* in 2022 at *UMTAS*. The data is obtained from the head of the Academic Administration Bureau of Student Affairs and Alumni in the form of overall registrant data. The registrant data obtained will be grouped, the aim is to find registrants who are eligible and ineligible as recipients with the criteria of *DTKS* Status (SDTK), *KIP* and *KKS* card recipients (KIP&KKS), average parent income per month (RPO), number of parent dependents (JTO) and student achievement (PS). The value of each criterion is obtained directly from the college in the form of an excel file. There are 191 people who register themselves as recipients of tuition fee assistance, but not all registrants are eligible as recipients of this assistance. Therefore, a method is needed to categorize the registrant data, which registrants are eligible and which registrants are not eligible**.** That data is declared eligible and will be ranked using the SAW method.

3.1 Result

The application of Fuzzy C-Means and SAW produces the expected solution in accordance with previous research, where the combination of these two methods is used in the case of scholarship acceptance. The combination of the two methods in determining the recipient of the UMTAS *KIP-Kuliah* produces a ranking of 30 alternatives as recipients of education fee assistance. Levels 1 to 30 can be seen in Figure 2.

In Figure 2 it is explained that the data in the first rank with a weight value of 0.87 is in alternative 57, the second rank with a weight value of 0.84 is in alternative 9, the third rank with a weight value of 0.81 is in alternative 21 and so on until rank 30 with a weight value of 0.68. The 30 alternatives with the highest weights will be used as decisions for tuition assistance recipients and the 31st to 72nd ranks are included in the decision as replacement alternatives.

The distribution of data included in ranks 1 to 30 is data with 29 (96.67%) DTKS data, and 1 (3.33%) not DTKS data. The number of data recorded by KIP or KKS was 26 (86.67%), and those not recorded by KIP or KKS were 4 (13.33%). The number of data on students with achievements is 14 (46.67%) and for students who do not have achievements 16 (53.33%). All data have parental income below five hundred thousand *(Rp. 500,000)* and have several dependents above 2 people. The results of this study indicate that the combination of the Fuzzy C-Means method and the SAW method is very objective and right on target because all 30 selected registrants are data that fall into the category of coming from poor/vulnerable families following the *KIP-Kuliah* guidelines.

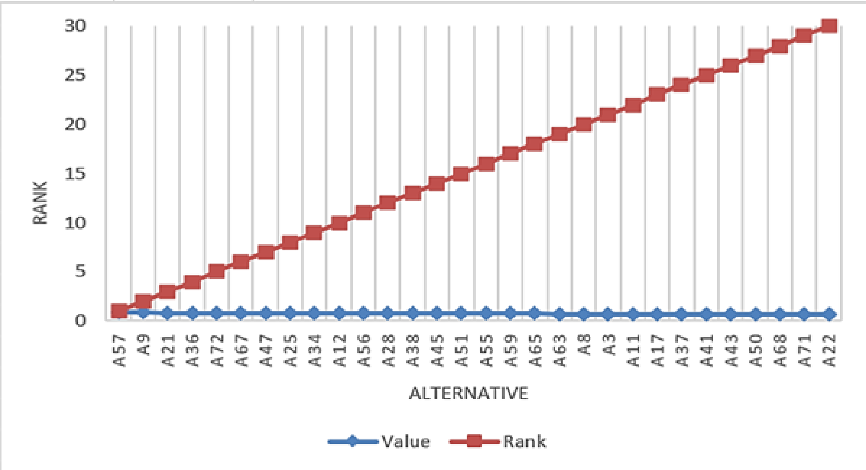

Figure 2. Rank 1 to 30 in Each Alternative

3.2 Discussion

3.2.1 Clustering Registrant Data Using Fuzzy C-Means Method

Fuzzy C-Means is one of the algorithms included in fuzzy logic, this method is a clustering technique that is determined based on the degree of membership of each data. The value of the membership degree in fuzzy has an interval value of [0,1] [30]. In this study, 2 clusters will be made, the initial determination of the cluster is very important because it affects the value of the degree of membership. The results of the calculation of the membership degree value can be seen in Table 4.

Table 4. Membership Degree Value μ_{ik}

Registrant Data	μ_{i1}	μ_{i2}	$(\mu_{i1})^2$	$(\mu_{i2})^2$
Data 1	0.49	0.51	0.24	0.26
Data 2	0.88	0.12	0.78	0.01
Data 3	0.52	0.48	0.27	0.23
…..	…..	…...	…..	…..
Data 189	0.55	0.45	0.30	0.21
Data 190	0.11	0.89	0.01	0.79
Data 191	0.70	0.30	0.49	0.09

Table 4 is the μ_{ik} kernel that can be derived randomly according to the random number generator η as an element of the initial partition matrix U. This calculation begins by forming a random membership with the assumption that each row of records is worth 1. Furthermore, the μ_{ik} value in each cluster is multiplied by a square of 2 so as to produce the membership value μ_{i1} and μ_{i2}, the results of this rank value are used as the next calculation with the value of each attribute of the tuition fee assistance registrant data on Table 3. Calculating the value of attribute data 1 in C1 and C2 which has the results in Table 5.

SDTK 0.24 * 1.00 = 0.24; KIP&KKS 0.24 * 0.00 = 0.00; RPO 0.24 * 1.00 = 0.24; JTO 0.24 * 0.17 = 0.04; PS 0.24 * 0.00 = 0.00.

SDTK 0.26 * 1.00 = 0.26; KIP&KKS 0.26 * 0.00 = 0.00; RPO 0.26 * 1.00 = 0.26; JTO 0.26 * 0.17 = 0.04; PS 0.26 * 0.00 = 0.00.

Table 5. Value Result $(\mu_{ik}{}^2)$*(X_{ij})

Data Pendaftar	SDTK	KIP&KKS	RPO	JTO	PS	SDTK	KIP&KKS	RPO	JTO	PS
Data 1	0.24	0.00	0.24	0.04	0.00	0.26	0.00	0.26	0.04	0.00
Data 2	0.24	0.16	0.24	0.04	0.00	0.17	0.17	0.26	0.04	0.00

Data 3	0.24	0.00	0.24	0.00	0.00	0.26	0.00	0.26	0.00	0.00
…..	…..	…..	…..	…..	…..	…..	…..	…..	…..	…..
Data 198	0.00	0.00	0.00	0.12	0.00	0.00	0.00	0.00	0.13	0.00
Data 190	0.00	0.00	0.00	0.12	0.00	0.00	0.00	0.00	0.13	0.00
Data 191	0.24	0.16	0.24	0.20	0.00	0.26	0.17	0.26	0.21	0.00
Total	39.20	8.88	39.20	14.22	4.40	26.67	9.287	40.98	14.87	4.60

To define a new cluster (V_{kj}) by means of the total value of each data attribute in each cluster divided by the total value of the membership degree. $\left(\frac{X_{ij}}{\mu_{i1}{}^2}\right)$ and $\left(\frac{X_{ij}}{\mu_{i2}{}^2}\right)$. The calculation results can be seen in Table 6.

Table 6. New Cluster (V_{kj})

Cluster (V_{kj})	SDTK	KIP&KKS	RPO	JTO	PS
1	0.59	0.13	0.59	0.21	0.07
2	0.44	0.15	0.67	0.24	0.08

3.2.2 Objective Function (FO)

Calculate the objective function to get the total P (P_t). This P_t is used to calculate the t-th iteration with the maximum iteration = 100 and the smallest expected error (ξ) = 0.01. if the smallest iteration or error has been obtained then the FO calculation is stopped. Calculating data 1 in C1 which has the results in Table 7.

$(X_1 - V_1 1)^2$ = (1.00 – 0.59)² = 0.17; $(X_2 - V_1 2)^2$ = (0.00 – 0.13)² = 0.02; $(X_3 - V_1 3)^2$ = (1.00 – 0.59)² = 0.17; $(X_4 - V_1 4)^2$ = (0.17 – 0.21)² = 0.00; $(X_5 - V_1 5)^2$ = (0.00 – 0.07)² = 0.00.

Counting data 1 in cluster 2 which has the results in Table 8.

$(X_1 - V_2 1)^2$ = (1.00 – 0.44)² = 0.32; $(X_2 - V_2 2)^2$ = (0.00 – 0.15)² = 0.02; $(X_3 - V_2 3)^2$ = (1.00 – 0.67)² = 0.11; $(X_4 - V_2 4)^2$ = (0.17 – 0.24)² = 0.01; $(X_5 - V_2 5)^2$ = (0.00 – 0.08)² = 0.01.

Table 7. Calculation Result (X_i-V_{kj})² in C1

Registrant Data	SDTK	KIP&KKS	RPO	JTO	PS	Total
Data 1	0.17	0.02	0.17	0.00	0.00	0.36
Data 2	0.01	0.28	0.17	0.00	0.00	0.47
Data 3	0.17	0.02	0.17	0.05	0.00	0.41
…..	…..	…..	…..	…..	…..	…..
Data 198	0.34	0.02	0.34	0.08	0.00	0.79
Data 190	0.34	0.02	0.34	0.08	0.00	0.79
Data 191	0.17	0.28	0.17	0.38	0.00	1.01

Table 8. Calculation Result (X_i-V_{kj})² in C2

Registrant Data	SDTK	KIP&KKS	RPO	JTO	PS	Total
Data 1	0.32	0.02	0.11	0.01	0.01	0.46
Data 2	0.05	0.26	0.11	0.01	0.01	0.44
Data 3	0.32	0.02	0.11	0.06	0.01	0.51
…..	…..	…..	…..	…..	…..	…..
Data 198	0.19	0.02	0.45	0.07	0.01	0.74
Data 190	0.19	0.02	0.45	0.07	0.01	0.74
Data 191	0.32	0.26	0.11	0.35	0.01	1.04

After calculating $\sum_{j=1}^{m}\left(X_{ij} - V_{kj}\right)^2$ in C1 and C2, to further calculate $C_i{*}(\mu_{ik})^2$ $\sum_{k=1}^{c}\left(\left[\sum_{j=1}^{m}\left(X_{ij} - V_{kj}\right)^2\right](\mu_{ik})^w\right)$ then add up the results of the overall count to get the value (P_t). The calculation results can be seen in Table 9.

0.36 * 0.244 = 0.09; 0.46 * 0.256 = 0.12.

Table 9. Calculation Result $C_i{*}(\mu_{ik})^{\wedge 2}$ and (P_t)

Registrant Data	C1*$(\mu_{ik})^{\wedge 2}$	C2*$(\mu_{ik})^{\wedge 2}$	P_t
Data 1	0.09	0.12	0.21
Data 2	0.36	0.01	0.37
Data 3	0.11	0.12	0.23

…..	…..	…..	…..
Data 198	0.24	0.15	0.39
Data 190	0.01	0.58	0.59
Data 191	0.50	0.09	0.59

The result for the FO value in the first iteration is **63.50**, if FO $P_t - P_{t-1}$ is smaller than the error or the specified iteration has reached the maximum then the iteration is stopped. if not then continue calculating the iteration by redetermining the cluster center using the new μ_{ik} function.

3.2.3 Change of Partition Matrix

The change in the partition matrix is to determine the new μ_{ik} value for the next iteration calculation with a maximum result equal to 1 in each calculation. Calculating the new μ_{ik} for data 1 and having the results in Table 10.
$(0.36)^{-1} + (0.46)^{-1}$ = 2.74 + 2.18 = 4.92; μ_{i1}= 2.74/4.92 = 0.56; μ_{i2}= 2.1/4.92 = 0.44

Table 10. New Membership Degree (μ_{ik})

Registrant Data	${\mu_{i1}}^{-1}$	${\mu_{i2}}^{-1}$	Total	(μ_{i1})	(μ_{i2})
Data 1	2.74	2.18	4.92	0.56	0.44
Data 2	2.14	2.29	4.43	0.48	0.52
Data 3	2.45	1.95	4.40	0.56	0.44
…..	…..	…..	…..	…..	…..
Data 198	1.26	1.36	2.61	0.48	0.52
Data 190	1.26	1.36	2.61	0.48	0.52
Data 191	0.99	0.96	1.95	0.51	0.49

The clustering process in this study uses Fuzzy C-Means, the clustering data is successfully obtained at the 20th iteration with an FO value of 83.68, because the error value obtained is smaller than the specified error value of 0.006, the iteration process is stopped. Data clustering results using the Fuzzy C-Means method in this study can be seen in Table 11.

Table 11. Results of Fuzzy C-Means Clustering

Registrant Data	μ_{i1}	μ_{i2}	Proximity	Cluster
Data 1	0.59	0.41	0.59	C1
Data 2	0.46	0.54	0.54	C2
Data 3	0.54	0.46	0.54	C1
…..	…..	…..	……	……
Data 189	0.06	0.94	0.94	C2
Data 190	0.06	0.94	0.94	C2
Data 191	0.73	0.27	0.73	C1

Data on recipients of education cost assistance are grouped into 2 clusters, namely C1 as many as 72 data and C2 as many as 119; when viewed as a whole, C1 is considered as a data group that deserves tuition assistance, and C2 as an uneligible data group. The grouping results declared eligible as assistance recipients on C1 data using the Fuzzy C-Means method are then given a ranking using the SAW method.

3.2.4 Ranking Registrant Data with SAW Method

The results of clustering using the Fuzzy C-Means method will be processed again using the SAW method for ranking as many as 72 data that are declared eligible as recipients of tuition fee assistance. The input parameters for calculating this method are taken from the data of all registrants who are declared eligible and then divided into 2, namely alternative data (A) and criteria data (C). alternative data is all registrant data and criteria data is an attribute of each alternative. Suitability rating data for each alternative on each criterion can be seen in Table 12.

Table 12. Suitability Rating Data for Each Alternative

Alternative	C1	C2	C3	C4	C5
A1	4	1	4	2	0
A2	4	1	4	1	0
A3	4	3	4	3	0
….	…..	…..	……	……	……
A70	4	3	4	2	0

A71	4	3	4	3	0
A72	4	3	4	6	0

The calculation in this method is that all values in each alternative on all criteria are match values so that the largest value is the best because all existing criteria values are assumed to be benefit criteria. The weight value given in this method for each criterion is C1 = 4; C2 = 4; C3 = 2; C4 = 3 and C5 = 2. so that after normalization it becomes W = 0.27; 0.27; 0.13; 0.20; 0.13. The calculation of the normalization matrix on alternative 1 in each criterion has the results in Table 13.

$r_1 1 = \left(\frac{(4)}{\max\{4;4;4;\ldots.;4;4;4\}}\right) = \frac{4}{4} = 1.00$; $r_1 2 = \left(\frac{(1)}{\max\{1;1;3;\ldots.;3;3;3\}}\right) = \frac{1}{4} = 0.25$; $r_1 3 = \left(\frac{(4)}{\max\{4;4;4;\ldots.;4;4;4\}}\right) = \frac{4}{4} = 1.00$; $r_1 4 = \left(\frac{(4)}{\max\{2;1;3;\ldots.;2;3;6\}}\right) = \frac{4}{6} = 0.33$; $r_1 5 = \left(\frac{(0)}{\max\{0;0;0;\ldots.;0;0;0\}}\right) = \frac{0}{3} = 0.00$

Table 13. Normalized Matrix r

Alternative	C1	C2	C3	C4	C5
A1	1.00	0.25	1.00	0.33	0.00
A2	1.00	0.25	1.00	0.17	0.00
A3	1.00	0.75	1.00	0.50	0.00
…..	…..	…..	…..	…..	…..
A70	1.00	0.75	1.00	0.33	0.00
A71	1.00	0.75	1.00	0.50	0.00
A72	1.00	0.75	1.00	1.00	0.00

The results of the data normalization calculation in Table 13 are the sum of the attribute values in each data. The process of calculating the ranking of alternative 1 on each criterion, the calculation results can be seen in Table 14.
$V_1 = (1.00*0.27) + (0.25*0.27) + (1.00*0.13) + (0.33*0.20) + (0.00*0.13) = 0.53$

Table 14. Ranking Results on Each Alternative

Alternative	Value Weight	Rank
A1	0.53	64
A2	0.50	69
A3	0.70	21
…..	…..	…..
A70	0.67	37
A71	0.70	29
A72	0.80	5

4. Conclusion

Grouping data on recipients of *KIP-Kuliah* tuition fee assistance using the Fuzzy C-Means method applies membership degrees to each attribute value in grouping the data. before entering into testing, all 191 data were normalized, testing this method was carried out by determining the number of clusters as many as 2, data attributes as many as 5 and the specified error value of 0.01. This test stops at the 20th iteration with an error value of 0.006, this test gets the results of grouping data in the first cluster (C1) as much as 72 data and the second cluster (C2) as much as 119 data. Data grouping in C1 is closer to the requirements of aid recipients than the C2 data group because C1 data consists of 100% *DTKS* recipients, 50% of *KIP* and *KKS* card owners, 100% of parents' income <750,000, 40.28% of parents' dependents >=2 people and 29.17% of outstanding applicants. Ranking the data in C1 using the SAW method was successfully carried out by getting a ranking of 1-72. The first rank with a weighted value of 0.87 is in alternative 57, the second rank with a weight value of 0.84 is in alternative 9, the third rank with a weight value of 0.81 is in alternative 21 and so on until rank 30 with a weight value of 0.68. Ranks 1-30 with the highest weight will be used as beneficiaries. In comparison, the 31st to 72nd rankings are included in the decision as a replacement alternative. Optimization the Fuzzy C-Means method with the SAW method in selecting recipients of education fee assistance is objective and right on target.

References

[1] F. Sabilah and S. Salahudin, “Public Policy Analysis on Education Budget Allocation: Case Study in Thirty-Eight Regencies/Municipalities in East Java, Indonesia,” *J. Stud. Pemerintah.*, vol. 13, no. 1, pp. 59–85, 2022. https://doi.org/10.18196/jgp.v13i1.12529

[2] Permendikbud No 10 Tahun 2020, “Peraturan Mentri Pendidikan dan Kebudayaan Republik Indonesia Nomor 10 Tahun 2020 Tentang Program Indonesia Pintar,” Menteri Pendidik. dan Kebud. RI, p. 82, 2020.

[3] R. Trisudarmo, E. Sediyono, and J. E. Suseno, “Combination of Fuzzy C-Means Clustering Methods and Simple Additive Weighting in Scholarship of Decision Support Systems,” *Proc. 1st Annu. Int. Conf. Nat. Soc. Sci. Educ. (ICNSSE 2020)*, vol. 547, no. Icnsse, pp. 161–169, 2021. https://doi.org/10.2991/assehr.k.210430.025
[4] W. Gunawan and B. S. P. Diwiryo, “Implementasi Algoritma Fuzzy C-Means Clustering Sistem Crowdfunding pada Sektor Industri Kreatif Berbasis Web,” *J. Edukasi dan Penelit. Inform.*, vol. 6, no. 2, p. 193, 2020. http://dx.doi.org/10.26418/jp.v6i2.38018
[5] M. A. Kasri and H. Jati, “Combination of K-Means and Simple Additive Weighting in Deciding Locations and Strategies of University Marketing,” *Khazanah*, vol. 6, pp. 132–141, 2020. https://doi.org/10.23917/khif.v6i2.11281
[6] N. D. Arianti and S. Moedjiono, “DSS for scholarship recipients using the fuzzy C-means clustering integrated with the simple additive weighting method,” Int. J. Sci. Technol. Res., vol. 8, no. 9, pp. 2048–2050, 2019.
[7] E. V. Manoppo, N. A. Laoh, I. Pemerintahan, D. Negeri, P. Provinsi, and S. Utara, “Strategi Pemanfaatan Data Terpadu Kesejahteraan Sosial (DTKS) dalam Penyaluran Bantuan Sosial RS-RTLH oleh Dinas Sosial Provinsi Sulawesi Utara,” vol. 4, no. April. pp. 25–39, 2022. https://doi.org/10.33701/jk.v4i1.2598
[8] A. Amin, R. N. Sasongko, and A. Yuneti, “Kebijakan Kartu Indonesia Pintar untuk Memerdekakan Mahasiswa Kurang Mampu,” *J. Adm. Educ. Manag.*, vol. 5, no. 1, pp. 98–107, 2022. https://doi.org/10.31539/alignment.v5i1.3803
[9] R. T. Allopa and T. Pabendon, “Program Kartu Keluarga Sejahtera: Efektifkah? (Bukti Empiris di Kampung Mulia Kencana Kabupaten Mimika),” *ESEOSI J. Econ. Reg. Sci.*, vol. 2, no. 2, pp. 139–155, 2022. https://doi.org/10.52421/jurnal-esensi.v2i2.223
[10] A. K. Wardhani, “Implementasi Algoritma K-Means untuk Pengelompokkan Penyakit Pasien pada Puskesmas Kajen Pekalongan,” J. Transform., vol. 14, no. 1, pp. 30–37, 2016.
[11] C. Ais, A. Hamid, D. Candra, and R. Novitasari, “Analysis of Livestock Meat Production in Indonesia Using Fuzzy C-Means Clustering,” *J. Ilmu Komput. dan Inf. (Journal Comput. Sci. Information)*, vol. 15, no. 1, pp. 1–8, 2022. https://doi.org/10.21609/jiki.v15i1.993
[12] F. R. Hariri, “Implementation of Fuzzy C-Means for Clustering the Majelis Ulama Indonesia (MUI) Fatwa Documents,” *J. Online Inform.*, vol. 6, no. 1, p. 79, 2021. https://doi.org/10.15575/join.v6i1.591
[13] D. L. Rahakbauw, V. Y. I. Ilwaru, and M. H. Hahury, “Implementasi Fuzzy C-Means Clustering Dalam Penentuan Beasiswa,” J. Ilmu Mat. dan Terap., vol. 11, pp. 1–12, 2017.
[14] Z. Li, S. Chen, L. Pei, J. Chu, and J. Song, “Teacher Allocation and Evaluation Based on Fuzzy C-Means Clustering,” *Mob. Inf. Syst.*, vol. 2022, no. 4, pp. 1–9, 2022. https://doi.org/10.1155/2022/8465713
[15] L. Wang and Y. Jiang, “Collocating Recommendation Method for E-Commerce Based on Fuzzy C-Means Clustering Algorithm,” *J. Math.*, vol. 2022, pp. 1–11, 2022. https://doi.org/10.1155/2022/7414419
[16] Y. Li, J. Gou, and Z. Fan, “Educational data mining for students performance based on fuzzy C-means clustering,” *J. Eng.*, vol. 2019, no. 11, pp. 8245–8250, 2019. https://doi.org/10.1049/joe.2019.0938
[17] M. Wulandariyaningsih, A. N. Hasmi, and S. Pancahayani, “Sistem Pendukung Keputusan Penentuan Uang Kuliah Tunggal Mahasiswa Dengan Metode Fuzzy C-Means (Studi Kasus Institut Teknologi Kalimantan),” *Teorema Teor. dan Ris. Mat.*, vol. 7, no. 1, p. 117, 2022. http://dx.doi.org/10.25157/teorema.v7i1.6701
[18] M. A. Pamungkas, H. Oktavianto, and R. Umilasari, “Perbandingan Fuzzy C-Means Dan K-Means Untuk Mengelompokkan Tingkat Buta Huruf Berdasarkan Provinsi Di Indonesia,” 2021.
[19] G. Setiadi and W. Hadikurniawati, “Implementasi Metode Hybrid AHP-SAW-TOPSIS Untuk Pemilihan Taman TOGA,” *J. Inform.*, vol. 9, no. 1, pp. 18–25, 2022. https://doi.org/10.31294/inf.v9i1.11901
[20] R. Fauzan, Y. Indrasary, and N. Muthia, “Sistem Pendukung Keputusan Penerimaan Beasiswa Bidik Misi di POLIBAN dengan Metode SAW Berbasis Web,” *J. Online Inform.*, vol. 2, no. 2, p. 79, 2018. https://doi.org/10.15575/join.v2i2.101
[21] R. Sovia, E. P. W. Mandala, and S. Mardhiah, “Algoritma K-Means dalam Pemilihan Siswa Berprestasi dan Metode SAW untuk Prediksi Penerima Beasiswa Berprestasi,” *J. Edukasi dan Penelit. Inform.*, vol. 6, no. 2, p. 181, 2020. http://dx.doi.org/10.26418/jp.v6i2.37759
[22] I. Septiana, M. Irfan, A. R. Atmadja, and B. Subaeki, “Sistem Pendukung Keputusan Penentu Dosen Penguji Dan Pembimbing Tugas Akhir Menggunakan Fuzzy Multiple Attribute Decision Making dengan Simple Additive Weighting (Studi Kasus: Jurusan Teknik Informatika UIN SGD Bandung),” *J. Online Inform.*, vol. 1, no. 1, p. 43, 2016. https://doi.org/10.15575/join.v1i1.10
[23] A. Muhariya, I. Riadi, and Y. Prayudi, “Cyberbullying Analysis on Instagram Using K-Means Clustering,” *JUITA J. Inform.*, vol. 10, no. 2, p. 261, 2022. http://dx.doi.org/10.30595/juita.v10i2.14490
[24] A. M. Fadhilah, M. I. Wahyuddin, and D. Hidayatullah, “Analisis Faktor yang Mempengaruhi Perokok Beralih ke Produk Alternatif Tembakau (VAPE) menggunakan Metode K-Means Clustering,” *J. JTIK (Jurnal Teknol. Inf. dan Komunikasi)*, vol. 5, no. 2, p. 219, 2020. https://doi.org/10.35870/jtik.v5i2.182
[25] R. Setiawan, “Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Promosi Mahasiswa Baru (Studi Kasus : Politeknik Lp3i Jakarta),” J. Lentera Ict, vol. 3, no. 1, pp. 76–92, 2016.
[26] D. P. M and A. Fadlil, “Penerapan Clustering K-Means untuk Pengelompokan Tingkat Kepuasan Pengguna Lulusan Perguruan Tinggi,” vol. 6, pp. 1693–1700, 2022. http://dx.doi.org/10.30865/mib.v6i3.4191
[27] Y. S. Siregar and P. Harliana, “Algoritma Fuzzy C-Means Pada Aplikasi Matlab Dalam Menentukan Dosen Pembimbing Tugas Akhir,” Semin. Nas. Unisla, pp. 213–217, 2018.
[28] N. Noviyanto, “Penerapan Data Mining dalam Mengelompokkan Jumlah Kematian Penderita COVID-19 Berdasarkan Negara di Benua Asia,” *Paradig. - J. Komput. dan Inform.*, vol. 22, no. 2, pp. 183–188, 2020. https://doi.org/10.31294/p.v22i2.8808
[29] Pelsri Ramadar Noor Saputra and A. Chusyairi, “Perbandingan Metode Clustering dalam Pengelompokan Data Puskesmas pada Cakupan Imunisasi Dasar Lengkap,” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 4, no. 6, pp. 5–12, 2020. https://doi.org/10.29207/resti.v4i6.2556
[30] A. M. Anter and M. Ali, “Feature selection strategy based on hybrid crow search optimization algorithm integrated with chaos theory and fuzzy c-means algorithm for medical diagnosis problems,” *Soft Comput.*, vol. 24, no. 3, pp. 1565–1584, 2020. https://doi.org/10.1007/s00500-019-03988-3