Implementation of K-Means and SAW Methods in Determining Non-Cash Food Assistance Recipients

Yunita[*1], Rizki Kurniati[2], Desty Rodiah[3], Allsela Meiriza[4], Luh Sri Mulia Eni[5]
[1,2,3,4,5]Fakultas Ilmu Komputer, Universitas Sriwijaya, Indonesia
E-mail: [*1]**yunita@ilkom.unsri.ac.id**, [2]rizki.kurniati@ilkom.unsri.ac.id,
[3]destyrodiah@gmail.com, [4]allsela@unsri.ac.id, [5]luhsrimuliaeni3@gmail.com

Abstract

Determination of prospective non cash food assistance recipients, especially in Air Talas village, still uses a manual system so that in the process of determining the recipient there is a risk that the recipient will be inaccurate, so that the village government needs a system that can assist the process of determining prospective non cash food assistance recipients. This study aims to implement the K-Means and SAW methods in determining recipients of non cash food assistance in Air Talas village. The benefits of this research can help the Air Talas village government in determining and recommending prospective non cash food assistance recipients in accordance with established criteria, making it easier to filter, group, and rank appropriate population data according to criteria. In addition, this research is also useful for providing convenience to the community through data collection, clustering, and ranking in a transparent, real, and fast and accurate manner using decision support system software. The K-Means clustering method and the Simple Additive Weighting Ranking method were used in this study with data collection techniques through interviewing sources, in this case the village government, the social section of the community, and through collecting village archive data and relevant journals. The research location is Air Talas village with 316 data used. The results of the study are clustering data as much as 77 data obtained from feasible clusters. The cluster data was then tested using the accuracy value and obtained a value of 80%. Then the research is also in the form of ranking data using clustered data which obtains an accuracy value of 64%.

Keywords *— Non Cash Food Assistance, K-Means, Simple Additive Weighting*

1. INTRODUCTION

Community social protection will be provided to Beneficiary Families (KPM) from low-income groups/poor and vulnerable families in the form of food social assistance. Food social assistance aims to meet their food needs and reduce the burden of spending on poor and vulnerable families. The Food Social Assistance Program (formerly Rastra) was changed to Non-Cash Food Assistance (BPNT) in 2017. The Rastra Grant program was fully transformed into the Food and Social Assistance Program which was channeled through the Innatura Program and the Rastra Social Assistance Program in 2018[1]. Food and social assistance programs in all districts/cities were implemented at the end of 2019 with a cashless system or BPNT. BPNT is the government's effort to change the form of assistance to non-cash (cashless). That is, using an electronic card that is forwarded directly to the KPM.

This social assistance is channeled to KPM using the banking system, which can then be used to obtain nine staple goods (sembako) at e-Warong, so that KPM can meet their food needs. This is an effort from the Government to provide KPM access to basic ingredients with other nutritional content. The Food Poverty Line has a major influence on the formation of the Poverty Line (according to data from the Central Bureau of Statistics (BPS)). The stability of food prices has an impact on reducing the poverty rate. Through the Sembako program, it is hoped that it can reduce the expenditure burden of poor families in terms of food, so that some of the basic needs of the poor can be fulfilled[2]. The distribution of BPNT in the Rambang Dangku sub-district, Air Talas Village, has several obstacles, which can be seen from the right target indicators, which are still high in Exclusion Error and Inclusion Error. The exclusion error, for example, is that there are still households that should receive BPNT, but are not registered as BPNT recipients, so the community does not receive the BPNT social assistance. Meanwhile, the Inclusion Error, on the other hand, is that there are households that are not entitled to receive BPNT, instead receiving the assistance. The problem currently being faced is that the determination of the recipient of the BPNT funds must compare the survey results one by one so that it is found who is most entitled to receive the aid funds. In this way it can cause a relatively high complexity both in terms of time, accuracy of results and can affect the target object of the recipient of assistance. The reason is that there is no system that can assist in determining the decision to determine recipients of aid funds based on criteria determined by the government.

In order to make the data processing and calculation more effective, in this study the K-Means and Simple Additive Weighting (SAW) methods were applied for clustering and ranking data based on predetermined criteria. The implementation of the K-Means method is carried out because its clustering ability is easy to do even though it uses large data and the computational process is relatively short by determining the centroid value[3]. While the ranking process uses the SAW method because this method is easy to implement and simple in its calculations by multiplying the weighted criteria[4]. These two methods are combined with the aim of grouping the resulting data and displaying them on a list of beneficiaries according to the criteria. In this case, K-Means functions in the data sorting process, dividing the group with the existing alternative values into sub-groups with criterion values which then the sub-groups resulting from this sorting will be used in determining the final alternative results which are processed using the SAW method.

2. RESEARCH METHOD

The research methodology in this study is described as follows:

2.1. Data Mining

Data mining is defined as a technique of extracting confidential information from a database with a large collection of data so that an interesting pattern has not been found before. Data mining has several characteristics, namely related to the discovery of something hidden with a certain pattern that was not previously known, using large data so that the results are more reliable, and useful for making critical and strategic decisions[5].

2.2. K-Means

K-Means is a non-hierarchical clustering method used to group data into one or more clusters. In one cluster there are data with the same characteristics and data with different characteristics will be grouped in another cluster, so that the data in one cluster has a very small level of variation[6].

The steps that must be taken to carry out the clustering process using the K-Means method are[7]:

1. Determine the number of clusters to be formed and the maximum iteration.
2. Randomly assigns as many centroid values as needed.
3. Calculates the distance of each input data to all centroids to obtain the closest distance from the data to the centroid. To calculate this distance, the Eucledian Distance formula is used, as follows:

$$d(xi,\mu j) = \sqrt{(x_i - \mu_j)^2} \quad (1)$$

Grouping data into clusters with a minimum centroid distance from each data.

Calculate the new centroid. The new centroid is obtained by calculating the average value of the data in the same cluster using the following formula:

$$\mu_j(t+1) = \frac{1}{N_{Sj}} \sum_{j \in Sj} X_j \quad (2)$$

Information: j(t + 1) = new centroid in iteration (t + 1) NSj = lots of data in cluster Sj
Repeat steps 2 to 5 so that no more data moves to other clusters.

2.3. Simple Additive Weighting (SAW)

The Simple Additive Weighting (SAW) method is a method used in the ranking process based on preference values. The SAW method, also known as the weight addition method, works by adding up the weights of the criteria on each alternative, then normalizing the decision matrix (X) on a comparison scale with existing alternative ratings[8].

There are several steps or procedures that must be carried out to rank using the SAW method[8] , namely:

1. Calculating the normalization of the alternative matrix, so that the normalized matrix R is obtained. The normalization of the matrix is carried out using the following equation:

$$r_{ij} = \begin{cases} \frac{x_{ij}}{\max x_{ij}}, & benefit \\ \frac{\min x_{ij}}{x_{ij}}, & cost \end{cases} \quad (3)$$

Information:
r_{ij} = normalized matrix [i][j]
x_{ij} = decision matrix [i][j]
max xi = maximum value of each column of the decision matrix
min xi = minimum value of each column of the decision matrix

2. Calculate the preference value, using the following equation:
Information:

$$V_i = \sum_{j=1}^{n} w_j \, r_{ij} \quad (4)$$

Information:
V_i = alternative final value
w_j = weight value
r_{ij} = normalized criterion value
n = number of data

3. Sorts preference values as a result of ranking.

The Research Framework is described in Figure 1.

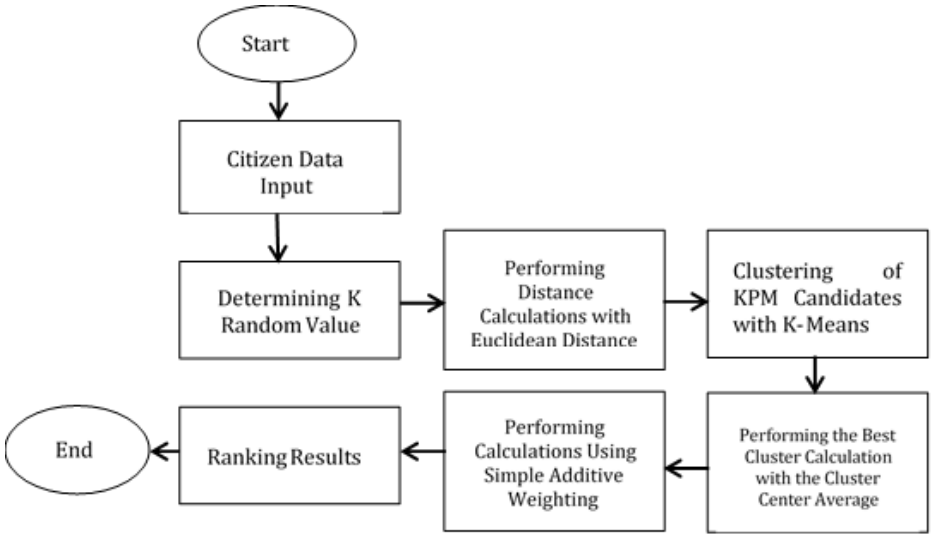

Figure 1. Research Framework

3. RESEARCH RESULTS AND DISCUSSION

3.1. Data Transformation

Before the clustering and ranking process is carried out, it is necessary to transform preprocessed data as categorical data into postprocessed data in the form of numeric data. The amount of data used in the clustering process is 316 data. The transformation of the criteria data is shown in table 1 and the weight of the criteria is shown in table 2.

Table 1. Criteria Data Transformation

Kondisi	Pendapatan Keluarga	Bobot	Nilai
0	4 Juta Rupiah atau Lebih	20	0
1	3 Juta – 3,9 Juta	20	20
2	2 Juta – 2,9 Juta	20	40
3	1,1 Juta – 1,9 Juta	20	60
4	500 Ribu – 1 Juta	20	80
5	Tidak Memiliki Pendapatan	20	100

Kondisi	Kepemilikan Aset	Bobot	Nilai
1	Toko Grosir	10	10
2	Warung	10	20
3	Kebun	10	30
4	Rumah	10	40
5	Menumpang	10	50

Kondisi	Tanggungan	Bobot	Nilai
0	Tidak ada tanggungan	15	0
1	1 – 2 Orang	15	15
2	3 – 4 Orang	15	30
3	5 – 6 Orang	15	45
4	> 6 Orang	15	60

Kondisi	Pekerjaan	Bobot	Nilai
0	PNS	5	0
1	Polisi	5	5
2	Petani	5	10
3	Pegawai	5	15
4	Wiraswasta	5	20
5	Pedagang	5	25
6	Sopir	5	30
7	IRT	5	35
8	Buruh	5	40
9	Tidak Bekerja	5	45

Table 2. Criteria Weight Data

Kondisi	Keterangan	Bobot (%)	Nilai
1	Pekerjaan	5	Benefit
2	Pendapatan Per Bulan	20	Cost
3	Kepemilikan Aset	10	Benefit
4	Tanggungan	15	Benefit

Source: [9]

3.2. Clustering Process

The clustering process in this study uses the K-Means method to divide the data on prospective BPNT recipients into several clusters, which in this study are divided into 3 clusters, namely the eligible clusters, less feasible, and not eligible to receive BPNT.

3.3. K-Means Clustering

From 316 population data of prospective BPNT recipients after clustering using the K-Means method, the results are shown in tables 3 and 4:

Table 3. Initial Centroid Value (Random)

Cluster	Kriteria			
	Pekerjaan	Pendapatan	Kepemilikan Aset	Tanggungan
C1	40	80	40	45
C2	10	20	70	30
C3	40	80	50	15

Table 4. Final Centroid Value

Atribut	*Cluster*		
	C1	C2	C3
Pekerjaan	38,240740740741	10,972972972973	35
Pendapatan	57,407407407407	32,886486486486	80
Kepemilikan Aset	41,666666666667	70,594594594595	45,844755844156
Tanggungan	25,833333333333	21,72972972973	18,701298701299
Mean	40,787037037	34,045945946	44,8863636363

After analyzing the final centroid value, cluster C1 has an average value of 40.787037037, cluster C2 has an average value of 34.045945946, and cluster C3 has an average final centroid value of 44.8863636363. Clusters with an average final centroid value will be used as a feasible cluster whose data will be ranked using the SAW method which in this study obtained a C3 cluster value greater than the C1 and C2 cluster values so that the data in the C3 cluster will be ranked.

3.4. Ranking with SAW

From the data obtained from the results of clustering C3 cluster as a cluster that deserves to be ranked. In cluster C3 there are 77 data on prospective BPNT recipients who are ready to rank using the SAW method. The ranking results using SAW as shown in table 5.

Table 5. SAW Ranking Results

Nama	Rank	Nama	Rank	Nama	Rank	Nama	Rank
Puji Hartono	1	M. Sabar	21	Jamaludin	41	Wagina	61
Wayan Yuliansa	2	Sodikin	22	Rahmat Hidayat	42	Satinem	62
Roni Azhar	3	Heri Saiful U.	23	Al Budiman	43	Rukiah	63
Bustamin Bakri	4	Gd. Sudi	24	Sudirman	44	Tina Marleni	64
Gede Edi	5	Rengki Anggel	25	Rojak	45	Ngatjem	65
M. Ramadhan	6	Angga Jetara	26	Wy. Gunarta	46	Handriani	66
Rusmanto	7	Panera	27	Zulkarnain Cahyadi	47	Er-Fahani	67
Rahmad	8	Agus Heriyanto	28	Gd. Sandat	48	Gd. Ginarda	68

Nama	Rank	Nama	Rank	Nama	Rank	Nama	Rank
Mulyadi	9	Gd. Sukarta	29	Ny. Redia Laksana	49	Mawisah	69
Abdul Hadi	10	Gd. Suryawan	30	I ngh. Wirta	50	Tugimin	70
Vedri	11	Md. Kris Permana	31	Agus Arip Pudin	51	Ranggi Noveli	71
Ely Setiawan	12	Km. Alit Udayana	32	Made Ria Respartane	52	Ketut Dinastra	72
Julianto	13	Km. Tri Swarna	33	Zainuddin	53	Nyoman Gatri	73
Tri Joko N.	14	Rizal Ahmadi	34	Km. Dauh	54	Wy. Diasa	74
Abdul Kodir	15	Kd. Aryana Suitra	35	Ketut Wiyadi	55	Wy. Wirya	75
Tukiman	16	Luh Suri	36	Gd. Sastra Wiguna	56	Wy. Buda Ada	76
I Gede Gemode	17	Topik Hidayat	37	Nengah Budiarta	57	Ketut Sumeriji	77
Ny. Astawa	18	Mukti	38	Md. Susila	58		
Ny. Serina	19	Dian Reza	39	Kadek Suprama	59		
Kd. Sudiartha	20	Bambang	40	Nopiyanti	60		

3.5. K-Means Method Accuracy Testing

In this study to determine the accuracy of the K-Means clustering method used accuracy calculations. In general the accuracy is calculated as shown in equation 5:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (5)$$

Table 6. Confusion Matrix

		Nilai sebenarnya	
		TRUE	*FALSE*
Nilai prediksi	*TRUE*	*True Positive* (TP)	*False Positive*(FP)
	FALSE	*False Negatif* (FN)	*True Negative* (TN)

From the results of clustering using K-Means, the data that deserved to receive BPNT were 77 data from 316 data. The target of BPNT recipients is 50 data and the actual data for BPNT recipients is 120, so the calculation is obtained as follows:

Tabel 7. Confusion Matrix K-Means

		Nilai sebenarnya	
		TRUE	*FALSE*
Nilai prediksi	*TRUE*	77 (TP)	48 (FP)
	FALSE	43 (FN)	268 (TN)

$$Accuracy = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

$$Accuracy = \frac{77 + 268}{77 + 268 + 48 + 43}$$

$$Accuracy = \frac{345}{436}$$

$$= 0{,}80$$

$$= 80\%$$

3.6. SAW Ranking Accuracy Test

From the ranking results using SAW data compared with the actual recipient data by using the number of target data as much as 50 data obtained 32 data from 50 valid data and 18 invalid data so that the calculation of SAW accuracy is obtained as follows:

$$Accuracy = \frac{\sum data\ valid}{\sum \text{data uji perbandingan}}$$

$$= \frac{32}{50}$$

$$= 0{,}64$$

$$= 64\%$$

4. CONCLUSION

Based on the discussion that has been described regarding the application of the method K-Means and SAW in determining BPNT recipients, it can be concluded that the K-Means method was successfully implemented by dividing the data into 3 clusters so as to produce 77 data that deserve to be ranked. The results of clustering are tested by calculating the accuracy value which shows the accuracy of prediction with an accuracy value of 80%. In addition, ranking using the Simple Additive Weighting (SAW) method was also successfully implemented in this study by obtaining a ranking comparison value between the actual data and the system output as many as 32 valid data and 18 invalid data, so that when calculated the accuracy obtained a value of 64%.

From the results of clustering and ranking tests using the method K-Means and SAW, there are several results that have the exact same final total value so that the ranking determination is still based on the typical cluster sequence so that researchers have suggestions that other researchers who want to develop this research can use, namely to develop this research using other ranking methods such as the Promethee and the Electre method..

5. REFERENCES

[1] E. Y. Yunus, "IMPLEMENTASI PROGRAM BANTUAN PANGAN NON TUNAI (BPNT) DI KECAMATAN KANIGARAN KOTA PROBOLINGGO," REFORMASI, vol. 9, no. 2, 2019, doi: 10.33366/rfr.v9i2.1454.

[2] Kemensos, "Pedoman Umum Program Sembako," 2020. https://docs.google.com/viewer?docex=1&url=https://kemensos.go.id/uploads/topics/15828061583725.pdf (accessed Jan. 22, 2022).

[3] R. Akhsani, A. Sukmana, H. D. Ramandita, and Kusrini Kusrini, "PENERAPAN K-MEAN PADA METODE SAW UNTUK SISTEM PENDUKUNG KEPUTUSAN PEMILIHAN KARYAWAN TERBAIK," Semnasteknomedia, vol. 6, no. 1, pp. 1–2, 2018, Accessed: Jan. 22, 2022. [Online]. Available: https://ojs.amikom.ac.id/index.php/semnasteknomedia/article/download/2054/1863.

[4] A. Andreas, A. Wiryadinata, and H. Agung, "Penerapan Algoritma Simple Additive Weighting untuk Membantu Dalam Menentukan Calon Penerima Bantuan Pangan Non Tunai," KALBISCIENTIA J. Sains dan Teknol., vol. 6, no. 1, 2020, doi: 10.53008/kalbiscientia.v6i1.56.

[5] R. A. Indraputra and R. Fitriana, "K-Means Clustering Data COVID-19," J. Tek. Ind., vol. 10, no. 3, 2020.

[6] K. Pamungkas, T. S.; Nugroho, A. S.; Wasiso, I.; Anggoro, T.; Kusrini, "Decision Support System for Direct Target Cash Recipients Using the AHP and K-Means Method," Res. J. Comput. Inf. Syst. Technol. Manag., vol. 3, no. 2, pp. 45–54, 2020.

[7] Y. Yunita, R. Effendi, and D. P. Rini, "Metode K-Means & SAW dalam Seleksi Penerima Dana Zakat pada Badan Amil Zakat," Teknomatika, vol. 9, no. 2, pp. 163–174, 2019.

[8] N. D. Saksono, Y. A. Sari, and R. K. Dewi, "Rekomendasi Lokasi Wisata Kuliner Menggunakan Metode K-Means Clustering Dan Simple Additive Weighting," J. Pengemb. Teknol. Inf. dan Ilmu Komput. Univ. Brawijaya, vol. 2, no. 10, 2018.

[9] M. Belanawane S., Kelembagaan Penyediaan Dan Penyaluran Komoditas Bantuan Pangan Non-Tunai (BPNT). 2020.