

Implementation Fuzzy C-Means on Decision Support System BPNT (Bantuan Pangan Non-Tunai) Ministry of Social Affairs Indonesia

Aji Setiawan, Jordan Nur Akbar

Darma Persada University
Jl. Terusan Casablanca, Pondok Kelapa, Jakarta, Indonesia
Tel: (021) - 8649059
E-mail: aji_setiawan@ft.unsada.ac.id

Received October 23, 2019; Revised November 18, 2019; Accepted December 8, 2019

Abstract

Decision Support System can be an alternative solution to determine the candidate's decision. Bantuan Pangan Non-Tunai (BPNT) are selected based on criteria established by the Ministry of Social Affairs of the Republic of Indonesia. BPNT recipients are conducted by the government to help someone who is less able to meet their daily needs. The occurrence of errors in determining the eligibility of prospective beneficiaries is a significant problem, based on these problems there needs to be an information system that can provide a valid BPNT recommendation and one of which uses a grouping method with the Fuzzy C-Means (FCM) algorithm. System development using the waterfall method. The results of system implementation and testing showed that 90% of the system was following what expected according to the results of the test with the system build.

Keywords: Decision Support System, Fuzzy C-Means (FCM), BPNT

1. INTRODUCTION

Poverty is one of the problems that must be considered by both the central and regional governments. Based on the results of data published by the Central Statistics Agency (BPS) [1] in particular, the province of DKI Jakarta in the period 2017 - 2019 recorded an increase in the poverty line of DKI 1 to 1.1% each year. KEMSOS has determined criteria that can use as guidelines by stakeholders in determining potential recipients of Non-Cash Food Assistance (BPNT). In the field conditions, often the invalid recipients of assistance such as poverty and so on are not correctly applied, and technical errors occur where BPNT recipients do not match the criteria determined by KEMENSOS. So far, the calculation of eligibility for beneficiaries has been

carried out manually without using a method and only using proposals. To make the management and calculation process more effective, a decision support system is created by applying a technique, the Fuzzy C-Means (FCM) method. The basic concept of FCM by finding a weighted sum of the performance ratings on each alternative performance on all attributes, as a record of the Decision Support System method that is built only as an alternative input for stakeholders, the final decision is still made by the authorities.

The concept of a decision support system or Decision Support System (DSS) designed to support all stages of decision making ranging from identifying problems, selecting relevant data and determining the approach used in the decision-making process, to evaluating the selection of alternatives based on existing data knowledge [2] Decision support systems have used in various aspects of life such as for trading purposes [3], health [4], plantations [5], to the field of transportation [6]. The decision support system that will build helps assess each low-income family makes changes to the criteria, and change the parameter values. This medium is useful to facilitate decision-makers related to the selection problem of BPNT recipients, especially in the Pondok Kelapa Jakarta, so that the final result will get the most deserving of assistance.

2. RELATED WORKS

The need for application design regarding the selection process for the acceptance of Non-Cash Food Assistance and the absence of applications used in the Pondok Kelapa is the reason for the design of the "Decision Support System for the Selection of Acceptance of BPNT (Non-Cash Food Assistance)". This application is web-based and uses the Fuzzy C-Means method. This method requires several supporting parameters to determine the decisions in the calculation process. Using this method is expected to provide results to determine the recipients of Non-Cash Food Assistance. The results of the decision formulated with the C-Means algorithm, some preliminary studies have been carried out such as [7] making a comparison between K-Means with C-Means research results found that comparing algorithms based on the quality of grouping and performance depends on the time complexity between the various number of clusters selected by the user.

Yohanes, in his research, successfully demonstrated that Fuzzy C-Means is better in terms of computing to find the degree of membership of each cluster in grouping data. The results of clustering for each data and the number of clusters in each algorithm varies according to its complexity [8]. Some development of C-Means done by modifying the merging of clustering techniques between GSA and C-Means, as done by Mulyanto et al. [9], getting the conclusion that gravitational search algorithm (GSA) is combined into FCM to find the optimal cluster center by minimizing objective functions FCM. The results show that the proposed method of gravitational search algorithm fuzzy c-means (GSA-FCM) can show more optimal results than the

FCM algorithm. And the last reference, Awangga, in his research used Fuzzy K-Means in determining the operational activities of the ministry of energy and mineral resources which proved to be able to show optimal activity ratings [10].

3. ORIGINALITY

Poverty is one of the problems that must be solved seriously, based on our research background, we try to develop new approaches to reduce the problem of poverty. With a system based on the Fuzzy c-means (FCM) approach in previous studies, the application uses more K-means clusters. This Object Research for program BPNT under the social ministry (KEMENSOS).

The approach of FCM is basically to record the average value for each collection, which will later be used as the center of the cluster where the main requirement is the center value of the cluster is still inaccurate. Each data has a participation level for each cluster, by fixing the central value cluster and membership each data regularly until it finds a value that approaches the central cluster. FCM Algorithm [11]:

- a) Input the data into cluster x , in the form of matrix $n \times m$

$$x = \begin{bmatrix} x_{11} & \dots & x_{1m} \\ \vdots & \ddots & \vdots \\ x_{n1} & \dots & x_{nm} \end{bmatrix} \quad (1)$$

Information : n is a number of samples of data, m is a attribute of each data, X_{ij} is a sample data to- i ($i = 1, 2, \dots, n$), and attribute j ($j = 1, 2, \dots, m$).

- b) Determine, explain about cluster and iteration with details.

Information : c = The number of clusters, w = Rank, MaxIter = Maximum iterations, ξ = Least Expected Error, 0 = Initial Objective function, Early iterations = $t = 1$;

- c) Generating the random numbers μ_{ik} , $i = 1, 2, \dots, n$; $k = 1, 2, \dots, c$; as elements of the matrix U . μ_{ik} initial partition is the degree of membership that refers to how likely the data could be members into a cluster. Calculates the sum of each column (attribute):

$$Q_j = \sum_{k=1}^c \mu_{ik} \quad (2)$$

Information: C = Total Cluster, Q_j = number of degrees of membership values each, column = 1 with $i = 1, 2, \dots, m$, k = attribute to 1, μ_{ik} = degree of membership Q_j is the sum of the values of degrees of membership each column = 1 to $j = 1, 2, \dots, m$

Calculate :

$$\mu_{ik} = \frac{\mu_{ik}}{Q_j} \quad (3)$$

- d) Calculate the center of the cluster k - k: K_j , with $k = 1, 2, \dots, c$; and $j = 1, 2, \dots, m$

$$V_{kj} = \frac{\sum_{i=1}^n ((\mu_{ik})^w * X_{ij})}{\sum_{i=1}^n (\mu_{ik})^w} \quad (4)$$

$$x = \begin{bmatrix} x_{11} & \dots & x_{1m} \\ \vdots & \ddots & \vdots \\ x_{n1} & \dots & x_{nm} \end{bmatrix} \quad (5)$$

Information: V_k = Cluster Center, n = Total Sample Data, e = Rank, X_{ij} = sample data to- i ($i = 1, 2, \dots, n$), attribute j ($j = 1, 2, \dots, m$)

- e) Calculate the objective function at iteration = TPT:

$$Pt = \sum_{i=1}^n \sum_{k=1}^c ([\sum_{j=1}^m (X_{ik} - V_{kj})^2] (\mu_{ik})^w) \quad (6)$$

Information: P_t = Objective Function, c = Number of Cluster, n = Total Sample Data, m = attribute of each data, V_{kj} = Cluster Center, X_{ij} = sample data to- i ($i = 1, 2, \dots, n$), attribute j ($j = 1, 2, \dots, m$), μ_{ik} = degree of membership t is the calculated iteration, if iteration starts from 1 then at the beginning of calculation the value of t is 1. The iteration will be repeated in accordance with the provisions of the iteration that is running.

- f) Calculate the changes of matrix partition:

$$\mu_{ik} = \frac{[\sum_{j=1}^m (X_{ik} - V_{kj})^2]^{\frac{-1}{w-1}}}{\sum_{k=1}^c [\sum_{j=1}^m (X_{ik} - V_{kj})^2]^{\frac{-1}{w-1}}} \quad (7)$$

Information : μ_{ik} = degree of membership, X_{ik} = sample data to- i ($i = 1, 2, \dots, n$), attribute k (with $k = 1, 2, \dots, c$), V_{kj} = Cluster Center, c = Number of Cluster, m = attribute of each data, w = exponent value This step is to change the parameter values used in the fuzzy c-means algorithm. The goal is to see the effect of the specified parameter values. In this study, the parameters used are height, weight, and age.

- g) Check the stops condition.

If: ($| Pt - Pt-1 | < \epsilon$) or ($t > \maxIter$) then stops. If not: $t = t + 1$, repeat step 4 Iterations will repeat if certain values or conditions have not reached.

The condition is if: ($| Pt - Pt - 1 | < \epsilon$) or ($t > \text{MaxIter}$), then stop which Pt is the center of the iteration cluster to t less than the expected error value or if t (number of iterations) is greater than Maximum iteration. However, if iteration will repeat with $t + 1$, then it will repeat the 4th process and count the center of the cluster again.

4. SYSTEM DESIGN

The development of this research uses the waterfall method approach. Waterfall describes the steps of activities undertaken, such as steps down the waterfall with several stages of study including [12].

1. Analysis Requirement

In this study, the method based on data processing, the data obtained is the data of prospective beneficiaries who have been surveyed by the team. Data collected from interviews with the object of research in the area of Pondok Kelapa, East Jakarta.

2. System Design

Entity Relationship Diagram (ERD) used to represent the subject domain decomposition of a system into an entity.

3. Coding & Testing

In this stage, we make the application using program codes that are in line with the original purpose of web-based programming.

4. Integration & Testing

In this stage, we do application socialization and also testing, so we know whether there is an error or not after it has done.

5. Operation & Maintenance

In this stage, the development and maintenance of the system that builds whether there will be errors or not.

5. EXPERIMENT AND ANALYSIS

1. Analysis Requirement

In this study, the method based on data processing, the data obtained is the data of prospective beneficiaries who have been surveyed by the team. Data collected from interviews with the object of research in the area of Pondok Kelapa, East Jakarta. The data will process by clustering the feasibility of the recipient of the aid object into 2 clusters, namely the clusters is accepted, and the clusters rejected. The clustering process uses the Fuzzy C-Means (FCM) method. At this stage, the authorization of users and the functions of each user is determined.

2. System Design

Entity Relationship Diagram (ERD) consists of nodes, which represent entity types, and edges, which represent relationships [13]. ERD table can be seen in Figure 1.

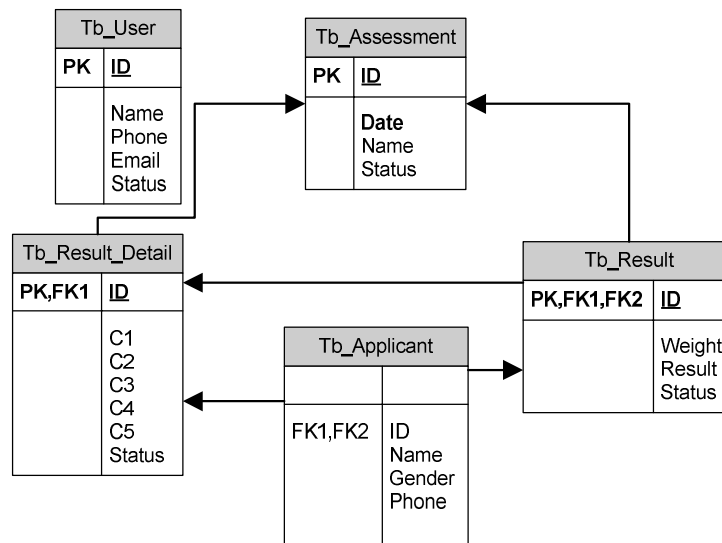


Figure 1. ERD Table

In the testing phase of the FCM method, the first step taken is to determine the number of clusters. The condition is accepted or agreed to be entered in cluster 1 (C1), and if the state is rejected, it is included in cluster 2 (C2) as in the data table 1.

Table 1. Determination of the Number of Clusters & Iterations

Number of Clusters	MaxIteration (t)
2	2

Cluster determination aims to group results based on the closest value of each cluster. Furthermore, determining the rank value in each variable that will measure in this study the work, income, dependents, housing, and property variables as in table 2.

Table 2. Determination of Rank Value (W) Work

No	Occupations	Weight (W)
1	Permanent Employee	0,1
2	Entrepreneurs	0,2
3	Entrepreneurship	0,3
4	Motorcycle Taxis (Ojek)	0,4
5	Driver	0,5
6	Scavengers	0,6
7	Labor	0,7
8	Masons	0,8
9	Doesn't work	0,9

Based on the results of the interview, it founded that the conditions for prospective recipients of non-cash food assistance were as follows: 1. Indonesian citizens, 2. The main occupation of the head of the family (laborers, scavengers, masons, motorbike drivers (ojek), taxi drivers,

public transportation drivers, not working), 3. The income of the head of the family under 1 million rupiah, does not have his own house (rent). From this data, a cluster of values created for each job. The income or salary component is also a consideration of the appropriateness of whether or not assistance given to the community; for this reason, it is also necessary to provide a rank based on monthly salary.

Table 3. Determination of Rank Value (W) Revenue

No	Income	Weight (W)
1	More than 1 million rupiah	0,1
2	less than 1 million rupiah	0,7

Next is the dependency component as an assessment of the eligibility of aid recipients. The number of dependents divided into three clusters, namely 1-2, 3-4, and more than 5, and each cluster is given a rank value (w).

Table 4. Determination of Rank Value (W) Dependents

No	Dependents	Weight(W)
1	1 - 2 person	0,1
2	3 - 4 person	0,7
3	More than 5 person	0,9

Status of homeownership divided into two types, namely own house or rent, each given a rank value. In table 6 is the valuation of the asset component, where the intended property is the jewelry used. These assets use in calculating the eligibility of participants.

Table 5. Determination of Rank Value (W) House Status

No	House Status	Weight (W)
1	Own	0,3
2	Rent	0,7

Table 6. Determination of Rank Value (W) Assets

No	Assets	Pangkat (W)
1	Have jewelry	0,35
2	Don't have jewelry	0,75

The next step is to normalize the data of 2 potential beneficiaries.

Table 7. Normalization of Prospective Beneficiaries

No	Applicant	Occupation	Salary	Dependent	House	Assets
1	Tito Nurahman	0,1	0,1	0,7	0,3	0,35
2	Surya Adi Jaya	0,6	0,7	0,1	0,7	0,75

Next, generate the partition matrix ($n \times c$) with random values formula $\text{rand}()$, and proceed with normalizing the partition matrix.

Table 8. Partition Matrix

No	C1	C2	SUM
1	0,002	0,942	0,944
2	0,857	0,356	1,213

Normalization is done by dividing each element by line. The results of this division will form the normalization of the partition matrix.

Table 9. Partition Matrix Normalization

No	C1	C2	SUM
1	0,002118644	0,997881356	1
2	0,706512778	0,293487222	1

If you get the normalized value of the partition matrix, then calculate the center of the cluster and generate matrix v.

Table 10. Cluster Center

No	Xi	X2	X3	X4	X5	C1	C2	C1 ²	C2 ²
1	0,1	0,1	0,7	0,3	0,35	0,2946	0,7053	0,0868	0,4975
2	0,6	0,7	0,1	0,7	0,75	0,2917	0,7082	0,0851	0,5016
Tot	0,7	1	0,4	1	1,1	0,5863	1,4136	0,1719	0,9991

Table 11. Generating Matrix V

No	Cluster	X1	X2	X3	X4	X5
1	C1	0.3475107 9699576	0.49601727 519322	0.20099568 120169	0.49800863 759661	0.54800863 759661
2	C2	0.3510326 0877108	0.50165217 403373	0.19958695 649157	0.50082608 701686	0.55082608 701686

Table 12. Calculating the Objective Function (P)

No	Xi	X2	X3	X4	X5	Dx11	Dx12	Dx13	Dx14	Dx15
1	0,1	0,1	0,7	0,3	0,35	0,0612	0,1568	0,0098	0,0392	0,0392
2	0,6	0,9	0,1	0,7	0,75	0,0637	0,1632	0,0102	0,040	0,0408

Table 13. Formation of Partition Matrix μ

No	C1	C2	C1 ⁻¹	C2 ⁻¹	SUM
1	0,3063	0,3150	3,2646	3,1737	6,4384
2	0,3187	0,3099	3,1372	3,2265	6,3638

Table 14. Update Partition Matrix μ

No	C1	C2	Cluster	Result
1	0.047575049964145	0.04893856544539	1	A
2	0.050088505415272	0.048700944883815	2	B

Based on the results of calculations with the FCM algorithm, the results show that Surya Adi Jaya is entitled to receive assistance while Tito Nur Rahman is not allowed to BPNT.

2. Implementation

The results of the design process and script tested to see the functionality of each function. The following are the results of the implementation of the system build.

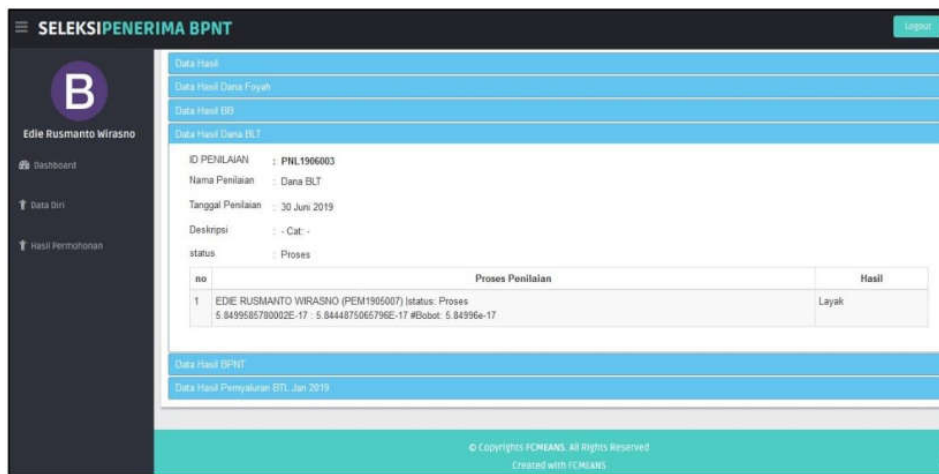


Figure 2. FCM calculation results page

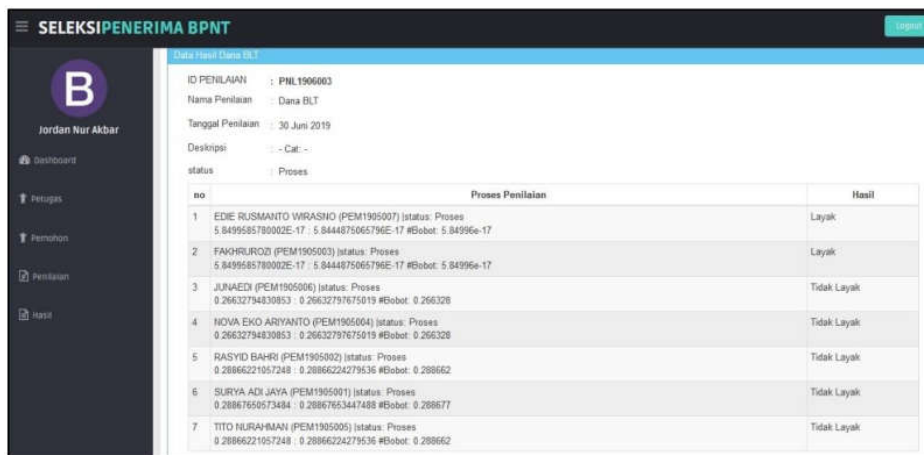


Figure 3. Recapitulation page for BPNT Recipient Candidates

Table 15. Recapitulation Acceptance BPNT

Cases	Manual Test					Result	System Test					Result	Information
	W	S	D	H	A		W	S	D	H	A		
Case 1	0,1	0,1	0,7	0,3	0,35	Not feasible	0,1	0,1	0,7	0,3	0,35	Not feasible	Valid
Case 2	0,6	0,7	0,1	0,7	0,75	Worthy	0,6	0,7	0,1	0,7	0,75	Worthy	Valid
Case 3	0,7	0,1	0,7	0,3	0,35	Not feasible	0,7	0,1	0,7	0,3	0,35	Not feasible	Valid
Case 4	0,5	0,7	0,7	0,7	0,35	Worthy	0,5	0,1	0,1	0,7	0,75	Not feasible	Invalid
Case 5	0,3	0,7	0,7	0,3	0,35	Not feasible	0,3	0,7	0,7	0,3	0,35	Not feasible	Valid
Case 6	0,4	0,1	0,1	0,7	0,75	Not feasible	0,4	0,1	0,1	0,7	0,75	Not feasible	Valid
Case 7	0,1	0,1	0,9	0,3	0,35	Not feasible	0,1	0,1	0,9	0,3	0,35	Not feasible	Valid

Information : W = work, H = home, D = dependents, S = salary, A = assets

6. CONCLUSION

The conclusions obtained from research on the application of FCM for the selection of BPNT participants are as follows (1). The making of this application uses the waterfall system development method, which begins with interviewing the needs of the user, then continues with the design of the overall appearance than the process of implementing the system into coding. (2) The results of system testing conducted by daily performing officers on Pondok Kelapa, the results show that 90% application is useful between manual testing with the system so that it can conclude that it can use as consideration for BPNT recipient decisions in the East Jakarta. (3) For further research, it can be developed with other fuzzy method or machine learning approaches to compare the test results with FCM (4) Can conduct cross-sectoral research to look for patterns of relationship between the number and profile of beneficiary participants.

REFERENCES

- [1] Badan Pusat Statistik, **Kemiskinan dan Ketimpangan**, *Badan Pusat Statistik - Kemiskinan dan Ketimpangan*, 2019.
- [2] B. A. Alyoubi, **Decision Support System and Knowledge-based Strategic Management**, in *Procedia Computer Science*, 2015.
- [3] S. L. Chan and W. H. Ip, **A dynamic decision support system to predict the value of customers for new product development**,

- Decis. Support Syst.*, 2011.
- [4] J. Mysiak, C. Giupponi, and P. Rosato, **Towards the development of a decision support system for water resource management**, *Environ. Model. Softw.*, 2005.
- [5] H. Navarro-Hellín, J. Martínez-del-Rincon, R. Domingo-Miguel, F. Soto-Valles, and R. Torres-Sánchez, **A decision support system for managing irrigation in agriculture**, *Comput. Electron. Agric.*, 2016.
- [6] M. Nourinejad and M. J. Roorda, **A dynamic carsharing decision support system**, *Transp. Res. Part E Logist. Transp. Rev.*, 2014.
- [7] T. Velmurugan and T. Santhanam, **A comparative analysis between K-Medoids and fuzzy C-Means clustering algorithms for statistically distributed data points**, *J. Theor. Appl. Inf. Technol.*, 2011.
- [8] Yohanes, **Analisis Perbandingan Algoritma Fuzzy C-Means dan K-Means**, *Annu. Res. Semin. 2016*, 2016.
- [9] A. Mulyanto and R. S. Wahono, **Penerapan Gravitational Search Algorithm untuk Optimasi Klasterisasi Fuzzy C-Means**, *J. Intell. Syst.*, 2015.
- [10] R. M. Awangga, S. F. Pane, and K. Tunnisa, **Collaboration FMADM And K-Means Clustering To Determine The Activity Proposal In Operational Management Activity**, *Emit. Int. J. Eng. Technol.*, 2019.
- [11] R. Suganya and R. Shanthi, **Fuzzy C-Means Algorithm-A Review**, *Int. J. Sci. Res. Publ.*, 2012.
- [12] A. Powell-Morse, **Waterfall Model: What Is It and When Should You Use It?**, *Airbrake*, 2016. .
- [13] R. J. Wieringa, **Entity-Relationship Diagrams**, in *Design Methods for Reactive Systems*, 2007.