Zero : Jurnal Sains, Matematika, dan Terapan
E-ISSN : 2580-5754; P-ISSN : 2580-569X
Volume 10, Number 1, 2026
DOI: 10.30829/zero.v10i1.28831
Page: 351-364

A Decision Support System for Selecting Beneficiaries of the Free Nutritious Meal Program Using Fuzzy Possibilistic C-Means

[1] M. Husen Al Farisy
Department of Mathematics Education, Universitas Nahdlatul Ulama Sunan Giri, 62115, Indonesia

[2] M. Ivan Ariful Fathoni
Department of Mathematics Education, Universitas Nahdlatul Ulama Sunan Giri, 62115, Indonesia

[3] Anisa Fitri
Department of Mathematics Education, Universitas Nahdlatul Ulama Sunan Giri, 62115, Indonesia

[4] Nabila Alisa Rohmah
Department of Mathematics Education, Universitas Nahdlatul Ulama Sunan Giri, 62115, Indonesia

[5] Siti Nur Laili Zuhri
Department of Mathematics Education, Universitas Nahdlatul Ulama Sunan Giri, 62115, Indonesia

Article Info

Article history:

Accepted 15 May 2026

Keywords:

Clustering Algorithm;
Decision Support System;
Free Nutritious Meal Program;
Fuzzy Possibilistic C-Means;
Social Assistance Targeting.

ABSTRACT

The Free Nutritious Meal Program (MBG) aims to improve students' nutritional status, but beneficiary selection is challenged by data uncertainty and overlapping criteria. This study proposes a decision support system based on the Fuzzy Possibilistic C-Means (FPCM) algorithm. The dataset consists of 200 students from Bojonegoro Regency using five indicators: nutritional status, household income, family size, participation in social assistance, and distance to school. The methodology includes data standardization, clustering using Fuzzy C-Means (FCM) and FPCM, and evaluation using Partition Coefficient, Silhouette Score, and Dunn Index. The results classify students into three groups: 52% not eligible, 18.5% moderately eligible representing borderline cases requiring further verification, and 29.5% eligible. FPCM achieves higher cluster clarity with a Partition Coefficient of 0.84 compared to 0.74 in FCM, while other evaluation metrics indicate comparable structural quality between methods. These findings indicate that FPCM provides a more interpretable and robust framework for decision-making under uncertainty.

Corresponding Author:

M. Husen Al Farisy,
Department of Mathematics Education
Universitas Nahdlatul Ulama Sunan Giri, Bojonegoro, Indonesia
Email: husenalfa@unugiri.ac.id

1. INTRODUCTION

Social protection programs are widely recognized as one of the most effective policy instruments to reduce poverty, improve welfare, and promote social justice in developing countries. Among these initiatives, food and nutrition assistance programs hold strategic importance, as malnutrition remains a major public health challenge affecting human capital development and long-term national productivity. In this regard, UNICEF emphasizes that nutritious food programs play a critical role in reducing childhood stunting and malnutrition, promoting healthier pregnancies, and improving children's learning outcomes and overall health status. In Indonesia, the Free Nutritious Meal Program (MBG) has been designed to address nutritional deficiencies among vulnerable groups such as school-age children, pregnant women, and breastfeeding mothers [1]. The program is also aligned with the national agenda of Indonesia Emas 2045, which envisions a healthy, educated, and productive population as the foundation of sustainable development.

However, the implementation of the MBG program faces a critical challenge: ensuring that assistance is delivered to the right beneficiaries. Mis-targeting, reflected in inclusion errors (benefits given to ineligible individuals) and exclusion errors (eligible individuals being left out), undermines the effectiveness of the program. These errors not only reduce fairness but also generate inefficiencies in resource allocation, particularly considering the massive 2025 budget allocation of IDR 71 trillion for approximately 19.47 million recipients. With such a large financial commitment, inaccurate targeting could jeopardize public trust, compromise the sustainability of the program, and even pressure other strategic sectors such as higher education and healthcare. More broadly, mistargeting in social assistance programs whether in the form of inclusion errors or exclusion errors represents a major challenge in ensuring equity and efficiency in Indonesia's social policy, particularly in the context of large fiscal commitments such as the 2025 MBG allocation [2].

Data-driven approaches, particularly clustering-based methods such as Fuzzy C-Means (FCM), have been widely applied in social assistance and educational decision-making contexts, including scholarship allocation, food voucher distribution, and welfare targeting [3], [4], [5]. These approaches help reduce subjectivity compared to manual selection procedures; however, discrepancies between algorithmic results and field decisions still occur due to overlapping socio-economic characteristics [6]. Despite its flexibility in handling heterogeneous data, FCM is sensitive to noise, overlapping class boundaries, and incomplete data, which may lead to unstable clustering results and biased estimations [7], [8], [9]. Furthermore, existing implementations often produce rigid or binary decisions, limiting their ability to support nuanced interpretation in borderline cases within complex policy environments [10], [11].

To address these limitations, the Fuzzy Possibilistic C-Means (FPCM) algorithm, which extends the possibilistic clustering framework originally introduced by Krishnapuram and Keller [12], was proposed by Pal et al. [13]. It combines fuzzy membership and possibilistic typicality to improve robustness against noise and overlapping clusters. Unlike conventional FCM, FPCM allows each data point to be evaluated not only in relative terms (membership) but also in absolute terms (typicality), making it more suitable for complex socio-economic data. Previous studies have demonstrated that FPCM provides more stable and accurate clustering results compared to conventional methods, particularly in datasets with uncertainty and overlapping characteristics [14], [15], [16], [17].

Despite these advances, previous studies on Fuzzy Possibilistic C-Means (FPCM) have primarily focused on improving clustering performance in technical domains such as medical and general data analysis. However, the integration of FPCM into decision support systems (DSS) for real-world social welfare targeting remains limited. In particular, existing approaches often do not address the need for interpretable and policy-relevant decision mechanisms that can accommodate uncertainty, overlapping eligibility criteria, and ambiguity in socio-economic data. This limitation is critical, as beneficiary selection in large-scale welfare programs requires not only accurate grouping but also transparent and graded decision support to minimize mis-targeting and improve policy effectiveness. Therefore, there is a need to extend FPCM beyond its conventional use as a clustering algorithm toward a decision support framework that explicitly incorporates uncertainty in beneficiary targeting.

Several alternative approaches have been proposed to address uncertainty in clustering and welfare targeting problems, including evidential clustering, rough-fuzzy hybrid models, and supervised machine learning methods. Evidential clustering frameworks are capable of representing uncertainty through belief assignments, while rough-fuzzy approaches explicitly model boundary regions in overlapping datasets. In contrast, supervised learning models such as decision trees and support vector machines may provide strong predictive performance when reliable labeled data are available. However, these approaches often face limitations in practical welfare targeting contexts, where labeled data are incomplete, interpretability is essential, and socio-economic boundaries are inherently ambiguous. In this regard, FPCM offers a suitable balance between robustness to noise, flexibility in handling overlapping characteristics, and interpretability through the integration of membership and typicality values.

In addition, previous Indonesian studies on welfare targeting and educational assistance have predominantly relied on conventional FCM approaches that produce rigid cluster assignments without explicitly incorporating typicality-based interpretation [5], [6], [8]. As a result, borderline cases are often insufficiently represented within the decision-making process. This study advances prior works by extending FPCM beyond a conventional clustering technique into a decision support framework that supports graded recommendations, identifies ambiguous cases requiring further verification, and enhances transparency in beneficiary selection under uncertainty.

In this study, five critical variables are employed to evaluate potential MBG beneficiaries: nutritional status (BMI), household income, family size, participation in other social assistance programs, and distance to school. These variables have been widely recognized in previous literature as important determinants of eligibility for welfare programs [8], [18]. By incorporating them into the FPCM framework, the proposed system seeks to produce more objective, accurate, and equitable classifications of beneficiaries.

The contribution of this research lies not only in demonstrating the technical advantages of FPCM but also in contextualizing its application within a high-impact national program. The objectives of this study are threefold: (1) to improve the accuracy and objectivity of MBG beneficiary selection; (2) to address uncertainty and noise in social data through FPCM-based clustering; and (3) to develop a decision support system that enhances transparency and can be adapted to other large-scale welfare programs. This research therefore contributes both methodologically, by advancing the application of FPCM in social decision-making contexts, and practically, by supporting the effectiveness and sustainability of Indonesia's flagship nutrition program.

2. RESEARCH METHOD

This study employs a computational experimental framework to develop and evaluate a decision support system (DSS) for beneficiary selection based on the Fuzzy Possibilistic C-Means (FPCM) algorithm. The methodology consists of five main stages: data collection and preprocessing, implementation of Fuzzy C-Means (FCM), implementation of FPCM, evaluation and comparison, and system interpretation. The experimental scenario is designed to compare the clustering performance and interpretability between FCM and FPCM in the context of beneficiary classification. Both algorithms are applied to the same standardized dataset, and the resulting cluster structures, membership values, and typicality values are analyzed to evaluate their suitability for decision support.

The selection of parameters in both FCM and FPCM plays a crucial role in determining clustering behavior and stability. The fuzzification coefficient is set to $m = 2$, which is widely adopted in the literature as it provides a balanced trade-off between cluster fuzziness and convergence stability. Lower values of m tend to produce hard clustering behavior, while higher values increase cluster overlap and may reduce interpretability. Therefore, $m = 2$ is commonly considered an optimal choice for maintaining meaningful fuzzy partitions.

In the FPCM algorithm, the possibilistic coefficient is set to $\eta = 2$ because it provides a balanced influence between membership and typicality components. The parameter η controls the spread of typicality values, where smaller values may lead to excessive sensitivity to noise, while larger values may overly smooth the typicality distribution. The choice of $\eta = 2$ ensures a stable representation of absolute data compatibility while preserving the robustness of the clustering process, as supported by previous studies [13], [17].

The parameter selection in this study follows commonly adopted settings in fuzzy clustering literature. The fuzzification coefficient was set to $m = 2$ because it provides a balanced trade-off between cluster fuzziness and stability, while avoiding excessively hard or overly diffuse partitions. Preliminary experimental trials with nearby parameter values produced comparable clustering structures, indicating that the selected parameters provide sufficiently stable results for the socio-economic dataset used in this study.

Both FCM and FPCM algorithms are implemented using iterative optimization procedures. The maximum number of iterations is set to 100, and the algorithms update cluster centers, membership values, and typicality values at each iteration until convergence is achieved. The stopping criterion is defined based on the convergence of the objective function. Iteration is terminated when the change in the objective function value between successive iterations falls below a predefined threshold ($\varepsilon = 10^{-5}$), or when the maximum number of iterations is reached. In this study, the convergence threshold was set to $\varepsilon = 10^{-5}$ to ensure stable iterative optimization while maintaining computational efficiency. The iterative process was terminated when the difference between successive objective function values fell below this threshold.

The dataset used in this study was collected from schools in Bojonegoro Regency that had not yet received the MBG program. A total of 200 student records were obtained through primary data collection with the support of school administrators. Five key variables were selected as the basis for determining eligibility: (1) student nutritional status measured by Body Mass Index (BMI), (2) household income, (3) family size, (4) participation in other social assistance programs, and (5) distance from home to school. These variables were chosen because they reflect critical socioeconomic and health indicators relevant to program targeting. To protect participant privacy, all student records were anonymized prior to analysis, and no personally identifiable information was included in the dataset. The data collection process was conducted with the permission and coordination of the participating schools, and the study was carried out solely for academic research purposes in accordance with institutional research ethics principles.

To ensure data validity, several measures were considered during the data collection process. Household income data were obtained through school administrative records and cross-checked with available supporting information provided by students or guardians. Although some degree of reporting bias may exist, efforts were made to ensure consistency and plausibility of the data. Prior to data standardization, the dataset was examined for extreme values in BMI and household income variables. Descriptive inspection indicated that the observed values remained within plausible socio-economic and health ranges; therefore, no winsorization or outlier removal procedures were applied. This approach was adopted to preserve the original data structure, while the robustness of the FPCM algorithm against noisy observations helped reduce the influence of potential extreme values during clustering.

Student nutritional status was assessed using Body Mass Index (BMI), calculated from height and weight measurements recorded by the schools. These measurements follow standard health assessment practices commonly used in educational settings, although minor variations in measurement conditions may occur.

It is acknowledged that, as with most real-world socio-economic datasets, the data may contain measurement errors and uncertainties. However, the use of the FPCM algorithm, which incorporates possibilistic typicality, is expected to reduce the impact of such imperfections by providing robustness against noise and ambiguity in the data.

The sampling process in this study follows a purposive sampling approach, where data are collected from schools in Bojonegoro Regency that had not yet received the MBG program. The selection criteria include students who meet the administrative requirements for potential program eligibility and have complete data across all five variables used in the analysis. This approach ensures that the dataset is relevant to the decision-making context being studied.

Although the sample is limited to a specific region, it reflects a realistic subset of potential beneficiaries within the targeted population. However, it is acknowledged that the dataset may not fully represent broader demographic variations. Therefore, the findings of this study should be interpreted within the context of the selected sample, and further studies with more diverse and large-scale data are recommended to enhance generalizability. Prior to clustering, exploratory examination of variable relationships was conducted to identify potential redundancy among socio-economic indicators. The analysis indicated no excessively dominant correlations that would substantially distort the distance-based clustering process, allowing all selected variables to be retained in the model.

The dataset used in this study consists of 200 student records collected from schools in Bojonegoro Regency. While this sample size may be considered limited for large-scale clustering applications, this study is designed as a case-based and exploratory investigation to evaluate the feasibility of an FPCM-based decision support system (DSS) in a localized context. In clustering analysis, the primary objective is to identify structural patterns within the available data rather than to achieve statistical generalization. Therefore, the dataset is sufficient to demonstrate the applicability and interpretability of the proposed approach.

Nevertheless, it is acknowledged that a larger and more diverse dataset would further improve cluster stability and generalizability, particularly for broader policy implementation. Future research is therefore recommended to extend this study using larger-scale datasets to enhance the robustness of the proposed DSS in real-world applications.

2.1 Data Standardization

Data standardization, often referred to as scaling or normalization, constitutes a fundamental preprocessing stage in data analysis and machine learning. It is designed to transform variables into a comparable scale, thereby facilitating consistent analysis, comparison, and modeling. The purpose of this procedure is to ensure uniformity across all features within a dataset, typically by adjusting the distribution to have a mean of zero and a standard deviation of one. Such standardization enables fair comparison among variables and prevents any single feature from disproportionately affecting the analysis due to differences in scale [19].

In standard practice, each variable is usually converted into standardized values, commonly referred to as z-scores [20].

$$z = \frac{x_i - \mu}{\sigma} \tag{1}$$

2.2 Fuzzy C-Means (FCM)

As the first clustering step, the FCM algorithm was applied to group students into potential beneficiary categories. FCM assigns membership values between 0 and 1 to each student for all clusters, reflecting the degree of association with each group. The algorithm was implemented in MATLAB, and parameters such as the number of clusters and fuzzification coefficient were tuned iteratively to obtain stable results.

According to [21], the FCM algorithm is classified as a soft clustering technique, which operates through iterative adjustments of membership values and cluster centers. Mathematically, the centroid of each cluster in FCM is determined using Equation (2).

$$v_{kj} = \frac{\sum_{k=1}^{n}\left((\mu_{ik})^{w} x_{ij}\right)}{\sum_{k=1}^{n}(\mu_{ik})^{w}} \tag{2}$$

The symbol v_{kj} denotes the j-th element of the centroid for cluster k. The symbol n represents the total number of data points. The term μ_{ik} indicates the degree of membership of data point i in cluster k, with values ranging between 0 and 1. The parameter w corresponds to the fuzzification coefficient, typically set to a value greater than 1. Finally, x_{ij} denotes the j-th attribute of the i-th data point.

The term μ_{ik} represents the membership strenght of the i-th data point in cluster k. Its value lies between 0 and 1, where values approaching 1 denotes stronger association and those nearer to 0 indicate weaker affiliation. The computation of this membership degree is specified in Equation (3).

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

The summation symbol, $\sum_{j=1}^{m}\left(x_{ij} - v_{kj}\right)$, is used to measure the cumulative difference between each data component and the corresponding centroid component across all dimensions. This operation is applied to every data point in the dataset. Here, x_{ij} denotes the value of the i-th data point in the j-th dimension, while v_{kj} represents the centroid value of cluster k in the same dimension. The parameter w serves as the fuzzification coefficient, which determines the level of fuzziness in clustering. Typically set to a value greater than 1, larger values of w lead to more overlapping cluster boundaries. Using this formulation, the FCM algorithm iteratively updates cluster centers and membership degrees, allowing more flexible segmentation of data compared to traditional hard clustering techniques.

To determine precise cluster centers and membership values, an iterative procedure is employed to minimize the objective function. The formulation of this objective function in the FCM method is expressed in Equation (4).

$$J_w(U,V) = \sum_{i=1}^{n}\sum_{k=1}^{c}(\mu_{ik})^w d_{ik}^2 \tag{4}$$

In the FCM algorithm, Equation (4) defines the objective function as the summation of weighted squared distances between data points and cluster centroids, incorporating membership degrees. In this context, n denotes the number of data points, c represents the number of clusters, μ_{ik} is the degree of membership of data point i to clusterk, d_{ik} refers to the Euclidean distance between data point i and the centroid of cluster k, and w is the fuzzification parameter that determines the extent of clustering fuzziness. The optimization goal is to minimize $J_w\ (U,V)$ through iterative updates, thereby producing accurate cluster centers and membership values. Smaller values of $J_w\ (U,V)$ indicate improved clustering performance.

The Euclidean distance equation utilized in this study is defined by Equation (5):

$$d_{ik} = \sqrt{\sum_{j=1}^{m}\left(x_{ij} - v_{kj}\right)^2} \tag{5}$$

The symbol m represents the number of dimensions, x_{ij} denotes the j-th attribute of data pointi, and v_{kj} corresponds to the j-th attribute of the centroid of cluster k. This expression quantifies the dissimilarity between data points and their corresponding cluster centroids.

2.3 Fuzzy Possibilistic C-Means (FPCM)

Following the initial clustering, the FPCM algorithm was applied to improve accuracy and robustness. FPCM integrates fuzzy membership (relative association) and possibilistic membership (absolute belonging), making it less sensitive to noisy or uncertain data [12], [13]. The algorithm was executed in MATLAB, with parameter optimization to maximize clustering performance. The expected output was more precise identification of eligible beneficiaries compared to FCM. The cluster's centre in this method can be found using equatian (6).

$$v_{kj} = \frac{\sum_{i=1}^{n}\left((\mu_{ik}^{w} + t_{ik}^{\eta})x_{ij}\right)}{\sum_{i=1}^{n}\mu_{ik}^{w} + t_{ik}^{\eta}} \tag{6}$$

The matrix T, known as the absolute distinctiveness matrix, is obtained from the integration of membership values (μ_{ik}) and cluster centroids (v_{kj}) in the FCM algorithm. Its computation within the FPCM framework is formally defined by Equation (7).

$$T = \begin{pmatrix} t_{11} & \cdots & t_{1k} \\ \vdots & \ddots & \vdots \\ t_{i1} & \cdots & t_{ik} \end{pmatrix} \tag{7}$$

With

$$t_{ik} = \left[\frac{\left[\sum_{j=1}^{m}\left(x_{ij} - v_{kj}\right)^{2}\right]^{\frac{1}{\eta-1}}}{\sum_{i=1}^{n}\left[\sum_{j=1}^{m}\left(x_{ij} - v_{kj}\right)^{2}\right]^{\frac{1}{\eta-1}}}\right]^{-1} \tag{8}$$

The objective function of FPCM algorithm is defined by:

$$J_{w,\eta}(U,T,V) = \sum_{k=1}^{c}\sum_{i=1}^{n}(\mu_{ik}^{w} + t_{ik}^{\eta})\sqrt{\sum_{j=1}^{m}\left(x_{ij} - v_{kj}\right)^{2}} \tag{9}$$

All clustering procedures were implemented in MATLAB. To support reproducibility, the source code used in this study is publicly available at GitHub [22].

Computational complexity is an important aspect in evaluating the feasibility of the proposed approach for practical implementation. Let n denote the number of data points, c the number of clusters, d the number of variables, and t the number of iterations. In the Fuzzy C-Means (FCM) algorithm, the dominant computational cost arises from the iterative update of membership values and cluster centers, resulting in a time complexity of $O(ncdt)$. The Fuzzy Possibilistic C-Means (FPCM) algorithm extends this process by additionally computing typicality values, which introduces extra computational steps. As a result, FPCM also has a time complexity of $O(ncdt)$, although with a slightly higher constant factor due to the simultaneous update of both membership and typicality matrices. Despite this additional overhead, the increase in computational cost is not asymptotically significant. Given the dataset size used in this study ($n = 200$), both algorithms remain computationally efficient and suitable for decision support applications. This indicates that the proposed DSS framework is feasible for moderate-sized datasets, while further optimization may be required for large-scale implementations.

To provide a more rigorous representation, the proposed decision support system (DSS) is formally defined as a mapping from the input data space to a set of decision recommendations. Let $X \in \mathbb{R}^{n\times d}$ represent the standardized dataset, where n denotes the number of observations and d represents the number of variables. The FPCM algorithm produces two key outputs: the membership matrix $U = [u_{ik}]$, which reflects the relative degree of association of each data point to each cluster, and the typicality matrix $T = [t_{ik}]$, which represents the absolute compatibility of each data point with the cluster.

Based on these outputs, the DSS is defined through a decision function that assigns each data point to a cluster according to the maximum membership value, while simultaneously considering the corresponding typicality value as a measure of confidence. In this framework, decision recommendations are not treated as rigid classifications but are structured into three interpretative levels. Data points with both high membership and high typicality are considered strongly recommended (eligible), those with high membership but relatively low typicality are categorized as borderline cases requiring further verification, and those with low membership values are classified as not eligible. This formalization ensures that the DSS provides graded and interpretable recommendations, aligning the clustering results with practical decision-making processes in the presence of uncertainty.

3. RESULT AND ANALYSIS

3.1 Data Input and Standardization

In this study, the dataset consists of five variables that serve as the basis for determining MBG program beneficiaries. The variables are defined as follows: x_1 = student nutritional status (BMI), x_2 = household income, x_3 = family size, x_4 = participation in other social assistance programs, and x_5 = distance from home to school. Together, these variables capture both socioeconomic and health-related aspects that are essential in assessing eligibility for the program. Table 1 provides the detailed data collected from 200 students in Bojonegoro Regency.

Table 1. Student Data

Subject	x_1	x_2	x_3	x_4	x_5
1	17.80	2000000	4	0	8
2	19.07	3000000	4	0	7
3	22.95	600000	4	1	0,5
4	19.98	800000	4	1	4
...	...	...	...	...	...
197	21.38	2500000	4	0	26
198	28.75	4000000	4	0	13
199	23.49	2500000	5	0	12
200	27.19	1500000	4	0	1

The data standardization stage is conducted to harmonize measurement units across the dataset. As discussed in Section 2, the research variables are expressed on different numerical scales, which may affect distance-based clustering procedures. To address this issue, z-score normalization is applied to transform the data into a common scale with zero mean and unit variance. This preprocessing step helps ensure that all variables contribute proportionally during the clustering process.

Table 2. Standardized Dataset (Z-score)

Subject	x_1	x_2	x_3	x_4	x_5
1	-0.94	-0.04	-0.34	-0.53	-0.56
2	-0.69	0.66	-0.34	-0.53	-0.67
3	0.07	-1.01	-0.34	1.88	-1.41
4	-0.51	-0.88	-0.34	1.88	-1.01
...	...	...	...	...	...
197	-0.24	0.31	-0.34	-0.53	1.49
198	1.22	1.35	-0.34	-0.53	0.01
199	0.18	0.31	0.74	-0.53	-0.10
200	0.91	-0.39	-0.34	-0.53	-1.35

The sample size used in this study consists of 200 observations, which may be considered limited for large-scale generalization. However, it is important to note that the primary objective of clustering analysis is to identify structural patterns within the available data rather than to perform statistical inference on a broader population. Therefore, the dataset is considered sufficient for a case-based and exploratory analysis aimed at evaluating the feasibility of the proposed decision support system. Nevertheless, it is acknowledged that larger and more diverse datasets would further enhance cluster stability and generalizability in future studies.

In terms of data quality, household income data were obtained from school administrative records and supported by information provided by students or guardians. While this may introduce potential reporting bias, efforts were made to ensure data consistency through cross-checking procedures and plausibility verification. Furthermore, the use of the FPCM algorithm contributes to mitigating the impact of such imperfections, as the incorporation of typicality values allows the model to reduce sensitivity to noisy or uncertain data.

To ensure the reliability of the dataset for clustering purposes, preprocessing procedures were applied. Data standardization using z-score normalization was performed to eliminate scale differences across variables. In addition, the dataset was examined for completeness prior to analysis, and only records with complete information across all variables were included. This approach minimizes the influence of missing data and ensures consistency in the clustering process. Although formal statistical reliability tests are more commonly associated with inferential studies, the preprocessing steps implemented in this study are sufficient to support the robustness of clustering-based analysis.

3.2 Visualization of Clustering Results

Figure 1 presents the visualization of clustering results obtained using the Fuzzy Possibilistic C-Means (FPCM) algorithm. The data are projected onto two principal components to facilitate visual interpretation. Each point represents an observation in the dataset, and the distribution of points reflects the cluster structure identified by the algorithm. This visualization is intended to provide a general overview of the clustering pattern and the relative separation among clusters in a reduced two-dimensional space.

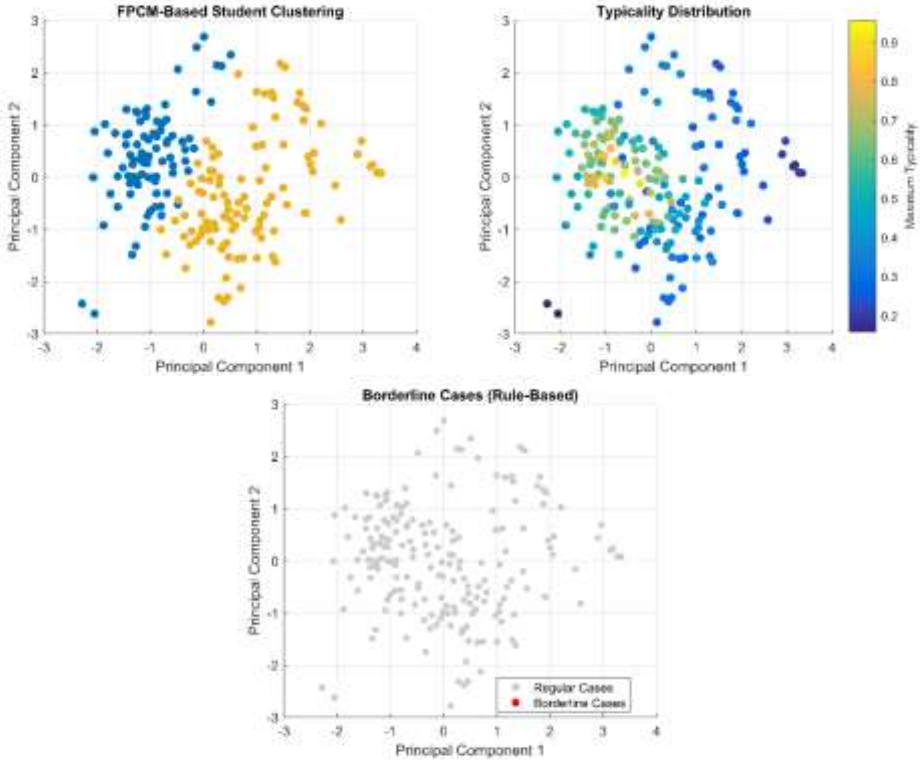

Figure 1. Clustering Visualization

As illustrated in Figure 1, the clustering visualization provides an overview of the distribution, typicality patterns, and borderline cases identified by the FPCM algorithm. Panel (a) illustrates the grouping of students into three clusters corresponding to not eligible, moderately eligible, and eligible categories. The absence of rigid boundaries between clusters reflects the overlapping nature of socio-economic data, which is typical in social assistance contexts.

Panel (b) shows the distribution of typicality values, representing the absolute degree of suitability of each student within a cluster. Higher typicality values indicate students who strongly represent cluster characteristics, while lower values indicate weaker compatibility.

Panel (c) highlights borderline cases, defined as students with relatively high membership values but low typicality. These cases are not directly classified as eligible or non-eligible; instead, they are explicitly flagged for further verification. This visualization demonstrates the role of FPCM in supporting a decision support system by distinguishing clear cases from ambiguous ones rather than enforcing rigid decisions.

3.3 Distribution of Students Across Clusters

To quantify the clustering outcome, the number and proportion of students in each cluster are summarized in Table 3.

Table 3. Distribution of Students Across Clusters (FPCM)

Cluster	Number of Student	Percentage
Not Eligible	104	52 %
Moderately Eligible	37	18.5 %
Eligible	59	29.5 %
Total	200	100 %

The results show that the majority of students are classified as not eligible (52%), followed by eligible students (29.5%), while a smaller proportion falls into the moderately eligible cluster (18.5%). Although the moderately eligible group represents a minority, its presence is crucial, as it captures borderline cases that cannot be reliably handled by a binary classification scheme. This distribution highlights the importance of a graded decision framework within the proposed decision support system, where ambiguous cases are explicitly identified for further consideration rather than forced into rigid categories.

3.4 Cluster Characteristics Based on Standardized Variables

The socio-economic and nutritional profiles of each cluster were examined using the mean standardized values of the input variables. The results are presented in Table 4.

Table 4. Mean Standardized Variable Values per Cluster

Cluster	BMI	Income	Family Size	Aid Status	Distance
Not Eligible	0.42	0.05	0.28	0.49	0.43
Moderately Eligible	-0.16	0.05	0.10	-0.53	-0.11
Eligible	-0.64	-0.11	-0.56	-0.53	-0.68

The eligible cluster is characterized by lower BMI and slightly lower household income compared to the overall average, along with smaller family size. In addition, students in this cluster tend to live closer to school and show lower participation in other social assistance programs. These patterns indicate that eligibility in the proposed DSS is not determined by a single socio-economic indicator, but rather by a combined profile of nutritional status, economic condition, and household characteristics, thereby supporting the interpretability of the clustering results from a policy perspective.

3.5 Membership and Typiclity Analysis

Beyond cluster size and centroids, the effectiveness of the DSS is reflected in the joint analysis of membership and typicality values. Table 5 summarizes the average membership and typicality for each cluster.

Table 5. Average Typicality per Cluster

Cluster	Avg. Typicality	Std. Typicality
Not Eligible	0.330	0.077
Moderately Eligible	0.607	0.074
Eligible	0.670	0.116

Table 5 summarizes the average typicality values for each cluster. The not eligible cluster exhibits a low average typicality (0.330) with relatively small variability, indicating a consistent lack of representativeness. In contrast, the moderately eligible cluster shows a higher average typicality (0.607) with limited dispersion, reflecting a group of borderline cases with relatively stable characteristics. The eligible cluster attains the highest average typicality (0.670), but also the largest standard deviation (0.116), suggesting heterogeneity within this group. From a decision support perspective, these results imply that eligibility recommendations should be prioritized based on typicality strength rather than treated as uniform decisions.

3.6 Quantitative Evaluation of Clustering Performance

To objectively evaluate clustering performance, the Partition Coefficient (PC) is employed. The PC measures the degree of cluster clarity, where higher values indicate more distinct and less ambiguous membership distributions.

Table 6. Quantitative Comparison of Clustering Performance

Method	Partition Coefficient (PC)
FCM	0.74
FPCM	0.84

The results show that the FPCM algorithm achieves a higher PC value (0.84) compared to FCM (0.74). This indicates a noticeable improvement in cluster clarity, suggesting that FPCM produces more well-defined and stable clustering structures. In addition to the Partition Coefficient (PC), further evaluation is conducted using additional internal validation metrics to provide a more comprehensive assessment of clustering performance.

Table 7. Quantitative Comparison of Clustering Performance

Method	Silhouette	Dunn Index
FCM	0.058	0.0546
FPCM	0.129	0.0543

The results indicate that the FPCM algorithm achieves a higher Silhouette Score, suggesting slightly better separation and compactness under classical distance-based evaluation. However, the FCM algorithm attains a comparable Dunn Index, indicating that the overall cluster structure remains similarly well-defined in terms of inter-cluster separation relative to cluster diameter.

It should be noted that the Silhouette Scores obtained in this study remain relatively low for both methods, indicating limited cluster separation in the standardized feature space. However, this outcome is consistent with the nature of socio-economic and welfare-related data, where population characteristics often form continuous and overlapping distributions rather than well-separated groups. In the context of beneficiary targeting, strict boundaries between eligible and non-eligible individuals are rarely observed in practice, as many cases exhibit intermediate or ambiguous conditions. Therefore, the relatively low Silhouette values do not necessarily indicate poor clustering performance, but instead reflect the inherent complexity and gradual structure of real-world social vulnerability data. From a decision support perspective, this characteristic further supports the relevance of FPCM, which is specifically designed to accommodate uncertainty and overlapping cluster memberships.

These findings suggest that while FCM achieves comparable performance in certain metrics, FPCM provides superior or equal structural quality while offering additional advantages through its typicality-based interpretation. This reinforces the role of FPCM not only as a clustering method, but as a more informative decision support framework capable of handling uncertainty and identifying borderline cases.

This improvement can be attributed to the incorporation of possibilistic typicality in FPCM, which reduces the influence of noise and overlapping data points. As a result, FPCM is better able to capture the underlying structure of the data, particularly in the presence of uncertainty. Therefore, the findings support the use of FPCM as a more robust approach for clustering in decision support systems involving complex socio-economic data.

3.7 Discussion from a Decision Support Perspective

The primary objective of the proposed decision support system (DSS) is not to produce definitive eligibility labels, but to provide structured, interpretable, and graded recommendations that can assist policymakers in managing uncertainty inherent in social assistance data. The clustering results obtained using FPCM demonstrate how such a DSS framework can support decision-making beyond binary eligibility rules. This perspective is consistent with recent studies indicating that decision support systems based on fuzzy approaches are effective in handling uncertainty and improving interpretability in welfare-related decision-making contexts [23].

First, the distribution of students across clusters (Table 3) reveals that more than half of the students are classified as not eligible (52%), while approximately one-third are identified as eligible (29.5%). Importantly, a non-negligible proportion of students (18.5%) falls into the moderately eligible cluster. Although this group represents a minority, its presence is significant from a DSS perspective, as it captures cases that cannot be reliably resolved through strict binary classification. In practical policy contexts, these cases often correspond to students whose socio-economic and nutritional profiles do not clearly meet or fail eligibility criteria, thereby requiring additional consideration. This finding extends previous clustering-based welfare studies, which typically rely on conventional FCM approaches that produce rigid cluster assignments without explicitly capturing borderline cases [24].

Second, the cluster characteristics based on standardized variables (Table 4) indicate that eligibility emerges from a combination of indicators, rather than from a single dominant factor. The eligible cluster is associated with lower BMI, slightly lower household income, smaller family size, closer distance to school, and lower participation in other social assistance programs. This profile suggests that the DSS does not simply replicate existing assistance mechanisms (as reflected by aid status), but instead integrates multiple dimensions of vulnerability. From a policy standpoint, this is desirable, as it reduces over-reliance on any single proxy variable and enhances the interpretability of recommendations.

Third, the typicality analysis (Table 5) provides a crucial layer of decision support that is not available in conventional fuzzy clustering approaches. The not eligible cluster exhibits low average typicality with limited variability, indicating relatively stable non-eligibility recommendations. In contrast, the moderately eligible cluster shows intermediate typicality values with low dispersion, reflecting a group of consistently borderline cases. Most notably, while the eligible cluster attains the highest average typicality, it also shows the greatest variability. This heterogeneity implies that not all eligible students represent equally strong candidates for assistance. This result aligns with prior research highlighting that the incorporation of typicality in FPCM enhances robustness in the presence of noise and overlapping data, which is not addressed in standard fuzzy clustering methods [25].

Recent studies have further emphasized the growing role of fuzzy and welfare-oriented clustering frameworks in supporting socio-economic policy analysis and targeted assistance programs. In particular, clustering approaches have increasingly been applied to welfare indicator analysis and public decision-making contexts characterized by overlapping vulnerability conditions and uncertainty [26]. These developments reinforce the relevance of interpretable clustering methods such as FPCM for social assistance targeting and decision support applications.

From a DSS perspective, this finding has important implications. Rather than treating all eligible students as a homogeneous group, the system enables prioritization based on typicality strength. Students with higher typicality values can be considered stronger recommendations, while those with lower typicality (despite being classified as eligible) may warrant further verification or phased intervention. This graded recommendation mechanism enhances transparency and aligns algorithmic output with realistic policy workflows, where resource constraints and administrative discretion play critical roles.

Overall, the integration of cluster distribution, standardized variable profiles, and typicality measures demonstrates that the proposed FPCM-based DSS supports graded, interpretable, and policy-relevant decision-making. Instead of replacing human judgment, the system structures complex socio-economic information into actionable insights, thereby improving the robustness and accountability of MBG beneficiary selection under data uncertainty.

From an implementation perspective, the proposed DSS can support local education authorities and social welfare offices by functioning as an initial screening and prioritization tool for MBG beneficiary selection. The clustering outputs can be integrated with existing school administrative databases and social assistance records to assist policymakers in identifying students with varying levels of eligibility. In practical applications, threshold calibration may be adjusted according to regional budget capacity, policy priorities, or the availability of nutritional assistance resources. This flexibility allows the DSS to support adaptive and transparent allocation strategies rather than relying solely on rigid eligibility rules.

The moderately eligible cluster has important practical implications within the proposed framework. Instead of being directly categorized as eligible or non-eligible, students in this group may undergo additional verification procedures, such as targeted household surveys, manual administrative review, or phased assistance mechanisms. For example, local authorities may prioritize full assistance for students with high typicality values, while moderately eligible cases could receive conditional or temporary support pending further assessment. Such an approach enables policymakers to manage uncertainty more effectively while reducing the risk of exclusion and inclusion errors in welfare targeting.

While the proposed FPCM-based DSS demonstrates advantages in handling uncertainty and providing interpretable recommendations, it has several limitations. In datasets with highly overlapping characteristics, membership and typicality values may become less discriminative, leading to ambiguity in decision outcomes. In addition, the performance of FPCM depends on parameter selection and data quality, which may affect clustering stability.

Compared to alternative approaches, clustering-based methods such as FPCM are suitable when labeled data are unavailable and interpretability is required. However, supervised machine learning methods, such as decision trees or support vector machines, may provide higher predictive accuracy when labeled datasets are available. Therefore, FPCM is better positioned as an exploratory and interpretative decision support tool, while future research may consider hybrid approaches to improve performance.

4. CONCLUSION

This study proposes a decision support system (DSS) based on the Fuzzy Possibilistic C-Means (FPCM) algorithm for beneficiary selection under conditions of uncertainty and overlapping eligibility criteria. The clustering results indicate that 52% of students are classified as not eligible, 18.5% as moderately eligible, and 29.5% as eligible, reflecting the presence of both clear and borderline cases within the dataset. From a methodological perspective, the main contribution of this study lies in the formal integration of FPCM into a DSS framework, where clustering results are not only used for grouping but also for supporting decision interpretation through typicality-based evaluation. This approach enables the identification of varying levels of eligibility rather than enforcing rigid classifications, thereby enhancing transparency in decision-making.

The evaluation results show that FPCM achieves a higher Partition Coefficient (0.84) compared to FCM (0.74), indicating improved cluster clarity, while additional validation metrics demonstrate comparable structural quality between the two methods. These findings suggest that the strength of FPCM lies not solely in numerical superiority, but in its ability to handle uncertainty and support more nuanced decision processes. Future research may focus on extending this framework through the incorporation of larger datasets, additional contextual variables, and hybrid approaches that combine clustering with supervised learning techniques to further improve decision support performance. In addition, the proposed framework has the potential to be adapted for other national welfare targeting programs in Indonesia, such as PKH and BPNT, where uncertainty and overlapping socio-economic conditions similarly affect beneficiary selection processes.

ACKNOWLEDGEMENT

The authors would like to express their sincere gratitude to the Ministry of Education, Culture, Research, and Technology of the Republic of Indonesia for the continuous support provided through the National Competitive Research Program. The authors also extend appreciation to the Universitas Nahdlatul Ulama Sunan Giri for its facilities, collaboration, and academic environment that have greatly contributed to the completion of this study.

5. REFERENCES

[1] U. Indonesia, 'Pusat unggulan nasional program makan bergizi gratis', UNICEF Indonesia, 2025. [Online]. Available: https://www.unicef.org/indonesia/id/gizi/laporan/pusat-unggulan-nasional-program-makan-bergizi-gratis

[2] M. R. K. M. Muluk and S. Sajida, 'How does Indonesia's 2025 budget reform reflect the logic of austerity governance?', *Asian Review of Public Administration*, vol. 33, no. 1, pp. 95–110, 2025, doi: 10.64423/ARPA.V33I1.76.

[3] A. N. Anwar, 'Implementasi fuzzy C-mean (FCM) untuk menentukan penerima beasiswa', *Jurnal Ilmu Komputer*, vol. 6, no. 1, pp. 84–87, 2023, [Online]. Available: https://jurnal.pranataindonesia.ac.id/index.php/jik/article/view/155

[4] D. R. Anamisa, M. Yusuf, and M. A. Syakur, 'Determination system of food vouchers for the poor based on fuzzy c-means method', *J. Phys. Conf. Ser.*, vol. 953, no. 1, p. 012132, 2018, doi: 10.1088/1742-6596/953/1/012132.

[5] A. Setiawan and J. N. Akbar, 'Implementation of fuzzy C-means in decision support systems BPNT (Bantuan Pangan Non-Tunai) Ministry of Social Affairs Indonesia', *EMITTER International Journal of Engineering Technology*, vol. 7, no. 2, pp. 559–569, 2019, doi: 10.24003/EMITTER.V7I2.444.

[6] N. N. Rahmawati and M. I. A. Fathoni, 'Penentuan Penerima Kip Kuliah Mahasiswa S1 Unugiri Menggunakan Fuzzy C-Means Clustering', *Jurnal Pendidikan Matematika dan Matematika*, vol. 6, no. 2, pp. 121–130, 2022, doi: 10.36526/tr.v%vi%i.2197.

[7] J. C. Bezdek and J. M. Keller, 'Streaming Data Analysis: Clustering or Classification?', *IEEE Trans. Syst. Man Cybern. Syst.*, vol. 51, no. 1, pp. 91–102, 2021, doi: 10.1109/TSMC.2020.3035957.

[8] P. S. Saputra, G. R. Dantes, and I. G. A. Gunadi, 'Perbandingan algoritma Fuzzy C-Means dan algoritma Naive Bayes dalam menentukan Keluarga Penerima Manfaat (KPM) berdasarkan status sosial ekonomi', *JST (Jurnal Sains dan Teknologi)*, vol. 10, no. 1, pp. 64–74, 2021, doi: 10.23887/jstundiksha.v10i1.23340.

[9] I. H. Sarker, 'Machine learning: Algorithms, real-world applications and research directions', *SN Comput. Sci.*, vol. 2, no. 3, p. 160, 2021, doi: 10.1007/s42979-021-00592-x.

[10] R. J. Hathaway and J. C. Bezdek, 'Fuzzy c-means clustering of incomplete data', *IEEE Transactions on Systems, Man, and Cybernetics Part B: Cybernetics*, vol. 31, no. 5, pp. 735–744, 2001, doi: 10.1109/3477.956035.

[11] C. Wu and J. Zhang, 'Robust semi-supervised spatial picture fuzzy clustering with local membership and KL-divergence for image segmentation', *International Journal of Machine Learning and Cybernetics*, vol. 13, no. 4, pp. 963–987, 2022, doi: 10.1007/s13042-021-01429-y.

[12] R. Krishnapuram and J. M. Keller, 'A possibilistic approach to clustering', *IEEE Transactions on Fuzzy Systems*, vol. 1, no. 2, pp. 98–110, 1993, doi: 10.1109/91.227387.

[13] N. R. Pal, K. Pal, J. M. Keller, and J. C. Bezdek, 'A possibilistic fuzzy c-means clustering algorithm', *IEEE Transactions on Fuzzy Systems*, vol. 13, no. 4, pp. 517–530, 2005, doi: 10.1109/TFUZZ.2004.840099.

[14] Y. Gu, T. Ni, and Y. Jiang, 'Deep possibilistic C-means clustering algorithm on medical datasets', *Comput. Math. Methods Med.*, vol. 2022, no. 1, p. 3469979, 2022, doi: 10.1155/2022/3469979.

[15] M. B. Naghi, L. Kovács, and L. Szilágyi, 'A generalized fuzzy-possibilistic c-means clustering algorithm', *Acta Universitatis Sapientiae, Informatica*, vol. 15, no. 2, pp. 404–431, 2023, doi: 10.2478/ausi-2023-0023.

[16] D. Kurniasari, V. Kurniawati, A. Nuryaman, M. Usman, and R. K. Nisa, 'Implementation of fuzzy c-means and fuzzy possibilistic c-means algorithms on poverty data in indonesia', *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, vol. 18, no. 3, pp. 1919–1930, 2024, doi: 10.30598/barekengvol18iss3pp1919-1930.

[17] Ö. Özdemir and A. Kaya, 'Comparison of FCM, PCM, FPCM and PFCM algorithms in clustering methods', *Afyon Kocatepe University Journal of Science and Engineering*, vol. 19, no. 1, pp. 92–102, 2019, doi: 10.35414/akufemubid.429540.

[18] A. W. Maghfiroh, N. Ulinnuha, and A. Fanani, 'Penerapan Fuzzy C-Means dalam mengelompokkan kabupaten/kota berdasarkan fasilitas pelayanan kesehatan di Jawa Timur', *Informasi: Jurnal Ilmiah Bidang Teknologi Informasi dan Komunikasi*, vol. 4, no. 1, pp. 8–14, 2019, doi: 10.25139/inform.v4i1.1444.

[19] T. A. S. Srinivas, Y. Sravanthi, Y. V Kumar, and I. V. D. Srihith, 'Data standardization: Key to effective data integration', *Advanced Innovations in Computer Programming Languages*, vol. 6, no. 1, pp. 1–4, 2023.

[20] R. E. Walpole, R. H. Myers, S. L. Myers, and K. Ye, *Probability and Statistics for Engineers and Scientists*. London, U.K.: Pearson, 2022.

[21] Y. Shan, S. Li, F. Li, Y. Cui, and M. Chen, 'Dual-level clustering ensemble algorithm with three consensus strategies', *Sci. Rep.*, vol. 13, no. 1, p. 22617, 2023, doi: 10.1038/s41598-023-49947-9.

[22] M. H. Al Farisy, 'FPCM2025', 2025. [Online]. Available: https://github.com/husenalfa123/FPCM2025

[23] D. Mawadah, N. Puspitasari, and A. Sari, 'Decision support system for community welfare assessment using fuzzy logic Mamdani in Ponorogo Regency', *Jurnal Riset Informatika*, vol. 7, pp. 340–350, 2025, doi: 10.34288/jri.v7i4.407.

[24] R. D. Putri and D. P. Sari, 'Regional clustering in Sumatera based on welfare indicators using Fuzzy C-Means', *Journal of Advanced Information and Computing*, vol. 9, no. 5, pp. 2399–2409, 2025.
[25] Z. Zhang, 'Application of fuzzy decision support systems in risk assessment of Southeast Asian labor market', *International Journal of Computational Intelligence Systems*, vol. 17, no. 1, p. 153, 2024, doi: 10.1007/s44196-024-00556-y.
[26] C. J. Zhang, S. A. Esmaeili, and J. Morgenstern, 'Welfare-Centric Clustering', *arXiv preprint arXiv:2508.10345*, 2025.