SCIENTIFIC OASIS

Decision Making: Applications in Management and Engineering

Journal homepage: www.dmame-journal.org
ISSN: 2560-6018, eISSN: 2620-0104

A Novelty Decision-Making Based on Hybrid Indexing, Clustering, and Classification Methodologies: An Application to Map the Relevant Experts Against the Rural Problem

Muhammad Faisal[1*], Titik Khawa Abd Rahman[2], Ida Mulyadi[1*], Komang Aryasa[3], Irmawati[4], Musdalifa Thamrin[5]

[1] Department of Computer Science, STMIK Profesional, Makassar, Indonesia
[2] Department of Ph.D ICT, AeU University, Kuala Lumpur, Malaysia
[3] Department of Informatics Engineering, DIPA University, Makassar, Indonesia
[4] Department of Computer and Network Engineering, Politeknik Negeri Ujung Pandang, Indonesia
[5] Department of Informatics Management, STMIK Profesional, Makassar, Indonesia

ARTICLE INFO

Article history:
Received 31 October 2023
Received in revised form 29 January 2024
Accepted 8 February 2024
Available online 19 February 2024

Keywords:
Rural; Indexing; Clustering; Classification; Recommendation; Village Assistants.

ABSTRACT

Exploring the potential of technology in addressing Sustainable Development Goals (SDGs) within intricate rural contexts holds paramount significance. Sustainable rural development holds profound significance for both developed and developing nations. This study was conducted to develop a methodological framework for placing experts with strategically relevant competencies to meet the specific needs of villages, thus enabling the practical application of their expertise in generating innovative solutions to problems in a village. This study entails several pivotal phases. Firstly, it constructs a Community Standard of Living Index (CSLI) using Delphi and Rank Reciprocal (RR). Secondly, it establishes village clustering through a hybrid Fuzzy C-Means (FCM), Self-Organizing Map (SOM), and Xie-Beni (XB). Thirdly, a classification of village development levels is created using Tsukamoto and Smallest of Maximum (SM). Finally, recommendations for placing experts in villages, aligning their skills with identified needs using the Cosine Similarity (CS). The results obtained are compared with factual data of each village to obtain relevant conclusions, where an accuracy value of 0.95 indicates a high success rate in the test results of the proposed technique. This study has the potential to significantly enhance decision-making by introducing opportunities for the development of hybrid methodologies in expert mapping for rural issues.

1. Introduction

Integrating indexing, clustering, and classification methodologies often involves using advanced technologies such as machine learning, data mining, and artificial intelligence. This is the motivation of this study to explore the potential of technology in solving problems related to the Sustainable Development Goals in complex rural areas because sustainable rural development has significance

* *Corresponding author.*
E-mail address: muhfaisal@stmikprofesional.ac.id; idamulyadi@stmikprofesional.ac.id

https://doi.org/10.31181/dmame7220241023

for many developed and developing countries [1, 2]. Indonesia is one of the developing countries that implements the SDGs Program by mapping and placing experts in village planning and development. However, the skills in knowledge possessed by accompanying experts are not yet relevant to the problems happening in a village [3]. Therefore, this study is essential to build an expert mapping technique according to village needs so that the skills can be applied to create ideas and solutions related to problems in a village.

This research uses methods that can be a tool in overcoming the problems found, including indexing, clustering, and classification techniques related to mapping the placement of village assistance experts. Research on the application of the Fuzzy Delphi method was conducted by Yang *et al*., [4], who propose the Fuzzy Delphi method to overcome uncertainty in weighted data, where this study contributes to group decision-making, uncertainty modelling, and heterogeneous data processing. Research by Chatterjee and Chakraborty [5] proposed Reciprocal methods explored to solve machine equipment rating problems to help improve the overall efficiency of manufacturing processes, where reciprocal methods can be successfully used to solve low and high dimensional MCDM problems in real-time manufacturing environments. Previous research applied the Self-Organizing Map method to analyze the way elements appear in coal, where the results state that the Self-Organizing Map algorithm is a reliable and intuitive method to analyze the way elements appear in coal [6]. Previous research on the Fuzzy C-Mean method was used to modernize performance and shorten the complexity involved in image detection in predicated image segmentation [7], where the proposed method fully works automatically and is tested on different types of brain images. The research discussed applying fuzzy clustering methods and Xie-Beni validity to identify interactive factors, promote livestock welfare, increase the productivity index [8], and increase property values, where research results help assist decision-making in livestock containment systems. The fuzzy set theory was created to address the crisp gap value [9] adopted in systems to handle unconfirmed fractional situations or parameter uncertainty [10]. This study applied fuzzy logic to the village classification process by combining the Tsukamoto method with the Smallest of Maximum. The Fuzzy Tsukamoto method can adapt to the crisp value gap to follow the system output given by the expert [11]. In fuzzy theory, the Smallest of Maximum principles determines the predicate value by selecting the smallest domain value corresponding to the maximum membership value. In the defuzzification process, the output is generated through a transformation into a numeric value [12,13].

By Utilizing clustering or classification techniques, many village issues find resolution as diverse data points within village contexts are organized and categorized efficiently. These methods facilitate the streamlined analysis of complex datasets, unveiling patterns and correlations crucial for informed decision-making. Through data science, communities gain deeper insights into their challenges, enabling the crafting of tailored strategies. This approach, in turn, fosters sustainable development and enhances resident's quality of life by systematically addressing complex village-related issues and developing targeted practical solutions. Studies relevant to this study are shown in Table 1.

Table 1
Research gap summary

Publications	Contributions	Research Gap
Morales *et al*., [14]	Implement the Self-Organizing Map and Fuzzy C-Means methodology to classify families in the community based on sociodemographic information obtained through surveys.	Propose Optimizing the method by using the xie-beni method to build index validity so that hybrid SOM, FCM, and Xie-Beni can produce better clustering.

Decision Making: Applications in Management and Engineering
Volume 7, Issue 2 (2024) 132-171

Publications	Contributions	Research Gap
Mupedziswa *et al.,* [15]	Stating that the declining trend of community development initiatives is due to the lack of community participation in decision-making does not bode well for successful community development.	Propose increasing community participation in decision-making related to village development through practical village mentoring programs.
Lathief *et al.,* [16]	Previous research proposed a forecasting system by combining a validated FCM algorithm using the Xie Beni index method to find the best number of clusters.	Propose hybrid SOM, FCM, and XIe-eni in the clustering process so that the sluggishness related to the initial weight on FCM can be overcome by using convergence values in SOM.
Nugraha *et al.,* [17]	Previous research suggested that it is better to first carry out an influence analysis of factors in community-based village development programs in determining alternative policy strategies.	Propose Identify problems in a village by using multi-object optimization so that village development programs can be appropriately implemented.
Nur Fitria *et al.,* [18]	Resulting in the status of village classification using the C5.0 Algorithm to the variables contained in the Village Building Index, including Health, Education, Settlements, Access to Credit, Natural Disasters	Propose combining 38 criteria, CSLI, DVI, and HDI indicators as determining factors with a broader reach and relevant to building village classification using Tsukamoto - Smallest of Maximum based on expert knowledge.
Pattanayek *et al.,* [19]	The measurement of people's living standards using the literacy rates(LTRs), Employment Status, Scheduled Tribe Ratio, PCFGP, and population density(PD) factors are calculated using the iterative average correlation method.	Propose combining 38 criteria with a more comprehensive range that are relevant for measuring people's living standards using the combined Fuzzy Delphi with Rank Reciprocal method.
Salima and Ilham [20]	Comparison of the algorithm used by the Central Bureau of Statistics of Indonesia with the Tree algorithm for classification of the rural status	Propose a hybrid indexing, clustering and classification technique to build the status of village development by involving methods relevant to applying the techniques used.

The summary of previous research in Table 1 helps to understand unmet research needs as well as areas that require further understanding. The limitation of previous research is that the problem-solving process still uses a single method, so weaknesses are still found in the method used. Furthermore, the variables used are not built independently, so they do not support elements of objectivity and decision-making flexibility. Therefore, this research integrates all entities and theories used into a hybrid technique to produce decision-making on the placement of village assistance experts, where the objective proposed is as follows:

- To build a Community Standard of Living Index based on expertise and community feedback to produce an objective village ranking.
- To build village clustering using a hybrid methodology.
- To build a classification of village development levels by combining independently built CSLI indicators with indicators sourced from the Central Bureau of Statistics (CBS) of Indonesia.

To produce recommendations for placing experts in villages based on skills relevant to needs.

2. Methodology

This section provides details of the methodology used in research using a quantitative approach. The discussion following the theory and algorithms is needed at the analysis, system development, and testing stages to justify validity during the research.

2.1. Identification and Validation Criteria

At the core of any effective survey-based research lies the premise that the questions posed can effectively capture and describe the practices, conditions, experiences, personal attributes, or viewpoints of the respondents [21]. Hence, interviews were conducted with various sources, including local government representatives, experts in regional development planning, and village leaders. These interviews were the preliminary phase for constructing a questionnaire and gathering information. The act of posing questions is a fundamental element of the learning process, intricately woven into the fabric of critical thinking [22]. Some studies use the questionnaire method to collect the necessary data through data collection methods, either as part of structured interviews or through self-filling [23]. Questioning allows us to effectively navigate sources of information to identify and access the information needed in a given situation [24]. Therefore, questions are formed in a questionnaire format based on 38 criteria, as shown in Table A.1.

All criteria were obtained through an interview process with experts from representatives of academics, government, researchers and village officials, where each expert considered the material needed to accommodate problems in a village through responses from the community. In this study, 16 experts were involved in validating criteria and determining approval weights using fuzzy scales as follows: Totally Agree = 5; Agree = 4; Simply Agree = 3; Less Agree = 3; Disagree = 1. In other parts of the questionnaire, experts are allowed to add other criteria or sub-criteria that they consider essential. The questionnaire document is accompanied by a cover letter and an expert personal profile form, then provided to experts in Word format and via email or paper version. All assessments given by experts will be processed through software developed by the author to apply the Delphi method, as previous research stated that the advantage of the Delphi method is that it can help experts understand and reach a consensus [25].

2.2. Fuzzy Delphi

Fuzzy Delphi provides a robust framework that can handle a lack of precision and clarity, as incomplete or inaccurate information is considered a problem in decision-making [26, 27]. The performance of the MOO algorithm is assessed based on the balance between convergence and diversity of non-dominated solutions, which are measured using evaluation criteria different from quality performance indicators. The threshold value is calculated using the following Eq. (1):

$$d(\overline{m},\overline{n}) = \sqrt{\frac{1}{3} * [\,(m1 + m2 + m3)]} \tag{1}$$

The Data is based on the triangular fuzzy number, where it aims to get threshold value (d). The first condition to be followed is that threshold value (d) must be less or equal to 0.2. The aggregate calculation process is a fuzzy evaluation process to determine the value of the fuzzy score that is obtained using the following Eq. (2):

$$A_{Max} = \frac{1}{3} * (M1 + M2 + M3) \Rightarrow \bar{A} = \begin{bmatrix} \overline{A1} \\ \overline{A2} \\ \overline{A..n} \end{bmatrix} A = 1..m \quad (2)$$

The Fuzzy Delphi method allows research to act independently, adapt its dynamics to research objectives and make strategic decisions because the objective opinion of a group of specialists is always of higher quality than the opinion of one individual.

2.3. Rank Reciprocal

The process of assigning weights to criteria employs the rank reciprocal method, which calculates the magnitude of weight values by considering that the order of indices holds varying degrees of importance. The Rank Reciprocal method is calculated using the following Eq. (3) [28]:

$$W_j(RR) = \frac{[1 / \sum_{k=1}^{n}(1/r_k)]}{\sum_{k=1}^{n}[1 / \sum_{k=1}^{n}(1/r_k)]} \quad (3)$$

where wj is the weight of the jth alternative or criterion (j = 1,2,...,n), rk is the rank assigned to the kth alternative or criterion (k = 1,2,...,j), and n is the number of alternatives or criteria.

2.4. Village Scoring

Village scoring is an assessment or evaluation process carried out on certain villages or rural areas to measure or assess various aspects relevant to the progress and welfare of the village. The first step in the village scoring process entails collecting information, suggestions, and community complaints over specific periods. This data collection is conducted using two distinct methods. The village scoring is calculated using the following Eq. (4):

$$SV_j = \frac{\sum Cw * Sw}{\sum Respondents} \quad (4)$$

where SV_j is the village score, the Cw is the criteria weight, and the Sw is the weight of the subcriteria. The assessment results by communities are related to conditions in each village. Stages weighting to each criterion is done using the Rank Reciprocal method.

2.5. Clustering SOM, FCM, and Xie-Beni

This research employed the clustering process to determine which village groups should receive priority regarding mentoring. The selection criteria included value weight, the number of samples, and the population size. The presence of uncertainty related to the characteristics of data objects adds complexity when trying to ascertain the optimal number of clusters, especially when dealing with data objects characterized by high dimensions, fluctuating data sizes, and differing densities [29]. Hence, in this study, cluster index validation was conducted by considering the smallest cluster value to identify the optimal number of clusters. Fuzzy-based and Self-Organizing Map clustering techniques are comparatively more efficient than traditional approaches when uncovering hidden structures within datasets because the segments generated by SOM possess a greater capacity to offer valuable insights for making data-driven decisions [30]. The combined use of SOM and FCM clustering has proven to be a powerful tool for determining cluster boundaries [31].

It is discerned that within the context of employing a hybrid Self-Organizing Map, Fuzzy C-Means, and Xie-Beni clustering methodologies, the evaluation is conducted utilizing parameters such as CSLI

Score, Total Family Size, and the count of residential units. Two main considerations drive the adoption of a hybrid approach. Firstly, it facilitates the generation of a consolidated final weight value through the SOM process, which is then transformed into an initial cluster value within the FCM framework, thereby maintaining stability in the cluster input. Secondly, the cluster validity process relies on the Xie-Beni index in a cluster comparison procedure, with the smallest index assisting in identifying the most suitable cluster, as illustrated by the flowchart in Figure 1.

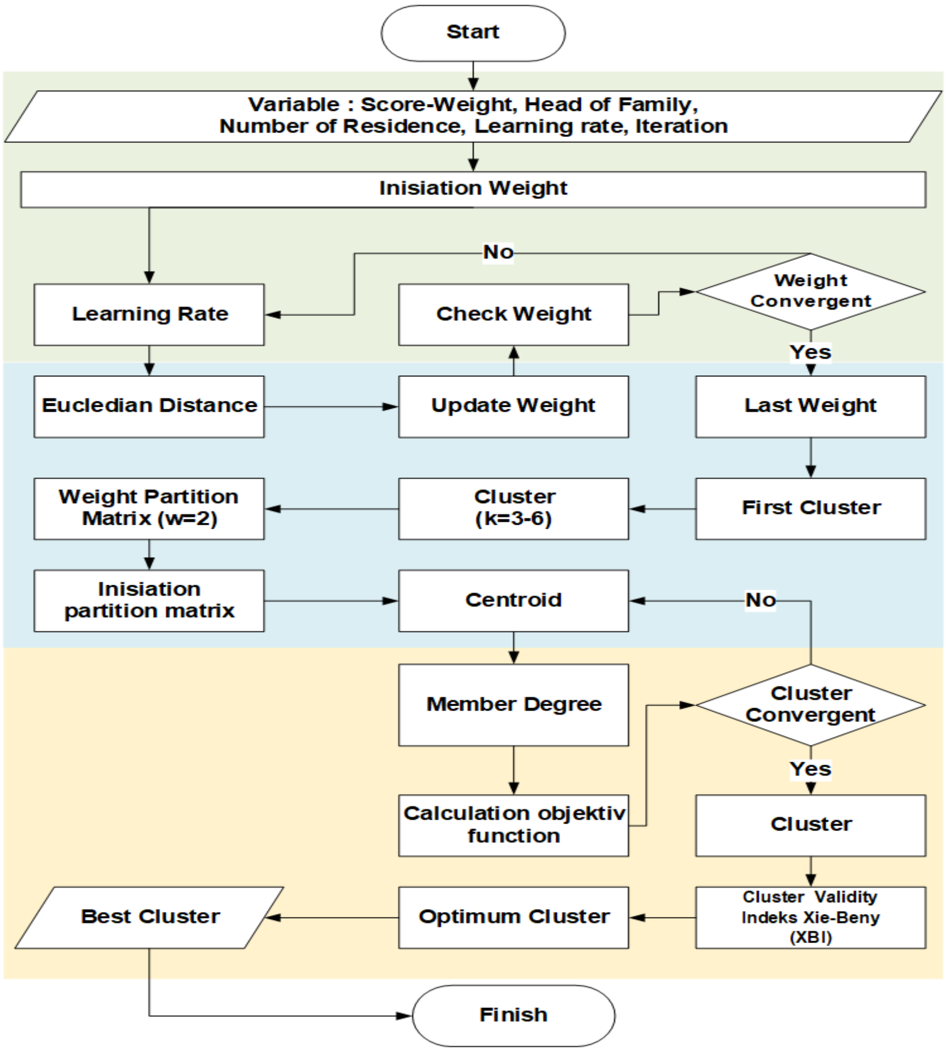


Fig. 1. Hybrid SOM, FCM, and Xie-Beni technique

Validation FCM using the Xie-Beni index method helps determine the optimal number of clusters, which improves the accuracy of FCM. Furthermore, the Xie-Beni method performs well for cluster validation and has the same result [16]. Xie-Beni was determined based on the objective function, and the square of the cluster centres minimum distances [32]. The Xie-Beni index is calculated using the following Eq. (5):

$$XB = \frac{\Sigma_{j=1}^{c}\Sigma_{j=1}^{n}u_{ij}^{m}\ ||xj-vi||^2}{n*min_{i,j}||\ v_i-v_j||^2} \tag{5}$$

Xie-Beni index can calculate the compactness and separation between fuzzy clusters [33] so that when the Xie-Beni index is applied to the clustering method, it can form an optimal cluster area.

2.6. Classification Tsukamoto and Smallest of Maximum

A trapezoidal curve is a mathematical curve consisting of two parallel segment lines connected by two other segment lines that form right angles, where its main function is to describe the relationship between variables in a given context [34]. The formula is written:

$$\begin{cases} 0 : x \le a \\ (x\text{-}a) \;/\; (b\text{-}a) : a \le x \le b \\ 1 : b \le x \le c \\ (d - x) \;/\; (d - c) : c \le x \le d \\ 0 : x \ge d \end{cases} \tag{6}$$

Defuzzification is an effective process to get one number from the output of a fuzzy set [35]. As an explicit, the value used in the defuzzification process is the value that exists at the composition stage of each rule. The Smallest of Maximum methodology is calculated using the following Eq. (7):

$$\mathrm{SM(Z)} = \frac{\mathrm{Min(\,Zi)}}{\sum \alpha - predikat}, \text{ i = 1, 2, 3.} \tag{7}$$

The fuzzification process in this study involves employing a membership function based on the variables of the Developing Village Index (DVI), Human Development Index (HDI), and Community Standard of Living Index. Previous research stated that each consequence of a rule in the IF-THEN form must be articulated through a fuzzy set represented by a membership function [36].

2.7. Text Mining Cosine Similarity

Text mining is obtaining high-quality information from the text to extract helpful information from a collection of texts or documents [37]. Using text mining algorithms, analyzing the focal point of a scientific study can be carried out and represent the problems to be explored [38]. The cosine similarity is calculated using the following Eq. (8):

$$CS = \frac{\sum_{i=1}^{n} Aj \; x \; Bj}{\sqrt{\sum_{i=1}^{n} (Ai)^2 \; x \; \sqrt{\sum_{i=1}^{n} (Bi)^2}}} \tag{8}$$

Text similarity measurement aims to find similarities between text documents, which is the basis for most information extraction, retrieval, and text mining problems.

2.8. Research Design

Hybrid algorithm research has received more attention because the principle limitations of one algorithm still have some gaps in decisions [39], so more efficient techniques are always needed to improve application performance [40]. There are several ways that research design can contribute to theory development in the field. Firstly, a well-designed study can provide evidence to support or refute existing theories, helping to refine and advance them. This is particularly important in scientific research, where theories are continually tested and refined through empirical evidence. Secondly, research design can also help to generate new theories by identifying patterns or relationships that were previously unknown or unexplored. The research design can contribute to developing new methods or techniques for collecting and analyzing data. By developing innovative research designs and methods, this research can contribute to the overall advancement of the field and enable others to build the research. Overall, the theoretical contribution of research design lies in its ability to help researchers generate new knowledge and advance existing theories through rigorous and systematic. Figure 2 illustrates the research design.

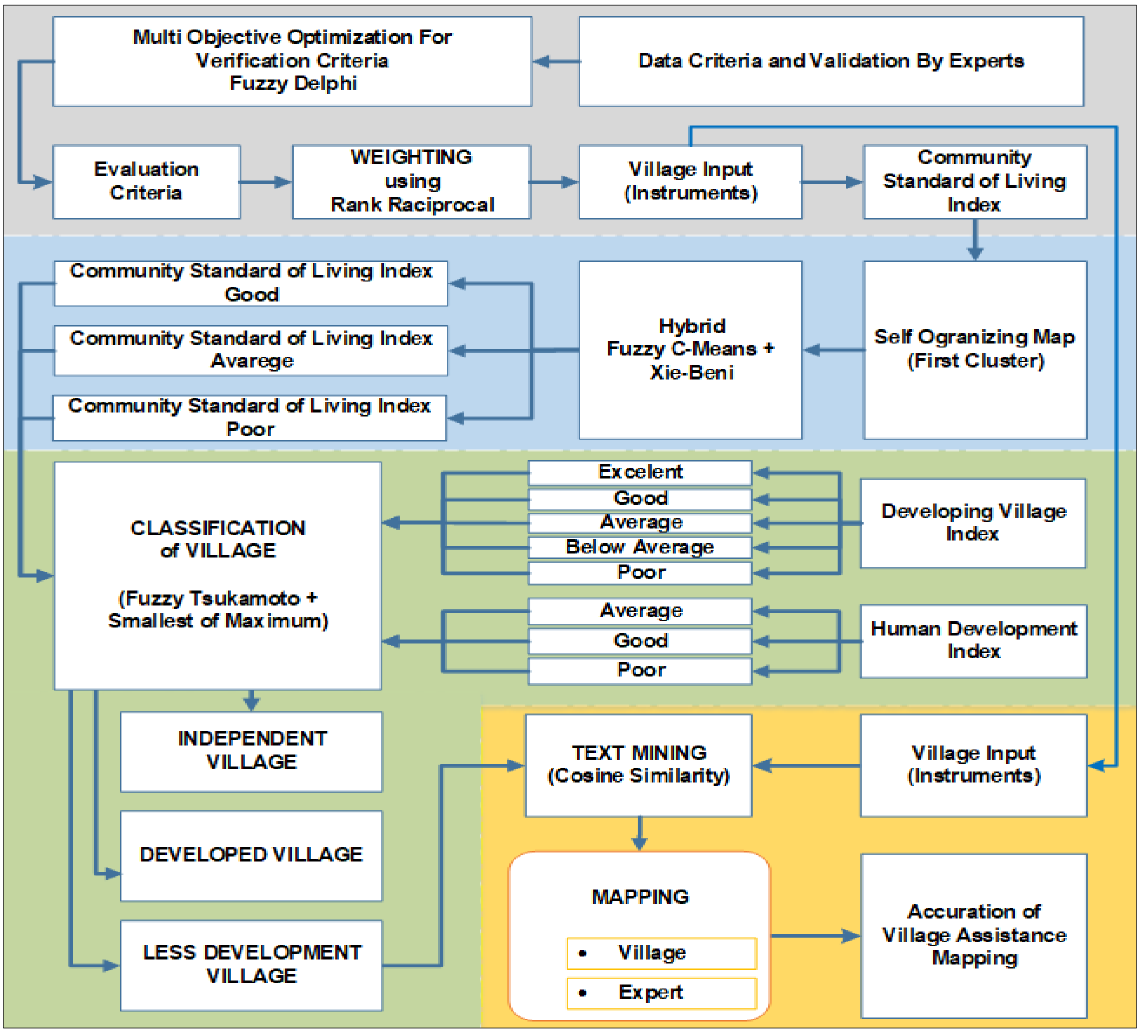


Fig. 2. Research design

Aware of Figure 2 illustrates that this research consists of several critical processes, including:

- Identify criteria and rank villages using multi-object optimization techniques Fuzzy Delphi and Rank Reciprocal.
- Build a clustering Community Standard of Living index for each village using hybrid Fuzzy C-Means, Self-Organizing Map, and Xie-Beni.
- Build a village classification based on the Developing Village Index (DVI), Human Development Index (HDI), and Community Standard of Living Index of each village using improvements to Fuzzy Tsukamoto and Smallest of Maximum.
- Determine the field of expertise of village assistants relevant to problems in a village through text mining using the Cosine Similarity algorithm.

The entire explanation of the stages of the process above is outlined in this study, while matters related to the technique of applying each methodology will be explained in the next section.

3. Illustrative By Case Study

This section presents the results of the research studies conducted, divided into several sections, each presenting findings related to the research objective. This problem is divided into indexing, clustering, classification, recommendation, and discussion subsections.

3.1. Identification and Determining of Criteria

In the first stage, all the experts provide the agreement value for all criteria from a scale of 1 to 5, then convert into weights for each criterion based on the fuzzy scale. Each approval weight consists of membership sets: 1 [0 ; 0 ; 0.2], 2 [0 ; 0.2 ; 0.4], 3 [0.2 ; 0.4 ; 0.6], 4 [0.4 ; 0.6 ; 0.8], 5 [0.6 ; 0.8 ; 1] , which is used to calculate the threshold criteria value, as shown in Table 2.

Table 2

Determining the average of each criteria

Experts	Average(M1, M2, M3)														
	C1			C2			C7			C37			C38		
1	0.6	0.8	1	0.4	0.6	0.8	1	0.6	0.8	0.4	0.6	0.8	0.6	0.8	1
2	0.6	0.8	1	0.4	0.6	0.8	1	0.6	0.8	0.6	0.8	1	0.6	0.8	1
3	0.6	0.8	1	0.6	0.8	1	1	0.4	0.6	0.6	0.8	1	0.6	0.8	1
4	0.6	0.8	1	0.6	0.8	1	1	0.4	0.6	0.6	0.8	1	0.4	0.6	0.8
5	0.6	0.8	1	0.6	0.8	1	1	0.4	0.6	0.2	0.4	0.6	0.2	0.4	0.6
6	0.6	0.8	1	0.6	0.8	1	1	0.6	0.8	0.6	0.8	1	0.6	0.8	1
7	0.6	0.8	1	0.6	0.8	1	1	0.6	0.8	0.6	0.8	1	0.2	0.4	0.6
8	0.6	0.8	1	0.6	0.8	1	1	0.6	0.8	0.2	0.4	0.6	0.6	0.8	1
9	0.6	0.8	1	0.6	0.8	1	1	0.6	0.8	0.4	0.6	0.8	0.2	0.4	0.6
10	0.6	0.8	1	0.6	0.8	1	0.8	0.6	0.8	0.2	0.4	0.6	0.6	0.8	1
11	0.6	0.8	1	0.6	0.8	1	0.8	0.6	0.8	0.4	0.6	0.8	0.6	0.8	1
12	0.6	0.8	1	0.6	0.8	1	0.8	0.6	0.8	0.2	0.4	0.6	0.4	0.6	0.8
13	0.6	0.8	1	0.6	0.8	1	0.8	0.4	0.6	0.4	0.6	0.8	0.2	0.4	0.6
14	0.6	0.8	1	0.6	0.8	1	1	0.6	0.8	0.2	0.4	0.6	0.2	0.4	0.6
15	0.6	0.8	1	0.6	0.8	1	0.8	0.4	0.6	0.2	0.4	0.6	0.6	0.8	1
16	0.4	0.6	0.8	0.6	0.8	1	0.8	0.4	0.6	0.6	0.8	1	0.4	0.6	0.8
Fuzzy Average	0.58	0.78	0.98	0.57	0.77	0.97	0.92	0.52	0.72	0.4	0.6	0.8	0.43	0.65	0.83
	m1	m2	m3	m1	m2	m3	m3	m1	m2	m1	m2	m3	m1	m2	m3

Based on Table 2, the threshold value for each criterion is obtained using Eq. (1). The threshold value for each criterion is shown in Table 3.

Table 3

Threshold value of each criteria

Expert	C1	C2	..	C7	..	C37	C38
1	0.0125	0.175	..	0.075	..	0.0	0.162
2	0.0125	0.175	..	0.075	..	0.2	0.162
3	0.0125	0.2	..	0.125	..	0.2	0.162
4	0.0125	0.2	..	0.125	..	0.2	0.037
5	0.0125	0.2	..	0.125	..	0.2	0.2
6	0.0125	0.025	..	0.075	..	0.2	0.162
7	0.0125	0.025	..	0.075	..	0.2	0.2
8	0.0125	0.025	..	0.075	..	0.2	0.162
9	0.0125	0.025	..	0.075	..	0.0	0.2
10	0.0125	0.025	..	0.075	..	0.2	0.162
11	0.0125	0.025	..	0.075	..	0.0	0.162
12	0.0125	0.025	..	0.075	..	0.2	0.037
13	0.0125	0.025	..	0.125	..	0.0	0.2
14	0.0125	0.025	..	0.075	..	0.2	0.2
15	0.0125	0.025	..	0.125	..	0.2	0.162
16	0.1875	0.025	..	0.125	..	0.2	0.037
Average per criteria	0.0234	0.043	..	0.093	..	0.15	0.162
Average all criteria			0.097				

In determining the threshold value, if all criteria are ≤ 2, all criteria are declared to be in consensus, or all experts have agreed on each decision. Finally, determine each criterion ranking so that each criterion order based on the fuzzy score can be known using Eq. (2). The evaluation result is shown in Table 4.

Table 4
Result the result of evaluating the criteria

Code	Criteria	Score	Weight	Rank
C1	Welfare	12.600	0.23652	1
C2	Residential	12.400	0.11826	2
C3	Land of residence	12.200	0.07884	3
C10	Get drinking water	12.200	0.05913	4
C4	Widest flooring	12.000	0.04730	5
C15	Installed electrical power	12.000	0.03942	6
C8	The widest-quality roof	11.867	0.03378	7
C5	Type of walls	11.800	0.02956	8
C6	Roofs	11.600	0.02628	9
C7	Wall quality	11.600	0.02365	10
C16	Fuel for cooking	10.800	0.02150	11
C31	Gold worth 10 grams	10.800	0.01478	12
C26	Ship	10.600	0.01244	13
C17	Gas of 5.5kg	10.200	0.01075	14
C38	KUR program participants	10.200	0.01028	15
C14	Sources of illumination	10.000	0.00946	16
C19	Computer	10.000	0.01971	17
C25	Tractor	10.000	0.01819	18
C11	Defecation facilities	9.800	0.01689	19
C21	Motorcycle	9.800	0.01576	20
C28	Air conditioning	9.800	0.01391	21
C9	Drinking water	9.600	0.01314	22
C13	Fecal landfill	9.600	0.01182	23
C37	Raskin program participants	9.600	0.01126	24
C12	Toilets	9.600	0.00985	25
C24	Motor Ownership	9.600	0.00909	26
C23	Boat	9.400	0.00876	27
C27	Refrigerator	9.400	0.00844	28
C32	Ownership of Land	9.400	0.00815	29
C33	Home Locations	9.400	0.00788	30
C22	Car	9.400	0.00762	31
C35	Dormitori	9.200	0.00739	32
C29	Water heater	9.000	0.00716	33
C30	Television	9.000	0.00695	34
C20	Bicycles	9.000	0.00675	35
C18	Telephone	8.800	0.00657	36
C34	Household members joint businesses	8.800	0.00639	37
C36	PKH program	8.400	0.00622	38

After the evaluation of the criteria was carried out, there was a change in the ranking position on several criteria. The initial ranking is determined based on the interview of government officers.
The weight of the criteria and instruments is determined using the Rank Reciprocal given by Eq. (3), in which the number of criteria is 38.

$$W_j(RR) = \frac{1/R_j}{\sum_{k=1}^{n}(1/R_k)}$$

$$Rj = 38$$

W_j : 1..38 $= \frac{1}{1} \dots \dots \frac{1}{38}$

W_j : 1..38 = 1.00 … … 0.263

W_j : 1..38 = 1.00+…+…+0.263= 4.2279

W_j : 1 $= \frac{1}{4.2279} = 0.2365$

W_j : 38 $= \frac{0.0263}{4.2279} = 0.00622$

Village input is an accumulated score obtained by each village based on answers from respondents from each village, and then the village input score will be used as a guideline for village index provisions.

3.2. Village Ranking

This investigation applies a case study involving 900 respondents from 30 villages. The structured format of input from the community is illustrated in Table A.2, where each response from representatives of each village is calculated using Eq. (4). The village ranking procedure is carried out as a reference for evaluating the level of community welfare in each village. After the data collection stage, the next stage is to calculate the assessment by the community. Participants from a specific village completed the questionnaire, whereby the weight for each response was determined by multiplying the criterion weight with the corresponding response provided in the instrument. Finally, all weights according to answers are calculated and set as response scores. The result of the village ranking is shown in Table 5.

Table 5
Village ranking

Village	Score	Ranking
Babana (A)	0.52550	1
Galung (B)	0.47335	2
Manase (C)	0.45890	3
Tuppu (D)	0.45794	4
Mesakada (E)	0.39561	5
Ampoe (F)	0.38065	6
Binanga (G)	0.37964	7
Lembang (H)	0.36849	8
Letta (I)	0.36471	9
Kariango (J)	0.35804	10
Buttu (K)	0.34530	11
Allo (L)	0.33649	12
Riattang (M)	0.33464	13
Sidrap (N)	0.33390	14
Tanete (O)	0.33390	15
Rubae (P)	0.33390	16
Sekkang (Q)	0.33073	17
Sawitto (R)	0.32924	18
Kire (S)	0.32778	19
Kombiling (T)	0.32100	20
Sidodadi (U)	0.31983	21
Polman (V)	0.31860	22

Village	Score	Ranking
Pajalele (W)	0.31336	23
Rungkeke (X)	0.31064	24
Tabbae (Y)	0.31047	25
Bone (Z)	0.30667	26
Jeneponto (AA)	0.29003	27
Masale (AB)	0.26416	28
Sudiang (AC)	0.21900	29
Abdesir (AD)	0.20524	30

Table 5 illustrates the sequential arrangement of villages determined by their ranking order, as derived from the village score or Community Standard of Living Index grounded on community feedback.

3.3. Village Clustering

This section shows the proposed clustering method using a numerical example and a case study in the South Sulawesi province.

3.3.1. Input clustering

This research employed a combined method that incorporated the Self-Organizing Map, Fuzzy C-Means, and Xie-Beni techniques to group villages according to CSLI criteria for development. The outcomes of this clustering analysis can potentially guide policy choices aimed at improving community welfare. The following sections provide further details on these discoveries. The utilization of weights has demonstrated its exceptional capacity to dynamically enhance the overall performance of the clustering technique [41]. The stages of clustering include:

i. Initiation input and neuron output.

The Input variables used are CSLI Score, Head of Family, and Number of Residences, while the output neurons include CLSI-Good (cluster 1), CSLI-Average (cluster 2), and CSLI-Poor (cluster 3), the learning rate value is 0.6. The input data for each village data is shown in Table A.3.

ii. The formula of criteria normalization.

Criteria normalization significantly improves both processing efficiency and accuracy [42]. After making a comparison of the input data on the criteria, the largest value is obtained from each criterion, namely: Max (VS) = 0.5255, Max (VP) = 1398, Max (VR) = 4198. The next process is calculating the normalization criteria using Eq. (9).

$$r_{ij} = \frac{X_{ij}}{Max(X_{ij})} \quad (9)$$

Written in example $0.629363 = \frac{0.33073}{0.5255}$. The result is shown in Table A.4.

iii. Finding the shortest distance from each output neuron to the input data using the formula:

Distance values are calculated using Eq. (10) : $D_j = \sum_j (w_{ji} - x_j)2$ (10)

The results are displayed in Table 6.

Table 6
The shortest distance value

Village	Distance			Minimum value	BMU
	X1	X2	X3		
Village - A	0.71480	0.18135	0.03643	0.03643	3
Village - B	0.28085	0.87867	0.57307	0.28085	1

In Table 6, the calculation process and data changes in the first iteration 1 are displayed, where it can be known that the selected cluster data is the one with the smallest value.

iv. Each weight w_{ij} is updated by neighbouring weights.

The Best Matching Unit (BMU) is a labelling based on clusters containing the smallest value used to calculate weight values until the data is convergent. If there are still changes to the cluster membership, the next iteration updates the layer weighting using Eq. (11):

$$W_{ij}(\text{new}) = W_{ij}(\text{old}) + \alpha\left(X_i\text{-}W_{ij}(\text{old})\right) \tag{11}$$

updates are made to weights with the closest distance to the data. If the specified learning rate (α) is 2.28882E-06, then the change in weight to the Village-A data in iteration 18 is written as follows:

$$= \begin{bmatrix} 0.59656 \\ 0.91906 \\ 0.62338 \end{bmatrix} + (2.2882\text{E-}06) * \left(\begin{bmatrix} 0.525 \\ 599 \\ 2434 \end{bmatrix} - \begin{bmatrix} 0.59656 \\ 0.91906 \\ 0.62338 \end{bmatrix} \right) = \begin{bmatrix} 0.59656 \\ 0.92043 \\ 0.62895 \end{bmatrix}$$

After calculating the weight of all village data, the Village-AD weight was found in the last iteration, as shown in Table 7.

Table 7
Weight on the last iteration

Village	W1	W2	W3
	0.596580062	0.684005088	0.754091366
Village-AD	0.918972384	0.425102021	0.462436764
	0.623388681	0.254118616	0.288860264

Convergence is the stage at which the SOM network has managed to find the optimal representation of a given data [43]. Convergence data as input clustering are shown in Table A.5. The weight in the last iteration is considered a convergent state, indicating no change in the weight data.

3.3.2. Determining membership degree

Once all the required components for the village clustering procedure are gathered, the final stage involves computing the objective value. The Fuzzy C-means clustering approach is an effective clustering method that has been successfully applied to a number of real-world problems [44]. This computation is executed with predefined parameters: weight = 2, iterations = 100, and an epsilon value set to 0.000001. The initial weight for the Fuzzy C-Means was established using the convergent value found in Table A.5. The next step is to determine the membership value using Eq. (12):

$$Qj = \sum_{k=1}^{c} \mu\ ik \tag{12}$$

The outcomes of membership value calculations are presented in Table 8.

Table 8
Membership value

MiU C1			MiU C2			MiU C3		
0.01729	0.00136	0.01717	0.00011	0.00001	0.00011	0.01706	0.00134	0.01695
0.01613	0.03957	0.07272	0.09704	0.23797	0.43736	0.32776	0.80380	1.47726
0.00157	0.00074	0.00372	0.00035	0.00017	0.00084	0.00886	0.00421	0.02106
0.04074	0.00240	0.00905	0.00014	0.00001	0.00003	0.00201	0.00012	0.00045
0.07458	0.01536	0.00324	0.00316	0.00065	0.00014	0.00014	0.00003	0.00001
0.26825	0.08415	0.01705	0.02640	0.00828	0.00168	0.00108	0.00034	0.00007
0.00042	0.00241	0.00426	0.01388	0.08015	0.14185	0.04349	0.25105	0.44430
0.06261	0.01628	0.00135	0.00424	0.00110	0.00009	0.00003	0.00001	0.00000
0.09987	0.02831	0.00163	0.00803	0.00227	0.00013	0.00003	0.00001	0.00000
0.30447	0.10789	0.01668	0.03823	0.01355	0.00209	0.00091	0.00032	0.00005
0.00784	0.00282	0.00153	0.00101	0.00036	0.00020	0.00030	0.00011	0.00006

MiU C1			MiU C2			MiU C3		
0.00352	0.00206	0.00177	0.00120	0.00070	0.00060	0.00089	0.00052	0.00045
0.00001	0.00002	0.00003	0.00030	0.00517	0.00892	0.00089	0.01540	0.02660
0.00002	0.00001	0.00002	0.00052	0.02451	0.04154	0.00150	0.07043	0.11940
0.00062	0.00072	0.00092	0.00083	0.00096	0.00124	0.00138	0.00160	0.00207
0.00001	0.00001	0.00002	0.00061	0.02748	0.04573	0.00170	0.07610	0.12665
0.01582	0.00630	0.00124	0.00251	0.00100	0.00020	0.00010	0.00004	0.00001
0.00000	0.00001	0.00002	0.00034	0.00937	0.01494	0.00086	0.02383	0.03802
0.06870	0.02666	0.00004	0.01034	0.00401	0.00001	0.00001	0. 00001	0. 00001
0.00001	0.00001	0.00001	0.00005	0.01958	0.03145	0.00013	0.05053	0.08117
0.00831	0.00444	0.00142	0.00237	0.00126	0.00040	0.00024	0.00013	0.00004
0.18510	0.07847	0.00533	0.03326	0.01410	0.00096	0.00015	0.00006	0.00000
0.15503	0.06547	0.00434	0.02765	0.01168	0.00077	0.00012	0.00005	0.00000
0.00001	0.00001	0.00001	0.00054	0.03408	0.05217	0.00127	0.07984	0.12219
0.05303	0.02596	0.00079	0.01271	0.00622	0.00019	0.00001	0.00001	0.00001
0.25714	0.13548	0.01603	0.07138	0.03761	0.00445	0.00100	0.00053	0.00006
0.31870	0.23018	0.07649	0.16625	0.12008	0.03990	0.01836	0.01326	0.00441

A fuzzy cluster centroid signifies the data points distributed within each class. Consequently, establishing a cluster centre is carried out before determining the objective value using Eq. (13):

$$Vkj = \frac{\Sigma_{i=1}^{n}((\mu ik)^{w} * Xij)}{\Sigma_{ij}^{n}(\mu ik)^{w}} \tag{13}$$

The results of the centroid value calculation are shown in Table A.6. The membership degree is utilized to account for the level of support in describing the uncertain effect of the evaluation factor on engineering stability [45]. Membership degrees are updated in each iteration of Fuzzy C-Means based on the distance between the data point and the cluster centre. The results of calculating the membership degree are shown in Table 9.

Table 9
Membership degree

C1	C2	C3	Degree Value
0.00284	0.01732	0.00166	0.02182
0.04355	0.04971	0.15855	0.25181
0.00148	0.01644	0.00186	0.01978
0.00794	0.01595	0.00131	0.02519
0.01482	0.00627	0.00086	0.02195
0.01014	0.00048	0.00017	0.01079
0.00110	0.00170	0.02659	0.02939
0.01401	0.00723	0.00090	0.02215
0.01077	0.00152	0.00027	0.01256
0.01519	0.00098	0.00038	0.01654
0.00079	0.00942	0.00026	0.01047
0.00011	0.00418	5.00E-05	0.00434
0.00041	0.00898	0.00042	0.00981
0.00361	0.02225	0.00974	0.03560
0.00056	0.00148	0.01464	0.01668
0.00023	0.00691	0.00018	0.00733
0.00038	0.00096	0.01246	0.01379
0.00274	0.01304	0.00060	0.01637
0.00397	0.01897	0.01555	0.03849
0.01270	0.00354	0.00052	0.01677
0.00185	0.00576	0.02120	0.02881
0.00091	0.00949	0.00028	0.01067
0.00031	0.00077	0.01146	0.01254

C1	C2	C3	Degree Value
0.01302	0.00571	0.00077	0.01950
0.01793	0.00142	0.00060	0.01996
0.00042	0.00069	0.01548	0.01659
0.05282	0.01406	0.00960	0.07648
Objective			0.79111

Table 9 elucidates that the revised membership degrees are subsequently employed to recompute the cluster centre, and this iterative process persists until convergence is attained.

3.3.3. Cluster validity index

Based on the membership degree value accumulation, the objective value is calculated using Xie-Beni, written by Eq. (14):

$$P_t = \sum_{i=1}^{n}\sum_{k=1}^{c}\left(\left[\sum_{j=1}^{n}(x_{ik} - v_{kj}\right](\mu_{ik})^{w}\right) \tag{14}$$

Subtracting the new objective value from the previous objective = 0, the result obtained in the first iteration is 1.283007700. Table A.7 displays the objective values between 1 and 28 for each iteration. Based on the data shown in Table A.7, it is known that the iteration process stops at objective values 7.396630273826E-7. The optimal cluster is determined by comparing the distance objective value to the number of clusters using Eq. (14) until the optimum number of clusters with the smallest value is determined. The result is determined that the optimum number of clusters to be used is 3. The results of testing the number of 3 to 6 clusters are shown in Table 10.

Table 10
Result of testing optimum cluster using Xie-Beni

Cluster	Xie-Beni
3	7.396630273826E-7
4	8.1562379394828E-7
5	7.9070423741889E-7
6	9.6305179772549E-7

3.3.4. Determining convergence cluster

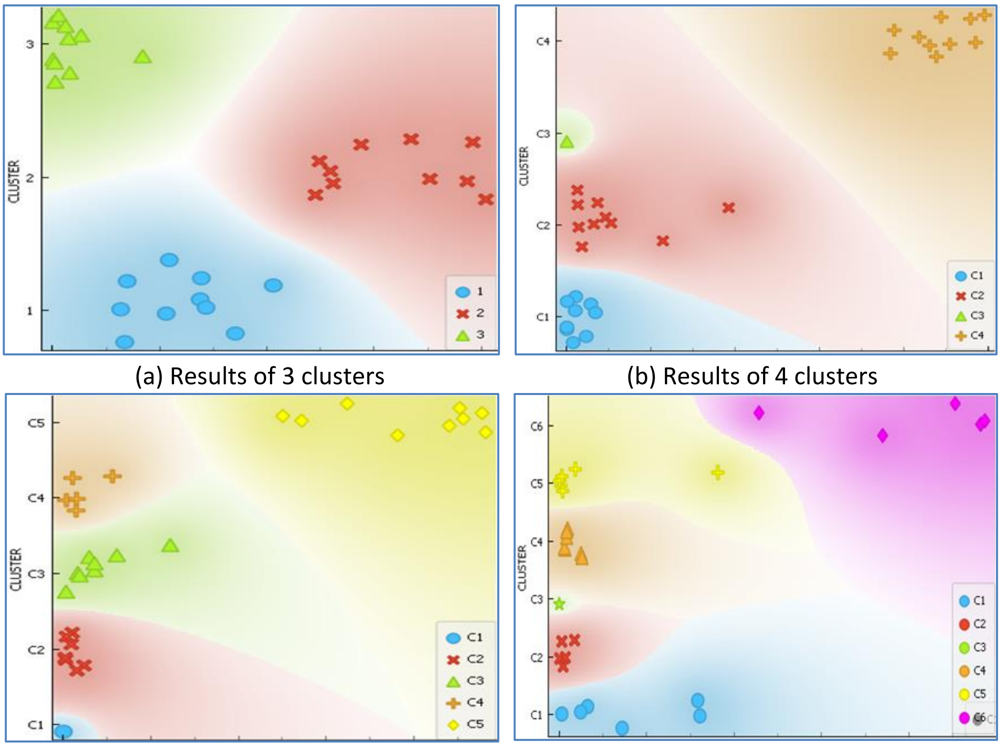
The cluster centre is moved iteratively to a high-density region by maximizing the validity index of the new cluster, and the convergent criteria are met [46]. In Fuzzy C-Means, the calculation of membership values assigned by FCM to specific clusters is based on convergent decisions, where clusters with high membership values indicate clusters with greater significance. FCM is often found to be more effective, robust, and consistent in its performance for some instances or applications compared to other clustering algorithms. In order to discern the convergent cluster associated with each village, the latest membership values can be calculated using the following Eq. (15):

$$\mu_{ik} = \frac{\left[\sum_{j=1}^{m}(x_{ik} - v_{kj})^2\right]^{\frac{-1}{w-1}}}{\sum_{k=1}^{c}\left[\sum_{j=1}^{m}(x_{ik} - v_{kj})^2\right]^{\frac{-1}{w-1}}}. \tag{15}$$

It is written in example:

Convergence C1 = 6.28169
= ((0.25856 - 0.53494)2 + (0.02031 - 0.239995)2 + (0.25688 - 0.071016)^2)(-1 / (2 - 1)) .
Convergence C2 = 4.78471
= ((0.25856 -0.25398)2 + (0.02031 - 0.40144)2 + (0.25688 - 0.5093)2)(-1 / (2 - 1)).
Convergence C3 = 2.36558
= ((0.25856 - 0.115)2 + (0.02031 - 0.43328)2 + (0.25688 - 0.7381)2)(-1 / (2 - 1)).

In this study, an optimum clustering analysis was carried out to obtain the value of the number of clusters, where testing was carried out on clusters 3 to 6. Based on the result of analyst optimum clusters, it can be seen that cluster 3 in visual (a) is the best cluster to implement because it can build every data collected generally based on distance with the centroid. Results are shown in Figure 3.



(a) Results of 3 clusters (b) Results of 4 clusters

(c) Results of 5 clusters (d) Results of 6 clusters

Fig. 3. Visualization results of comparison clusters 3 to 6

Calculating convergent values involves data in Table A.5 and Table A.6, where the final results are shown in Table A.8. The conclusive outcomes of the members in the 28th iteration were derived and served as a reference for delineating the clusters corresponding to the 30 villages, as delineated in Table A.9. The primary purpose of using indexes is to search for the optimal number of unknown clusters [47], It is known that often clustering process proposes only one optimal number of clusters. Based on the result of analyst optimum clusters, it can be seen that cluster 3 in visual (a) is the best cluster to implement because it can build every data collected generally based on distance with the centroid.

In hybrid clustering, the aggregation function creates a new grouping similar to all first clusters [48]. Table 11 shows previous research related to hybrid methodologies.

Based on the information in Table 11, it can be seen that hybrid techniques are widely used to solve clustering problems. Likewise, this study uses hybrid SOM, FCM, and Xie-Beni techniques to solve village clustering problems. Findings underscore the efficacy of the employed methodology, particularly the synergistic utilization of the self-organizing map (SOM) and fuzzy C-Means (FCM) techniques. This integration leverages the inherent topological structure of SOM and the adaptive capability of FCM in assigning fuzzy memberships to data points, thereby augmenting the precision of clustering. Furthermore, the inclusion of Xie-Beni integration facilitates an objective evaluation and validation of clustering outcomes through quantitative measurement of clustering quality. However, a notable constraint of the hybrid SOM, FCM, and Xie-Beni approach lies in its susceptibility to the quality of initial data, as the presence of noise in the initial dataset can markedly influence clustering accuracy, potentially resulting in erroneous conclusions.

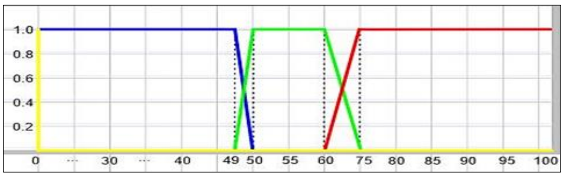
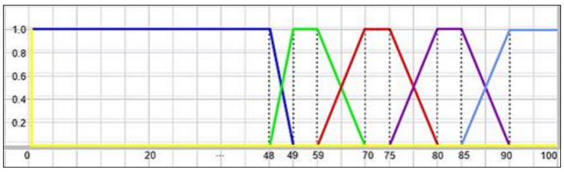
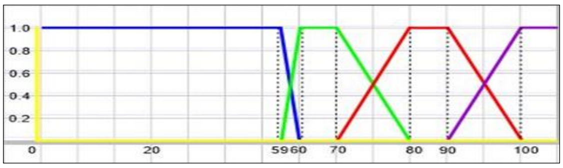
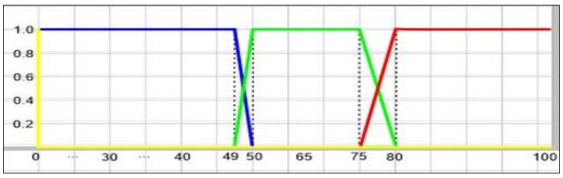
Table 11
Summary of application hybrid technique of past few review papers

Author - Algorithm	Objective	Application	Dataset For Testing	Measure
Liang [49], Fuzzy C-means-user Collaborative filtering.	Solve data distribution problems and improve recommendation accuracy.	Value Analysis and Realization of Artistic Intervention in Rural Revitalization.	IRIS dataset.	Test experiment to determine the efficacy of the FCM-UserCF algorithm.
Shen *et al*.,[50], K-Means and PCA.	Identify accident-prone spots on rural highways.	Clustering techniques using PCA and K-means.	Spatial data of traffic accidents on rural roads	Verified by FCM, Xie-Beni.
Yuan *et al.*,[51], Markov clustering (MCL).	Propose a rural cluster strategy Propose a rural cluster strategy.	Integration of Urbanization and Rural.	Statistical Yearbook of Laizhou City	Comparison of theoretical groupings with actual data from villages.
Ilbeigipour *et al.,* [52], K-means + Self-Organizing Map.	Testing the relationship between death and recovery of COVID-19 patients.	Unsupervised machine learning techniques.	COVID-19 Datasets.	Silhouette, Davis Boldin (DB) calculations, and the average size of optimal intra-cluster distance.
Karthik *et al*.,[53], SDS + K-Means.	Finding optimal clustering points.	Data Leak Prevention in Social Media.	Social Media Database.	True Positive Rate (TPR).
Jin *et al.,* [54], CAABC + K-means.	Optimal clustering.	Generate initial points for K-means.	Iris, Balance-Scale, Wine, Abalone, Musk, Pendigits.	Rosenbrock, Rastrigin, Alpine and Ackley CMC
Faisal and Rahman [55], Hybrid Delphi, Rank Reciprocal, SOM, FCM, Xie-Beni.	Mapping the relevant experts against the rural problem	Hybrid indexing, clustering methodologies.	Dataset of Central Bureau of Statistics Indonesia.	Confusion Matrix, Accurasy, F1 Score, Data actual condition.

3.4 Classification Village

Determining indicators in fuzzy logic is essential because indicators serve as the inputs or outputs of a fuzzy system and play a crucial role in the decision-making process. Indicators related to humanitarian conditions are priorities in achieving several targets of the sustainable development goals [56]. The purpose of forming indicators in the Tsukamoto method is to describe and measure the membership level of an input or output variable in fuzzy sets so that it is easier to apply fuzzy thinking in system modelling. Fuzzification captures and represents the inherent uncertainty in data, enabling a fuzzy system to work effectively with imprecise or ambiguous information. Fuzzy rules are particularly beneficial when data is inaccurate or uncertain because they facilitate probabilistic decision-making by indicating potential occurrences of various conditions [57]. The fuzzification process involves finding the degree of membership to the linguistic terms describing the calculated attribute function [58]. Table 12 shows the function of fuzzy set membership against four indicators in this study.

Table 12
The function of fuzzy set membership for each indicator

Indicator	Attribute	Curve	Fuzzy Set
CSLI-Cluster	Poor Average Good		x1 ≤ 49 50 - 75 ≥ 75
DVI	Poor Below-Average Average Good		x2 ≤ 49 50 - 70 71 - 80 81 - 90 ≥ 90
HDI|	Poor Average Good Excellent		x3 ≤ 60 60 - 80 81 - 90 ≥ 90
Village Development Level (VDL)	Less-Development Developed Independent		z ≤ 49 50 - 80 ≥ 80

This study uses four indicators. Firstly, the CSLI-Cluster Indicator is used to measure the level of welfare of each family in each village. The CSLI-Cluster indicator is obtained from the previous clustering process, consisting of 3 clusters as attributes: CSLI-Good, CSLI-Average, and CSLI-Poor. Secondly, DVI indicators measure aspects of development such as education, health, economy, infrastructure, and environment in each village. Thirdly, HDI indicators assist governments and stakeholders in designing development policies and programs that focus on improving overall human well-being to achieve sustainable development. Finally, VDL is the indicator targeted for the classification process and becomes a guideline for conclusions about the level of development of a village.

3.4.1 Rule evaluation

Fuzzy modelling relies heavily on rule-based fuzzy models, which hold significant influence and find wide-ranging applications in the field of system modelling [59]. In fuzzy Tsukamoto, rule-based knowledge functions are used as heuristics to guide the fuzzy inference process so that strategies for linking numerical data into linguistic variables can be described.

Table 13 uses rule-based knowledge to display the rules used in the fuzzy Tsukamoto classification process.

Table 13
List of rules based on Fuzzy Tsukamoto

No	Rules
1	IF CSLI = GOOD AND DVI= EXCELLENT AND HDI= EXCELLENT THEN DVL = INDEPENDENT
..	..
27	IF CSLI = AVERAGE AND DVI= GOOD AND HDI= AVERAGE THEN DVL = DEVELOPED
..	..
35	IF CSLI = AVERAGE AND DVI= AVERAGE AND HDI= AVERAGE THEN DVL = LESS DEVELOPMENT
..	..
60	IF CSLI = POOR AND DVI= POOR AND HDI= POOR THEN DVL = LESS DEVELOPMENT

Practical exploration of a rule-based knowledge base can be accomplished through the use of an IF clause [60]. Fuzzy systems are typically regarded as a base of knowledge, an inference module, and a defuzzifier [61]. Thus, fuzzy rules are heuristics in decision-making based on the given numerical data.

3.4.2 Input classification

The fuzzy process involves the conversion of inputs and outputs within a system into linguistic terms, enabling the inference of rules that can represent the behaviour of that system [62]. The village assessment phase entails a systematic procedure aimed at gathering the respective worth of each village according to predefined metrics. The valuation of the CSLI-Cluster indicator is derived from the village clustering procedure. In contrast, the VDI and HDI indicators are sourced from Central Bureau of Statistics of South Sulawesi. The assessment data is shown in Table 14.

Table 14
Input classification for each village

Village	CLSI - Cluster	DVI	HDI
A	90	71	69
B	60	85	69.69
C	60	90	69
D	60	87	66.7
E	90	80	69.71
F	90	71	69.69
G	60	84	69.69
H	90	82	69.71
I	30	72	69
J	60	76	69.7
K	30	75	69.7
L	90	68	69.69
M	60	75	69.7
N	90	87	69.7
O	90	64	67
P	30	77	67
Q	90	83	67
R	90	62	67.01
S	30	66	67
T	30	62	67
U	30	61	67
V	30	64	67
W	30	69	67
X	30	78	67
Y	30	73	67

Village	CLSI - Cluster	DVI	HDI
Z	60	68	65
AA	30	67	65
AB	90	71	65
AC	90	70	65
AD	60	70	65

In Table 14, the input value in the CSLI-Cluster indicator consists of 3 parts: CSLI-Good = 90, CSLI-Average = 60, and CSLI-Poor = 30, so each village value is determined based on the clustering results. As an example for the Village - J. After knowing the value on the CSLI Cluster indicator = x1 = 60, then it is stated that the set µPoor consists of [0(n1), 0(n2), 49(n3), 50(n4)]. Then, tracking values based on conditions is carried out in a fuzzy set:

x1 = 60
IF x1 ≤ 0, then µPoor = 0
IF x1 >= 0 AND x1<= 49, then µPoor = 1
IF x1 ≥ 49 AND x1≤ 50, then µPoor = (50 - x1) / (50 - 49)
IF x1 ≥ 50, then µPoor = 0.

Based on the predetermined statement, the eligible conditions are x1 ≥ 50, so the value of Poor membership = 0. The next step is to establish the set µAverage consisting of [59(n1), 70(n2), 75(n3), 80(n4)]. Then, the trace based on conditions is determined in Table 22:

x2 = 76
IF x2 ≤ 59 then µAverage = 0
IF x2 ≥ 59 AND x2 <= 70, then µAverage = (x2 - 59) / (70 - 59)
IF x2 ≥ 70 AND x2 <= 75, then µAverage = 1
IF x2 ≥ 75 AND x2 ≤ 80, then µAverage = (80 - x2) / (80 - 75).
IF x2 ≥ 80 then µAverage = 0.

Based on the predetermined statement, the eligible conditions are x2 ≥ 75 AND x2 ≤ 80, so the µAverage membership value = 0.8.

The next step is to establish the set µGood consisting of [75(n1), 80(n2), 85(n3), 90(n4)].

x2 = 76
IF x2 ≤ 75, then µAverage = 0
IF x2 ≥ 75 AND x2<= 80, then µGood = (x2– 75) / (80– 75)
IF x2 ≥ 80 AND x2 <= 85, then µGood = 1
IF x2 ≥ 85 AND x2 ≤ 90, then µGood = (90 - x2) / (90- 85)
IF x2 ≥ 90, then µGood = 0.

Based on the predetermined statement, the eligible conditions are x2 ≥ 75 AND 76<= 80, so the µGood membership value is 0.2.

Based on the predetermined statement, the eligible conditions are 76 ≥ 85, so the membership value of µGood = 0. After calculating all conditions for each item in the Village–J DVI indicator, the resulting fuzzification value is [0 - 0.8 - 0 - 0.2 - 0].

The next step is to calculate the fuzzification value on the HDI indicator, as it is known that the value on the HDI indicator = x3 = 69.7, and the set µPoor = [0(n1), 0(n2), 59(n3), 60(n4)]. Then, tracking values based on conditions is carried out :

x3 = 69.7
IF x3 ≤ 59, then µAverage = 0
IF x3 ≥ 59 AND x3 <= 60, then µBelowAverage = (x3 - 59) / (60 – 59)
IF x3 ≥ 60 AND x3<= 70, then µBelowAverage = 1

IF x3 ≥70 AND x3 ≤ 80, then μBelowAverage = (80 - x3) / (80 - 70)
IF x3 ≥ 80, then μBelowAverage = 0.

Based on the predetermined statement, the eligible conditions are 69.7 ≥ 60 AND x3 ≤ 70, so the membership value of μBelowAverage = 1. The results of fuzzification are shown in Table 15.

Table 15
The result of fuzzification for each village

	CSLI			DVI					HDI			
	P	A	G	P	A	BA	G	E	P	A	G	E
VILLAGE	0	49	60	0	59	48	75	85	0	59	70	90
	0	50	75	0	70	49	80	90	0	60	80	10
	49	60	100	48	75	59	85	100	59	70	90	10
	50	75	100	49	80	70	90	100	60	80	100	100
VILLAGE - J	0	1	0	0	0.8	0	0.2	0	0	1	0	0
VILLAGE - AD	0	1	0	0	1	0	0	0	0	1	0	0

After calculating all conditions for each item in the DVI indicator, the resulting fuzzification value for VILLAGE - J is [0 - 1 - 0 - 0].

3.4.3 Fuzzy implication

Implications are critical operations in fuzzy function and approximate reasoning [63]. Based on Data in Table 23, it was carried out, finding the minimum value based on the rule. It is known that fuzzy implication functions are one of the leading operators of fuzzy logic defined by binary functions to satisfy the monotonicity of some variable limit conditions calculated using Eq. (16):

$$\alpha_i=\mu_{Ai}(x)\cap \mu_{Bi}(x)=\mathrm{Min}\{\mu_{Ai}(x).\ \mu_{Bi}(x)\} \tag{16}$$

The results of implementing the fuzzy implication function are shown in Table 16.

Table 16
Result of fuzzy implication for each village

Villages		[V-A]	[V-A]	[V-J]	[V-J]	[V-Y]	[V-Y]	..	[V-AD]	[V-AD]
Membership		MIU	Z	MIU	Z	MIU	Z	..	MIU	Z
Rules	R27	0	[49;80]	0.2	[49.2;79]	0	[49;80]	..	0	[49;80]
	..	..	..	..	..	..	..	..	..	..
	R35	0	[50]	0.8	[49.2]	0	[50]	..	1	[0.49]

Based on the data in Table 16, the results of fuzzy implication calculations are displayed and determined according to the Rule value. An example of calculating the value of implications for Village-J is given using the following rules 27 and 35:
Based on rule 27, the value of each attribute mentioned in the rule is set, and the smallest value is set as a choice, as follows:

Value of the Average attribute on the CSLI indicator = 1.
Value of the Good attribute on the DVI indicator = 0.2.
Value of the Average attribute on the HDI indicator = 1

Then, using Eq. (16), it is set fuzzy implication (α) = (min (1;0.2;1)) = 0.2.
In rule 27, the value of the DEVELOPED attribute in the Village Development Level indicator is calculated after obtaining the smallest membership value. Therefore, it is specified in the set μ(DEVELOPED) = [49(n1);50(n2);75(n3);80(n4)], LB = 0, and UB = 100.

The next step is calculating the fuzzy value using the membership degree (0.2) against the lower limit using the membership function set in the Village Development Level indicator in Table 12.

$$\alpha = 0.2 = \frac{Z - 50}{50 - 49} = 49 + (\alpha * (N2 - N1)) = 49 + (0.2 * (50 - 49)) = 49.2.$$

Next, calculate the fuzzy value using the membership degree (0.2) against the upper limit using the membership function set in the Village Development Level indicator in Table 12.

$$\alpha = \frac{80 - Z}{80 - 75} = -0.2 * (N4 - N3) + N4 = -0.2 * (80 - 75) + 80 = 79.$$

Finally, if the minimum value on MiU = 1 or the value of z > LB, or z < UB, then the value of z is expressed as the value of membership degree in the DEVELOPED attribute. Therefore, based on the calculation results above, the fuzzy values in rule 27 are 49.2 and 79.

In rule 35, the value of the Less Development attribute in the Village Development Level indicator is calculated after obtaining the smallest membership value. Therefore, it is specified in the set μ(LESS-DEVELOPMENT) = [0(n1);0(n2) ;49(n3);50(n4)], LB = 0, and UB = 100.

The next step is calculating the membership degree for value 0.8 using the function set on the VDL indicator in Table 12.

$$\alpha = 0.8 = \frac{Z - 0}{0 - 0} = 0 + (\alpha * (0 - 0)) = 0 + (0.8 * (0 - 0)) = 0.$$

Next, the calculation of the membership degree for value 0.8 using the membership function is written in Table 20.

$$\alpha = \frac{50 - Z}{50 - 49} = -0.8 * (N4 - N3) + N4 = -0.8 * (50 - 49) + 50 = 49.2$$

Finally, the condition is set if the minimum value on Miu = 1 or the value of Z > LB, or z < UB, then the value of z is expressed as the value of membership degree in the Less Development attribute. Therefore, based on the calculation results above, the fuzzy value in Rule-35 = 49.2.

3.4.4. Defuzzification

Defuzzification is an effective process to get one number from the output of a fuzzy set as an explicit output value. It is known that the value used in the defuzzification process is the value that exists at the composition stage of each rule. Implementing SM Defuzzification using Eq. (7) is described through Village-J data. Table A.10 exhibits information attributed to Village-J. However, the data employed for the defuzzification procedure is limited to the rule section where the a-predicate value exceeds 0, specifically within the 27th and 35th rules. Fuzzy Tsukamoto and the Smallest of Maximum methods are applied in the village classification process so that villages in the category of less development can be selected in large numbers. Data on selected villages using a combination of Tsukamoto and SM methods are shown in Table 17.

Table 17
Classification of villages using Fuzzy Tsukamoto and Smallest of Maximum

Village	Score	Status
A - F - I - M - W - Y - A B - AC - AD	49	LESS DEVELOPMENT
B - C - E - G - H - Q	50	DEVELOPED
D - N - P - X - V	49.4	LESS DEVELOPMENT
J - T- AA	49.2	LESS DEVELOPMENT
L - U - Z	49.18	LESS DEVELOPMENT
K	24.5	LESS DEVELOPMENT
O	49.54	DEVELOPED
R	49.72	DEVELOPED
S	49.36	LESS DEVELOPMENT

Figure 4 elucidates that the amalgamation of Tsukamoto and Smallest of Maximum (SoM) culminated in the categorization of villages, with 22 villages identified under the classification of "Less Development," Tsukamoto and Weighted Average (WA) encompassing 8 villages, Mamdani and Mean of Maximum (MoM) comprising 20 villages, and Mamdani and Smallest of Maximum accounting for 20 villages. Consequently, it is asserted that the combination of Tsukamoto and Smallest of Maximum is an optimal choice for discerning villages earmarked as priority groups for developmental interventions. In assessing the reliability of the proposed technique, a comparative analysis was conducted with the Mamdani and Smallest of Maximum methods.

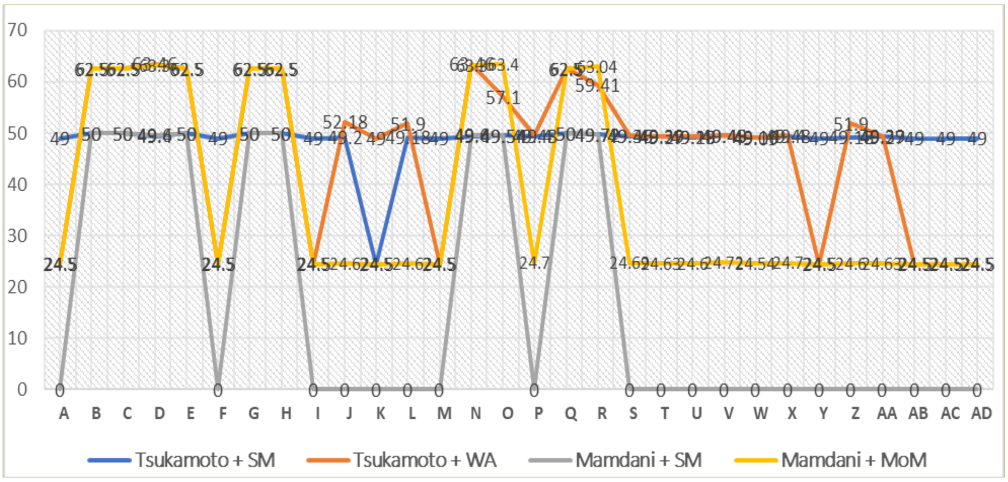


Fig. 4. Comparison of village classification

The combination of the Fuzzy Tsukamoto and Smllest of Maximum methods has the advantage of being adaptable to handle specific cases that require emphasis on the smallest maximum value as well as adapting to data dynamics and changing conditions, thus ensuring a model or system remains relevant and works effectively. The synergistic application of the Fuzzy Tsukamoto and Smallest of Maximum methods presents several advantages. Nevertheless, it is imperative to acknowledge certain limitations, notably the increased complexity in model interpretation resulting from the intricate combination of the Fuzzy Tsukamoto rule and the Smallest of Maximum principles. This intricacy introduces challenges in effecting adjustments and necessitates more meticulous parameter settings.

3.5. Mapping Experts to Villages

This section describes the procedure results to recommend assistance personnel to a respective village according to the field of knowledge needed. The results of the classification indicate that villages in the Less Development category should be given priority for development, as many of these villages have low levels of welfare. The process of recommending experts to be placed in "Less Development" group villages was determined using a dataset of community feedback, employing the cosine similarity method. The dataset of community comments is shown in Table 18.

Table 18
Dataset of community comments

ID	Comments
Query	*agricultural*
D1	irrigation in our village is impaired resulting in suboptimal agricultural yields
D2	village assets have not reached its maximum potential
D3	Unlawful logging has occurred in the eastern forest region
D4	village assistance for the development of UKM programme in our village

The input data is the data from the query from the database, which is then determined as a Dataset. The next step is to build the TFIDF value shown in Table 19.

Table 19
Building TF and IDF value

No	Term	TF					DF	IDF
		Q	D1	D2	D3	D4		
1	UKM	0	0	0	0	1	1	0.69897
2	Unlawful	0	0	0	1	0	1	0.69897
3	agricultural	1	1	0	0	0	2	0.39794
4	assets	0	0	1	0	0	1	0.69897
5	assistance	0	0	0	0	1	1	0.69897
6	development	0	0	0	0	1	1	0.69897
7	eastern	0	0	0	1	0	1	0.69897
8	for	0	0	0	0	1	1	0.69897
9	forest	0	0	0	1	0	1	0.69897
10	hava	0	0	1	1	0	2	0.39794
11	impaired	0	1	0	0	0	1	0.69897
12	in	0	2	0	1	1	3	0.221849
13	irrigation	0	1	0	0	0	1	0.69897
14	is	0	1	0	0	0	1	0.69897
15	its	0	0	1	0	0	1	0.69897
16	logging	0	0	0	1	0	1	0.69897
17	maximum	0	0	1	0	0	1	0.69897
18	not	0	0	1	0	0	1	0.69897
19	occurred	0	0	0	1	0	1	0.69897
20	of	0	0	0	0	1	1	0.69897
21	our	0	1	0	0	1	2	0.39794
22	potential	0	0	1	0	0	1	0.69897
23	programme	0	0	0	0	1	1	0.69897
24	reached	0	0	1	0	0	1	0.69897
25	region	0	0	0	1	0	1	0.69897
26	resulting	0	1	0	0	0	1	0.69897
27	suboptimal	0	1	0	0	0	1	0.69897
28	the	0	0	0	1	1	2	0.39794
29	village	0	1	1	0	2	3	0.221849
30	yields	0	1	0	0	0	1	0.69897

Based on the data in Table 19, it is known that the term frequency (tf) value is obtained from the number of words or the number of occurrences of words in each document, while the document frequency (df) value is obtained from calculating the number of documents containing certain words using the Eq. (17):

IDF= $tfij\ x$ log n/df (17)

which n = number of documents is 5, df = 2, IDF 0.39794 = log (5 / 2).

Table 20
Counting of TFIDF weight for term

No	Term	Q	D1	D2	D3	D4
1	UKM	0	0	0	0	0.69897
2	Unlawful	0	0	0	0.69897	0
3	agricultural	0.39794	0.39794	0	0	0
4	assets	0	0	0.69897	0	0
5	assistance	0	0	0	0	0.69897
6	development	0	0	0	0	0.69897
7	eastern	0	0	0	0.69897	0
8	for	0	0	0	0	0.69897
9	forest	0	0	0	0.69897	0
10	have	0	0	0.39794	0.39794	0
11	impaired	0	0.69897	0	0	0
12	in	0	0.443698	0	0.221849	0.221849
13	irrigation	0	0.69897	0	0	0
14	is	0	0.69897	0	0	0
15	its	0	0	0.69897	0	0
16	logging	0	0	0	0.69897	0
17	maximum	0	0	0.69897	0	0
18	not	0	0	0.69897	0	0
19	occurred	0	0	0	0.69897	0
20	of	0	0	0	0	0.69897
21	our	0	0.39794	0	0	0.39794
22	potential	0	0	0.69897	0	0
23	programme	0	0	0	0	0.69897
24	reached	0	0	0.69897	0	0
25	region	0	0	0	0.69897	0
26	resulting	0	0.69897	0	0	0
27	suboptimal	0	0.69897	0	0	0
28	the	0	0	0	0.39794	0.39794
29	village	0	0.221849	0.221849	0	0.443698
30	yields	0	0.69897	0	0	0

The TFIDF value is generated based on the calculation using Eq. (18):

$wij = tfij\ x\ idfi$ (18)

It is written by example: Wij = 0.39794 = 1 * 0.39794, which calculates the TFIDF weight value involving the data in Tables 19 and 20. Subsequently, the following procedure involves computing the dot product value, as depicted in Table 21.

Table 21
Counting dot product value

No	Term	D1	D2	D3	D4
1	UKM	0	0	0	0
2	Unlawful	0	0	0	0
3	agricultural	0.158356	0	0	0
4	assets	0	0	0	0
5	assistance	0	0	0	0
6	development	0	0	0	0

No	Term	D1	D2	D3	D4
7	eastern	0	0	0	0
8	for	0	0	0	0
9	forest	0	0	0	0
10	have	0	0	0	0
11	impaired	0	0	0	0
12	in	0	0	0	0
13	irrigation	0	0	0	0
14	is	0	0	0	0
15	its	0	0	0	0
16	logging	0	0	0	0
17	maximum	0	0	0	0
18	not	0	0	0	0
19	occurred	0	0	0	0
20	of	0	0	0	0
21	our	0	0	0	0
22	potential	0	0	0	0
23	programme	0	0	0	0
24	reached	0	0	0	0
25	region	0	0	0	0
26	resulting	0	0	0	0
27	suboptimal	0	0	0	0
28	the	0	0	0	0
29	village	0	0	0	0
30	yields	0	0	0	0
SUM (Q * D) :		0.158356	0	0	0

The result of determining TFIDF is used to measure the similarity of two vectors using Eq. (19):

$$(TF/IDF(Q,D))^2 : \tag{19}$$

Given the following example: 0.158356 = *0.39794 * 0.39794.* Then, calculate vector values with keywords and documents according to the vector TFIDF, which is normalized, as shown in Table 22.

Table 22
The result of TFIDF normalize

No	Term	VQ	V_{D1}	V_{D2}	V_{D3}	V_{D4}
1	UKM	0	0	0	0	0.488559
2	Unlawful	0	0	0	0.488559	0
3	agricultural	0.158356	0.158356	0	0	0
4	assets	0	0	0.488559	0	0
5	assistance	0	0	0	0	0.488559
6	development	0	0	0	0	0.488559
7	eastern	0	0	0	0.488559	0
8	for	0	0	0	0	0.488559
9	forest	0	0	0	0.488559	0
10	have	0	0	0.158356	0.158356	0
11	impaired	0	0.488559	0	0	0
12	in	0	0.196867	0	0.049217	0.049217
13	irrigation	0	0.488559	0	0	0
14	is	0	0.488559	0	0	0

No	Term	VQ	V_{D1}	V_{D2}	V_{D3}	V_{D4}
15	its	0	0	0.488559	0	0
16	logging	0	0	0	0.488559	0
17	maximum	0	0	0.488559	0	0
18	not	0	0	0.488559	0	0
19	occurred	0	0	0	0.488559	0
20	of	0	0	0	0	0.488559
21	our	0	0.158356	0	0	0.158356
22	potential	0	0	0.488559	0	0
23	programme	0	0	0	0	0.488559
24	reached	0	0	0.488559	0	0
25	region	0	0	0	0.488559	0
26	resulting	0	0.488559	0	0	0
27	suboptimal	0	0.488559	0	0	0
28	the	0	0	0	0.158356	0.158356
29	village	0	0.049217	0.049217	0	0.196867
30	yields	0	0.488559	0	0	0
$\sum VQD_{ij}$		0.15835625	3.49415121	3.13892748	3.29728374	3.49415121
$\sqrt{\sum VQD_{ij}}$		0.39794001	1.86926488	1.77170186	1.81584243	1.86926488

TFIDF vector normalized values are generated using Eq. (20):

$$VQDij = (WTFIDFij)^2 \tag{20}$$

The final stage calculates the similarity value between the keyword and the document using Eq. (8), as shown in Table 23.

Table 23
Result of cosine similarity

ID	Comments	Cosine	Village
D1	irrigation in our village is impaired resulting in suboptimal agricultural yields	0.21288	Babana
D2	village assets have not reached its maximum potential	0	Kariango
D3	Unlawful logging has occurred in the eastern forest region	0	Tuppu
D4	village assistance for the development of UKM programme in our village	0	Rubae

Cosine similarity can be applied to sets of vectors, which can represent various types of data, such as text vectors for documents or feature vectors for numerical data. The calculation results indicate that the highest cosine similarity value is observed in the document "D1", with a value of 0.21288. The results of testing on villages in the Less Development group are shown in Table 24.

Table 24
Recommendation Expert for each village

No	Key	Village	Demographi	Cosine
1	Economic	A	Plains	0.1992
2	Economic	D	Mountains	0.1777
3	Economic	F	Plains	0.1180
4	Social	I	Coast	0.1651
5	Social	M	Mountains	0.1165
6	Social	K	Mountains	0.1347
7	Agriculture	L	Mountains	0.0653
8	Agriculture	J	Mountains	0.1970

No	Key	Village	Demographi	Cosine
9	Agriculture	N	Mountains	0.0824
10	Marine	P	Coast	0.1604
11	Marine	S	Coast	0.1329
12	Forestry	AC	Mountains	0.1126
13	Forestry	T	Plains	0.2347
14	Forestry	U	Mountains	0.1282
15	Regional Planning	V	Plains	0.0947
16	Regional Planning	W	Plains	0.0601
17	Entrepreneurship	X	Mountains	0.1614
18	Entrepreneurship	Y	Plains	0.1367
19	Entrepreneurship	Z	Coast	0.1785
20	Computer	AA	Coast	0.2244
21	Marine	AB	Coast	0.0688
22	Computer	AD	Coast	0.1949

The calculation result in Table 24 is used to measure the accuracy of village assistance placement recommendations based on feedback from respondents with actual data, as shown in Table 25.

Table 25
Recommendation expert for each village

Village	Field of Expertise								
	Demography	Economic	Agriculture	Forestry	Social	Entrepreneur	Regional Planning	Computer	Marine
Sidodadi (U)	Mountains	√	X	X	√	X	X	X	X
Tuppu (D)	Mountains	√	√	√	X	X	X	X	X
Sudiang (AC)	Mountains	√	√	X	√	√	X	X	X
Rungkeke (X)	Mountains	√	X	X	√	X	X	X	X
Riattang (M)	Mountains	X	√	√	X	√	X	X	X
Buttu (K)	Mountains	X	√	X	√	X	X	X	X
Kariango (J)	Mountains	X	√	√	X	X	X	X	X
Buttu (L)	Mountains	X	X	X	X	√	X	√	X
Sidrap (N)	Mountains	X	X	X	√	√	X	X	X
Babana (A)	Plains	√	√	√	X	X	X	X	X
Ampoe (F)	Plains	√	X	X	√	√	√	√	X
Pajalele (W)	Plains	√	√	X	√	√	X	X	X
Kombiling (T)	Plains	√	√	X	X	X	X	X	X
Tabbae (Y)	Plains	X	X	X	X	√	X	X	X
Polman (V)	Plains	X	√	X	√	X	√	√	X
Letta (I)	Coast	X	X	X	X	√	X	X	X
Rubae (P)	Coast	X	√	X	√	X	X	√	√
Kire (S)	Coast	X	X	X	X	X	X	X	√
Bone (Z)	Coast	X	X	X	√	X	X	X	√
Jeneponto (AA)	Coast	X	X	X	√	X	X	X	X
Masale (AB)	Coast	√	√	X	√	X	X	X	X
Abdesir (AD)	Coast	√	X	X	X	X	X	X	√
Number		10	11	4	11	8	2	4	4

The data presented in Table 25 is the outcome of observations and interactions with respondents to understand the current scientific needs of experts in villages designated as priority development areas. Based on these findings, it is evident that most villages classified as Less Development are situated in mountainous regions. Apart from that, the greatest need for experts is in agriculture and social services, as shown in Table 25, with 11 experts for each. This suggests that individuals with lower levels of welfare are more likely to reside in less developed areas.
Based on the hybrid technique that has been built, testing of expert needs was carried out through respondents spread across several villages in the South Sulawesi region of Indonesia. It is known that a confusion matrix with multifaceted views serves a fundamental role in evaluating classification performance [64]. Testing of results is evaluated against 1214 data records using the confusion matrix testing method to determine the accuracy, recall, precision, and F1 Score value. Table 26 shows testing data from respondents.

Table 26
Record of testing data

Field Of Expertise	AGR	SOC	ECO	FOR	MAR	ENT	COM	REG
Agriculture (Arg)	151	1	1	1	1	0	1	1
Social (Soc)	1	150	0	0	1	1	0	1
Economic (Eco)	2	2	141	2	2	1	1	0
Forestry (For)	1	0	1	134	1	1	1	0
Marine (Maret)	2	0	2	1	142	1	1	2
Entrepreneurship (Ent)	0	0	2	2	1	144	1	1
Computer (Com)	1	1	2	1	2	0	147	0
Regional Planning (Reg)	1	2	1	1	1	0	1	150

Based on the data in Table 26, a validation of the recommendation results with actual data was carried out using the confusion matrix method as follows [65]:

Precision = TP/(TP+FP) (21)
Recall = TP/(TP+FN) (22)
F1 Score = 2(Precision x Recall) / (Precision + Recall) (23)
Accuracy = (TP) / (TP + FP + FN) (24)

Table 27
Result of Confusion Matrix

	AGR	SOC	ECO	FOR	MAR	ENT	COM	REG
TP	151	150	141	134	142	144	147	150
FP	6	4	10	5	9	7	7	7
FN	8	6	9	8	9	4	6	5
Precision Eq. (21)	10.785	15	7.421	10.307	7.888	13.090	11.307	12.5
Recall Eq. (22)	0.949	0.961	0.94	0.943	0.940	0.972	0.960	0.967
F1 Score Eq. (23)	1.745	1.807	1.668	1.729	1.680	1.811	1.771	1.796
Accuracy Eq. (24)	0.95							

This research results demonstrate commendable performance in data classification, as evidenced by the accuracy value 0.95 presented in Table 27. Based on the rural classification outcomes as a foundation for guiding expert placement, a comparative analysis is undertaken between the hybrid

Tsukamoto+SM+CS approach and C.50 [18], ID3, C4.5, and CART [20] algorithms to assess their respective performances. A comprehensive comparison is provided in Table 28.

Table 28
Result of the comparison accuracy of classification algorithms

Methodology	Accuracy
ID3	83.81%
C4.5	80.62%
CART	87.40%
C.50	82.35%
Tsukamoto + SM + CS	95 %

The comparison results indicate the superiority of the methodology proposed in this research. This research has substantial implications for advancing decision-making methodologies by integrating hybrid indexing, clustering, and classification techniques. Implementing these proposed techniques in case studies focused on mapping rural experts can enhance responsiveness to challenges in village development. A comprehensive approach utilizing hybrid methods offers extensive application potential, not only confined to rural settings but also presenting opportunities for adaptation across diverse domains. Further assessment of practical implementation will be necessary to gauge the effectiveness and applicability of these methods in real-world problem-solving scenarios.

4. Conclusions

This study concludes that the hybrid indexing, clustering, and classification methodology can be a model to support decision-making regarding the placement of experts in a village. Applying Fuzzy Delphi and Reciprocal Rank theory involving experts and communities can effectively and objectively produce CSLI ranking decisions for each village. Hybrid SOM, FCM, and XB methodologies can group villages into Good, Average, and Bad CSLI clusters. Applying the Fuzzy Tsukamoto and the Smallest of Maximum methods in the Classification process can identify villages included in the Less Development level. The role of the community as a respondent in the village assessment process facilitates the search for information about expert needs in a village, which is done by processing community comments using the cosine similarity algorithm to produce a recommendation for the placement of experts who have skills relevant to problems in a village. The results of the study of 30 villages showed that the recommended fields were agricultural science in 11 villages, social in 11 villages, economics in 10 villages, entrepreneurship in 8 villages, marine in 3 villages, forestry in 5 villages, computer in 4 villages, and regional planning in 2 villages. The results of testing the application of the methodology using 1214 comment data from the public showed an accuracy of 0.95, so it can be concluded that the government can apply the proposed method as a decision-making tool. The limitation of this study is that the measurements taken during this study can change over time. Therefore, the results only describe the conditions under which the study was conducted. The implications of this research can significantly contribute to the decision-making process, thus opening up opportunities for developing hybrid methodologies in the context of expert mapping of rural problems. Future work in this research includes conducting more comprehensive development and modelling, testing systems in different contexts to assess their effectiveness, and optimizing the variables and elements involved to produce better performance.

Author Contributions

Conceptualization, M.F., and T.K.A.R.; methodology, M.F., and T.K.A.R; software, I.M., K.A., and L.R.; validation, M.T and T.K.A.R.; formal analysis, M.F.,T.K.A.R., and I.M; investigation, M.T and K.A.; resources, M.F., and T.K.A.R.; data curation, I.M., K.A., and M.T.; writing—original draft preparation, M.F.,T.K.A.R., and K.A.; writing—review and editing, M.F.,T.K.A.R., and M.T.; visualization, M.F. and K.A.; supervision,M.F., T.K.A.R. and I.M.; project administration, I.M., K.A., and L.R and M.T.; funding acquisition, M.F. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Data Availability Statement

The data produced and examined in this study are not accessible to the public because they contain sensitive information. However, interested parties can obtain them by contacting the corresponding author.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

We would like to convey our heartfelt gratitude to our research team and the supervisor for their invaluable contributions to this study.

Appendix A

Table A.1
List of criteria

Code	Criteria	Descriptions
C1	Welfare Status	Referring to social welfare conditions of people in rural areas
C2	Residential	Referring to something to do with where to live
C3	Land Of Residence	Referring to the country or region where a person lives or resides
C4	Widest Flooring	Referring to flooring materials that come in larger widths
C5	Type Of Walls	Referring to various types of walls used in construction
C6	Roofs	Referring to The type of roof used in construction
C7	Wall Quality	Referring to The quality of a wall in construction
C8	The Widest-Quality Roof	Referring to different aspects of a roof
C9	Drinking Water	Referring to water that is safe and suitable for human consumption
C10	Get Drinking Water	Referring to obtaining drinking water by family
C11	Defecation Facilities	Referring to the means of sewage disposal in the village
C12	Toilets	Referring to sanitation fixtures used for the disposal of human waste
C13	Fecal Landfill	Referring to a site or facility where human fecal waste
C14	Sources Of Illumination	Referring to the various illuminated spaces
C15	Electrical Power	Fefers to the electrical energy being transferred or converted
C16	Fuel For Cooking	Referring to the energy sources used to generate heat for cooking
C17	Gas Of 5.5kg	Referring to a gas cylinder with a weight of 5.5 kilograms
C18	Telephone	Referring to a communication device for voice communication
C19	Computer	Computer ownership in a family
C20	Bicycles	Bicycle ownership in a family
C21	Motorcycle	Motorcycle ownership in a family
C22	Car	Car ownership in a family

Code	Criteria	Descriptions
C23	Boat	Boat ownership in a family
C24	Outboard Motor	Outboard Motor ownership in a family
C25	Tractor	Tractor ownership in a family
C26	Ship	Ship ownership in a family
C27	Refrigerator	Refrigerator ownership in a family
C28	Air Conditioning	Air Conditioning ownership in a family
C29	Water Heater	Water Heater ownership in a family
C30	Television	Television ownership in a family
C31	Gold Worth 10 Grams	Gold Worth 10 Grams ownership in a family
C32	Land	Land ownership in a family
C33	Home Locations	Home Location ownership in a family
C34	Household Members Joint	Household Members Joint business ownership in a family
C35	Dormitory	Dormitory Location ownership in a family
C36	PKH Program Participants	Status as a PKH aid recipient
C37	RASKIN Program Participants	Status as a RASKIN aid recipient
C38	KUR Program Participants	Status as a KUR aid recipient

Table A.2
Format of community feedback

No	Criteria	Weight(WC)	Subcriteria	Weight(WS)	WC * WS
1	Welfare	0.23652	Welfare in 11%-20%	0.27272	0.06450
2	Residential	0.11826	Privately Owned	0.43795	0.05179
3	Land of residence	0.07884	Privately Owned	0.48000	0.03784
4	Get drinking water	0.05913	Not buying	0.18181	0.01075
5	Widest flooring	0.04730	Tiles	0.08535	0.00403
6	Electrical power	0.03942	1300 watt	0.13605	0.00536
7	Widest-quality roof	0.03378	Height quality	0.6666	0.02252
8	Type of walls	0.02956	Wall	0.38565	0.01140
9	Roofs	0.02628	Seng	0.05693	0.00149
10	Wall quality	0.02365	Height quality	0.66667	0.01576
11	Fuel for cooking	0.02150	Gas>3kg	0.17673	0.00380
12	Gold worth 10 grams	0.01971	Yes	0.66666	0.01314
13	Ship	0.01819	No	0.33333	0.00606
14	Gas of 5.5kg	0.01689	Yes	0.66666	0.01126
15	KUR participants	0.01570	No	0.33333	0.00525
16	Sources of illumination	0.01473	PLN Electricity	0.54545	0.00806
17	Computer	0.01393	No	0.33333	0.00463
18	Tractor	0.01310	No	0.3333	0.00438
19	Defecation facilities	0.01249	Privately Owned	0.48000	0.00597
20	Motorcycle	0.01186	Yes	0.66667	0.00788
21	Air conditioning	0.01123	No	0.3333	0.00375
22	Drinking water	0.01071	Plumbing meter	0.10716	0.00115
23	Fecal landfill	0.01024	Earthen pit	0.13654	0.00139
24	Raskin participants	0.00985	No	0.33333	0.00328
25	Toilets	0.00941	Cemplung	0.16000	0.00151
26	Motor Ownership	0.00909	No	0.3333	0.00303
27	Boat	0.00876	No	0.33333	0.00292
28	Refrigerator	0.00840	Yes	0.66667	0.00563
29	Ownership of Land	0.00816	Yes	0.66667	0.00543
30	Home Locations	0.00784	No	0.33333	0.00262
31	Car	0.00760	Yes	0.66667	0.00508
32	Dormitori	0.00731	No	0.33333	0.00246
33	Water heater	0.00717	No	0.33333	0.00238
34	Television	0.00697	Yes	0.66667	0.00463

No	Criteria	Weight(WC)	Subcriteria	Weight(WS)	WC * WS
35	Bicycles	0.00680	Yes	0.66667	0.00450
36	Telephone	0.00670	No	0.33333	0.00219
37	Members businesses	0.00693	No	0.33333	0.00213
38	PKH program	0.00624	No	0.33333	0.00207
Comments :					
Please help for the repair of Hajj Mukti road, bonto cinde village no drainage, thank you.					
Expert Advice in the Field of Social Sciences; Economics; Computer; Entrepreneurship; Tourism					
SCORE					0.35218

Table A.3
Input data of village criteria

Village	CSLI	Head of Family	Number of Residence
A	0.52550	599	2434
B	0.47335	1398	4198
C	0.45890	698	2858
D	0.45794	429	1675
E	0.39561	320	1340
F	0.38065	293	500
G	0.37964	1100	4089
H	0.36849	319	1394
I	0.36471	350	1000
J	0.35804	162	709
K	0.34530	561	1743
L	0.33649	501	2285
M	0.33464	912	1945
N	0.33390	1172	3518
O	0.33390	845	2972
P	0.33390	1014	3395
Q	0.33073	755	2110
R	0.32924	1044	3351
S	0.32778	523	1518
T	0.32100	836	3172
U	0.31983	347	1200
V	0.31860	976	3250
W	0.31336	519	1808
X	0.31064	212	920
Y	0.31047	335	720
Z	0.30667	1009	3418
AA	0.29003	373	1291
AB	0.26416	200	739
AC	0.21900	1065	3527
AD	0.20524	582	2543

Table A.4
Normalization of criteria weight

Village	CSLI Score			Head of Family			Number of Residents		
	VS	Max.VS	NVS	VP	Max.VP	NVP	VR	Max.VR	NVR
A	0.52550		1	599		0.42846	2434		0.57979
B	0.47335		0.90104	1398		1	4198		1
C	0.45890		0.85803	698		0.49928	2858		0.68080
D	0.45794		0.87136	429		0.30686	1675		0.39899
E	0.39561		0.75280	320		0.22889	1340		0.31919
F	0.38065		0.72426	293		0.20958	500		0.11910
G	0.37964		0.72235	1100		0.78683	4089		0.97403
H	0.36849		0.70104	319		0.22818	1394		0.33206
I	0.36471		0.69400	350		0.25035	1000		0.23820
J	0.35804		0.68125	162		0.11587	709		0.16888
K	0.34530		0.65708	561		0.40128	1743		0.41519
L	0.33649		0.64015	501		0.35836	2285		0.54430
M	0.33464		0.63672	912		0.65236	1945		0.46331
N	0.33390		0.63539	1172		0.83834	3518		0.83801
O	0.33390		0.63539	845		0.60443	2972		0.70795
P	0.33390	*0.*5255	0.63539	1014	1398	0.72532	3395	4198	0.80871
Q	0.33073		0.62930	755		0.54005	2110		0.50262
R	0.32924		0.62645	1044		0.74678	3351		0.79823
S	0.32778		0.62359	523		0.37410	1518		0.36160
T	0.32100		0.61084	836		0.59799	3172		0.75559
U	0.31983		0.60856	347		0.24821	1200		0.28585
V	0.31860		0.60627	976		0.69814	3250		0.77417
W	0.31336		0.59619	519		0.37124	1808		0.43068
X	0.31064		0.59105	212		0.15164	920		0.21915
Y	0.31047		0.59067	335		0.23962	720		0.17151
Z	0.30667		0.58344	1009		0.72174	3418		0.81419
AA	0.29003		0.55185	373		0.26680	1291		0.30752
AB	0.26416		0.50256	200		0.14306	739		0.17603
AC	0.21900		0.41674	1065		0.76180	3527		0.84016
AD	0.20524		0.39048	582		0.41630	2543		0.60576

VS = Village Score
VP = Total of Community
VR = Total of Patriarch
NVS = Normalization value of Variable VS
NVP = Normalization value of for Variable VP
NVR = Normalization value of for Variable VR

Table A.5
Input clustering data

Village	X1	X2	X3
A	0.25856	0.02031	0.25688
B	0.25269	0.61969	1.13890
C	0.11612	0.05525	0.27616
D	0.34408	0.02029	0.07644
E	0.42092	0.08669	0.01831
F	0.64493	0.20231	0.04099
G	0.07469	0.43116	0.76306
H	0.39708	0.10328	0.00859
I	0.46395	0.13152	0.00756

Village	X1	X2	X3
J	0.67274	0.23839	0.03686
K	0.19862	0.07146	0.03875
L	0.15211	0.08887	0.07640
M	0.08727	0.13132	0.18921
N	0.03230	0.39223	0.64303
O	0.01000	0.17288	0.29850
P	0.00617	0.29046	0.49242
Q	0.08515	0.09876	0.12748
R	0.00675	0.30176	0.50219
S	0.25102	0.10002	0.01968
T	0.00762	0.21080	0.33625
U	0.40956	0.15891	0.00026
V	0.00068	0.26952	0.43297
W	0.20259	0.10811	0.03456
X	0.56990	0.24159	0.01640
Y	0.53720	0.22687	0.01504
Z	0.00517	0.32423	0.49623
AA	0.37571	0.18392	0.00561
AB	0.63590	0.33504	0.03963
AC	0.04575	0.50833	0.60437
AD	0.68306	0.49335	0.16395

Table A.6
Cluster centroid

C1	C2	C3
0.534940	0.253985	0.115000
0.239995	0.401444	0.433285
0.071016	0.509299	0.738098

Table A.7
Objective value each iteration using 3 clusters

Iteration	Objective
1	1.283007700
2	0.231746383
3	0.157153983
28	7.396630273826E-7

Table A.8
Convergent cluster

C1	C2	C3
6.28169	4.78471	2.36558
0.73302	2.2521	4.66532
3.97416	5.1757	2.80663
11.80259	2.9346	1.51324
25.45775	2.71733	1.36647
69.35099	2.42833	1.21931
1.37495	10.2644	443.96007
24.04235	2.7771	1.38763
48.00088	2.71233	1.33956
49.60822	2.35216	1.18935
7.01477	2.99966	1.5949
5.90226	3.38429	1.79262
4.42109	4.92083	2.54304

C1	C2	C3
1.65831	14.90027	56.9387
3.01372	6.40180	3.67536
2.17521	13.51089	10.79905
4.43564	3.76069	2.05908
2.13351	14.06226	11.81061
9.72394	3.02488	1.54873
2.86313	7.87485	4.49414
36.62835	2.92272	1.41552
2.39626	11.44253	7.51929
7.74118	3.18407	1.64372
237.66576	2.71514	1.30799
302.05784	2.81707	1.34468
2.13420	14.69692	12.1277
30.50965	3.16619	1.49994
49.47663	2.69645	1.30065
1.67848	15.66789	35.3225
10.55239	3.20699	1.52451

Table A.9
The result of clustering for each village

Village	Cluster-1	Cluster-2	Cluster-3	Decision
A	0.130	0.794	0.076	C2
B	0.173	0.197	0.630	C3
C	0.075	0.831	0.094	C2
D	0.315	0.634	0.052	C2
E	0.675	0.286	0.039	C1
F	0.940	0.044	0.015	C1
G	0.037	0.058	0.905	C3
H	0.632	0.327	0.041	C1
I	0.857	0.121	0.022	C1
J	0.918	0.059	0.023	C1
K	0.076	0.900	0.025	C2
L	0.026	0.962	0.012	C2
M	0.042	0.915	0.043	C2
N	0.004	0.007	0.988	C3
O	0.101	0.625	0.274	C2
P	0.034	0.089	0.878	C3
Q	0.032	0.944	0.025	C2
R	0.027	0.069	0.904	C3
S	0.167	0.797	0.036	C2
T	0.103	0.492	0.404	C2
U	0.757	0.211	0.031	C1
V	0.064	0.200	0.736	C3
W	0.085	0.889	0.026	C2
X	0.988	0.009	0.003	C1
Y	0.996	0.003	0.001	C1
Z	0.025	0.061	0.914	C3
AA	0.667	0.293	0.039	C1
AB	0.899	0.071	0.030	C1
AC	0.026	0.042	0.933	C3
AD	0.691	0.184	0.125	C1

Table A.10
Defuzzification weighted average

Rule	α-Predicate	Z	μ	∑(z / μ)
R1	0	75	1	0
R27	0.2	49.2;79	2	12.82
R35	0.8	49.2	1	39.36
R60	0	50	1	0
Defuzzification value				49.2

References

[1] Harbiankova, A., & Gertsberg, L. (2022). Information Model for Sustainable Rural Development. Energies, 15(11), 4009. https://doi.org/10.3390/en15114009

[2] Bennich, T., Persson, Å., Beaussart, R., Allen, C., & Malekpour, S. (2023). Recurring patterns of SDG interlinkages and how they can advance the 2030 Agenda. One Earth. https://doi.org/10.1016/j.oneear.2023.10.008

[3] Purba, J. T., Budiono, S., Hariandja, E. S., & Pramono, R. (2023). Sustainability Strategy to Alleviate Poverty Through Education, Energy, GRDP, and Special Funds: Evidence from Indonesia. International Journal of Sustainable Development and Planning, 18(5), 1397–1406. https://doi.org/10.18280/ijsdp.180510

[4] Yang, X., Xu, Z., & Xu, J. (2023). Large-scale group Delphi method with heterogeneous decision information and dynamic weights. Expert Systems with Applications, 213. https://doi.org/10.1016/j.eswa.2022.118782

[5] Chatterjee, S., & Chakraborty, S. (2023). Application of the R method in solving material handling equipment selection problems. Decision Making: Applications in Management and Engineering, 6(2), 74–94. https://doi.org/10.31181/dmame622023391

[6] Xu, N., Zhu, W., Wang, R., Li, Q., Wang, Z., & Finkelman, R. B. (2023). Application of self-organizing maps to coal elemental data. International Journal of Coal Geology, 277, 104358. https://doi.org/10.1016/j.coal.2023.104358

[7] Sahu, P., Kumar Sahoo, B., Kumar Mohapatra, S., & Kumar Sarangi, P. (2023). Segmentation of encephalon tumor by applying soft computing methodologies from magnetic resonance images. Materials Today: Proceedings, 80, 3371–3375. https://doi.org/10.1016/j.matpr.2021.07.255

[8] Mota, V. C., Damasceno, F. A., & Leite, D. F. (2018). Fuzzy clustering and fuzzy validity measures for knowledge discovery and decision making in agricultural engineering. Computers and Electronics in Agriculture, 150, 118–124. https://doi.org/10.1016/j.compag.2018.04.011

[9] Uma G, & S, N. (2023). An Investigative Study on Quick Switching System using Fuzzy and Neutrosophic Poisson Distribution. Neutrosophic Systems with Applications, 7, 61–70. https://doi.org/10.61356/j.nswa.2023.28

[10] Abdelhafeez, A., Mahmoud, H., & Aziz, A. S. (2023). Identify the most Productive Crop to Encourage Sustainable Farming Methods in Smart Farming using Neutrosophic Environment. Neutrosophic Systems with Applications, 6, 17–24. https://doi.org/10.61356/j.nswa.2023.34

[11] Dore, P., Chakkor, S., El Oualkadi, A., & Baghouri, M. (2023). Real-time intelligent system for wind turbine monitoring using fuzzy system. E-Prime - Advances in Electrical Engineering, Electronics and Energy, 3, 100096. https://doi.org/10.1016/j.prime.2022.100096

[12] Kurniawan, S., Ranggadara, I., Suhendra, & Ratnasari, A. (2020). Smallest of Maximum to find α-predicate for DeterminingCattle Health Conditions. International Journal of Advanced Trends in Computer Science and Engineering, 9(5), 8245–8251. https://doi.org/10.30534/ijatcse/2020/192952020

[13] Olajide, A. O., Busayo, A. T., & Olawale, E. T. (2017). Analyzing the Effects of the Different Defuzzification Methods in the Evaluation of Javacomponents' Customizability for Reusability. International Journal of Advanced Engineering, Management and Science, 3(9), 962–971. https://doi.org/10.24001/ijaems.3.9.10

[14] Morales, R., Marín, L. G., Roje, T., Caquilpan, V., Sáez, D., & Nuñez, A. (2024). Microgrid planning based on computational intelligence methods for rural communities: A case study in the José Painecura Mapuche community, Chile. Expert Systems with Applications, 235, 121179. https://doi.org/10.1016/j.eswa.2023.121179

[15] Mupedziswa, R., Malinga, T., & Nuggert Ntshwarang, P. (2021). Standard of Living, Well-Being and Community Development: The Case of Botswana. In Improving Quality of Life - Exploring Standard of Living, Wellbeing, and Community Development. IntechOpen. https://doi.org/10.5772/intechopen.97680

[16] Lathief, M. F., Soesanti, I., & Permanasari, A. E. (2019). Combination of Fuzzy C-Means, Xie-Beni Index, and Backpropagation Neural Network for Better Forecasting Result. Proceedings of the International Conference on Creative Economics, Tourism and Information Management, Iccetim 2019, 72–77. https://doi.org/10.5220/0009858200720077

[17] Nugraha, L. F., Sulistyowati, L., Setiawan, I., & Noor, T. I. (2022). Alternative Community-Based Village Development Strategies in Indonesia: Using Multicriteria Decision Analysis. Agriculture, 12(11), 1903. https://doi.org/10.3390/agriculture12111903
[18] Nur Fitria, M. C., Debataraja, N. N., & Rizki, S. W. (2022). Classification of Village Status in Landak Regency Using C5.0 Algorithm. Tensor: Pure and Applied Mathematics Journal, 3(1), 33–42. https://doi.org/10.30598/tensorvol3iss1pp33-42
[19] Pattanayek, S. K., Mookherjee, S., & Mondal, D. (2019). An Analysis of Standard of Living Index in West Bengal's Paschim Medinipur District: An Application of Average Correlation Method. Indian Journal of Human Development, 13(2), 159–182. https://doi.org/10.1177/0973703019851629
[20] Salima, N., & Ilham, A. (2022). Village Status Classification Based Tree Algorithm. Proceedings of the 3rd International Conference of Science Education in Industrial Revolution 4.0, ICONSEIR 2021, December 21st, 2021, Medan, North Sumatra, Indonesia, 2–8. https://doi.org/10.4108/eai.21-12-2021.2317488
[21] Einola, K., & Alvesson, M. (2021). Behind the Numbers: Questioning Questionnaires. Journal of Management Inquiry, 30(1), 102–114. https://doi.org/10.1177/1056492620938139
[22] Friston, K. J., Parr, T., Yufik, Y., Sajid, N., Price, C. J., & Holmes, E. (2020). Generative models, linguistic communication and active inference. Neuroscience & Biobehavioral Reviews, 118(June), 42–64. https://doi.org/10.1016/j.neubiorev.2020.07.005
[23] Taherdoost, H. (2018). Sampling Methods in Research Methodology; How to Choose a Sampling Technique for Research. SSRN Electronic Journal, 5(2), 18–27. https://doi.org/10.2139/ssrn.3205035
[24] Watson, L. (2018). Educating for Good Questioning: a Tool for Intellectual Virtues Education. Acta Analytica, 33(3), 353–370. https://doi.org/10.1007/s12136-018-0350-y
[25] Danacı, M., & Yıldırım, U. (2023). Comprehensive analysis of lifeboat accidents using the Fuzzy Delphi method. Ocean Engineering, 278, 114371. https://doi.org/10.1016/j.oceaneng.2023.114371
[26] Mohammed, R. T., Yaakob, R., Sharef, N. M., & Abdullah, R. (2021). Unifying The Evaluation Criteria Of Many Objectives Optimization Using Fuzzy Delphi Method. Baghdad Science Journal, 18(4), 1423. https://doi.org/10.21123/bsj.2021.18.4(Suppl.).1423
[27] Mohammed, R. T., Zaidan, A. A., Yaakob, R., Sharef, N. M., Abdullah, R. H., Zaidan, B. B., Albahri, O. S., & Abdulkareem, K. H. (2022). Determining Importance of Many-Objective Optimisation Competitive Algorithms Evaluation Criteria Based on a Novel Fuzzy-Weighted Zero-Inconsistency Method. International Journal of Information Technology & Decision Making, 21(01), 195–241. https://doi.org/10.1142/S0219622021500140
[28] Faisal, M., & Rahman, T. K. A. (2023). Optimally Enhancement Rural Development Support Using Hybrid Multy Object Optimization (MOO) and Clustering Methodologies: A Case South Sulawesi - Indonesia. International Journal of Sustainable Development and Planning, 18(6), 1659–1669. https://doi.org/10.18280/ijsdp.180602
[29] Ezugwu, A. E.-S., Agbaje, M. B., Aljojo, N., Els, R., Chiroma, H., & Elaziz, M. A. (2020). A Comparative Performance Study of Hybrid Firefly Algorithms for Automatic Data Clustering. IEEE Access, 8, 121089–121118. https://doi.org/10.1109/ACCESS.2020.3006173
[30] Arunachalam, D., & Kumar, N. (2018). Benefit-based consumer segmentation and performance evaluation of clustering approaches: An evidence of data-driven decision-making. Expert Systems with Applications, 111, 11–34. https://doi.org/10.1016/j.eswa.2018.03.007
[31] Lee, K. J., Yun, S. T., Yu, S., Kim, K. H., Lee, J. H., & Lee, S. H. (2019). The combined use of self-organizing map technique and fuzzy c-means clustering to evaluate urban groundwater quality in Seoul metropolitan city, South Korea. Journal of Hydrology, 569, 685–697. https://doi.org/10.1016/j.jhydrol.2018.12.031
[32] Afzal, A., Ansari, Z., Alshahrani, S., Raj, A. K., Kuruniyan, M. S., Saleel, C. A., & Nisar, K. S. (2021). Clustering of COVID-19 data for knowledge discovery using c-means and fuzzy c-means. Results in Physics, 29, 104639. https://doi.org/10.1016/j.rinp.2021.104639
[33] Sitepu, K. A. B., Sitompul, O. S., & Situmorang, Z. (2019). Analysis of Fuzzy C-Means and Analytical Hierarchy Process (AHP) Models Using Xie-Beni Index. 2019 International Conference of Computer Science and Information Technology (ICoSNIKOM), 1–6. https://doi.org/10.1109/ICoSNIKOM48755.2019.9111538
[34] Antczak, T. (2023). On optimality for fuzzy optimization problems with granular differentiable fuzzy objective functions. Expert Systems with Applications, 121891. https://doi.org/10.1016/j.eswa.2023.121891
[35] Sukheja, D., Shah, J. A., Madhu, G., Kautish, K. S., Alghamdi, F. A., Yahia, I. S., M. El-Kenawy, E. S., & Mohamed, A. W. (2022). New Decision-Making Technique Based on Hurwicz Criteria for Fuzzy Ranking. Computers, Materials & Continua, 73(3), 4595–4609. https://doi.org/10.32604/cmc.2022.029122
[36] Lan, L. T. H., Tuan, T. M., Ngan, T. T., Giang, N. L., Ngoc, V. T. N., & Van Hai, P. (2020). A New Complex Fuzzy Inference System With Fuzzy Knowledge Graph and Extensions in Decision Making. IEEE Access, 8, 164899–164921. https://doi.org/10.1109/ACCESS.2020.3021097

[37] Xie, X., Fu, Y., Jin, H., Zhao, Y., & Cao, W. (2020). A novel text mining approach for scholar information extraction from web content in Chinese. Future Generation Computer Systems, 111, 859–872. https://doi.org/10.1016/j.future.2019.08.033
[38] Sebestyén, V., Domokos, E., & Abonyi, J. (2020). Focal points for sustainable development strategies—Text mining-based comparative analysis of voluntary national reviews. Journal of Environmental Management, 263, 110414. https://doi.org/10.1016/j.jenvman.2020.110414
[39] Song, H., Bei, J., Zhang, H., Wang, J., & Zhang, P. (2024). Hybrid algorithm of differential evolution and flower pollination for global optimization problems. Expert Systems with Applications, 237, 121402. https://doi.org/10.1016/j.eswa.2023.121402
[40] Tan, W. H., & Mohamad-Saleh, J. (2023). A hybrid whale optimization algorithm based on equilibrium concept. Alexandria Engineering Journal, 68, 763–786. https://doi.org/10.1016/j.aej.2022.12.019
[41] Ishtaiwi, A., Alshahwan, F., Jamal, N., Hadi, W., & Abuarqoub, M. (2021). A dynamic clause specific initial weight assignment for solving satisfiability problems using local search. Algorithms, 14(1). https://doi.org/10.3390/a14010012
[42] Yin, C., Zhang, B., Liu, W., Du, M., Luo, N., Zhai, X., & Ba, T. (2022). Geographic Knowledge Graph Attribute Normalization: Improving the Accuracy by Fusing Optimal Granularity Clustering and Co-Occurrence Analysis. ISPRS International Journal of Geo-Information, 11(7), 360. https://doi.org/10.3390/ijgi11070360
[43] Mellor, S. (2022). A Construct Validity Study for the Union Intolerance Scale: Convergent-Discriminant Validity and Concurrent Criterion-Related Validity. Merits, 2(3), 210–223. https://doi.org/10.3390/merits2030015
[44] Wu, Z., Zhao, Y., Wang, W., & Li, C. (2023). Adaptive weighted fuzzy clustering based on intra-cluster data divergence. Neurocomputing, 552, 126550. https://doi.org/10.1016/j.neucom.2023.126550
[45] Deng, X., Yang, Y., & Jiang, W. (2023). Discrete choice models with Atanassov-type intuitionistic fuzzy membership degrees. Information Sciences, 622, 46–67. https://doi.org/10.1016/j.ins.2022.11.127
[46] Saberi, H., Sharbati, R., & Farzanegan, B. (2022). A gradient ascent algorithm based on possibilistic fuzzy C-Means for clustering noisy data. Expert Systems with Applications, 191, 116153. https://doi.org/10.1016/j.eswa.2021.116153
[47] Wiroonsri, N. (2023). Clustering performance analysis using a new correlation-based cluster validity index. Pattern Recognition, 1, 1–14. https://doi.org/10.1016/J.PATCOG.2023.10991
[48] Zhou, B., Lu, B., & Saeidlou, S. (2022). A Hybrid Clustering Method Based on the Several Diverse Basic Clustering and Meta-Clustering Aggregation Technique. Cybernetics and Systems, 1–27. https://doi.org/10.1080/01969722.2022.2110682
[49] Liang, J. (2022). Value Analysis and Realization of Artistic Intervention in Rural Revitalization Based on the Fuzzy Clustering Algorithm. Scientific Programming, 2022, 1–9. https://doi.org/10.1155/2022/3107440
[50] Shen, L., Lu, J., Long, M., & Chen, T. (2019). Identification of Accident Blackspots on Rural Roads Using Grid Clustering and Principal Component Clustering. Mathematical Problems in Engineering, 2019, 1–12. https://doi.org/10.1155/2019/2151284
[51] Yuan, W., Li, J., Liu, C., & Shang, R. (2022). How to Realize the Integration of Urbanization and Rural Village Renewal Strategies in Rural Areas: The Case Study of Laizhou, China. Land, 11(12), 2161. https://doi.org/10.3390/land11122161
[52] Ilbeigipour, S., Albadvi, A., & Akhondzadeh Noughabi, E. (2022). Cluster-based analysis of COVID-19 cases using self-organizing map neural network and K-means methods to improve medical decision-making. Informatics in Medicine Unlocked, 32, 101005. https://doi.org/10.1016/j.imu.2022.101005
[53] Karthik, J., Tamizhazhagan, V., & Narayana, S. (2021). WITHDRAWN: Data leak identification using scattering search K Means in social networks. Materials Today: Proceedings. https://doi.org/10.1016/j.matpr.2021.01.200
[54] Jin, Q., Lin, N., & Zhang, Y. (2021). K-Means Clustering Algorithm Based on Chaotic Adaptive Artificial Bee Colony. Algorithms, 14(2), 53. https://doi.org/10.3390/a14020053
[55] Faisal, M., & Rahman, T. K. A. (2023). Optimally Enhancement Rural Development Support Using Hybrid Multy Object Optimization (MOO) and Clustering Methodologies: A Case South Sulawesi - Indonesia. International Journal of Sustainable Development and Planning, 18(6), 1659–1669. https://doi.org/10.18280/ijsdp.180602
[56] Padilla-Rivera, A., do Carmo, B. B. T., Arcese, G., & Merveille, N. (2021). Social circular economy indicators: Selection through fuzzy delphi method. Sustainable Production and Consumption, 26, 101–110. https://doi.org/10.1016/j.spc.2020.09.015
[57] Mendel, J. M., & Bonissone, P. P. (2021). Critical Thinking About Explainable AI (XAI) for Rule-Based Fuzzy Systems. IEEE Transactions on Fuzzy Systems, 29(12), 3579–3593. https://doi.org/10.1109/TFUZZ.2021.3079503

[58] Fakir, M., Hicham, H., Chabi, M., & Sarfraz, M. (2020). Classification of Eyes Based on Fuzzy Logic. International Journal of Cognitive Informatics and Natural Intelligence, 14(4), 101–112. https://doi.org/10.4018/IJCINI.2020100106

[59] Zhu, X., Pedrycz, W., & Li, Z. (2022). A Granular Approach to Interval Output Estimation for Rule-Based Fuzzy Models. IEEE Transactions on Cybernetics, 52(7), 7029–7038. https://doi.org/10.1109/TCYB.2020.3025668

[60] Nowak-Brzezińska, A., & Wakulicz-Deja, A. (2019). Exploration of rule-based knowledge bases: A knowledge engineer's support. Information Sciences, 485, 301–318. https://doi.org/10.1016/j.ins.2019.02.019

[61] Ranasinghe, K., Sabatini, R., Gardi, A., Bijjahalli, S., Kapoor, R., Fahey, T., & Thangavel, K. (2022). Advances in Integrated System Health Management for mission-essential and safety-critical aerospace applications. Progress in Aerospace Sciences, 128, 100758. https://doi.org/10.1016/j.paerosci.2021.100758

[62] Ibrahim, E. A., Salifu, D., Mwalili, S., Dubois, T., Collins, R., & Tonnang, H. E. Z. (2022). An expert system for insect pest population dynamics prediction. Computers and Electronics in Agriculture, 198, 107124. https://doi.org/10.1016/j.compag.2022.107124

[63] Ji, W. (2021). Fuzzy implications in lattice effect algebras. Fuzzy Sets and Systems, 405, 40–46. https://doi.org/10.1016/j.fss.2020.04.021

[64] Xu, J., Zhang, Y., & Miao, D. (2020). Three-way confusion matrix for classification: A measure driven view. Information Sciences, 507, 772–794. https://doi.org/10.1016/j.ins.2019.06.064

[65] Faisal, M., Chaudhury, S., Sankaran, K. S., Raghavendra, S., Chitra, R. J., Eswaran, M., & Boddu, R. (2022). Faster R-CNN Algorithm for Detection of Plastic Garbage in the Ocean: A Case for Turtle Preservation. Mathematical Problems in Engineering, 2022, 1–11. https://doi.org/10.1155/2022/3639222