Indonesian Journal of Electrical Engineering and Computer Science
Vol. 32, No. 1, October 2023, pp. 506~516
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v32.i1.pp506-516

Success factors for citizen-based government decision making using K-means fuzzy learning vector quantization

Pradityo Utomo[1], Kusworo Adi[2], Oky Dwi Nurhayati[3]
[1]Department Doctoral of Information System, Diponegoro University, Semarang, Indonesia
[2]Department of Physic, Faculty of Science and Mathematics, Diponegoro University, Semarang, Indonesia
[3]Department Master of Information System, Diponegoro University, Semarang, Indonesia

Article Info

Article history:

Received Apr 19, 2023
Revised Jul 9, 2023
Accepted Jul 16, 2023

Keywords:

Citizen
Criteria
Decision-making
Government
K-Means FLVQ

ABSTRACT

The Indonesian government often needs assistance in making citizen-based decisions, for example selecting work program plans. Residents have their criteria in the forum to choose a work program plan. This study proposes the K-means fuzzy learning vector quantization (FLVQ) methods to select citizen-based government decision-making criteria. The K-means FLVQ method has never been used to assist government decision-making. However, citizen criteria can be a success factor for government decision-making. The selection of criteria begins with data collection from forum participants. The results of data collection get 11 criteria. Then, the K-means FLVQ method carries out labeling and classification. The addition of the K-means process in the selection criteria can provide optimal results. Citizens can give assessment criteria freely. Then the assessment of citizens is classified by FLVQ. The classification results obtained seven criteria, namely: i) urgency, ii) sustainability, iii) priority, vi) usability, v) prosperity, vi) comfortability, and vii) artistic. Governments can use these criteria to make decisions about planned work programs. The criteria selection algorithm was also evaluated using the confusion matrix method with an accuracy of 88% and an error of 12%.

Corresponding Author:

Pradityo Utomo
Department Doctoral of Information System, Diponegoro University
Imam Bardjo SH Street 5, Pleburan, Semarang, Indonesia
Email: pradityou@gmail.com

1. INTRODUCTION

Government is a group of organizations. The government must listen to citizens' opinions [1]. Citizen participation helps the government to make decisions [2]–[4]. Fan and Hui [5] said government interaction with citizens is vital in decision-making. Because decisions must be adjusted to the needs of citizens [5]. Morris *et al*. [6] argue that citizen-based decision-making can reduce problems. Kang *et al*. [7] also said that every government decision and justice impacts citizens' welfare. Therefore, Molinera *et al.* [8] argues that decision-makers often experience difficulties making decisions in government decision-making forums. Turner [9] argues that good governance can positively impact the success of a country's development. Social equality must be applied. This solution can avoid social inequality [10]. If these problems are resolved, citizens can strengthen the government to achieve national development [11], [12]. Based on the opinion of previous research, the government has not listened to the views of citizens optimally. The government often ignores the idea of citizens for decision-making. This research will accommodate the opinions of citizens for decision-making with information technology. Information technology can model mathematical calculations. Previous studies have not used information technology for problem-solving in development plan decision-making.

Several problems can be modeled into mathematical calculations [13]. The classification technique is a model of mathematical calculations. This technique can improve government efficiency and decision-making effectiveness [14]. Artificial intelligence is considered capable of helping government decision-making [15]. Artificial intelligence methods can be used for classification techniques, for example, neural networks [16]. Utomo *et al*. [17] use classification techniques to assist decision-making, especially in selecting criteria. These methods are neural networks, logistic regression, support vector classifiers (SVC), gradient boosting classifier, extra trees classifier, bagging classifier, AdaBoost classifier, Gaussian Naive Bayes (GNB), multi-layer perceptron (MLP) classifier, XG boost (XGB) classifier, light gradient boosting machine (LGBM) classifier, K nearest neighbor classifier, decision tree classifier, and random forest classifier. The study uses labeled data. Data has true and false labels [17]. Another classification method is neuro-fuzzy, Doctor *et al*. [18] used the neuro fuzzy method to select criteria. The criteria can be used for the selection of company employees. Other studies related to decision-making have used the neuro fuzzy method, for example, selecting agricultural service providers by Ren *et al*. [19], disease detection in the medical field by Kadhim [20], selecting construction machinery by Bozanic *et al*. [21], and manufacturing industries by Asadi *et al.* [22]. Combining the fuzzy and the neural network methods can provide optimal solutions for decision-making [19], [21]. Apart from the neuro fuzzy method, the fuzzy learning vector quantization (FLVQ) method combines fuzzy and neural networks [23]. Several applications have used the FLVQ method for classification. Amezcua and Melin [24] said that the advantage of the FLVQ method is that it can solve complex problems. Damayanti and Wediningsih [25] also proved the performance of the FLVQ method. The FLVQ method has 100% success in classifying tumors. Based on previous research, this study will use a classification technique to select citizen-based criteria. Citizen-based criteria are citizen opinions to make decisions on work program plans. The work program plan embodies the country's development. Selection of criteria using the FLVQ method. Classification techniques can be used for labeled data. Citizen-based criteria data does not yet have a label. This work aims to optimize the value of dynamic citizen-based criteria selection. After all citizen-based criteria data has been collected, the data is labeled. Labeling can use clustering techniques.

One of the clustering methods is the K-means method. In decision-making research, the K-means method can be combined with several methods. For example, in the manufacturing industry, Cao *et al.* [26] merged the K-means method with the multi-objective Runge Kutta optimizer (MORUN) and analytic hierarchy process (AHP) methods. In the medical field, Ilbeigipour *et al.* [27] combined the K-means method with the self organizing map (SOM) method. Xiong *et al*. [28] said the K-means method could solve large-scale group decision-making problems. The performance of the K-means method can also be seen in research other than decision-making. In the medical field, Wisky *et al*. [29] developed a machine-learning method using the K-means method to analyze infectious diseases. Rahman and Selvaperumal [30] combined the K-means method with the neuro fuzzy method for brain segmentation. Mukti *et al.* [31] said that the performance of the K-means method was better than other clustering methods. Kim *et al*. [32] also noted that the K-means method efficiently labeled clusters. Another research, Isnanto *et al.* [33] proved that the accuracy of the K-means method is better than other methods. The K-means method puts similar data into the same labels [34].

Based on previous research, the FLVQ method has never been used to support decision-making, especially in government. This study uses the FLVQ method to support decision-making. The FLVQ method is used to classify decision-making criteria. According to previous studies, classification methods can classify labeled data. This study uses unlabeled data. Because one of the objectives of this research is to obtain citizen-based criteria data. For this reason, data labels are not specified during data collection. Citizens can give value to each criterion freely. After the data is collected, labeling uses the K-means method. Then the labeled data is classified using the FLVQ method. The classification results are the selected criteria for reference in government decision-making. The K-means method and the FLVQ method have advantages. For this reason, this research is expected to produce citizen-based valid criteria data. The government can use citizen-based criteria to become a success factor. Hybrid K-means and FLVQ methods have never been used for citizen-based government decision-making. In addition to determining citizen-based criteria, this study also aims to determine the combined performance of the K-means method and the FLVQ method in government decision-making. The Confusion Matrix method can measure the performance of the proposed method.

Several confusion matrix tests are accuracy and error testing [35], [36]. Accuracy testing is done by comparing the results of a method's calculations with the expected results. Where the results should be able to use pre-existing data. The advantage of the confusion matrix method is it can test the validity of the results of a method [37].

2. METHOD

This research has four stages. Each stage of research uses the correct methods to get accurate results. The steps of the study are shown in Figure 1. Figure 1 describes the stages of the research. The steps of the study started with problem identification, data collection, determining data criteria, and evaluation. The problem identification stage was carried out by interviewing stakeholders. Then, the step continued with data collection. They collected data using observation and survey methods. After the data is collected, the next stage is determining the criteria data. The step of determining the criteria data consists of labeling and classification. Data labeling uses the K-Means method. Classification of data using the FLVQ method. The final stage of research is evaluation. The evaluation stage consists of recognition and testing. Recognition of using the Euclidean distance method. Testing using the confusion matrix method. The results of this study include the criteria selected for government decision-making and the success of the proposed method.

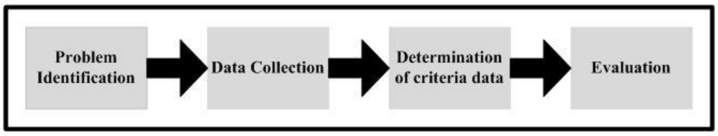

Figure 1. Research flow

2.1. K-means clustering

The research data is data that does not have a label. Therefore, as many as 775 data must be labeled. This research uses two labels, true and false. Labeling can use the K-means method. The K-means method can group similar data into data with the same label [34]. The calculation of the K-means method can be shown as [31], [38].

- Determine K (cluster center) as the number of clusters.
- Calculate the distance between vector data (x) and the cluster center in (1).

$$Dist(x_1, x_2) = \sqrt{\sum_{i=1}^{n} (x_{1i} - x_{2i})^2} \tag{1}$$

Where $Dist(x_1, x_2)$=the distance between vector data and the cluster center, x_1=cluster center vectors, and x_2=vector data.

- Determine the distance of each vector to the nearest center.
- Calculate the average value of all vectors in each cluster to become the new cluster center.
- Go back to step 1 with the new cluster center.
- If there is a change in the vertex point, then return to step 3. Otherwise, it is finished.

2.2. Fuzzy learning vector quantization

The FLVQ method is a combination of the fuzzy c-means (FCM) and learning vector quantization (LVQ) methods. The FLVQ method adopts the membership function of the FCM method [23].

$$u_{ij} = \left[\sum_{i=1}^{C} \left(\frac{\|x_i - w_j\|^2}{\|x_i - w_l\|^2} \right)^{1/(m_k - 1)} \right]^{-1} ; m_k > 1 \tag{2}$$

Where u=degree of membership, x=vector data, C=number of classes, w=weight of each iteration, m=weight rank, k=iteration. The FLVQ method also adopts the LVQ method regarding calculating the distance from the input vector. The FLVQ calculation algorithm can be shown as [23]:

- Initializing $C, mi, mf, N,$ and $k = 0$.
- Set $w_0 = \{w_1, w_2, \ldots, w_C\}$.
- Calculate iteration.

$$k = k + 1 \tag{3}$$

- $m = mi + k \left[{(mf - mi)}/{N} \right]$ (4)

- $$\alpha_{ij,k} = \left[\sum_{i=1}^{C}\left(\frac{\|x_i - w_{j,k-1}\|^2}{\|x_i - w_{l,k-1}\|^2}\right)^{1/(m-1)}\right]^{-m} ; 1 \le i \le M; 1 \le j \le C; \quad (5)$$

- $$\eta_{j,k} = \left(\sum_{i=1}^{M}\alpha_{ij,k}\right)^{-1}; 1 \le j \le C; \quad (6)$$

- $$w_{j,k} = w_{j,k-1} + \eta_{j,k}\sum_{i=1}^{M}\alpha_{ij,k}\left(x_i - w_{j,k-1}\right); 1 \le j \le C; \quad (7)$$

- $$E_k = \sum_{j=1}^{C}\left\|w_{j,k} - w_{j,k-1}\right\|^2 \quad (8)$$

- if $w < N$ and $E_k > \xi$, go back to step 3.

Where C=number of classes as a target, mi=weighted rank as a membership function, mf=weighted rank as a membership function, N=maximum iteration limit, ξ=error tolerance, w=weight of each iteration, η=learning rate for weight, E_k=error calculation for each iteration, k=iteration.

2.3. Euclidean distance

At the evaluation stage, the Euclidean distance method can match data similarities [39]. This method can match the input vector with the reference weight vector. The selected vector values are those that have a minimum Euclidean distance value [40]. Euclidean distance calculation can be calculated with (9) [40].

$$d_x(x, w_k) = \sqrt{(x - w_k)^T(x - w_k)} \quad (9)$$

Where d=distance, x=data value, w=reference weight.

2.4. Confusion matrix

The confusion matrix method is very suitable for validating classical methods [36]. One can find out the performance results of a method [37]. The matrix of this method can be realized in (10) [41].

$$M = \begin{bmatrix} TP & FN \\ FP & TN \end{bmatrix} \quad (10)$$

This method can calculate the accuracy in (11) and error in (12) of a method [35], [36].

$$Accuracy = \frac{TP+TN}{P+N} x100\% \quad (11)$$

$$Error = \frac{FP+FN}{P+N} x100\% \quad (12)$$

Where true positive (TP)=the amount of data from class true is recognized as true, true negative (TN)=the amount of data from the false class is recognized as false, false positive (FP)=the amount of data from the false class is recognized as true, false negative (FN)=the amount of data from true class is recognized as false, positive (P)=sum of TP and FN data, negative (N)=sum of FP and TN data.

2.5. The proposed method

The method proposed by this study is to obtain citizen-based criteria for government decision-making. Every citizen has the same opportunity to give a value to each criterion based on the results of observations. For this reason, the proposed method must be able to solve this problem. Because in previous studies, each piece of data already has a label. This study does not specify a label at the start. This is done to give freedom to citizens in the assessment. After the research obtained the residents' survey results, further research used methods for labeling and classification. Combining the K-means and FLVQ methods will likely solve this problem. Previously, this combination of methods had never been used for government decision-making research. The combination of the K-means and FLVQ methods is called the K-means FLVQ method. The method is described:

- Determine K.
- Calculate the distance with (1).
- Select the smallest distance.
- Calculate the average value of all vectors in each cluster to become the new cluster center.

- If the cluster is new, go back step 1. Otherwise, go to step 6.
- If there is a change in the vertex point, then return to step 3. Otherwise, go to step 7.
- Initializing $C, mi, mf, N,$ and $k = 0$, $w_{ij,0}$.
- Calculate k with (3).
- Calculate m with (4).
- Calculate α with (5).
- Calculate η with (6).
- Calculate w with (7).
- Calculate E with (8).
- if $w < N$ and $E_k > \xi$, go back to step 8. Otherwise, go to step 15.
- Get the value of final $w_{ij,k}$.
- Determine the criteria based the largest values $w_{ij,k}$.

3. RESULTS AND DISCUSSION

3.1. Problem identification

Before conducting research, the initial stage must identify the problem. Based on the results of identifying problems with several village heads, the village head gave a statement.

> "*We often experience difficulties in distributing program proposals from the community.*" (Pramudya Dewanto)

> "*The village government has tried to even out the program proposals, but sometimes there are hamlet areas that don't get the same proposals. It can cause envy and misunderstanding*." (Marquat)

Some participants should have expressed their opinions on selecting work program plans during the deliberations. Participants with high social strata dominate the considerations. In contrast, participants with lower social strata are reluctant to participants with high social strata. For this reason, this study proposes a method to solve the problem. The proposed method can select criteria fairly. Because the criteria in decision-making are the basis for citizen-based decision-making. Citizen-based criteria are a factor in the success of government decision-making in determining work program plans.

3.2. Data collection

This research uses primary data. Primary data were obtained from regions in Indonesia, especially in five sub-districts, nine villages, and 32 hamlets, to be precise. This study has two types of primary data, namely: i) the criteria of decision makers to choose a work program plan and ii) the results of a citizen-based criteria selection survey.

Data collection uses two methods, observation and survey. The observation method is used to obtain the criterion data. Criteria data is a citizen's parameter in selecting a work program. Implementation of observations simultaneously with the performance of deliberations on development plans. The results of the observation are the criteria data including (C1) artistic, (C2) usability, (C3) priority, (C4) urgency, (C5) sustainability, (C6) safety, (C7) suitability, (C8) prosperity, (C9) security, (C10) healthy, (C11) comfortability. After the criteria data is collected, the research surveys citizens to obtain an assessment of each criterion.

The assessment uses a Likert scale with a ratio of 1-9. The maximum value is 9, and the minimum value is 1. Assessment criteria are based on the level of importance [42]. The survey results using a Likert scale are as many as 775 data. Some survey results are shown in Table 1.

Table 1. List of survey results

ID	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
P1	8	5	9	8	8	7	9	6	9	6	8
P2	6	8	8	6	7	7	6	8	3	4	5
P3	8	6	7	8	7	5	5	8	3	8	5
…	…	…	…	…	…	…	…	…	…	…	…
P774	7	8	7	8	7	8	8	7	5	6	7
P775	7	8	7	8	8	7	5	6	7	8	7

3.3. Determination of criteria data

Table 1 has displayed survey data. Where the data is citizen-based, so the data does not have a label. At the forum, citizens and the government can assess each criterion. Based on the survey data, the research continued to determine the criteria data. Determination of criteria data using the K-means FLVQ method. The stage of determining the criteria data is divided into two: labeling and classification. Labeling using the K-means method. Labeling is made of two types, assuming that it will later have true and false data. The labeling stages produce labeled data, as shown in Table 2.

Table 2. List of data labeling results

ID	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	Label
P1	8	5	9	8	8	7	9	6	9	6	8	Cluster0
P2	6	8	8	6	7	7	6	8	3	4	5	Cluster0
P3	8	6	7	8	7	5	5	8	3	8	5	Cluster0
…	…	…	…	…	…	…	…	…	…	…	…	…
P774	7	8	7	8	7	8	8	7	5	6	7	Cluster0
P775	7	8	7	8	8	7	5	6	7	8	7	Cluster0

Table 2 shows labeled data. For data labeled cluster0 defines data as true, while data labeled cluster1 defines data as false. The number of data labeled cluster0 is 375 data, and the number of data labeled cluster1 is 400. Of the 375 data labeled cluster0 (True), 100 data are used for testing, and 275 data are used for training. For false data (cluster1), there are 400 data. From this data, 100 data are used for testing data, and 300 are used for training data. So that 575 data are used for training in the classification stage. The classification stage uses the FLVQ method. The FLVQ calculation must first determine the initial weight. The initial weight is set as:

$$w_{ij,0} = \begin{bmatrix} 9 & 9 & 9 & 9 & 9 & 9 & 9 & 9 & 9 & 9 & 9 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

From the initial weight matrix, it can be interpreted that the first row shows an assessment of true data, while the second row shows an evaluation of false data. Each column defines the assessment of each criterion (11 criteria). The first row is set to the maximum value of the Likert scale for each column because the community gives a maximum value of 9 for True data. For the second line, the minimum value for the data is set to False because there are people who give a minimum value of 1. The initial weight is not the same as one of the data patterns. If the initial weight pattern is precisely the same as one of the data, it will cause a computational error.

For the value of the weighted rank mi=50 and the weighted rank mf=20. A higher value of the weighted rank can provide maximum training results. Error tolerance is made 10^{-6}. This study uses the number of iterations of 250 iterations, 500 iterations, 750 iterations, 1,000 iterations, and 1,250 iterations. The final weight (true data) produced for each maximum iteration is shown in Table 3.

Table 3. The final weight shows true data

Weight	250 iterations	500 iterations	750 iterations	1000 iterations	1250 iterations
w[0,0]	5,27304348611581	5,27304348201509	5,27304348081568	5,27304348023049	5,27304347987935
w[0,1]	5,31478263157198	5,3147826197319	5,3147826162759	5,31478261458462	5,31478261356593
w[0,2]	5,36347826648209	5,3634782635485	5,36347826268957	5,36347826227068	5,36347826201949
w[0,3]	5,44173915492132	5,44173914224423	5,44173913854408	5,44173913673349	5,44173913564306
w[0,4]	5,39652175598482	5,39652174724629	5,39652174469449	5,39652174344647	5,39652174269534
w[0,5]	5,21391305502325	5,21391304904075	5,21391304729322	5,21391304643838	5,21391304592380
w[0,6]	5,2400000107732	5,2400000051704	5,24000000353305	5,24000000273312	5,24000000225234
w[0,7]	5,28695653262132	5,28695652697797	5,28695652532908	5,2869565245227	5,28695652403747
w[0,8]	5,20869565846108	5,20869565518205	5,20869565422255	5,20869565375425	5,20869565347316
w[0,9]	5,20869566280171	5,20869565728517	5,20869565567333	5,20869565488532	5,20869565441132
w[0,10]	5,27478261841165	5,27478261336749	5,2747826118932	5,27478261117248	5,27478261073902

Based on Table 3, it can be seen that the final weight values of the true groups. The selection of criteria can refer to the true final weight. From the calculation using the five types of shrimp paste, the maximum average analysis is carried out so that the value of the selection criteria can be seen from the high or low true weight value. The average calculation results using true weight values are shown in Figure 2.

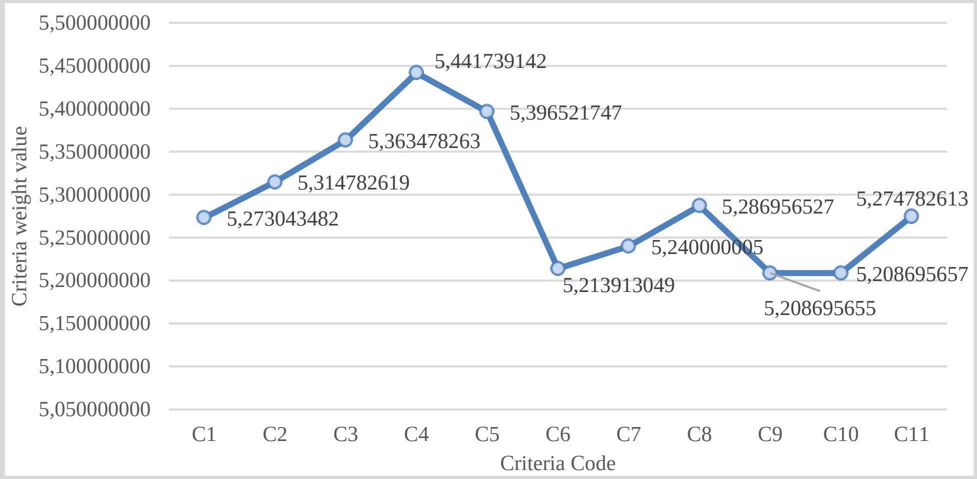

Figure 2. Average criterion value of true data

Figure 2 shows the high or low weight value of each criterion. The higher the value of the criterion (C[j]), the more the criterion is selected for use in decision-making. In this study, seven criteria were selected with high scores from eleven criteria. The seven criteria are urgency (C4), sustainability (C5), priority (C3), usability (C2), prosperity (C8), comfortability (C11), and artistic (C1).

3.4. Evaluation

The evaluation stage is divided into two, recognition and testing. Recognition using 200 data testing. Where 200 testing data are processed with the Euclidean distance to the final weight of each iteration, a total of 200 data consists of 100 true data and 100 false data. But the system does not correctly recognize some data. The list of results of recognition data using the Euclidean distance method is shown in Table 4.

Table 4. The results of the recognition with the Euclidean distance method

Iteration	Data type	The amount of data	The amount of data recognized as data true	The amount of data recognized as data false
250	True	100	91	9
	False	100	16	84
500	True	100	91	9
	False	100	16	84
750	True	100	91	9
	False	100	16	84
1000	True	100	92	8
	False	100	16	84
1250	True	100	92	8
	False	100	16	84

Based on Table 4, the recognition results can be used as a matrix reference for testing with the confusion matrix method. There are five types of matrices (M) which are the reference for calculating the confusion matrix.

- The matrix of the results of the recognition of the training weights of 250 iteration, 500 iterations, and 500 iterations.

$$M = \begin{bmatrix} 91 & 9 \\ 16 & 84 \end{bmatrix}$$

- The matrix of the results of the recognition of the training weights of 750 iteration and 1000 iterations.

$$M = \begin{bmatrix} 92 & 8 \\ 16 & 84 \end{bmatrix}$$

Testing uses these matrices to get accuracy and error. The test results in calculating the accuracy and error rate are shown in Table 5.

Table 5. Testing results

Iteration	Accuracy (%)	Error-rate (%)
250	87,5	12,5
500	87,5	12,5
750	87,5	12,5
1000	88	12
1250	88	12

Based on Table 5, the highest accuracy value is 88%, using training results with 1,000 iterations and 1250 iterations. In comparison, the accuracy value is 87.5% if the test uses the effects of training 250 iterations, 500 iterations, and 750 iterations. The highest error-rate value is 12.5% using the results of training 250 iterations, 500 iterations, and 750 iterations. The error-rate value is 12% in the test using the training results of 1,000 and 1,250 iterations.

3.5. Performance comparison

The performance of the proposed method is included in the good category. This is shown in Table 5. The combined method of K-means and FLVQ has the highest accuracy of 88% and an error of 12%. Based on these results, this study can refine other classification methods for decision-making in previous studies. In addition, the FLVQ K-means method can improve the accuracy and error of the methods in previous studies. This is shown in Table 6 and Table 7.

Table 6. Comparison of the accuracy of the classification methods on decision making

Method	Accuracy (%)
Neuro fuzzy [18]	71,6
Fuzzy neuro decision support system back pain [18]	83,6
K-means [43]	74,4
Albert LR [44]	63,6
Albert support vector machine [44]	50,3
Albert decision tree [44]	59,3
Albert Gaussian Naïve Bayes [44]	63
Albert K-nearest neigbor [44]	59,7
Albert random forest [44]	57,4
Albert gradient boosting decision tree [44]	62
Albert multi-layer perceptron [44]	55,7
Naïve Bayes [45]	81,6
Decision tree C4.5 [45]	83,2
Artificial neural network [45]	84,5
K-nearest neighbor [45]	85,4
Fuzzy Naïve Bayes [45]	81,1
Fuzzy multi-layer perceptron [45]	84,2
Fuzzy decision tree [45]	86,7
Random Forest [46]	81,8
Light GBM [46]	68,6
K-means FLVQ	88

Table 7. Comparison of the error of the classification methods on decision making

Method	Error (%)
Adaptive neuro fuzzy inference system [21]	14,5
Quadratic basic normal classifier [21]	15,2
K-nearest neighbor with k=3 [21]	28,5
K-means FLVQ	12

A comparison of the performance of the K-means FLVQ method with other methods is shown in Table 6 and Table 7. Table 6 displays the accuracy comparison, and Table 7 displays the error comparison. Based on performance comparisons, the K-means FLVQ method has better accuracy and error than other decision-making methods. In decisions-making, the K-means FLVQ method can improve the accuracy of other classification methods. In addition, the K-means FLVQ method can also reduce error values from other methods.

The K-means FLVQ method has labeling and classification capabilities, so the K-means FLVQ method is very suitable for solving this research problem. This research requires a method to select criteria based on citizens. Generally, selection techniques are synonymous with classification. But classification techniques require labeled data. This study does not specify a label at the outset. If this research determines data labels at the beginning of the study, then government decision-making cannot be called citizen-based. Because citizens have to fill in two types of data: true and false. This can reduce the freedom of citizens to give value. For this reason, this research gives freedom to citizens in assessing it. Citizens can provide a value to each criterion with a Likert scale of 1 – 9 freely. After that, the K-means FLVQ method can determine these criteria. Because these criteria are selected from the results of independent citizen assessments, these criteria are said to be citizen-based. Therefore, the selected criteria can be used as a success factor in government decision-making, especially in selecting work program plans.

4. CONCLUSION

Combining K-means and FLVQ methods can solve the problems in this study. The combination of these methods is called K-means FLVQ. First, the proposed method can select citizen-based criteria in government decision-making. Previously, the combination of K-means and FLVQ methods was not used for decision-making research. The proposed method can resolve cases of social inequality in the decision-making forum for work program plans. This study did not specify a label at the start. This is done to give freedom to all forum participants. Because most forum participants are residents, the results obtained are citizen-based. These results are the criteria for selecting work program plans for government decision-making. These criteria are urgency, sustainability, priority, usability, prosperity, comfortability, and artistic.

Second, the proposed method can improve other classification methods in decision-making. This is shown in the discussion of performance comparison. The K-means FLVQ method has an accuracy of 88% and an error of 12% in solving this research problem. The K-means FLVQ method shows better performance than other decision-making methods. For this reason, the K-means FLVQ method is successful in selecting criteria. Furthermore, these criteria can be used for government decision-making. For example, the government can use these criteria to select work program plans.

ACKNOWLEDGEMENTS

The authors would like to thank to the government of Madiun Regency, East Java, Indonesia, regarding the research data. The authors also thank to Diponegoro University and Merdeka Madiun University. Because of the support, the authors were able to finish this research.

REFERENCES

[1] A. Sharma and H. Shekhar, "Intelligent learning based opinion mining model for governmental decision making," *Procedia Computer Science*, vol. 173, no. 2019, pp. 216–224, 2020, doi: 10.1016/j.procs.2020.06.026.

[2] M. D. T. de Jong, S. Neulen, and S. R. Jansma, "Citizens' intentions to participate in governmental co-creation initiatives: Comparing three co-creation configurations," *Government Information Quarterly*, vol. 36, no. 3, pp. 490–500, 2019, doi: 10.1016/j.giq.2019.04.003.

[3] V. G.-Navarro, D. S.-Strelet, D. B.-Carrubi, and H. G. Gomez, "Media or information literacy as variables for citizen participation in public decision-making? A bibliometric overview," *Sustainable Technology and Entrepreneurship*, vol. 2, no. 1, p. 100030, 2023, doi: 10.1016/j.stae.2022.100030.

[4] J. P. Fernandes, N. Guiomar, and A. Gil, "Identifying key factors, actors and relevant scales in landscape and conservation planning, management and decision making: Promoting effective citizen involvement," *Journal for Nature Conservation*, vol. 47, no. June 2018, pp. 12–27, 2019, doi: 10.1016/j.jnc.2018.11.001.

[5] K. Fan and E. C. M. Hui, "Evolutionary game theory analysis for understanding the decision-making mechanisms of governments and developers on green building incentives," *Building and Environment*, vol. 179, no. February, p. 106972, 2020, doi: 10.1016/j.buildenv.2020.106972.

[6] J. C. Morris, M. W. McNamara, and A. Belcher, "Building resilience through collaboration between grassroots citizen groups and governments: two case studies," *Public Works Management and Policy*, vol. 24, no. 1, pp. 50–62, 2019, doi: 10.1177/1087724X18803116.

[7] K. Kang, M. Wang, and X. Luan, "Decision-making and coordination with government subsidies and fairness concerns in the poverty alleviation supply chain," *Computers and Industrial Engineering*, vol. 152, no. 5340, p. 107058, 2021, doi: 10.1016/j.cie.2020.107058.

[8] J. A. M.-Molinera, F. J. Cabrerizo, J. Mezei, C. Carlsson, and E. H.-Viedma, "A dynamic group decision making process for high number of alternatives using hesitant Fuzzy Ontologies and sentiment analysis," *Knowledge-Based Systems*, vol. 195, p. 105657, 2020, doi: 10.1016/j.knosys.2020.105657.

[9] R. Turner, "Investigating how governmentality and governance influence decision making on projects," *Project Leadership and Society*, vol. 1, no. May, p. 100003, 2020, doi: 10.1016/j.plas.2020.100003.

[10] A. Bächtiger and E. Beauvais, "Taking the goals of deliberation seriously: a differentiated view on equality and equity in deliberative designs and processes," *Journal of Deliberative Democracy*, vol. 12, no. 2, 2020, doi: 10.16997/jdd.254.

[11] B. W. Wirtz, J. C. Weyerer, and M. Rösch, "Open government and citizen participation: an empirical analysis of citizen expectancy towards open government data," *International Review of Administrative Sciences*, vol. 85, no. 3, pp. 566–586, 2019, doi: 10.1177/0020852317719996.

[12] N. T. Binh and N. Q. Giai, "Study on citizen participation in local governance in Vietnam," *NISPAcee Journal of Public Administration and Policy*, vol. 14, no. 2, pp. 87–109, 2021, doi: 10.2478/nispa-2021-0016.

[13] W. Shi, H. Wang, C. Chen, and Z. Kong, “Evolutionary game analysis of decision-making dynamics of local governments and residents during wildfires,” *International Journal of Disaster Risk Reduction*, vol. 53, no. October 2019, p. 101991, 2021, doi: 10.1016/j.ijdrr.2020.101991.
[14] C. H. Ku and G. Leroy, “A decision support system: Automated crime report analysis and classification for e-government,” *Government Information Quarterly*, vol. 31, no. 4, pp. 534–544, 2014, doi: 10.1016/j.giq.2014.08.003.
[15] D. Valle-Cruz, V. Fernandez-Cortez, and J. R. Gil-Garcia, “From E-budgeting to smart budgeting: exploring the potential of artificial intelligence in government decision-making for resource allocation,” *Government Information Quarterly*, vol. 39, no. 2, p. 101644, 2022, doi: 10.1016/j.giq.2021.101644.
[16] B. Jha, K. Ram, M. Rao, and Y. Sumanth, “Detailed dynamic modeling and vector control of doubly fed induction generator,” *International Journal of Applied Engineering Research*, vol. 5, no. 15, pp. 1765–1778, 2010, doi: 10.37622/000000.
[17] P. Utomo, K. Adi, O. D. Nurhayati, A. Wibowo, and B. Warsito, “Classification of multi-criteria group decision making in dynamic environments for resolving corruption cases,” in *Proceedings - 2021 International Seminar on Application for Technology of Information and Communication: IT Opportunities and Creativities for Digital Innovation and Communication within Global Pandemic, iSemantic 2021*, 2021, pp. 190–195, doi: 10.1109/iSemantic52711.2021.9573183.
[18] F. Doctor, H. Hagras, D. Roberts, and V. Callaghan, “A neuro-fuzzy based agent for group decision support in applicant ranking within human resources systems,” in *IEEE International Conference on Fuzzy Systems*, 2009, pp. 744–750, doi: 10.1109/FUZZY.2009.5277379.
[19] W. Ren, K. Wu, Q. Gu, and Y. Hu, “Intelligent decision making for service providers selection in maintenance service network: An adaptive fuzzy-neuro approach,” *Knowledge-Based Systems*, vol. 190, no. November 2019, p. 105263, 2020, doi: 10.1016/j.knosys.2019.105263.
[20] M. A. Kadhim, “FNDSB: A fuzzy-neuro decision support system for back pain diagnosis,” *Cognitive Systems Research*, vol. 52, no. August 2018, pp. 691–700, 2018, doi: 10.1016/j.cogsys.2018.08.021.
[21] D. Božanić, D. Tešić, D. Marinkovic, and A. Milić, “Modeling of neuro-fuzzy system as a support in decision-making processes,” *Reports in Mechanical Engineering*, vol. 2, no. 1, pp. 222–234, 2021, doi: 10.31181/rme2001021222b.
[22] S. Asadi, M. Nilashi, M. Iranmanesh, S. S. Hyun, and A. Rezvani, “Effect of internet of things on manufacturing performance: A hybrid multi-criteria decision-making and neuro-fuzzy approach,” *Technovation*, vol. 118, no. November, p. 102426, 2022, doi: 10.1016/j.technovation.2021.102426.
[23] N. B. Karayiannis and J. C. Bezdek, “An integrated approach to fuzzy learning vector quantization and fuzzy c-means clustering,” *IEEE Transactions on Fuzzy Systems*, vol. 5, no. 4, pp. 622–628, 1997, doi: 10.1109/91.649915.
[24] J. Amezcua and P. Melin, “A new fuzzy learning vector quantization method for classification problems based on a granular approach,” *Granular Computing*, vol. 4, no. 2, pp. 197–209, 2019, doi: 10.1007/s41066-018-0120-7.
[25] A. Damayanti and I. Werdiningsih, “Classification of tumor based on magnetic resonance (MR) brain images using wavelet energy feature and neuro-fuzzy model,” in *Journal of Physics: Conference Series*, 2018, vol. 974, no. 1, pp. 1–11, doi: 10.1088/1742-6596/974/1/012027.
[26] W. Cao, Y. Yu, J. Li, D. Wu, J. Ni, and X. Chen, “High stability multi-objective decision-making approach of dry hobbing parameters,” *Journal of Manufacturing Processes*, vol. 84, no. November, pp. 1184–1195, 2022, doi: 10.1016/j.jmapro.2022.10.077.
[27] S. Ilbeigipour, A. Albadvi, and E. A. Noughabi, “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*, vol. 32, p. 101005, 2022, doi: 10.1016/j.imu.2022.101005.
[28] K. Xiong, Y. Dong, and Q. Zha, “A large-scale consensus model to manage non-cooperative behaviors in group decision making: A perspective based on historical data,” *Expert Systems with Applications*, vol. 214, 2023, doi: 10.1016/j.eswa.2022.119163.
[29] I. A. Wisky, M. Yanto, Y. Wiyandra, H. Syahputra, and F. Hadi, “Machine learning classification of infectious disease distribution status,” *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 27, no. 3, pp. 1557–1566, 2022, doi: 10.11591/ijeecs.v27.i3.pp1557-1566.
[30] J. S. U. Rahman and S. K. Selvaperumal, “Integrated approach of brain segmentation using neuro fuzzy k-means,” *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 29, no. 1, pp. 270–276, 2023, doi: 10.11591/ijeecs.v29.i1.pp270-276.
[31] F. S. Mukti, A. Junikhah, P. M. A. Putra, A. Soetedjo, and A. U. Krismanto, “A Clustering optimization for energy consumption problems in wireless sensor networks using modified K-Means++ algorithm,” *International Journal of Intelligent Engineering and Systems*, vol. 15, no. 3, pp. 355–365, 2022, doi: 10.22266/ijies2022.0630.30.
[32] H. Kim, H. K. Kim, and S. Cho, “Improving spherical k-means for document clustering: Fast initialization, sparse centroid projection, and efficient cluster labeling,” *Expert Systems with Applications*, vol. 150, p. 113288, 2020, doi: 10.1016/j.eswa.2020.113288.
[33] R. R. Isnanto, O. D. Nurhayati, and T. P. Nan Cerah, “Area of mangrove forests calculated by color image segmentation using K-means clustering and region growing),” in *2020 5th International Conference on Informatics and Computing*, 2020, pp. 1–9, doi: 10.1109/ICIC50835.2020.9288530.
[34] N. Sureja, A. Vasant, and N. Chaudhari, “A hybrid shuffled frog-simulated annealing algorithm for clustering,” *International Journal of Intelligent Engineering and Systems*, vol. 14, no. 4, pp. 490–499, 2021, doi: 10.22266/ijies2021.0831.43.
[35] O. Okfalisa, R. Hafsari, G. Nawanir, S. Toto, and N. Yanti, “Optimizing placement of field experience program: an integration of moora and rule-based decision making,” *Pertanika Journal of Science and Technology*, vol. 29, no. 2, pp. 895–918, 2021, doi: 10.47836/pjst.29.2.11.
[36] H. G. Lewis and M. Brown, “A generalized confusion matrix for assessing area estimates from remotely sensed data,” *International Journal of Remote Sensing*, vol. 22, no. 16, pp. 3223–3235, 2001, doi: 10.1080/01431160152558332.
[37] S. Ruuska, W. Hämäläinen, S. Kajava, M. Mughal, P. Matilainen, and J. Mononen, “Evaluation of the confusion matrix method in the validation of an automated system for measuring feeding behaviour of cattle,” *Behavioural Processes*, vol. 148, no. March 2018, pp. 56–62, 2018, doi: 10.1016/j.beproc.2018.01.004.
[38] S. Periyasamy, S. Khara, and S. Thangavelu, “Balanced cluster head selection based on modified K-means in a distributed wireless sensor network,” *International Journal of Distributed Sensor Networks*, vol. 2016, pp. 1–11, 2016, doi: 10.1155/2016/5040475.
[39] D. Liu, G. Liu, and Z. Liu, “Some similarity measures of neutrosophic sets based on the euclidean distance and their application in medical diagnosis,” *Computational and Mathematical Methods in Medicine*, pp. 1–9, 2018, doi: 10.1155/2018/7325938.
[40] A. C. Bhensle and R. Raja, “An efficient face recognition using PCA and Euclidean distance classification,” *International Journal of Computer Science and Mobile Computing*, vol. 3, no. 6, pp. 407–413, 2014.

[41] A. Luque, A. Carrasco, A. Martín, and A. de las Heras, "The impact of class imbalance in classification performance metrics based on the binary confusion matrix," *Pattern Recognition*, vol. 91, no. February 2019, pp. 216–231, 2019, doi: 10.1016/j.patcog.2019.02.023.

[42] M. F. Hossain, Z. H. Adnan, and M. A. A. Hasin, "Improvement in weighting assignment process in analytic hierarchy process by introducing suggestion matrix and Likert scale," *International Journal of Supply Chain Management*, vol. 3, no. 4, pp. 91–95, 2014, doi: https://doi.org/10.59160/ijscm.v3i4.1000.

[43] Y. Khanorkar and P. V. Kane, "Selective inventory classification using ABC classification, multi-criteria decision making techniques, and machine learning techniques," in *Materials Today: Proceedings*, 2023, vol. 72, pp. 1270–1274, doi: 10.1016/j.matpr.2022.09.298.

[44] J. Li, Q. Huang, S. Ren, L. Jiang, B. Deng, and Y. Qin, "A novel medical text classification model with Kalman filter for clinical decision making," *Biomedical Signal Processing and Control*, vol. 82, p. 104503, 2023, doi: 10.1016/j.bspc.2022.104503.

[45] E. Zaitseva, J. Rabcan, V. Levashenko, and M. Kvassay, "Importance analysis of decision making factors based on fuzzy decision trees," *Applied Soft Computing*, vol. 134, no. January 2023, p. 109988, 2023, doi: 10.1016/j.asoc.2023.109988.

[46] H. Musbah, G. Ali, H. H. Aly, and T. A. Little, "Energy management using multi-criteria decision making and machine learning classification algorithms for intelligent system," *Electric Power Systems Research*, vol. 203, p. 107645, 2022, doi: 10.1016/j.epsr.2021.107645.

BIOGRAPHIES OF AUTHORS

Pradityo Utomo is a student at the Doctoral of Information Systems Department, Diponegoro University, Semarang, Indonesia. He earned a Master of Computer Science in 2016 at Gadjah Mada University, Yogyakarta, Indonesia. In addition, he earned a Bachelor of Computer degree in 2012 at Sebelas Maret University, Solo, Indonesia. His work he is a lecturer in the Informatics Management Department at Merdeka Madiun University, Indonesia. His research areas are intelligent systems, computer science, decision support systems, and information systems. He can be contacted at email: pradityou@gmail.com.

Kusworo Adi is a Professor at Diponegoro University, Semarang, Indonesia. His area of expertise is imaging and image processing. He has published more than 100 articles, with an H-index of 15. In his occupation, he is a lecturer in the Physics Department, Faculty of Science and Mathematics, Diponegoro University, Semarang, Indonesia. He is Deputy Dean of the Faculty of Science and Mathematics at Diponegoro University, Semarang, Indonesia. He can be contacted at email: kusworoadi@lecturer.undip.ac.id.

Oky Dwi Nurhayati is an Assistant Professor at Diponegoro University, Semarang, Indonesia. Her work she is a lecturer in the Master of Information System Department at Diponegoro University, Semarang, Indonesia. She is a secretary for the Master of Information Systems Department at Diponegoro University, Semarang, Indonesia. Her expertise is in artificial intelligence, image processing, multimedia, and computer engineering. She has had many publications since 2001, with an H-index of 16. She can be contacted at email: okydwi.n@gmail.com.