JOIV : Int. J. Inform. Visualization, 5(4) - December 2021 415-421

INTERNATIONAL JOURNAL
ON INFORMATICS VISUALIZATION

journal homepage : www.joiv.org/index.php/joiv

Feature-reduction Fuzzy c-means Clustering for Basketball Players Positioning

Yessica Nataliani [a,*]

[a] *Department of Information System, Satya Wacana Christian University, Diponegoro 52-60, Salatiga, 50711, Indonesia*
*Corresponding author: *yessica.nataliani@uksw.edu*

***Abstract*— One of the best-known clustering methods is the fuzzy c-means clustering algorithm, besides k-means and hierarchical clustering. Since FCM treats all data features as equally important, it may obtain a poor clustering result. To solve the problem, feature selection with feature weighting is needed. Besides feature selection by assigning feature weights, there is also feature selection by assigning feature weights and eliminating the unrelated feature(s). THE Feature-reduction FCM (FRFCM) clustering algorithm can improve the FCM clustering result by weighting the features and discarding the unrelated feature(s) during the clustering process. Basketball is one of the famous sports, both international and national. There are five players in basketball, each with a different position. A player can generally be in guard, forward, or center position. Those three general positions need different characteristics of players' physical conditions. In this paper, FRFCM is used to select the related physical feature(s) for basketball players, consisting of height, weight, age, and body mass index. to determine the basketball players' position. The result shows that FRFCM can be applied to determine the basketball players' position, where the most related physical feature is the player's height. FRFCM gets one incorrect player's position, so the error rate is 0.0435. As a comparison, FCM gets five incorrect player's positions, with an error rate of 0.2174. This method can help the coach decide the basketball new player's position.**

***Keywords*— Feature-reduction; clustering; fuzzy c-means; basketball; position.**

Manuscript received 24 Aug. 2021; revised 22 Sep. 2021; accepted 26 Nov. 2021. Date of publication 31 Dec. 2021.

I. INTRODUCTION

In pattern recognition, cluster analysis is unsupervised learning. Clustering is a method to find groups, such that the similar characteristics data will be in the same cluster, and data with different characteristics will be in the other clusters. Many areas use cluster analysis, such as business, image processing, education, etc. k-Means is the most popular clustering method, where each data point belongs to exactly one cluster. k-Means is extended to be fuzzy c-means (FCM), where each data point can be included into several clusters, depending on the cluster membership values [1], [2].

A dataset is defined by a number of points and dimensions (attributes, features). In general, the clustering process treats all features to be equally important. However, some features may be unrelated and affect the clustering performance. It is important to find the related features to obtain a better clustering result [3]. Feature selection can solve the problem by removing unrelated and redundant data to reduce the computation time, and the clustering performance can be improved [4], [5].

Furthermore, currently, high-dimensional data is widely used for data analysis. Of course, high-dimensional data needs more computation time. Therefore, the use of feature selection is very important [6]. Feature weight can be used to support feature selection, where each feature has its weight [7], [8]. Feature-weighted has been used for clustering algorithms, for example, sparse k-means, entropy-weighted k-means, feature-weighted k-means, feature-weighted FCM, weighted FCM with feature-weight learning, etc. [9]. Some literature applied feature selection and featured weight in clustering to discard unrelated features. There are feature-reduction FCM [9], feature-reduction k-means for multi-view data [10], feature-reduction scheme for possibilistic c-means [11], and dimensionality-reduction using PCA and k-means [12]. Basketball is one of the famous sports, both international and national. The National Basketball Association (NBA) is a well-known international basketball event. In Indonesia, one big event is named the Indonesian Basketball League (IBL). 12 teams are competing in IBL. One of them is Satya Wacana Saints [13]. This team consists of 23 students of Satya Wacana Christian University in Salatiga.

In basketball, there are five players, where each player can be on the shooting guard, point guard, small forward, power forward, or center position. Each position needs different players' physical conditions, e.g., height, weight, age, and body mass index (BMI). Some experts said that anthropometric or body measurements could determine the athlete's career, especially in basketball [14], [15]. Physical condition is important for coaches to choose a new player, besides his/her talent. Players' physical conditions affect the players' position in basketball. Unsuitable positions allow players not to play optimally [16], [17]. The players' physical condition can generally determine the position. If the player is tall and big, he can be in the center or power forward position. If the player is small and agile, he can be on the guard position [18].

In general, the players' positions can be divided into three positions, i.e., guard, forward, and center. Players in the guard position are more often outside the paint area. The team puts the smallest and most agile players for this position. Guards have less physical contact with opposing players than the forward and center positions. Guards usually are the brain of attack on a team. This position consists of two kinds, point guard and shooting guard. The second position is forward. A player in this position is a player whose job is to see an open position near the paint area, to break through the opponent's defense, or in other words, to drive inside. A forward is usually tall and strong because his main job is to defend and rebound. Players in this position must have a medium level of shooting accuracy. This position consists of two types, small forward and power forward. The last position is center. Players in this position are often called the big man, who is in charge of guarding their paint area and attacking the opponent's paint area. The center position is more likely to physically collide with opposing players in scoring or blocking the opponent's center position. This position is held by the tallest and biggest player [19].

Clustering can be used to find a suitable position for each player. Players with similar physical conditions will be grouped into the same position, and players with different physical conditions will be in different positions. In this paper, the grouping of the positions of Satya Wacana Saint basketball team based on their physical conditions with FCM will be discussed. The physical conditions used here are referred to as height, weight, age, and BMI. Furthermore, the feature-reduction method with FCM will be used to find the most important features that can determine the player's position, so it can help the coach decide whether a new basketball player is more suitable in what position.

There are some variants of feature-weighted clustering. Some of them are modifying the k-means and FCM clustering objective function into a new objective function by adding feature weight. Before some feature-weighted clustering is presented, FCM clustering will be described first.

A. FCM Clustering

Let $X = \{x_1, \ldots, x_n\}$ be a dataset on $\mathbb{R}^p$ with n is the number of data and p is the number of dimensions. The FCM objective function is defined in Eq. (1).

$$J(U,V) = \sum_{k=1}^{c} \sum_{i=1}^{n} \mu_{ik}^{m} d^2(x_i, v_k) \tag{1}$$

where c is the cluster number, n is the number of points, $\mu_{ik} = [0,1]$ is the membership value of the i^{th} points and the k^{th} cluster center, with $\sum_{k=1}^{c} \mu_{ik} = 1, \forall i$, x_i is the i^{th} point, v_k is the k^{th} cluster center, and $1 < m < \infty$ is the fuzzy exponent.

The Euclidean distance, $d(x_i, v_k) = \|x_i - v_k\|$, is the distance between the i^{th} point and the k^{th} cluster center. Since $d(x_i, v_k) = \sqrt{\sum_{j=1}^{p} (x_{ij} - v_{kj})^2}$, where p is the number of attributes (features or dimension), then Eq. (1) can be written as in Eq (2).

$$J(U,V) = \sum_{k=1}^{c} \sum_{i=1}^{n} \sum_{j=1}^{p} \mu_{ik}^{m} (x_{ij} - v_{kj})^2 \tag{2}$$

The updating formula for the membership value and the cluster center are shown in Eqs. (3) and (4),

$$\mu_{ik} = \frac{\left(\sum_{j=1}^{p} (x_{ij} - v_{kj})^2\right)^{-1/m-1}}{\sum_{t=1}^{c} \left(\sum_{j=1}^{p} (x_{ij} - v_{tj})^2\right)^{-1/m-1}} \tag{3}$$

$$v_{kj} = \frac{\sum_{i=1}^{n} \mu_{ik}^{m} x_{ij}}{\sum_{i=1}^{n} \mu_{ik}^{m}} \tag{4}$$

where $i = 1, \ldots, n$; $k = 1, \ldots, c$; $j = 1, \ldots, p$.

The FCM clustering algorithm can be described as follows, Input: points X, number of clusters c, and threshold $\varepsilon > 0$.

1) Initialize random cluster centers, $V_k^{(0)}$, $k = 1, \ldots, c$.
2) Let the iteration rate, $t = 1$.
3) Compute the membership values, $\mu_{ik}^{(t)}$ using Eq. (3).
4) Update the cluster centers, $V_k^{(t)}$ using Eq. (4).
5) If $\left\|V_k^{(t)} - V_k^{(t-1)}\right\| < \varepsilon$, then STOP, else go back to step c and $t = t + 1$.

Output: cluster points C_k, $k = 1, \ldots, c$ [1].
The stopping condition used in this algorithm is when there is no significant difference between the cluster center in the next iteration.

B. Feature-weighted Clustering

Some extensions from k-means for feature-weighted clustering are presented. Weighted k-means added the feature weights during clustering iteration processes. The objective function is $J(U,V,W) = \sum_{k=1}^{c} \sum_{i=1}^{n} \sum_{j=1}^{p} \mu_{ik} w_j^{\beta} (x_{ij} - v_{kj})^2$, where w_j is the feature weight of the j^{th} feature and β is a constant. Entropy weighted k-means is also an extension of k-means by adding a weight entropy term. The objective is to minimize the within-cluster distance and maximize the negative weight entropy. Its objective function is $J(U,V,W) = \sum_{k=1}^{c} \sum_{i=1}^{n} \sum_{j=1}^{p} \mu_{ik} w_{kj} (x_{ij} - v_{kj})^2 + \gamma \sum_{k=1}^{c} \sum_{j=1}^{p} \left(w_{kj} \log w_{kj}\right)$, where w_{kj} is the feature weight of the j^{th} feature in the k^{th} cluster and γ are a constant [9]. Sparse k-means is also used to select features by assigning feature weights in the interval [0,1] at first and then updating the feature's weight (s) with small weights into 0. Whether a feature weight is small or not is determined based on a threshold [20].

Besides k-means, FCM is also extended into some methods for feature-weighted clustering. Weighted FCM added feature-weight learning to improve the FCM performance, where the objective function is $J(U,V,W) = \sum_{k=1}^{c} \sum_{i=1}^{n} \sum_{j=1}^{p} \mu_{ik}^{m} w_j^2 (x_{ij} - v_{kj})^2$ [9]. The other method is simultaneous clustering and attributes discrimination (SCAD1). In SCAD1, each cluster has a different feature weight. SCAD1's objective function is $J(U,V,W) =$

$\sum_{k=1}^{c}\sum_{i=1}^{n}\sum_{j=1}^{p}\mu_{ik}^{m}w_{kj}(x_{ij}-v_{kj})^2+\sum_{k=1}^{c}\sum_{j=1}^{p}\delta_k w_{kj}^2$, where δ_k is a constant which indicates the importance of feature weights in each cluster [21]. All these methods consider feature-weighted clustering, so selecting the related feature(s) needs to be done manually.

Another modification from FCM objective function is for feature-reduction. The feature-reduction idea is to eliminate the unrelated feature(s) automatically. Some methods have been proposed by modifying the FCM clustering objective function. For example, feature-reduction for FCM clustering algorithm, where the objective function is $J(U,V,W)=\sum_{k=1}^{c}\sum_{i=1}^{n}\sum_{j=1}^{p}\mu_{ik}^{m}\delta_j w_j\left(x_{ij}-v_{kj}\right)^2+\frac{n}{c}\sum_{j=1}^{p}\left(w_j\log\delta_j w_j\right)$ [9]. Another method is feature-reduction fuzzy co-clustering algorithm (FRFCoC). The objective function is $J(U,V,W)=\sum_{k=1}^{c}\sum_{i=1}^{n}\sum_{j=1}^{p}\mu_{ik}v_{kj}\delta_j w_j\left(x_{ij}-v_{kj}\right)^2+\alpha_1\sum_{j=1}^{p}\left(w_j\log\delta_j w_j\right)+\alpha_2\sum_{k=1}^{c}\sum_{i=1}^{n}\left(\mu_{ik}\log\mu_{ik}\right)+\alpha_3\sum_{k=1}^{c}\sum_{j=1}^{p}\left(v_{kj}\log v_{kj}\right)$, where v_{kj} is the feature membership for the k^{th} cluster center and the j^{th} feature, α_1, α_2, α_3 are constants. Its objective function consists of four terms, one term is a modification from the FCM objective function and three terms are the entropy terms of feature weight, fuzzy membership, and feature membership [22].

C. Basketball Players' Position Analysis

Some analysis about basketball players' position has been presented in the literature. Pion *et al.* [23] found that artificial neural networks can provide a specific position characteristic in basketball. They mentioned that multivariate variance analysis could not predict the specific players' position accurately. Zhang *et al.* [24] used clustering to see the performance of NBA players according to their anthropometric features and their playing performance.

Erga and Nataliani have researched feature selection with FCM for basketball players' positioning. They combined four physical conditions, i.e., height, weight, age, and BMI, one by one to get the best clustering result, which was measured by the accuracy rate. Height and BMI are the best combinations to determine the player's position [19].

II. Materials and Method

Some datasets may contain unimportant features that affect the clustering results. These unimportant features should be discarded to make a better clustering result. Yang and Nataliani [9] proposed a method of feature-reduction for FCM clustering algorithm, abbreviated with FRFCM. In the FRFCM clustering, the features are weighted with feature weights. Feature(s) with small feature weights must be discarded during the clustering process. In this way, FRFCM clustering method improves FCM clustering. FRFCM automatically computes different feature weights of each feature by modifying the objective function of FCM in Eq. (2) and adding a feature-weighted entropy term, $\sum_{j=1}^{p}\left(w_j\log\delta_j w_j\right)$. This algorithm can eliminate the unrelated features with small feature weights, such that a better clustering result can be obtained and computation time can be decreased.

The objective function of FRFCM clustering is defined in Eq. (5).

$$J(U,V,W)=\sum_{k=1}^{c}\sum_{i=1}^{n}\sum_{j=1}^{p}\mu_{ik}^{m}\delta_j w_j(x_{ij}-v_{kj})^2+\frac{n}{c}\sum_{j=1}^{p}\left(w_j\log\delta_j w_j\right) \tag{5}$$

where p is the number of features, v_{kj} is the cluster center of the k^{th} cluster and the j^{th} feature, $w_j=[0,1]$ is the feature weight of the j^{th} feature, with $\sum_{j=1}^{p}w_j=1$. Here, a constant δ_j is used to handle the dispersion and variation of each feature. The formula of δ_j is shown in Eq. (6),

$$\delta_j=-\left|\frac{\mathrm{mean}(x)}{\mathrm{var}(x)}\right|_j \tag{6}$$

where $\mathrm{mean}(x)=\frac{\sum_{i=1}^{n}x_i}{n}$ and $\mathrm{var}(x)=\frac{\sum_{i=1}^{n}(x_i-mean(x))^2}{n-1}$, $\forall j$.

The updating formula for the membership, the cluster center, and the feature weight are computed as in Eqs. (7), (8), and (9), respectively,

$$\mu_{ik}=\frac{\left(\sum_{j=1}^{p}\delta_j w_j(x_{ij}-v_{kj})^2\right)^{-1/m-1}}{\sum_{t=1}^{c}\left(\sum_{j=1}^{p}\delta_j w_j(x_{ij}-v_{tj})^2\right)^{-1/m-1}} \tag{7}$$

$$v_{kj}=\frac{\sum_{i=1}^{n}\mu_{ik}^{m}x_{ij}}{\sum_{i=1}^{n}\mu_{ik}^{m}} \tag{8}$$

$$w_j=\frac{\frac{1}{\delta_j}\exp\left(\frac{-c\sum_{k=1}^{c}\sum_{i=1}^{n}\mu_{ik}^{m}\delta_j\left(x_{ij}-v_{kj}\right)^2}{n}\right)}{\sum_{y=1}^{p}\frac{1}{\delta_j}\exp\left(\frac{-c\sum_{k=1}^{c}\sum_{i=1}^{n}\mu_{ik}^{m}\delta_j\left(x_{ij}-v_{kj}\right)^2}{n}\right)} \tag{9}$$

The FRFCM clustering algorithm is described as follows,
Input: points X, number of clusters c, and threshold $\varepsilon>0$.

1) Initialize random cluster centers, $V_k^{(0)}$, $k=1,\dots,c$ and define the initialization of feature weight, $W_j^{(0)}=[1/p]_{1\times p}, j=1,\dots,p$.
2) Let the iteration rate, $t=1$.
3) Compute δ_j using Eq. (6)
4) Compute the membership values, $\mu_{ik}^{(t)}$ using Eq. (7).
5) Update the cluster centers, $V_k^{(t)}$ using Eq. (8).
6) Update the feature weights, $W_j^{(t)}$ using Eq. (9).
7) Discard the feature(s) if the feature weight is less than $1/\sqrt{np}$.
8) Adjust $W_j^{(t)}$ using $w_j'=\frac{w_j}{\sum_{y=1}^{p(new)}w_y}$, in order to keep $\sum_{j=1}^{p}w_j=1$.
9) If $\left\|\|W_j^{(t)}\|-\|W_j^{(t-1)}\|\right\|<\varepsilon$, then STOP, else go back to step 3) with $d=d^{(new)}$ and $t=t+1$. Since $d^{(new)}$ is obtained, then points, $X_i^{(t)}$ and the cluster center, $V_k^{(t)}$ need to be updated.

Output: cluster points X_c and feature weight [9].
The stopping condition used in this algorithm is when there is no significant difference between the feature weight of the current iteration and the next iteration. According to the algorithm, the flowchart of FRFCM algorithm is shown in Fig. 1.

This paper applies FRFCM to find the important and related feature(s) on basketball players' positioning. There are four features used to determine the position of a player, i.e., height, weight, age, and BMI.

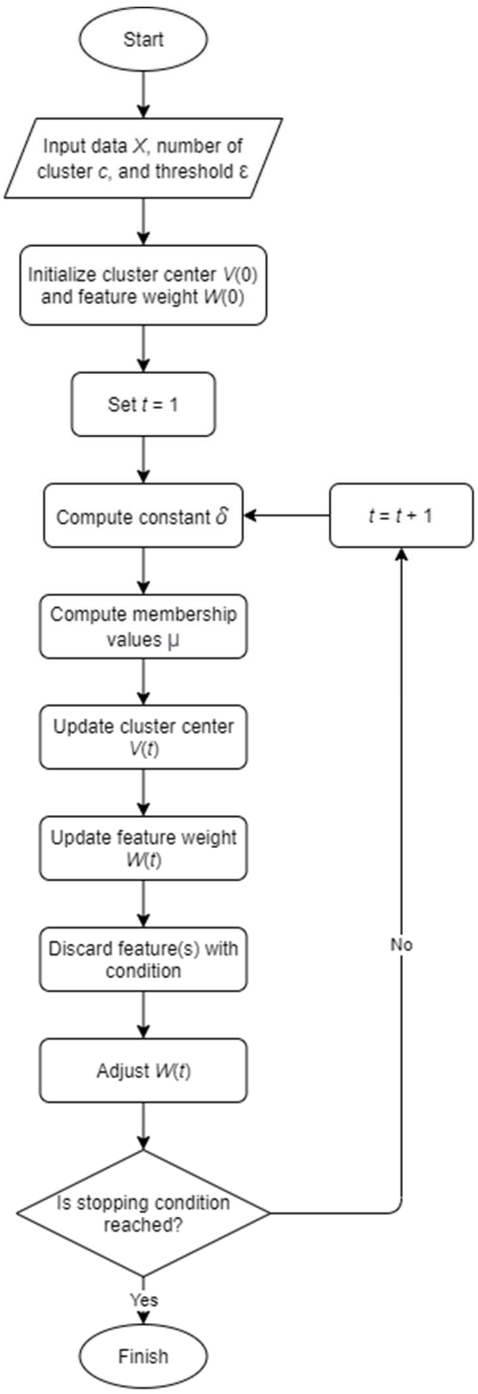

Fig. 1 Flowchart of FRFCM clustering algorithm

III. Results and Discussion

The data was collected from all 23 players of the Satya Wacana Saints Salatiga basketball team in 2021. There are two kinds of data collected, data of the physical conditions and data of the position of each player. Physical condition data consists of four features, features of height, weight, age, and BMI. Physical condition data is used to cluster players' positions using FCM, consisting of three positions, namely guard, forward, and center, while player position data compares the clustering results with actual conditions. Table I shows each player's physical condition and actual position, and Fig. 2 shows the matrix of scatter plots by a group for the basketball players' position according to their physical conditions.

TABLE I
BASKETBALL PLAYERS' DATA

Player's Name	Height	Weight	Age	BMI	Actual Position
Anjas Rusadi Putra	190	75	23	20	Forward
Antoni Erga	179	76	20	23	Guard
Ardian Ariadi	180	83	26	26	Guard
Aldi Falentino	171	70	20	24	Guard
Alexander Franklyn	184	82	20	25	Guard
Bryan Adha Elang	196	98	22	37	Center
David Liberty Nuban	190	80	22	22	Forward
Elyakim Tampa'i	175	75	23	25	Guard
Febrianus Gregory	181	76	21	23	Guard
Fransiscus Bryan	183	80	20	24	Guard
Henry Cornelis Lakay	196	96	22	25	Center
Raymond Putra Fajar	195	130	21	34	Center
Randy Ady Prasetya	202	77	23	19	Center
Mas Kahono Alif	192	79	19	21	Forward
Rian Sanjaya	178	73	22	23	Guard
Janson Kurniawan	178	69	21	22	Guard
M. Yassir Alkatiri	184	74	19	22	Forward
Martin	179	76	21	23	Guard
Steven Ray	178	75	21	24	Guard
Jody Sebastian	204	110	21	26	Center
Peter Surjantoro	171	72	19	24	Guard
Fauji	186	85	22	25	Forward
Ridho Pamungkas	189	85	24	25	Forward

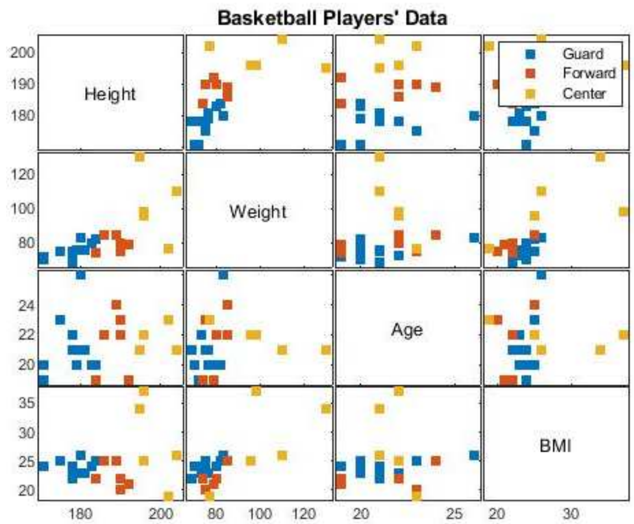

Fig. 2 Basketball players' data

All computations in this paper use the fuzzy exponent of $m = 2$. The error rate (ER) is used to measure the clustering performance. The formula of ER is $\mathrm{ER} = 1 - \frac{1}{n}\sum_{k=1}^{c} n(C_k)$, where $n(C_k)$ is the number of incorrect points in cluster k. The smaller the ER indicates, the better the clustering performance.

In the FCM clustering process, the first step is to determine the clusters, where $c = 3$, consisting of guard, forward, and center positions. After the cluster centers are initialized, then the membership is computed. The calculation of the cluster center and membership value are updated continuously until the stop condition is reached. Table II shows the final membership values of FCM clustering result. From Table II, for example, the first player, Anjas Rusadi Putra, according to FCM clustering result, he is more suitable on Cluster 2, with the membership value of 0.7130, than on Cluster 1 (with the membership value of 0.2635), and even more, does not suitable on the Cluster 3 (with the membership value of 0.0234). For the final cluster center of FCM clustering result, as shown in Table III, three cluster centers consist of four features.

TABLE II
MEMBERSHIP VALUES OF FCM CLUSTERING RESULT

Player's name	Cluster 1	Cluster 2	Cluster 3
Anjas Rusadi Putra	0.2635	0.7130	0.0234
Antoni Erga	0.9673	0.0304	0.0023
Ardian Ariadi	0.4992	0.4632	0.0377
Aldi Falentino	0.8587	0.1207	0.0206
Alexander Franklyn	0.2754	0.7050	0.0196
Bryan Adha Elang	0.1612	0.3517	0.4871
David Liberty Nuban	0.0290	0.9672	0.0037
Elyakim Tampa'i	0.9387	0.0548	0.0065
Febrianus Gregory	0.8965	0.0980	0.0056
Fransiscus Bryan	0.4623	0.5204	0.0173
Henry Cornelis Lakay	0.1611	0.5278	0.3111
Raymond Putra Fajar	0.0601	0.0854	0.8545
Randy Ady Prasetya	0.2238	0.6915	0.0847
Mas Kahono Alif	0.1061	0.8773	0.0166
Rian Sanjaya	0.9835	0.0151	0.0014
Janson Kurniawan	0.8962	0.0932	0.0106
M. Yassir Alkatiri	0.6974	0.2877	0.0149
Martin	0.9733	0.0249	0.0018
Steven Ray	0.9963	0.0034	0.0003
Jody Sebastian	0.0333	0.0646	0.9022
Peter Surjantoro	0.8693	0.1120	0.0187
Fauji	0.1220	0.8583	0.0197
Ridho Pamungkas	0.0653	0.9195	0.0152

TABLE III
CLUSTER CENTER OF FCM CLUSTERING RESULT

Cluster	Height	Weight	Age	BMI
Cluster 1	177.9485	74.4266	20.9808	23.4568
Cluster 2	189.6042	81.9087	21.8875	23.2556
Cluster 3	199.0460	115.3002	21.1878	30.3886

TABLE IV
FCM CLUSTERING RESULT

Player's name	Clustering results		Real position
	Cluster	Position	
Anjas Rusadi Putra	2	Forward	Forward
Antoni Erga	1	Guard	Guard
Ardian Ariadi	1	Guard	Guard
Aldi Falentino	1	Guard	Guard
Alexander Franklyn	**2**	**Forward**	**Guard**
Bryan Adha Elang	3	Center	Center
David Liberty Nuban	2	Forward	Forward
Elyakim Tampa'i	1	Guard	Guard
Febrianus Gregory	1	Guard	Guard
Fransiscus Bryan	**2**	**Forward**	**Guard**
Henry Cornelis Lakay	**2**	**Forward**	**Center**
Raymond Putra Fajar	3	Center	Center
Randy Ady Prasetya	**2**	**Forward**	**Center**
Mas Kahono Alif	2	Forward	Forward
Rian Sanjaya	1	Guard	Guard
Janson Kurniawan	1	Guard	Guard
M. Yassir Alkatiri	**1**	**Guard**	**Forward**
Martin	1	Guard	Guard
Steven Ray	1	Guard	Guard
Jody Sebastian	3	Center	Center
Peter Surjantoro	1	Guard	Guard
Fauji	2	Forward	Forward
Ridho Pamungkas	2	Forward	Forward

According to Table II, each data cluster can be determined by choosing the highest value of membership value. The clustering results of FCM are shown in Table IV, where 11 players are in the guard position, nine players are in the forward position, and three players are in the center position. Table IV shows that there are five players with incorrect positions (see the bold font), i.e., Alexander Franklyn, Fransiscus Bryan, Henry Cornelis Lakay, Randy Ady Prasetya, and M. Yassir Alkatiri. Therefore, the ER of the FCM clustering result is 0.2174.

Next, FRFCM is applied for this basketball player's position to see what feature(s) are related to determining the player's basketball position. Since the data has four features, then for the initialization, the feature weight is defined by $W_j^{(0)} = [0.25 \quad 0.25 \quad 0.25 \quad 0.25]$. For the computation of FCM, the same initialization of cluster centers with FCM, $V_k^{(0)}$, is used. After δ_j, $\mu_{ik}^{(1)}$, and $V_k^{(1)}$ are computed, then the feature weight, $W_j^{(1)}$, is updated. The result for the first iteration shows that the features weight of weight, age, and BMI are close to 0, while the feature weight of height is close to 1. The threshold for discarding features is defined by $1/\sqrt{np} = 1/\sqrt{(23)(4)} = 0.1043$. Therefore, the weight, age, and BMI features are discarded from the clustering process. The next step is adjusting $W_j^{(1)}$, where $w_j' = \frac{w_j}{\sum_{y=1}^{p(new)} w_y}$, such that the new $W_j^{(1)} = [1.00]$. The new W_j is used for the next iteration, along with the new cluster center and the new number of features. This process is repeated until the stop condition is reached. Table V shows the feature weights for each feature in each iteration. Tables VI and VII show the final membership values and the final cluster center of FRFCM clustering results, respectively. The final cluster center consists of just one feature, i.e., height, where the guard players (Cluster 1) have an average height of 177.8513, the forward players (Cluster 2) have an average height of 188.0854, and the center players (Cluster 3) have an average height of 197.9934.

TABLE V
FEATURE WEIGHTS IN EACH ITERATION

Iteration	Height	Weight	Age	BMI
Initialization	0.25	0.25	0.25	0.25
Iteration 1	≈1	2.1265e-27	1.6183e-44	2.9226e-43
Iteration 2	1.00	-	-	-

According to Table VI, the cluster of each data can be determined by choosing the highest value of membership value from Table VII. The clustering results of FRFCM is shown in Table VIII, where 14 players are on the guard position, five players are on the forward position, and four players are on the center position. As can be seen in Table VIII, there is just one player with incorrect position (see the bold font), i.e., M. Yassir Alkatiri. He tends to be tall but FRFCM put him on the guard position, so he will be more suitable to play on the forward position, because he has a tall and big body to break into the opponent's ring. Therefore, the ER of the FRFCM clustering result is 0.0435.

TABLE VI
MEMBERSHIP VALUES OF FRFCM CLUSTERING RESULT

Player's name	Cluster 1	Cluster 2	Cluster 3
Anjas Rusadi Putra	0.0245	0.9191	0.0565
Antoni Erga	0.9985	0.0012	0.0003
Ardian Ariadi	0.9692	0.0253	0.0055
Aldi Falentino	0.7782	0.1561	0.0657
Alexander Franklyn	0.3577	0.5864	0.0559
Bryan Adha Elang	0.0068	0.0329	0.9604
David Liberty Nuban	0.0245	0.9191	0.0565
Elyakim Tampa'i	0.9047	0.0709	0.0243
Febrianus Gregory	0.8891	0.0933	0.0176
Fransiscus Bryan	0.5583	0.3919	0.0498
Henry Cornelis Lakay	0.0068	0.0329	0.9604
Raymond Putra Fajar	0.0196	0.1114	0.869
Randy Ady Prasetya	0.0332	0.0941	0.8727
Mas Kahono Alif	0.0528	0.6323	0.3149
Rian Sanjaya	0.9943	0.0045	0.0012
Janson Kurniawan	0.9943	0.0045	0.0012
M. Yassir Alkatiri	0.3577	0.5864	0.0559
Martin	0.9985	0.0012	0.0003
Steven Ray	0.9943	0.0045	0.0012
Jody Sebastian	0.0541	0.138	0.808
Peter Surjantoro	0.7782	0.1561	0.0657
Fauji	0.0772	0.8913	0.0315
Ridho Pamungkas	0.0066	0.9835	0.0099

TABLE VII
CLUSTER CENTER OF FRFCM CLUSTERING RESULT

Cluster	Height
Cluster 1	177.8513
Cluster 2	188.0854
Cluster 3	197.9934

TABLE VIII
FRFCM CLUSTERING RESULTS

Player's name	Clustering results		Real position
	Cluster	Position	
Anjas Rusadi Putra	2	Forward	Forward
Antoni Erga	1	Guard	Guard
Ardian Ariadi	1	Guard	Guard
Aldi Falentino	1	Guard	Guard
Alexander Franklyn	1	Guard	Guard
Bryan Adha Elang	3	Center	Center
David Liberty Nuban	2	Forward	Forward
Elyakim Tampa'i	1	Guard	Guard
Febrianus Gregory	1	Guard	Guard
Fransiscus Bryan	1	Guard	Guard
Henry Cornelis Lakay	3	Center	Center
Raymond Putra Fajar	3	Center	Center
Randy Ady Prasetya	3	Center	Center
Mas Kahono Alif	2	Forward	Forward
Rian Sanjaya	1	Guard	Guard
Janson Kurniawan	1	Guard	Guard
M. Yassir Alkatiri	**1**	**Guard**	**Forward**
Martin	1	Guard	Guard
Steven Ray	1	Guard	Guard
Jody Sebastian	3	Center	Center
Peter Surjantoro	1	Guard	Guard
Fauji	2	Forward	Forward
Ridho Pamungkas	2	Forward	Forward

As a comparison, the result of Erga and Nataliani [19], using the combination of height and BMI, the ER is also 0.0435. As mentioned before, they combined four physical conditions, i.e., height, weight, age, and BMI, one by one to get the best clustering result. Since there are four features, there are 15 combinations in selecting the features, where the combinations consist of one feature only, two features, three features, and all four features. The advantage of FRFCM is that the algorithm can automatically find the related feature(s) during the clustering process by updating the feature weight and discarding the unrelated feature(s).

IV. CONCLUSION

The FRFCM clustering algorithm can group the basketball players' positions. FRFCM is done by weighting each feature with feature weight and discards the feature(s) with a small feature weight. The weighting and discarding processes are included in the clustering process simultaneously. There are four features of the players' physical condition, i.e., height, weight, age, and BMI. FRFCM finds the height feature as the most related physical condition to determine the players' position, especially for Satya Wacana Saints team. By comparing the clustering result with the actual position, FRFCM obtains only one incorrect position, such that the ER is 0.0435. Different from the previous research conducted by Erga and Nataliani [19], where height and BMI are the most important features, they found the related feature(s) by combining the related feature(s) one by one. This method can help the coach to decide the basketball player's position. For future works, FRFCM can be implemented on high-dimensional data related to more complex players' features. The features used are not limited to the physical conditions, but other measurements, for example, jumping ability, shooting score, rebound skill, can also be used to determine the player's position.

REFERENCES

[1] J. C. Bezdek, *A Primer on Cluster Analysis: 4 Basic Methods that (usually) Work*, 1st ed. Sarasota: First Edition Design Publishing, 2017.

[2] H. R. Vignesh, *Fuzzy C-mean Clustering using Data Mining*, Kindle. Munich: BookRix, 2019.

[3] K. Maheshwari and V. Sharma, "Optimization of Fuzzy C-Means Algorithm Using Feature Selection Strategies," *Adv. Intell. Syst. Comput.*, vol. 672, pp. 368–379, 2018.

[4] J. Cai, J. Luo, S. Wang, and S. Yang, "Feature Selection in Machine Learning: A New Perspective," *Neurocomputing*, vol. 300, pp. 70–79, Jul. 2018, doi: 10.1016/j.neucom.2017.11.077.

[5] E. Hancer, B. Xue, and M. Zhang, "A Survey on Feature Selection Approaches for Clustering," *Artif. Intell. Rev.*, vol. 53, no. 6, pp. 4519–4545, 2020, doi: 10.1007/s10462-019-09800-w.

[6] K. Mrudula, R. E. Keshava, and S. T. Hitendra, *Improvements over Fuzzy Clustering Methods for Large Datasets*, 1st ed. Republic of Moldova: LAP LAMBERT Academic Publishing, 2021.

[7] M. Dousthagh, M. Nazari, A. Mosavi, S. Shamshirband, and A. T. Chronopoulos, "Feature Weighting using a Clustering Approach," *Int. J. Model. Optim.*, vol. 9, no. 2, pp. 2–6, 2019, doi: 10.7763/IJMO.2019.V9.686.

[8] M. Hashemzadeh, A. Golzari Oskouei, and N. Farajzadeh, "New Fuzzy C-Means Clustering Method based on Feature-Weight and Cluster-Weight Learning," *Appl. Soft Comput. J.*, vol. 78, pp. 324–345, 2019, doi: 10.1016/j.asoc.2019.02.038.

[9] M. S. Yang and Y. Nataliani, "A Feature-Reduction Fuzzy Clustering Algorithm Based on Feature-Weighted Entropy," *IEEE Trans. Fuzzy Syst.*, vol. 26, no. 2, pp. 817–835, 2018, doi: 10.1109/TFUZZ.2017.2692203.

[10] M. S. Yang and K. P. Sinaga, "A Feature-Reduction Multi-view k-Means Clustering Algorithm," *IEEE Access*, vol. 7, pp. 114472–114486, 2019, doi: 10.1109/ACCESS.2019.2934179.

[11] M. S. Yang and J. B. M. Benjamin, "Feature-Weighted Possibilistic c-Means Clustering with a Feature-Reduction Framework," *IEEE Trans. Fuzzy Syst.*, vol. 29, no. 5, pp. 1093–1106, 2021, doi: 10.1109/TFUZZ.2020.2968879.

[12] A. Jamal, A. Handayani, A. A. Septiandri, E. Ripmiatin, and Y. Effendi, "Dimensionality Reduction using PCA and K-Means Clustering for Breast Cancer Prediction," *Lontar Komput. J. Ilm. Teknol. Inf.*, vol. 9, no. 3, pp. 192–201, 2018, doi: 10.24843/lkjiti.2018.v09.i03.p08.
[13] I. B. League, "IBL Team." https://iblindonesia.com/profile/team.
[14] K. Gryko, P. Stastny, A. Kopiczko, K. Mikołajec, O. Pecha, and K. Perkowski, "Can Anthropometric Variables and Maturation Predict The Playing Position in Youth Basketball Players?," *J. Hum. Kinet.*, vol. 69, no. 1, pp. 109–123, 2019, doi: 10.2478/hukin-2019-0005.
[15] I. Zaric, M. Dopsaj, M. Markovic, M. Zaric, S. Jakovljevic, and D. Beric, "Body composition characteristics measured by multichannel bioimpedance in young female basketball players: Relation with match performance," *Int. J. Morphol.*, vol. 38, no. 2, pp. 328–335, 2020, doi: 10.4067/S0717-95022020000200328.
[16] S. Liang and Y. Li, "Using Camshift and Kalman Algorithm to Trajectory Characteristic Matching of Basketball Players," *Complexity*, 2021, doi: 10.1155/2021/4728814.
[17] J. Pino-Ortega, C. D. Gómez-Carmona, F. Y. Nakamura, and D. Rojas-Valverde, "Setting Kinematic Parameters That Explain Youth Basketball Behavior: Influence of Relative Age Effect According to Playing Position," *J. Strength Cond. Res.*, vol. February, 2020, [Online]. Available: https://europepmc.org/article/med/32084109.
[18] J. L. Russell, B. D. McLean, F. M. Impellizzeri, D. S. Strack, and A. J. Coutts, "Measuring Physical Demands in Basketball: An Explorative Systematic Review of Practices," *Sport. Med.*, vol. 51, no. 1, pp. 81–112, 2021, doi: 10.1007/s40279-020-01375-9.
[19] A. Erga and Y. Nataliani, "Seleksi Fitur pada Pengelompokan Posisi Pemain Basket menggunakan Fuzzy C-Means," *JOINTECS (Journal Inf. Technol. Comput. Sci.*, vol. 6, no. 2, pp. 77–84, 2021, doi: 10.31328/jointecs.v6i2.2346.
[20] Š. Brodinová, P. Filzmoser, T. Ortner, C. Breiteneder, and M. Rohm, "Robust and Sparse k-Means Clustering for High-Dimensional Data," *Adv. Data Anal. Classif.*, vol. 13, no. 4, pp. 905–932, 2019, doi: 10.1007/s11634-019-00356-9.
[21] A. Saha and S. Das, "Clustering of fuzzy data and simultaneous feature selection: A model selection approach," *Fuzzy Sets Syst.*, vol. 340, pp. 1–37, 2018, doi: 10.1016/j.fss.2017.11.015.
[22] N. Van Pham, L. T. Pham, WitoldPedrycz, and L. T. Ngo, "Feature-Reduction Fuzzy Co-Clustering Approach for Hyper-Spectral Image Analysis," *Knowledge-Based Syst.*, vol. 216, p. 106549, 2021.
[23] J. Pion, V. Segers, J. Stautemas, J. Boone, M. Lenoir, and J. G. Bourgois, "Position-Specific Performance Profiles, Using Predictive Classification Models in Senior Basketball," *Int. J. Sport. Sci. Coach.*, vol. 13, no. 6, pp. 1072–1080, 2018, doi: 10.1177/1747954118765054.
[24] S. Zhang, A. Lorenzo, M. A. Gómez, N. Mateus, B. Gonçalves, and J. Sampaio, "Clustering performances in the NBA according to players' anthropometric attributes and playing experience," *J. Sports Sci.*, vol. 36, no. 22, pp. 2511–2520, 2018, doi: 10.1080/02640414.2018.1466493.