Fairness constraint of Fuzzy C-means Clustering improves clustering fairness

Xia Xu hsuhsia0620@gmail.com
Hui Zhang* zhanghui@swust.edu.cn
Chunming Yang yangchunming@swust.edu.cn
Xujian Zhao jasonzhaoxj@gmail.com
Bo Li libo.swust@qq.com
Southwest University of Science and Technology, Mianyang, China

Editors: Vineeth N Balasubramanian and Ivor Tsang

Abstract

Fuzzy C-Means (FCM) clustering is a classic clustering algorithm, which is widely used in the real world. Despite the distinct advantages of FCM algorithm, whether the usage of fairness constraint in the FCM could improve clustering fairness remains fully elusive. By introducing a novel fair loss term into the objective function, a Fair Fuzzy C-Means (FFCM) algorithm was proposed in this current study. We proved that the membership value was constrained by distance and fairness in the meantime during the optimization process in the proposed objective function. By studying the Fuzzy C-Means Clustering with fairness constraint problem and proposing a fair fuzzy C-means method, this study provided mechanism understanding in achieving the fairness constraint in Fuzzy C-Means clustering and bridged up the gap of fair fuzzy clustering.

Keywords: fair clustering, algorithm fairness, Fuzzy C-Means.

1. Introduction

Clustering is a classical unsupervised learning task that seeks to group data objects into different clusters, so as to maximize intra-cluster similarity and minimize inter-cluster similarity. Currently, several types of clustering algorithms are available, e.g., the partition-, density-, hierarchical-, spectral- and fuzzy-clustering. Since real-world problems are often fuzzy, fuzzy clustering is becoming more and more popular because of its advantage in dealing with problems using fuzzy mathematics. Instead of grouping objects into a certain cluster, fuzzy clustering partition objects non-uniquely (fuzzy), so that an object can belong to multiple clusters with different membership in the range of 0 to 1. As one of the most representative fuzzy clustering algorithms, Fuzzy C-Means (FCM) clustering Dunn (1973); Bezdek (2013) possesses some outstanding advantages, e.g., 1) simple structure and thus easily being programming; 2) much more approaching real-world problems; 3)Its objective function optimization is supported by nonlinear programming theory. FCM is widely utilized in pattern recognition Chuang et al. (1999), data mining Iyer et al. (2000), classification Hirota and Pedrycz (1999) and image segmentation Rezaee et al. (2000), among others.

* Corresponding author

Recently, the importance of designing fair algorithms has been caught by the machine learning community. Traditional machine learning algorithms failed to take the bias (against certain attributes, this paper refers to these attributes as sensitive attributes, and the value of sensitive attribute is called group.) into account, therefore their output may contain or even augment the bias. To date, considerable amounts of previous efforts have been proposed by the machine learning community to address the fair clustering task Chierichetti et al. (2017); Rösner and Schmidt (2018); Kleindessner et al. (2019b); Chen et al. (2019); Bera et al. (2019); Sam Abraham et al. (2020); Micha and Shah (2020). However, these works are mostly focused on center- or spectral- based methods. Despite the distinct advantages of the FCM algorithm, whether the usage of fairness constraint in the FCM could improve clustering fairness remains fully elusive. In this study, we questioned how the utilized of fairness constraint to enhance the fairness of Fuzzy C-Means Clustering.

To answer the above question, a fair Fuzzy C-Means algorithm was proposed in this paper. Considering that FCM was a fuzzy clustering based on the objective function, we innovatively introduced a fair loss term into the objective function, and optimized the objective function to obtain a fair result. By setting the weight of the fair loss term, our method achieved a trade-off between clustering quality and fairness. The fair loss function was defined as the square of the difference between the ratio of groups in clusters and the ratio of groups in the original dataset. To the best of our knowledge, this was the first result that introduced fair loss fuction into Fuzzy C-Means clustering.

2. Related Work

In recent years, the research community has done a lot of work in providing fairness guarantee for machine learning algorithms or studying fair variants of clustering algorithms. Plenty of relevant formulations of fairness have been proposed for supervised learning and specifically for classification tasks Dwork et al. (2012); Hardt et al. (2016); Kleinberg et al. (2017); Zafar et al. (2017). In this study, we emphasized the importance of fair clustering. Two work lines on fair clustering were available, i.e., the fellow-up work proposed by Chierichetti et al. (2017) and some independent works.

The most recognized notion in the field of fair clustering was proposed by Chierichetti et al. (2017). They modeled fairness based on disparate impact doctrine Feldman et al. (2015), which posited that any "sensitive attribute" must have approximately equal representation in the decision taken (through algorithms). To achieve fair clustering, they introduced the fairlet decomposition, which partitioned data objects into small and balance subsets (e.g., the fairlets). These subsets were clustered to get a fair result. Due to their method only worked in the scenery of two groups and time-consuming, the Chierichetti method was not so successful but he actually opened up ideas for further research, especially since he first defined a normal notion for fair clustering. Later, several follow-up works extended this idea. Schmidt et al. (2020) proposed a fast fairlets decomposition method. They defined a notion called "coreset", which was a representative subset of the original dataset. By solving fair clustering problem on the coreset, an approximate solution for the original dataset was provided. Backurs et al. (2019) suggested embedding the input data into Hierarchically well-Separated Tree to accelerate fairlet decomposition. Both methods could speed up the fairlet decomposition, however, the fairness notion used in both methods was proposed by Chierichetti et al. (2017), and hence only worked for two groups. Rösner and Schmidt (2018) considered a multiple groups variant of fair k-center clustering and developed a k-center constant factor approximation approach. Bercea et al. (2019) improved the result of Rösner and Schmidt (2018) and gave a bicriteria constant factor approximation algorithm for several classical clustering objectives. Kleindessner et al. (2019b) introduced fairness to spectral clustering by viewing the fairness notion as a constrained variant of spectral clustering. This method only worked for single sensitive attribute. Bera et al. (2019) extended the fairness notion as the upper and lower bound of the representation of a group in a cluster. In addition, they gave a general fair clustering framework. The framework consisted of two steps. Firstly, vanilla clustering was used to generate cluster centers, and then the linear programming method was used to fairly assign data objects to the cluster centers. Nevertheless, this framework was only suitable for K-center and K-median clustering. Ahmadian et al. (2019) explored the problem of preventing excessive representation of groups in each cluster, and gave an algorithm based on linear programming. Ahmadian et al. (2020) recently studied two variants of fair clustering, the one was minimum divergence clustering, and the other one was upper and lower bound clustering. Their algorithms work for multiple groups. Davidson and Ravi (2020) showed that for any clustering with two groups, linear programming could be used to calculate the most similar fair clustering. Quy et al. (2021) recently studied the fairness problem in the education domain and gave two fair method, namely hierarchical clustering and partitioning-based clustering.

Chen et al. (2019) considered the proportional centroid clustering problem and outlined an independent fairness notion. For clustering n points with k centers, any n/k points were entitled to form their own cluster if there is another center that was closer in distance for all n/k points. Kleindessner et al. (2019a) advised a simple k-center clustering algorithm with fairness constraint. In this method, a cluster was regarded as a summary of the original dataset, and fair summary was generated for each cluster. Jung et al. (2020) proposed the notion of individual fairness of clustering, which required each object somewhat to close a center, "somewhat" depended on the object's k nearest neighbors. Recently, Ghadiri et al. (2021) and Abbasi et al. (2021) independently proposed the social fair clustering problem. Makarychev and Vakilian (2021) improved and generalized the $O(\ell)$-approximation algorithms of the social fair clustering problem in Ghadiri et al. (2021) and Abbasi et al. (2021). The works most relevant to this study were Ziko et al. (2019) and Sam Abraham et al. (2020). Ziko et al. (2019) incorporated fairness constraints into the clustering steps by adding a fairness loss term into the objective function. They defined the fair loss as the KL divergence between the probability distribution of sensitive attributes in clusters and the probability distribution in the dataset. However, this method was designed only for one sensitive attribute. Sam Abraham et al. (2020) also used a similar idea, but works for multiple sensitive attributes. Unfortunately, they did not consider introducing fairness into fuzzy clustering. Fuzzy clustering was widely used in the real world, it was imperative to ensure that these algorithms were fair. In this current study, the proposed algorithm provided fair guarantee for fuzzy c-means clustering. In addition, fair loss was defined as the square of the difference between the ratio of groups in clusters and the ratio of groups in the original dataset.

3. Preliminaries

In this section, vanilla FCM Bezdek (2013) is formally described and the fuzzy C-means with fairness constraints problem is defined in order to introduce terminology and set the ground for our works.

Let P be a collection with n objects embedded in metric space (X, d), where $P := \{p_1, p_2, \cdots, p_n\}$ and $d : P^2 \to R \geq 0$. Let k be the number of clusters, u_{ij} is the membership value of object j to cluster center c_i and u_{ij} is subject to $u_{j1} + u_{j2} + \cdots + u_{jk} = 1$, $u_{ij} \in [0, 1]$. The task of vanilla FCM is to find a fuzzy partition matrix $U := [u_{ij}]$ and a set of cluster centers $\mathcal{C} := \{c_1, c_2, \cdots, c_k\}$ to minimize the objective function $J_{fcm} := \sum\limits_{i=1}^{k} \sum\limits_{j=1}^{n} u_{ij}^m \|p_j - c_i\|^2$.

In fair clustering task, one is additionally, there are two sets of attributes C_a and S_a defined over datasets. C_a denote the attributes that are relevant to the clustering task interest, such as the patient's symptoms and living environment in the case of disease clustering. S_a denote sensitive attributes, which may be expected to maintain fairness in the generated clusters, such as gender, race, religion, nationality, etc. What's more, the values of sensitive attribute are called groups (as mentioned in the section 1). Assuming that gender is a sensitive attribute, it may have two groups, named male and female. The fairness notion utilized in this paper is proportional fairness, which maintains the same proportion of sensitive attribute groups in clusters as they are in the original dataset. In other words, there are two groups of a sensitive attribute (Considering gender) with a ratio of 7:3 in the original dataset, and their ratios are expected to be 7:3 (ideally) in each generated clusters. Fuzzy c-means with fairness constraints problem is formally defined below.

Definition 1 *(Fuzzy C-Means clustering with fairness constraint problem.) Given l groups $P_1, P_2, \cdots, P_l$ as the values of the sensitive attribute, and $P_1 \cup P_2 \cup \cdots \cup P_l = P$. The fair fuzzy clustering problem can be described as finding a partition of P so as to minimize object function $J_{fcm} := \sum\limits_{i=1}^{k} \sum\limits_{j=1}^{n} u_{ij}^m \|p_j - c_i\|^2$ and make the ratio of each group in clusters as close as possible to its ratio in original dataset.*

4. Proposed Approach

We propose a Fair Fuzzy C-Means (FFCM) clustering method to attack the Fuzzy C-Means clustering with fairness constraint problem. In FFCM, a novel fair loss function is constructed to quantify the fair loss in clusters and an optimization method is designed for the objective function to achieve fair clustering. Firstly, the construction of the fair loss function is detailed in the next part of this section. After that, the optimization method is explained, and the validity of the proposed objective function formula is proved theoretically. In the last part of this section, the complexity of the proposed algorithm is analyzed.

4.1. Fair Loss Term Construction

For an ideal fair clustering, the ratio of groups in the generated clusters is expected to be the same as in the original dataset. With the intent of generating clusters that are as fair as possible, a natural definition of fair loss function is the difference between the ratio of groups in clusters and the ratio of groups in the original dataset. The greater the ratio difference, the greater the loss function score will be. The mathematical expression for the fair loss of a group in a cluster is

$$bias(C_{is}) = \left(\frac{|C_i \cap P_s|}{|C_i|} - \frac{|P_s|}{|P|} \right)^2, \quad s \in [l] \tag{1}$$

where $bias(C_{is})$ is the fair loss of group P_s in cluster C_i. Naturally, the entire fair loss of cluster C_i is the sum of all the group's loss score in cluster C_i. The entire fair loss can be written as

$$bias(C_i) = \sum_{s=1}^{l} bias(C_{is}) \tag{2}$$

It should be noted that different sensitive attributes may have different numbers of groups (considering the sensitive attributes of gender and age. Gender may have two groups, male and female, while age may have multiple groups, such as infants, children, youth, middle-aged and old). It can be observed from Eq. 2 that sensitive attributes with a large number of groups may produce a large fair loss score. This means that the sensitive attributes with more groups might dominate the fair loss term. In order to make each sensitive attribute have the same contribution to the fair loss term, the fair loss term is normalized in Eq. 3 by the number of groups.

$$Nbias(C_i) = \sum_{\omega=1}^{\mathcal{A}} \frac{\sum_{s=1}^{l_\omega} \left(\frac{|C_i \cap P_s|}{|C_i|} - \frac{|P_s|}{|P|} \right)^2}{|l_\omega|} \tag{3}$$

Where $\mathcal{A}$ is the number of sensitive attributes, and l_ω is the number of groups in the ω sensitive attribute.

Since fair loss must be related with membership value, the larger the distance from the object to the cluster center plus the fair loss score, the smaller the membership value of the object is expected to be. Thus, the fair loss is decomposed into each membership value. For an object p_j, it is assigned to clusters C_i, i from 1 to k, and the assignment of other objects is retained. Let $\underset{p_j \to C_1}{bias}, \underset{p_j \to C_2}{bias}, \cdots, \underset{p_j \to C_k}{bias}$ be the normalized fair loss score, which is calculated by Eq. 3.. These fair loss scores are introduced into the objective function so that the fair loss and distance jointly determine the membership value. The overall objective function is shown in Eq. 4

$$\begin{aligned} J_{ffcm} \;\; &= \;\; \sum_{i=1}^{k} \sum_{j=1}^{n} u_{ij}^m \left(\|p_j - c_i\|^2 + \eta \underset{p_j \to C_i}{bias} \right), \\ &s.\,t. \quad \sum_{i=1}^{k} u_{ij} = 1 \end{aligned} \tag{4}$$

where η is a hyperparameter, which denotes the weight of fair loss term.

4.2. Objective Function Optimization and FFCM Algorithm

After the fair loss is introduced into the objective function of FFCM, the task of Fuzzy C-Means with fairness constraint is reduced to identifying a partition that minimizes J_{ffcm} as much as possible. As in the vanilla fuzzy c-means, the Lagrange Multiplier Method is used to find the minimum value of the objective function. The Lagrange function of FFCM is constructed as

$$\mathcal{L} = \sum_{i=1}^{k} \sum_{j=1}^{n} {u_{ij}}^{m} \left(\|p_j - c_i\|^2 + \eta \underset{p_j \to C_i}{bias} \right) - \sum_{j=1}^{n} \lambda_j \left(\sum_{i=1}^{k} u_{ij} - 1 \right) \tag{5}$$

where the first part in Eq. 5 is the objective function of FFCM, and the second part is the inherent constraint of membership value (i.e., the sum of the membership values of an object to all clusters is equal to 1).

By calculating the partial derivative of the Lagrange function with respect to u_{ij}, the expression of membership value u_{ij} is obtained as

$$u_{ij} = \frac{\frac{1}{\left(\|p_j - c_i\|^2 + \eta \underset{p_j \to C_i}{bias} \right)^{\frac{1}{m-1}}}}{\frac{1}{\sum_{\tau=1}^{k} \left(\|p_j - c_\tau\|^2 + \eta \underset{p_j \to C_\tau}{bias} \right)^{\frac{1}{m-1}}}} \tag{6}$$

Theorem 2 *The membership value is constrained by distance and fair loss term at the same time during the optimization process.*

Proof For any object p_j and cluster center C_i, when p_j is assigned to C_i, the fair loss of this assignment is $\underset{p_j \to C_i}{bias} = \sum_{\omega=1}^{\mathcal{A}} \sum_{s=1}^{l_\omega} (|C_i \cap P_s|/|C_i| - |P_s|/|P|)/|l_\omega|$. Let d_{ij} be the distance from p_j to the cluster center c_i, Eq. 6 can be reduced to

$$u_{ij} = \frac{\frac{1}{\left(d_{ij} + \eta \underset{p_j \to C_i}{bias} \right)^{\frac{1}{m-1}}}}{\frac{1}{\sum_{\tau=1}^{k} \left(d_{i\tau} + \eta \underset{p_j \to C_\tau}{bias} \right)^{\frac{1}{m-1}}}} \tag{7}$$

We guarantee that $\sum_{\tau=1}^{k} \left(d_{\tau j} + \eta \underset{p_j \to C_\tau}{bias} \right)$ is a constant. Observing Eq. 7, it is easy to find that when the fair loss $\underset{p_j \to C_i}{bias}$ is large, $1 \Big/ \left(d_{ij} + \underset{p_j \to C_i}{bias} \right)$ will be small. Therefore, the membership value u_{ij} will also be small. Conversely, when $\underset{p_j \to C_i}{bias}$ is small, $1 \Big/ \left(d_{ij} + \underset{p_j \to C_i}{bias} \right)$ will be large, and correspondingly u_{ij} will also be large. ■

Algorithm 1 Fair Fuzz C-means Clustering.

Input: Collection P with n object, the number of clusters k, the end condition ε, the maximum number of iterations $\mathcal{T}$ and the weight of fair loss η.

Output: The labels of objects.

```
Initialize the membership value u_ij;
while |J_ffcm^(t) − J_ffcm^(t−1)| > ε and t < T do
    Update the cluster centers c_i by Eq. 9;
    for p_j ∈ P, j = 1 to n do
        the object p_j is assigned to clusters C_i, i from 1 to k,
        Calculate the fair loss term by Eq. 3;
        Update the membership value u_ij by Eq. 6;
    end for
    Calculate the value of the objective function J_ffcm by Eq. 4;
end while
return The labels of each object.
```

Lemma 3 *For any object p_j, the* $\sum_{\tau=1}^{k}\left(d_{\tau j}+\eta \underset{p_j\to C_\tau}{bias}\right)$ *is a constant.*

Proof If p_j is assigned to clusters $C_1, C_2, \cdots, C_k$, the fair loss of these assignments are $\underset{p_j\to C_1}{bias}, \underset{p_j\to C_2}{bias}, \cdots, \underset{p_j\to C_k}{bias}$. Then we have:

$$\sum_{\tau=1}^{k}\left(d_{\tau j}+\eta \underset{p_j\to C_\tau}{bias}\right) = (d_{1j}+d_{2j}+\cdots+d_{kj})+\eta\left(\underset{p_j\to C_1}{bias}+\underset{p_j\to C_2}{bias}+\cdots+\underset{p_j\to C_k}{bias}\right) \tag{8}$$

Because $d_{\tau j}$ and $\underset{p_j\to C_\tau}{bias}, \forall\tau\in[k]$ are all constants, their sum $\sum_{\tau=1}^{k}\left(d_{\tau j}+\eta\left(\underset{p_j\to C_\tau}{bias}\right)\right)$ is also a constant. ■

Similar to the membership value u_{ij}, by calculating the partial derivative of the Lagrange function with respect to c_i, the expression of the cluster center c_i is obtained as

$$c_i = \frac{\sum_{j=1}^{n} u_{ij} p_j}{\sum_{j=1}^{n} u_{ij}} \tag{9}$$

The FFCM algorithm is summarized in Algorithm 1. Algorithm 1 resembles the working of vanilla FCM except a fair loss is considered at each step that allows it to be fairer than vanilla FCM. Specifically, the membership value is initialized first, and proceeds iteratively. In each iteration, three steps are performed. First, update the cluster center through the membership value. Then, we traverse each object p in round-robin fashion, the object is assigned to each cluster respectively and the fair loss score of the present assignment is calculated. Finally, the membership value is updated based on the fair loss score and distance. The updating process will stop until the objective function converges or the maximum iteration threshold is reached.

4.3. Complexity Analysis

The complexity of FFCM is determined by fair loss calculation, membership value update and cluster center update. First, considering the complexity of the fair loss calculation. For a data object p, when it is assigned to each cluster, the ratio of groups needs to be re-stated. In each cluster, the fair loss needs to be calculated for $\mathcal{A}$ sensitive attributes and l groups of each sensitive attribute. The fair loss calculation method is shown in Eq. 1, which can be regarded as a simple calculation completed in a constant time. Therefore, the complexity of fair loss calculation is $\mathcal{O}\left(nk^2|\mathcal{A}|\,|l| + k\,|\mathcal{A}|\,|l|\right)$. Next, we consider the complexity of membership value update and cluster centers update. According to Eq. 5, the complexity of membership value update is $\mathcal{O}(nk^2)$. And according to Eq. 6, the complexity of cluster centers update is $\mathcal{O}(nk)$. In summary, the complexity of t iterations is $\mathcal{O}(tn^3k^4\,|\mathcal{A}|\,|l| + nk^3\,|\mathcal{A}|\,|l| + nk^2)$.

5. Experiments

In this section, we performed empirical evaluations of the proposed algorithm. Firstly, we outlined the datasets and settings in the experiment. Then, we illustrated the measurements. Finally, we reported the experimental results and the analysis of experimental results.

5.1. Datasets and Settings

We considered six real-world datasets, which were popular in the fair clustering task. (1) Diabetic Chierichetti et al. (2017) recorded information related to patients with diabetes. (2) Census1990 Meek et al. (2002) consisted of the 1990 U.S. census records. (3) Creditcard Yeh and hui Lien (2009) contained information about the card holders from a certain credit card in Taiwan. (4) Bank Moro et al. (2014) contained instances related to the direct telemarketing activities of Portuguese banking institutions. (5) Adult Zhou and Chen (2002) was also known as census, which contained examples of the 1994 U.S. Census. (6) Athlete Böhm et al. (2020) contained bio data on Olympic athletes and medal results from Athens 1896 to Rio 2016. We subsampled all datasets to 1000 records, except Creditcard. Creditcard was subsampled to 600 records. For each of these datasets, we chose numerical attributes to represent points in Euclidean space. In addition, we also set sensitive attribute for each dataset and created groups based on their values. More information about the dataset settings was shown in Table 1.

In the proposed method, the weight of the fair loss term η needed to be set. Based on empirical observation, we set η to 10^2 for Diabetic and Census1990, 10^3 for Athlete, 10^5 for Bank, 10^7 for Creditcard and Adult. We analyzed the sensitivity of η in section 5.4. The maximum number of iterations $\mathcal{T}$ was set to 10. Our codes are available on GitHub for public use.[1]

1. https://github.com/author's-name/author's-name-FuzzyC-MeansWithFairnessConstraint

Table 1: Clustering attributes and sensitive attributes.

Datasets	Clustering attributes	Sensitive attributes	Groups
Diabetic	age,time-in-hosp, num-medications,num-outpatient, num-emergency,num-inpatient	gender	male, female, unknown
Creditcard	age, bill-amt 1—6, limit-bal, pay-amt 1—6	education	7 groups
Bank	age, balance, duration	marital	married, single, divorced
Adult	age, education-num, final-weight, capital-gain, hours-per-week	race	asian-pac-isl, Amer-ind, white, black, other
Census1990	dAncstry1, dAncstry2, iAvail, iCitizen, iClass, dDepart,iFertil, iDisabl1, iDisabl2, iEnglish, iFeb55, dHispanic, dHour89	iSex	male, female
Athlete	Age, Height, Weight, Year	Season	summer, winter

5.2. Measurements

In fair clustering, two sets of metrics were mainly concerned. One was clustering quality, which was related to clustering interest (i.e., C_a, as described in Section 3). The other was fairness, which was related to fair interest (S_a). In this subsection, we described these metrics.

Clustering quality measured the rationality of clustering, it could include: (1) Clustering Cost (**Cost**), it measured the deviation of objects from cluster centers. The smaller the clustering cost, the better the clustering result. One of the clustering cost presentation was objective function, and the objective function of FCM Bezdek (2013) was

$$J_{fcm} := \sum_{i=1}^{k} \sum_{j=1}^{n} u_{ij}^{m} \|p_j - c_i\|^2 \tag{10}$$

(2) Silhouette Coefficient (**SC**) Rousseeuw (1987), it described the cohesion and separation of clusters. It lies in the range of $[-1, 1]$, the closer the value to 1, the better the clustering result and vice versa.

Fairness measured the fairness of attributes in the generated output, it could include: (1) **balance** Bera et al. (2019), which described the lowest level fairness of groups in a cluster. It was defined as: Let $f(P_s) = |P_s|/|P|$ be the ratio of group s in the entire dataset, and let $f(C_{is}) = |P_s \cap C_i|/|C_i|$ be the ratio of group s in cluster i. The $fairness(C_i) = min(f(C_{is})/f(P_s),\ f(P_s)/f(C_{is}))$, $i \in [k], s \in [l]$ was the balance in cluster i. (2) Euclidean distance of distribution vectors (**Ed**) Sam Abraham et al. (2020), which measured the unfairness of clustering. Let S be a sensitive attribute, which could take on l groups. The distribution of these groups in the dataset P produced a l-length distribution vector P_s. Similarly, the distribution of these groups in each cluster yielded a l-length distribution vector C_s. By calculating the Euclidean distance between the representations P_s and C_s to get the quantized intra-cluster unfairness. (3) Wasserstein distance of distribution vectors (**Wd**) Wang and Davidson (2019), The definition of Wd metric was similar to Ed, except that Euclidean distance was replaced by Wasserstein distance.

Table 2: Fairness comparison – FFCM vs. FCM

Cluster number	Datasets	Ed			Wd		
		FFCM	FCM	impr(%)	FFCM	FCM	impr(%)
k = 4	Diabetic	**0.4283**	0.7745	**44.6963**	**0.2045**	0.3476	**41.1644**
	Census1990	**0.1452**	0.3980	**63.5148**	**0.0884**	0.2370	**62.6912**
	Creditcard	**0.3105**	0.3789	**18.0532**	**0.1204**	0.1424	**15.4326**
	Bank	**0.2623**	0.6007	**56.3373**	**0.1402**	0.2689	**47.8452**
	Adult	**0.2345**	0.3270	**28.2810**	**0.0774**	0.1016	**23.8213**
	Athlete	**0.3258**	0.5606	**41.8921**	**0.2303**	0.3964	**41.8921**
k = 6	Diabetic	**0.7565**	1.5480	**51.1299**	**0.4071**	0.8046	**49.4120**
	Census1990	**0.2088**	0.6472	**67.7311**	**0.1252**	0.3576	**64.9790**
	Creditcard	**0.4000**	0.5026	**20.4046**	**0.1585**	0.1925	**17.6503**
	Bank	**0.3599**	0.5638	**36.1701**	**0.1490**	0.1968	**24.2674**
	Adult	**0.3874**	0.5775	**32.9252**	**0.1332**	0.1804	**26.1849**
	Athlete	**0.4575**	0.8619	**46.9213**	**0.3235**	0.6094	**46.9213**
k = 8	Diabetic	**0.6330**	1.6767	**62.2456**	**0.4283**	1.0256	**58.2363**
	Census1990	**0.5643**	1.8573	**69.6189**	**0.3006**	0.8933	**66.3466**
	Creditcard	**0.6782**	0.9104	**25.5040**	**0.2728**	0.3540	**22.9374**
	Bank	**0.3868**	0.8890	**56.4952**	**0.1984**	0.3719	**46.6364**
	Adult	**0.5010**	0.7830	**36.0159**	**0.1618**	0.2444	**33.8049**
	Athlete	**0.5832**	1.1865	**50.8454**	**0.4124**	0.8390	**50.8454**

5.3. Clustering Results Comparison with FCM

Table 2 displayed the fairness metrics of FFCM and FCM, and the best of all results were highlighted in boldface. The *impr* column represented the fairness improvement percentage by FFCM. As could be seen from Table 2, the performance of FFCM in fairness metrics significantly surpassed FCM. Specifically, FFCM performed the best on the Census1990, and the percentage of fairness improvement was about 65%, while it performed slightly poorly on the Creditcard, and the percentage of fairness improvement was still more than 20%. In addition, it was worth noting that when $k = 6$, the overall performance of FFCM was better than when $k = 4$. Whereas $k = 8$, the performance of FFCM was stronger. This indicated that FFCM benefitted from higher flexibility (with higher k) in the process of reducing fairness loss. Figure 1 presented the balance of FFCM and FCM in each cluster, and the balance of different methods was marked by different colors. Since the goal of the FFCM objective function was to minimize the overall fairness loss, and no specific constraint to ensure good performance on the lowest level fairness (i.e., *balance*). Therefore, in order to obtain a lower fairness loss, FFCM might sacrifice fairness in one or several clusters (the cluster 1 of Adult in Figure 1). Although the balance metric of FFCM implied that such trends were not widely prevalent, it was a direction to meliorate FFCM. Table 3 presented the results of clustering quality, and the *decr* column indicates the clustering

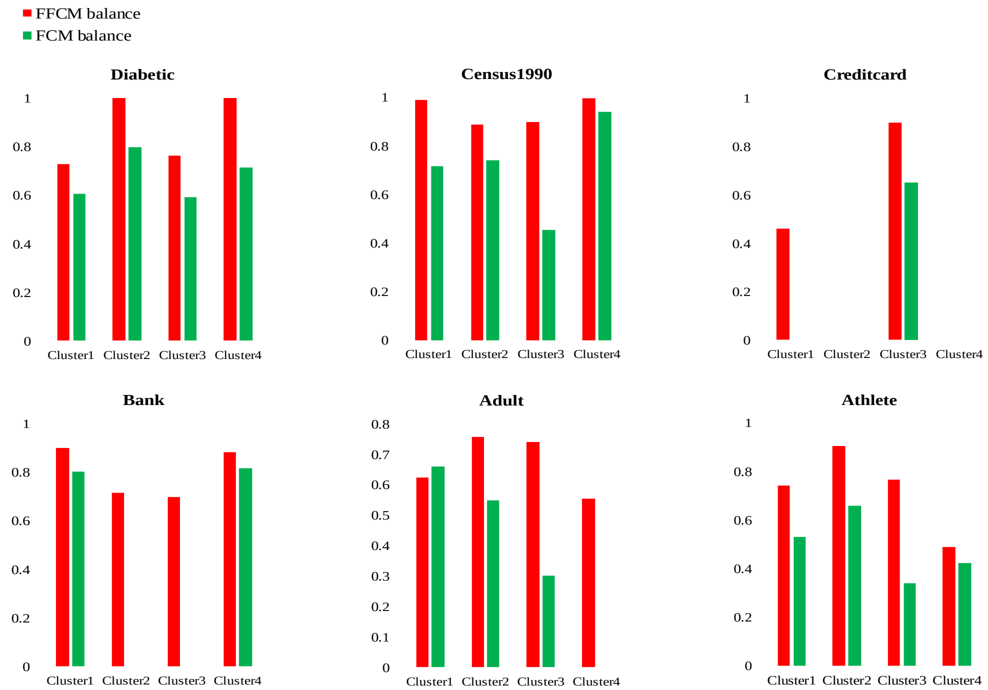


Figure 1: The balance of FFCM and FCM in clusters (k=4)

Table 3: Clustering quality comparison – FFCM vs. FCM

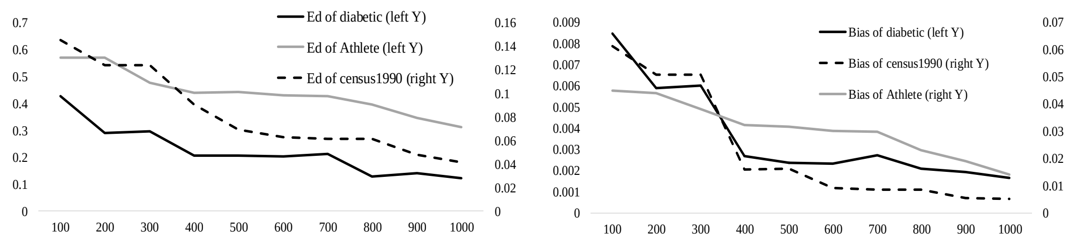
Clusters number	Datasets	SC			Cost		
		FFCM	FCM	decr(%)	FFCM	FCM	decr(%)
k = 4	Diabetic	0.2129	**0.3245**	34.3762	11530.20	**7619.95**	51.3159
	Census1990	0.3849	**0.4356**	11.6223	13198.64	**12101.05**	5.7647
	Creditcard	0.3356	**0.3681**	8.8241	60734399.00	**55858552.77**	8.7289
	Bank	0.3612	**0.6407**	43.6250	922940.18	**560447.64**	64.6791
	Adult	0.5434	**0.5591**	2.8085	17681552.15	**17178839.57**	2.9263
	Athlete	0.2169	**0.2663**	18.5702	23433.9640	**15485.7225**	8.4599
k = 6	Diabetic	0.1875	**0.3086**	39.2395	10684.31	**7263.53**	47.0952
	Census1990	0.2531	**0.2987**	15.2615	10990.44	**10951.39**	9.4878
	Creditcard	0.3302	**0.3868**	14.8517	52442935.66	**45661436.31**	14.8517
	Bank	0.4307	**0.5348**	19.4512	577817.18	**426868.76**	35.3618
	Adult	0.5234	**0.5581**	6.2172	12220439.55	**11276031.07**	8.3754
	Athlete	0.2206	**0.2724**	19.0199	14463.9656	**13232.7524**	9.3043
k = 8	Diabetic	0.1503	**0.3231**	53.4927	9235.05	**6018.09**	53.4549
	Census1990	0.3185	**0.3468**	13.9433	9953.08	**9746.08**	12.3845
	Creditcard	0.3022	**0.3617**	16.4626	43072356.90	**41601196.60**	15.5556
	Bank	0.2633	**0.4309**	38.8918	562538.63	**353313.56**	59.2180
	Adult	0.5005	**0.5418**	7.6275	10184113.67	**8337259.17**	22.1518
	Athlete	0.1028	**0.2739**	62.4609	14194.2831	**11915.2959**	19.1266

quality degradation percentage by FFCM. Due to FFCM needed to be held accountable for fairness, it was expected to perform poorly on clustering quality metric. We focused on the difference between the improvement of fairness and the degradation of clustering quality by FFCM. Table 4 displayed these results, the *avg.impr* row was the average improvement

Table 4: Comparison of fairness improvement and quality degradation by FFCM

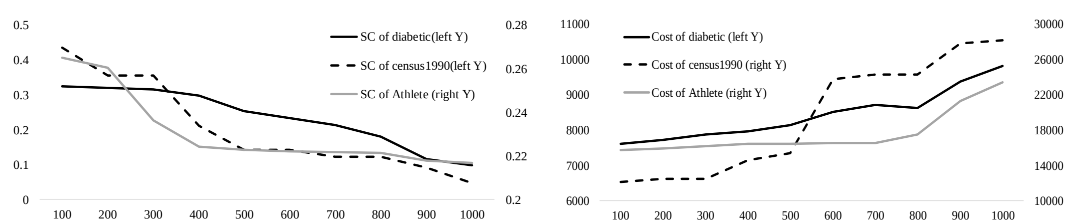
Datasets	Diabetic	Census1990	Creditcard	Bank	Adult	Athlete	agv.whole
avg.impr(%)	51.1474	65.8136	20.0000	44.6253	30.1722	46.5529	42.3511
avg.decr(%)	46.4958	17.1160	13.1772	43.5378	8.3511	22.8236	25.7356
margin(%)	4.6512	48.6975	6.8198	1.1087	21.8211	23.7293	16.616

percentage of Wd and Ed, and the *avg.decr* row was the average degradation percentage of Cost and SC. Their values were the average values when the number of clusters $k = 4, 6, 8$. It might be seen that compared to the degradation of clustering quality, FFCM had higher fairness margins. In general, the above results illustrated that FFCM provided fairer clustering. Despite FFCM reduced clustering quality, it brought fairer margin than the degradation of clustering quality.



(a) (b)

Figure 2: The change of fairness metrics with different η



(a) (b)

Figure 3: The change of clustering quality metrics with different η

5.4. Sensitivity Analysis to η

FFCM attempted to achieve a trade-off between clustering quality and fairness by changing the value of the only hyper-parameter η (i.e., the weight of fair loss function). The FFCM formula was expected that the greater the η, the smaller the fairness loss of clustering, and the better the performance on the fairness metrics. We observed such desired trends across all experimental datasets. Figures 2 and Figures 3 presented the changes of fairness metrics and clustering quality metrics on Diabetic, Census1990 and Athlete (Considering that the setting of η have the similar values) when η varied from 100 to 1000. For the metrics that widely varied in range, both sides of the y-axis were used to plot them, and the axis used for each metric was indicated in the figure. The Bias in Figure 2 was the whole fair loss of clustering. The fact could be obtained from Figures 2 and Figures 3 that with the increase of η, the clustering generated by FFCM tends to be fair (bias and Ed metric decrease gradually) and the clustering quality degrades accordingly (SC metric decreased and Cost metric increased). Although their change was not stable, the direction of change was in line with our expectations.

6. Conclusions And Future Directions

Although the fair clustering algorithm has attracted considerable attention, the fairness of fuzzy clustering has yet remained elusive. In this paper, we studied Fuzzy C-Means clustering with fairness constraint problem and proposed a fair fuzzy C-means method to bridge the gap of fair fuzzy clustering. By introducing a novel fairness loss term in the objective function, the membership value was constrained by fairness and distance at the same time during the optimization process, we gave theoretical proof for this constraint. We evaluated the performance of the proposed algorithm on real-world datasets, the empirical evaluation illustrated that FFCM significantly improved clustering fairness. We depicted two directions to enhance the performance of FFCM. Firstly, FFCM did not guarantee the lowest level of fairness, so it might sacrifice fairness in several clusters to minimize the fairness loss function. We considered adding a constraint in the loss function to prevent this phenomenon. Secondly, due to FFCM was a fair variant based on FCM, it was sensitive to the initial cluster centers and easily fell into the local optimal value. Future efforts should replace the vanilla FCM algorithm with a robust variant to improve the performance of FFCM.

References

Mohsen Abbasi, Aditya Bhaskara, and Suresh Venkatasubramanian. Fair clustering via equitable group representations. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, FAccT '21, page 504–514, New York, NY, USA, 2021. Association for Computing Machinery. ISBN 9781450383097. doi: 10.1145/3442188.3445913.

Sara Ahmadian, Alessandro Epasto, Ravi Kumar, and Mohammad Mahdian. Clustering without over-representation. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, KDD '19, page 267–275, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450362016. doi: 10.1145/3292500.3330987.

Sara Ahmadian, Alessandro Epasto, Ravi Kumar, and Mohammad Mahdian. Fair correlation clustering. In *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*, volume 108, pages 4195–4205. PMLR, 2020.

Arturs Backurs, Piotr Indyk, Krzysztof Onak, Baruch Schieber, Ali Vakilian, and Tal Wagner. Scalable fair clustering. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97, pages 405–413. PMLR, 2019.

Suman Bera, Deeparnab Chakrabarty, Nicolas Flores, and Maryam Negahbani. Fair algorithms for clustering. In *Advances in Neural Information Processing Systems*, volume 32, pages 4954–4965. Curran Associates, Inc., 2019.

Ioana O. Bercea, Martin Groß, Samir Khuller, Aounon Kumar, Clemens Rösner, Daniel R. Schmidt, and Melanie Schmidt. On the Cost of Essentially Fair Clusterings. In *Approximation, Randomization, and Combinatorial Optimization*, volume 145, pages 18:1–18:22, Dagstuhl, Germany, 2019. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. ISBN 978-3-95977-125-2. doi: 10.4230/LIPIcs.APPROX-RANDOM.2019.18.

James C Bezdek. *Pattern recognition with fuzzy objective function algorithms*. Springer Science & Business Media, 2013.

Matteo Böhm, Adriano Fazzone, Stefano Leonardi, and Chris Schwiegelshohn. Fair clustering with multiple colors. *arXiv preprint arXiv:2002.07892*, 2020.

Xingyu Chen, Brandon Fain, Liang Lyu, and Kamesh Munagala. Proportionally fair clustering. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97, pages 1032–1041. PMLR, 2019.

Flavio Chierichetti, Ravi Kumar, Silvio Lattanzi, and Sergei Vassilvitskii. Fair clustering through fairlets. NIPS'17, page 5036–5044, Red Hook, NY, USA, 2017. Curran Associates Inc. ISBN 9781510860964.

Kai-Hsiang Chuang, Ming-Jang Chiu, Chung-Chih Lin, and Jyh-Horng Chen. Model-free functional mri analysis using kohonen clustering neural network and fuzzy c-means. *IEEE Transactions on Medical Imaging*, 18(12):1117–1128, 1999. doi: 10.1109/42.819322.

Ian Davidson and SS Ravi. Making existing clusterings fairer: Algorithms, complexity results and insights. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 3733–3740, 2020. doi: 10.1609/aaai.v34i04.5783.

J. C. Dunn. A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters. *Journal of Cybernetics*, 3(3):32–57, 1973. doi: 10.1080/01969727308546046.

Cynthia Dwork, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Richard Zemel. Fairness through awareness. In *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference*, ITCS '12, page 214–226, New York, NY, USA, 2012. Association for Computing Machinery. ISBN 9781450311151. doi: 10.1145/2090236.2090255.

Michael Feldman, Sorelle A. Friedler, John Moeller, Carlos Scheidegger, and Suresh Venkatasubramanian. Certifying and removing disparate impact. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '15, page 259–268, New York, NY, USA, 2015. Association for Computing Machinery. ISBN 9781450336642. doi: 10.1145/2783258.2783311.

Mehrdad Ghadiri, Samira Samadi, and Santosh Vempala. Socially fair k-means clustering. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, FAccT '21, page 438–448, New York, NY, USA, 2021. Association for Computing Machinery. ISBN 9781450383097. doi: 10.1145/3442188.3445906.

Moritz Hardt, Eric Price, and Nati Srebro. Equality of opportunity in supervised learning. In *NIPS*, pages 3315–3323, 2016.

K. Hirota and W. Pedrycz. Fuzzy computing for data mining. *Proceedings of the IEEE*, 87 (9):1575–1600, 1999. doi: 10.1109/5.784240.

Naresh S. Iyer, Abraham Kandel, and Moti Schneider. Feature-based fuzzy classification for interpretation of mammograms. *Fuzzy Sets and Systems*, 114(2):271–280, 2000. ISSN 0165-0114. doi: https://doi.org/10.1016/S0165-0114(98)00175-4.

Christopher Jung, Sampath Kannan, and Neil Lutz. Service in Your Neighborhood: Fairness in Center Location. In *1st Symposium on Foundations of Responsible Computing (FORC 2020)*, volume 156, pages 5:1–5:15, Dagstuhl, Germany, 2020. Schloss Dagstuhl–Leibniz-Zentrum für Informatik. ISBN 978-3-95977-142-9. doi: 10.4230/LIPIcs.FORC.2020.5.

Jon Kleinberg, Sendhil Mullainathan, and Manish Raghavan. Inherent Trade-Offs in the Fair Determination of Risk Scores. In *8th Innovations in Theoretical Computer Science Conference (ITCS 2017)*, volume 67, pages 43:1–43:23, Dagstuhl, Germany, 2017. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. ISBN 978-3-95977-029-3. doi: 10.4230/LIPIcs.ITCS.2017.43.

Matthäus Kleindessner, Pranjal Awasthi, and Jamie Morgenstern. Fair k-center clustering for data summarization. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97, pages 3448–3457. PMLR, 2019a.

Matthäus Kleindessner, Samira Samadi, Pranjal Awasthi, and Jamie Morgenstern. Guarantees for spectral clustering with fairness constraints. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97, pages 3458–3467. PMLR, 2019b.

Yury Makarychev and Ali Vakilian. Approximation algorithms for socially fair clustering. *CoRR*, abs/2103.02512, 2021.

Christopher Meek, Bo Thiesson, and David Heckerman. The learning-curve sampling method applied to model-based clustering. *J. Mach. Learn. Res.*, 2:397–418, 2002. ISSN 1532-4435. doi: 10.1162/153244302760200678.

Evi Micha and Nisarg Shah. Proportionally Fair Clustering Revisited. In *47th International Colloquium on Automata, Languages, and Programming (ICALP 2020)*, volume 168, pages 85:1–85:16, Dagstuhl, Germany, 2020. Schloss Dagstuhl–Leibniz-Zentrum für Informatik. ISBN 978-3-95977-138-2. doi: 10.4230/LIPIcs.ICALP.2020.85.

Sérgio Moro, Paulo Cortez, and Paulo Rita. A data-driven approach to predict the success of bank telemarketing. *Decision Support Systems*, 62:22–31, 2014. ISSN 0167-9236. doi: https://doi.org/10.1016/j.dss.2014.03.001.

Tai Le Quy, Arjun Roy, Gunnar Friege, and Eirini Ntoutsi. Fair-capacitated clustering. *arXiv preprint arXiv:2104.12116*, 2021.

M.R. Rezaee, P.M.J. van der Zwet, B.P.E. Lelieveldt, R.J. van der Geest, and J.H.C. Reiber. A multiresolution image segmentation technique based on pyramidal segmentation and fuzzy clustering. *IEEE Transactions on Image Processing*, 9(7):1238–1248, 2000. doi: 10.1109/83.847836.

Clemens Rösner and Melanie Schmidt. Privacy Preserving Clustering with Constraints. In *45th International Colloquium on Automata, Languages, and Programming (ICALP 2018)*, volume 107, pages 96:1–96:14, Dagstuhl, Germany, 2018. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. ISBN 978-3-95977-076-7. doi: 10.4230/LIPIcs.ICALP.2018.96.

Peter J Rousseeuw. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics*, 20:53–65, 1987.

Savitha Sam Abraham, Deepak Padmanabhan, and Sowmya S. Sundaram. Fairness in clustering with multiple sensitive attributes. In *International Conference on Extending Database Technology: Proceedings*, pages 287–298, 2020.

Melanie Schmidt, Chris Schwiegelshohn, and Christian Sohler. Fair coresets and streaming algorithms for fair k-means. In *Approximation and Online Algorithms*, pages 232–251, Cham, 2020. Springer International Publishing.

Bokun Wang and Ian Davidson. Towards fair deep clustering with multi-state protected variables. *arXiv preprint arXiv:1901.10053*, 2019.

I-Cheng Yeh and Che hui Lien. The comparisons of data mining techniques for the predictive accuracy of probability of default of credit card clients. *Expert Systems with Applications*, 36(2, Part 1):2473–2480, 2009. ISSN 0957-4174. doi: https://doi.org/10.1016/j.eswa.2007.12.020.

Muhammad Bilal Zafar, Isabel Valera, Manuel Gomez Rogriguez, and Krishna P. Gummadi. Fairness Constraints: Mechanisms for Fair Classification. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, volume 54, pages 962–970, Fort Lauderdale, FL, USA, 2017. PMLR.

Zhi-Hua Zhou and Zhao-Qian Chen. Hybrid decision tree. *Knowledge-Based Systems*, 15(8): 515–528, 2002. ISSN 0950-7051. doi: https://doi.org/10.1016/S0950-7051(02)00038-2.

Imtiaz Masud Ziko, Eric Granger, Jing Yuan, and Ismail Ben Ayed. Clustering with fairness constraints: A flexible and scalable approach. *arXiv preprint arXiv:1906.08207*, 2019.