Life Isn’t Fair, But We Can Try To Be: Algorithmic Fairness in Clustering

Sophie Boileau, Avery Hall, Victor Huang, Jeremiah Mensah, Armira Nance, Muno Siyakurima, Brie Sloves
Advised by Professor Layla Oesper, Department of Computer Science, Carleton College

Fairness
Clustering algorithms identify similar groups in data. In certain applications, like partitioning electoral precincts or resource allocation, these groupings should be done in a fair way. However, the definition of fairness may vary.

k Means Clustering
Unsupervised machine learning task that aims to group similar objects together.

1. Randomly initialize k centroids
2. Assign each data point to its closest centroid to form clusters
3. Calculate new centroids
4. Repeat steps (2) and (3) until convergence

The algorithm is as follows:

Fair Clustering through Fairlets
Now, the data set contains protected classes. We are tasked with clustering this data fairly.

Let X be a set of points in a metric space, which we cluster into disjoint subsets C = \{c_1, ..., c_k\}. The balance of cluster Y is

\[
\text{balance}(Y) = \min \left( \frac{\text{square}(Y)}{\text{triangle}(Y)}, \frac{\text{triangle}(Y)}{\text{square}(Y)} \right) \in [0, 1],
\]

and the balance of a clustering C is

\[
\text{balance}(C) = \min_{c \in C} \text{balance}(c).
\]

Fairlets are minimal sets that preserve the balance of protected attributes from the overall dataset. Any dataset can be decomposed into fairlets, then clustered according to traditional algorithms.

Most clustering algorithms of this form use Euclidean distance as a metric of similarity, and new centroids are calculated by computing the center of mass of the clusters.

Principal Component Analysis
Dimensionality reduction technique to facilitate visualization of data.

1. Standardize the data
2. Compute the covariance matrix
3. Compute eigenvalues of the covariance matrix to identify principal components
4. Create a feature vector to decide which principal components to keep
5. Recast the data to the principal component axes

Socially Fair k Means Clustering
A more “human-centric” algorithm that is at odds with algorithms that prioritize proportionality. Rather than minimizing the average clustering cost over an entire dataset, we minimize the average clustering cost across different demographic groups in the dataset.

From data containing demographic groups A and B, clusters are selected in such a way that minimizes the objective function

\[
\phi(C, U) = \max_{c \in C} \min_{\{a \in A, b \in B\}} p(a|c, b) \mu(a, b),
\]

where C = \{c_1, ..., c_k\} is the set of centers for the clusters given by the partition U = \{U_1, ..., U_k\}. Here, we define the cost function to be

\[
\Delta(C, U) = \sum_{i=1}^{k} \sum_{j \in U_i} ||p - c_i||^2,
\]

data points p. Finding the solution to this problem is easily translated into a linear, convex problem. In order to do so, we define

\[
a_i = \{a \in A\}, \quad \mu_i = \{\mu_{a,b} \in \mu\}, \quad i = |\mu_i|.
\]

Also, let

\[
M^A = \{\mu_{a,b} \in \mu| a \in A\} \quad \text{and} \quad M^B = \{\mu_{a,b} \in \mu| b \in B\}.
\]

Then, the optimal solution will solve the following program:

\[
\min_{\theta} \text{ such that } \Delta(M^A, U \cap A) + \sum_{i \in [k]} \mu_i \theta \leq \theta
\]

\[
\Delta(M^B, U \cap B) + \sum_{i \in [k]} \mu_i \theta \leq \theta
\]

\[
0 \leq s_i, \leq t_i, \quad \forall i \in [k]
\]

For each cluster U_i, let \mu_i be the mean of A \cap U_i, and let \mu_i be the mean of B \cap U_i. Then, the optimal fair center c_i is on the line segment connecting \mu_i and \mu_i.

The only points that need to be checked are those in the set

\[
Z = \left\{x : x = (1 - \gamma)\mu_i + \gamma \mu^B, \quad \gamma \in [0, 1]\right\}.
\]

The Pareto-optimal solution is found in finite time and can be initialized tactfully with an approximation algorithm.

Intersectionality
So far, we have assumed that each data point can only belong to one demographic group or protected attribute. There are plenty of human-centered applications where this is not the case:

- Gender is not binary
- Many people identify with more than one racial group
- The experiences of a disabled, impoverished individual cannot be encapsulated by separately analyzing a disabled individual and an impoverished individual

A recommendation for future research lies in fair clustering algorithms for datasets with intersectional observations.

References