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

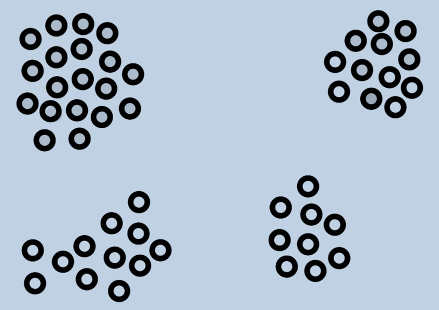
Sophie Boileau, Avery Hall, Victor Huang, Jeremiah Mensah, Armira Nance, Muno Siyakurima, Brie Sloves
Advised by Professor Layla Oesper, Department of Computer Science, Carleton College

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.

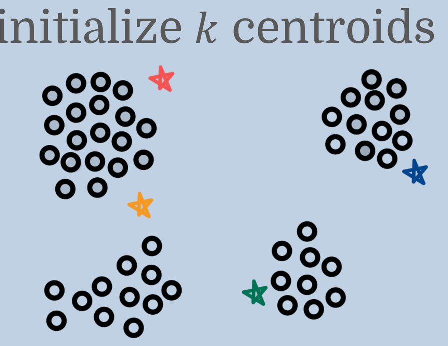
k Means Clustering

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



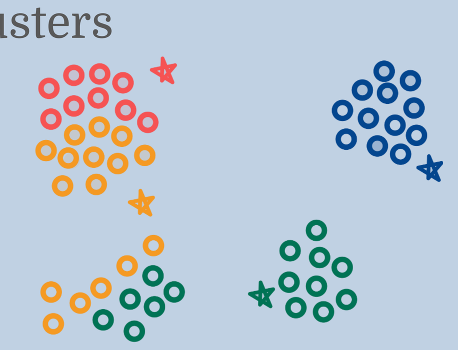
The algorithm is as follows:

1. Randomly initialize k centroids

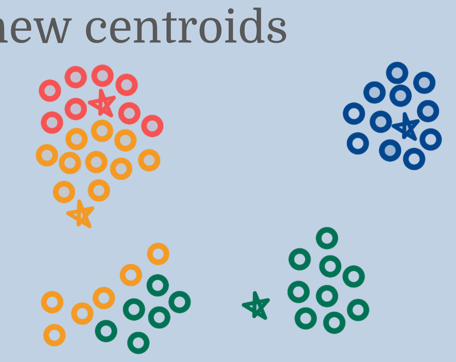


for this example, let k=4

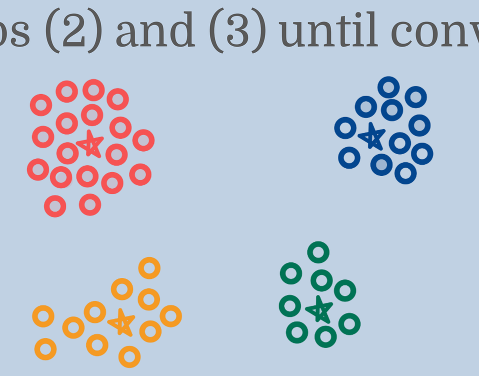
2. Assign each data point to its closest centroid to form clusters



3. Calculate new centroids



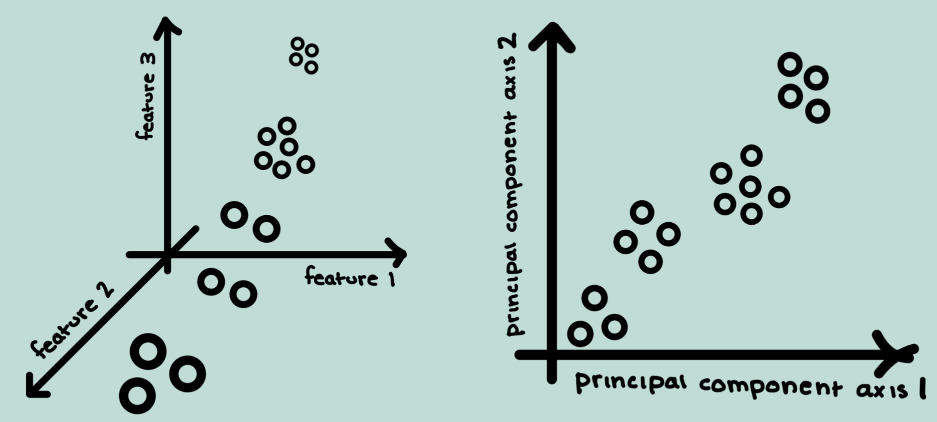
4. Repeat steps (2) and (3) until convergence



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

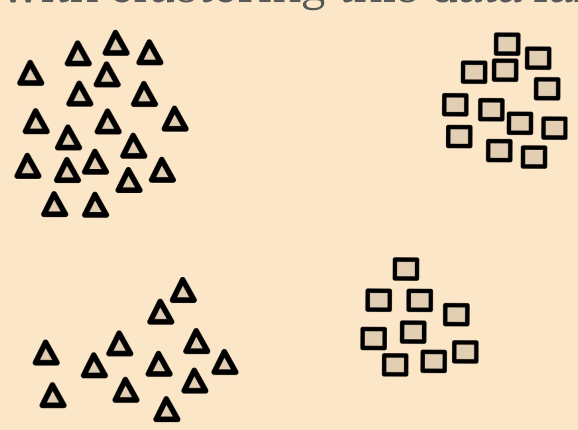
Dimensionality reduction technique to facilitate visualization of data.



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

Fair Clustering through Fairlets

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



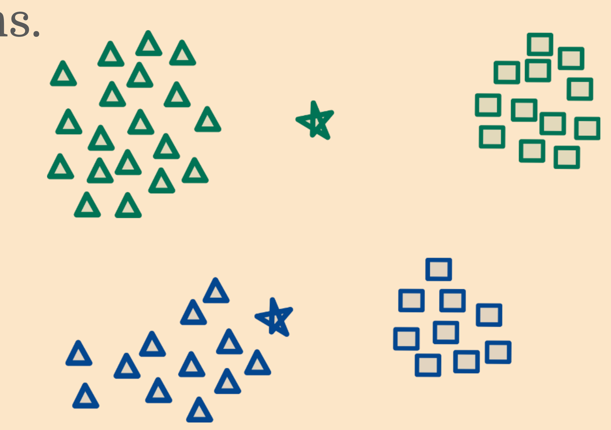
Let X be a set of points in a metric space, which we cluster into disjoint subsets $\mathcal{C} = \{C_1, \dots, C_k\}$. The *balance* of cluster Y is

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

and the balance of a clustering $\mathcal{C}$ is

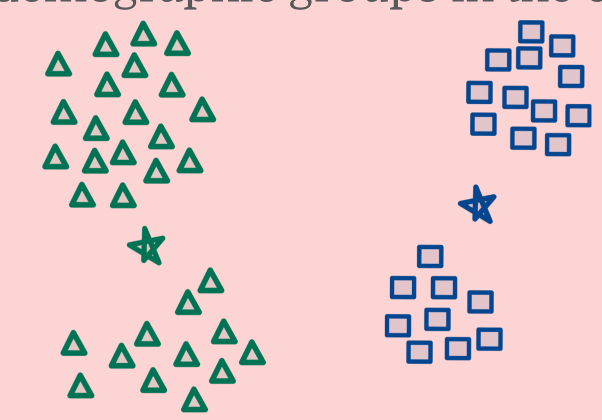
$$balance(\mathcal{C}) = min_{C \in \mathcal{C}} balance(C).$$

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.



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

$$\phi(C, \mathcal{U}) := \max\{\frac{\Delta(C, \mathcal{U} \cap A)}{|A|}, \frac{\Delta(C, \mathcal{U} \cap B)}{|B|}\},$$

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

$$\Delta(C, \mathcal{U}) := \sum_{i=1}^{k} \sum_{p \in U_i} ||p - c_i||^2,$$

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

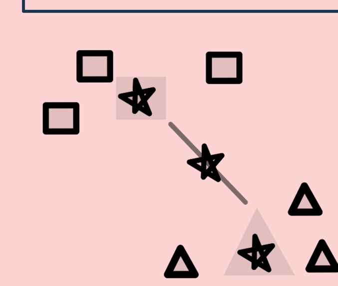
$$\alpha_i = \frac{|A \cap U_i|}{|A|}, \quad \beta_i = \frac{|B \cap U_i|}{|B|}, \quad \text{and } l_i = ||\mu_i^A - \mu_i^B||.$$

Also, let

$$M^A = \{\mu_1^A, \dots, \mu_k^A\} \quad \text{and } M^B = \{\mu_1^B, \dots, \mu_k^B\}.$$

Then, the optimal solution will solve the following program:

$$\boxed{\begin{gathered} \min \theta \text{ such that} \\ \frac{\Delta(M^A,\ \mathcal{U} \cap A)}{|A|} + \sum_{i \in [k]} \alpha_i x_i^2 \le \theta \\ \frac{\Delta(M^B,\ \mathcal{U} \cap B)}{|B|} + \sum_{i \in [k]} \beta_i (l_i - x_i)^2 \le \theta \\ 0 \le x_i \le l_i, \quad \forall i \in [k] \end{gathered}}$$



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

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

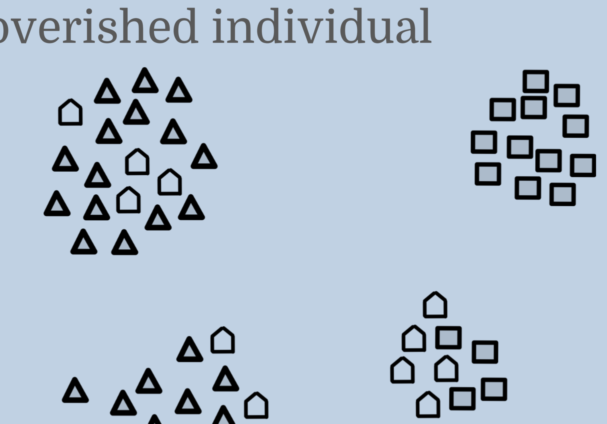
$$Z = \left\{x : x_i = \frac{(1-\gamma)\beta_i l_i}{\gamma\alpha_i + (1-\gamma)\beta_i}, \gamma \in [0,1]\right\}.$$

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

Chhabra, Anshuman et al. "An Overview of Fairness in Clustering." *IEEE*, vol. 9, 2021, https://ieeexplore.ieee.org/document/9541160.

Chierichetti, Flavio et al. "Fair Clustering Through Fairlets." *31st Conference on Neural Information Processing Systems*, 2017, https://proceedings.neurips.cc/paper_files/paper/2017/file/978fce5bcc4eccc88ad48ce3914124a2-Paper.pdf.

Chouldechova, Alexandra and Roth, Aaron. "A Snapshot of the Frontiers of Fairness in Machine Learning." *Communications of the ACM*, vol. 63, no. 5, 2020, https://cacm.acm.org/magazines/2020/5/244336-a-snapshot-of-the-frontiers-of-fairness-in-machine-learning/fulltext?mobile=false.

Ghadiri, Mehrdad et al. "Socially Fair k Means Clustering." *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 2021, https://dl.acm.org/doi/pdf/10.1145/3442188.3445906.

Oesper, Layla. *Algorithmic Fairness in Clustering: A Study in Replication.* Carleton College, Fall 2023, https://www.cs.carleton.edu/cs_comps/2324/fairnessInClustering/index.php.

Whitfield, Brennan and Pierre, Sadrach. "A Step by Step Explanation of Principal Component Analysis." *builtin*, https://builtin.com/data-science/step-step-explanation-principal-component-analysis.