Potential Bias in Predictive Clustering

Griffin Berlstein and Dr. Shion Guha

grberlstein@vassar.edu shion.guha@marquette.edu

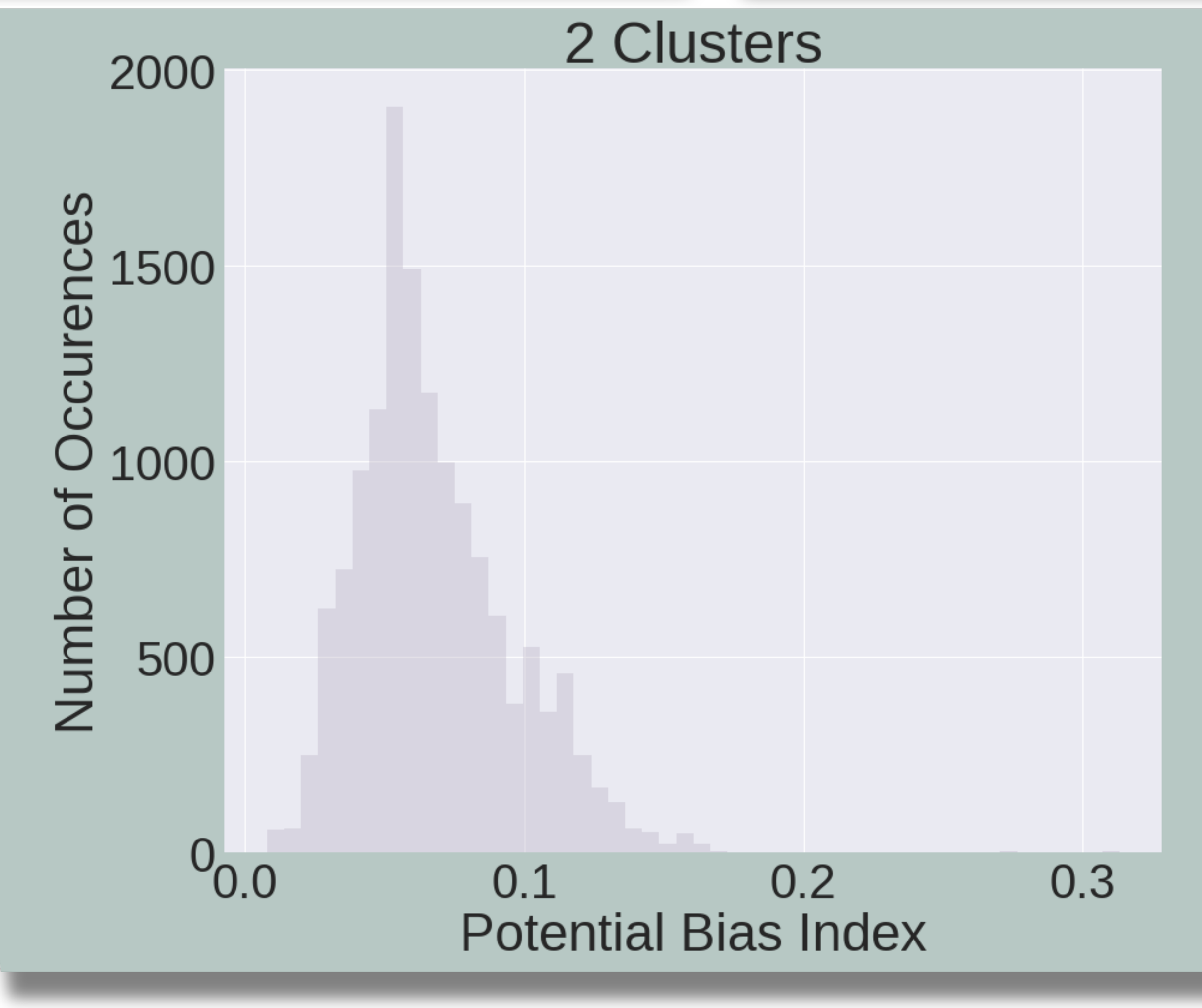
Background

Black-boxed algorithms are used to drive decisions that impact people's lives. This inscrutability creates the potential for misapplication that can have harmful effects. Potentially harmful algorithms, such as predictive crime analysis, must be closely examined to avoid bias.

Method

After implementing Haversine distance we compared the two metrics:

1. Set Starting Points
2. Cluster
3. Compute Potential Bias Index
4. Repeat for Each Month



K-Means Clustering

1. Initialize K cluster centers
2. Assign each point to the nearest center
3. Recompute cluster centers as mean of points
4. Repeat 2 & 3 until convergence

This assumes:

1. There are K clusters in the data
2. A euclidean distance metric is appropriate

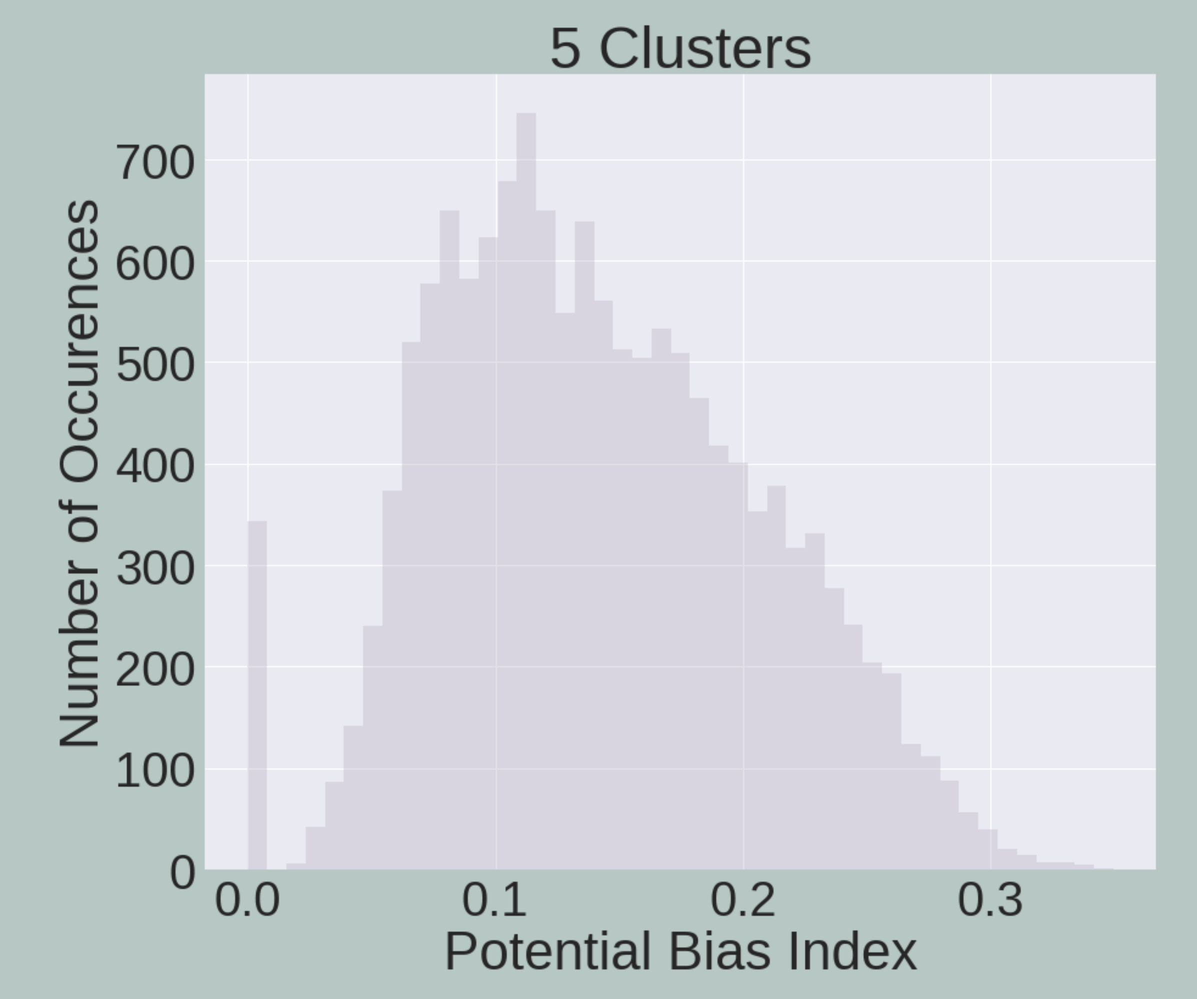
Euclidean distance misrepresents points on a sphere as it does not account for curvature.

Potential Bias Index

$$\mathrm{D} = 1 - \sum_{e=0}^{k-1} \left(\frac{O}{n_e} * \frac{O}{n_g} \right)$$

The Potential Bias Index for a geodesic cluster is D (disimilairity) * the probability of being a minority within a geodesic cluster

k is number of clusters, O is points that overlap between g and e, the geodesic and euclidean clusters.
n is total number of points in the cluster.



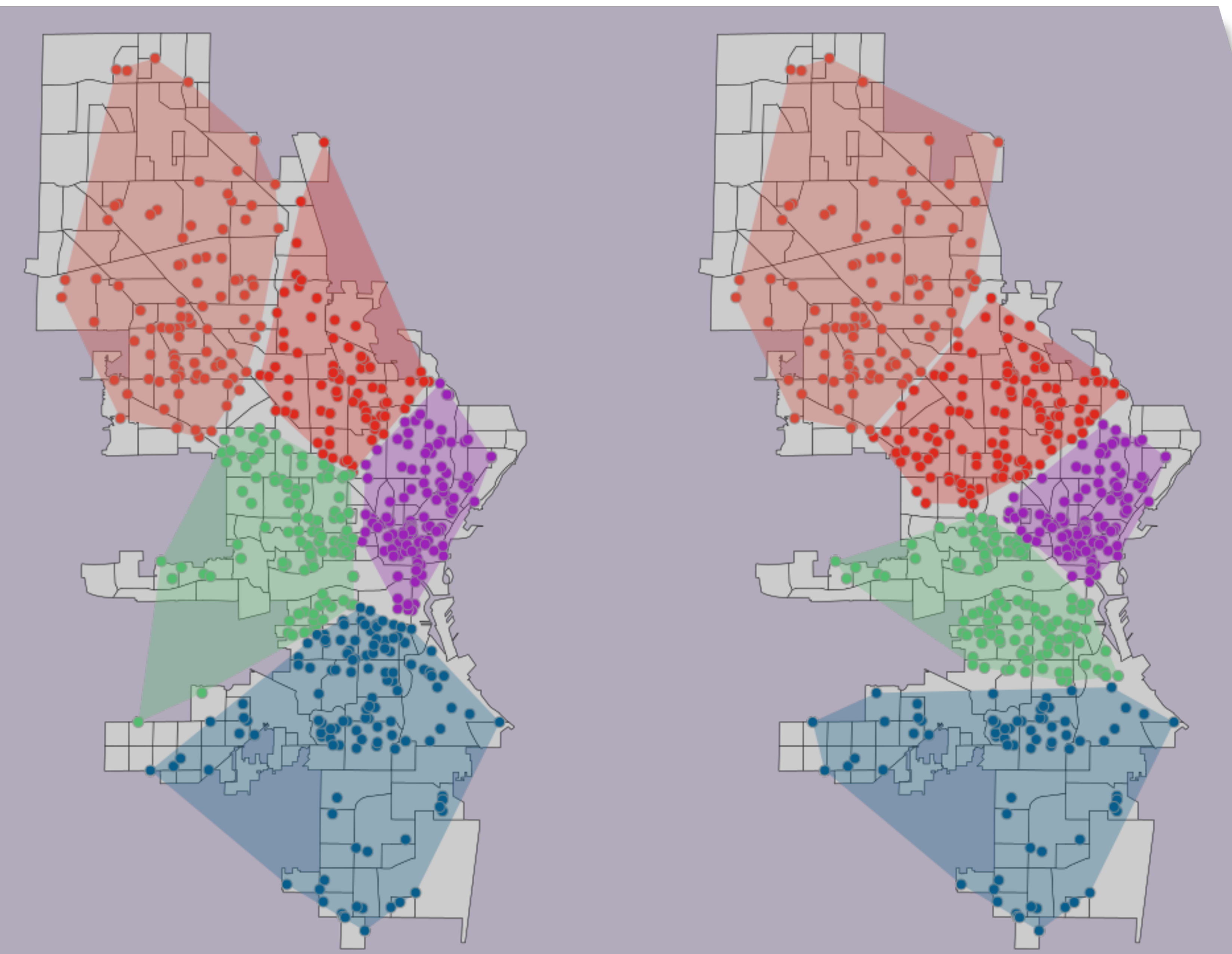


Figure 1: Euclidean (left) and Geodesic (right) clusterings from the same starting points on a map of theft in Milwaukee March 2015. K = 5.

Results

- Using a geodesic metric produced visibly different results
- Higher values of K tended to produce higher potential bias values
- These values suggest areas of potential misclassification
- Since these cluster are used to allocate police resources, misclassifcation runs the risk of causing unequal policing in areas with higher a minority population

Acknowledgements

Research was funded by the National Science Foundation (NSF Award ACI-1461264) and the REU was organized by Dr. Dennis Bylow and Dr. Kim Factor. Special thanks to my research partner Justin Miller

References

1. Cathy O'Neil. 2016. Weapons of math destruction: How big data increases inequalityand threatens democracy. Crown Books
2. Brent Daniel Mittelstadt, Patrick Allo, Mariarosaria Taddeo, Sandra Wachter, and Luciano Floridi. 2016. The ethics of algorithms: Mapping the debate
3. Batya Friedman and Helen Nissenbaum. 1996. Bias in computer systems.
4. Melanie Feinberg. 2017. A Design Perspective on Data. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. ACM, 2952–2963.
5. Christian Sandvig, Kevin Hamilton, Karrie Karahalios, and Cedric Langbort. 2014. Auditing algorithms: Research methods for detecting discrimination on internet platforms.