Reducing Discrimination in Learning Algorithms for Social Good in Sociotechnical Systems

Katelyn Morrison ~ Kmorrison@pitt.edu

University of Pittsburgh

IJCAI-PRICAI YOKOHAMA 2020

Primary Questions & Hypotheses

- How does the location of a bike station relate to the neighborhood's socioeconomic status and demographics?
  - *Agreeing with previous research [1-4], I predict the bike stations in Pittsburgh will be located primarily in medium to high-income neighborhoods.*
- How can we develop a model to help urban planners fairly allocate smart mobility resources?
  - *If we investigate how to employ model-based reinforcement learning paired with Bayesian Optimization, it could be possible.*

Exploratory Analyses

Datasets:

- 2020 FFIEC Census Report
- Healthy Ride Bike Station Locations
- Allegheny County Poor Housing Conditions
- Allegheny County 2010 Census Tracts

Analyses:

- ***Figure 1*** shows approximately 5% of bike stations are available in tracts with poor housing conditions.
- ***Figure 2*** reveals a significant outlier circled in purple of several bike stations in low-income tracts that makes up 3 universities.

Bike Station Location in Relation to Poor Housing Conditions in Pittsburgh, PA

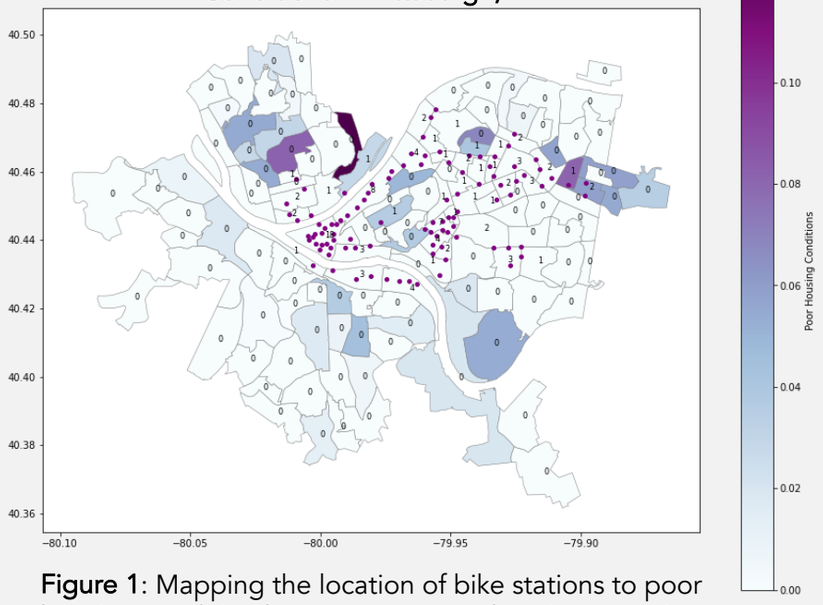

Figure 1: Mapping the location of bike stations to poor housing condition by census tract and station capacity.

Bike Station Location in Relation to 2015 Median Household Income in Pittsburgh, PA

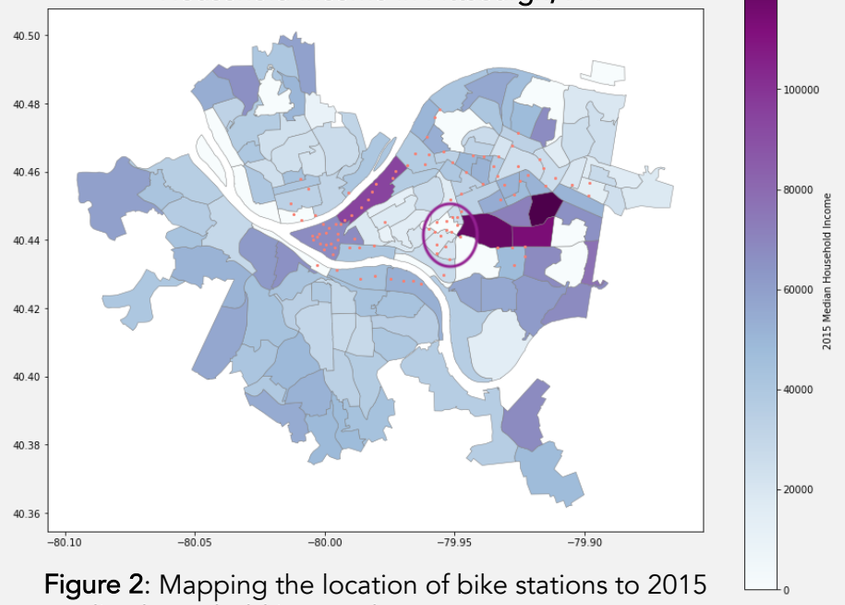

Figure 2: Mapping the location of bike stations to 2015 median household income by census tract.

Fairness in Station Location

Potential Model:

- Model-based Reinforcement Learning paired with Bayesian Optimization.
- Predict the optimal location that maximizes predicted demand while minimizing discrimination against demographics and income with following reward function, **e.g.**:

$$R_{L_K} = d_{L_K} + \frac{1}{4} POI_{L_K} - \left(\frac{medHHinc_L}{10000}\right) + (\%\ minority\ pop_L)$$

Future Work

- Identify more datasets that may be useful in exploratory analyses and training a model.
- Identify other kinds of models that could be used.

References

[1] Cardona, Mateo, et al. "Analisis de la red de ciclorutas de manizales (colombia) a partir de criterios de accesibilidad territorial urbana y cobertura de estratos socioeconomicos". Revista Espacios, 38(28), 2017.

[2] Garcia, Zuluaga , et al. "Propuesta metodologica para el diagnostico y planificacion urbana de una red de ciclorrutas. Caso estudio: Manizales." PhD thesis, Universidad Nacional de Colombia- Sede Manizales, 2017.

[3] Mooney, Stephen J, et al. "Freedom from the station: Spatial equity in access to dockless bike share." Journal of transport geography, 74:91–96, 2019.

[4] Liu, Yijia, et al. "Scooter Equity and Demand Analysis." UPenn MUSA, Master of Urban Spatial Analytics (MUSA) - University of Pennsylvania, 11 May 2020, pennmusa.github.io/MUSA_801.io/project_14/index.html#5_model_building.