6^{th} International Conference on Computational Social Science IC^2S^2
July 17-20, 2020, MIT Media Laboratory, Boston, MA, USA

Ageism in Traffic Policing Activity

Keywords: predictive bias, algorithmic fairness, ageism, policing activity, stop outcome prediction

Extended Abstract

Ageism—prejudice on the grounds of a person's age—is a frequently overlooked and relatively unexplored bias that can impact collective societal, cultural and political perceptions. While not as extensively studied as race and gender, age proves to be a concern for one's eligibility and competency for certain roles or privileges in society.

In this work, we investigate age-induced biases in traffic stop policing outcomes through the lens of algorithmic fairness in prosecution trends. We find that ageism indeed appears to play a role in traffic stop outcomes, with police being more likely to give harsher penalties to younger individuals than to older individuals. The segment of population age 20-29 is over-represented across all penalties (warning, citations, and arrests), while the opposite is true for the population age 50+. We present an age-debiased traffic-stop outcome prediction model, and demonstrate that debiasing the data for age results in substantially improved algorithmic fairness metrics with negligible change in predictive performance.

Data. We leverage the Stanford Open Policing Project (OPP) data set [3], which includes an aggregation of police traffic-stop data, including features such as driver race and vehicle type. Each traffic stop results in one of the following outcomes in order of most to least severe: arrest, citation, warning (verbal or written), or no consequence. We select three states whose data contains the required features for our analysis: California (CA)—a historically Democratic state on the West coast, Massachusetts (MA)—a Democratic state on the East coast, and Tennessee (TN)—a historically Republican state in the South.[1] These three states cover a broad political spectrum and represent three of the four major U.S. regions.

Across all three states, there is a consistent negative correlation between age and the severity of outcomes (cf., Figure 1). Younger subjects ($\leq$ 25 y/o) are more likely to receive a harsher outcome, and older subjects ($\geq$ 75 y/o) are nearly twice as likely to receive a warning compared to their younger counterparts. Conversely, younger subjects are twice as likely to be issued a citation and 7 times more likely to be arrested than older subjects. We find that older subjects

[1] Overall, 11 states contained sufficient relevant information, but we chose to focus on CA, MA and TN.

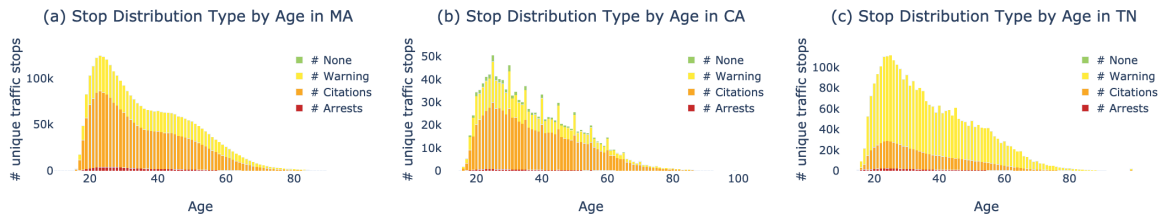

Figure 1: Traffic-stop distributions for (a) Massachusetts, (b) Tennessee, and (c) California. These graphs show the outcome type makeup per age.

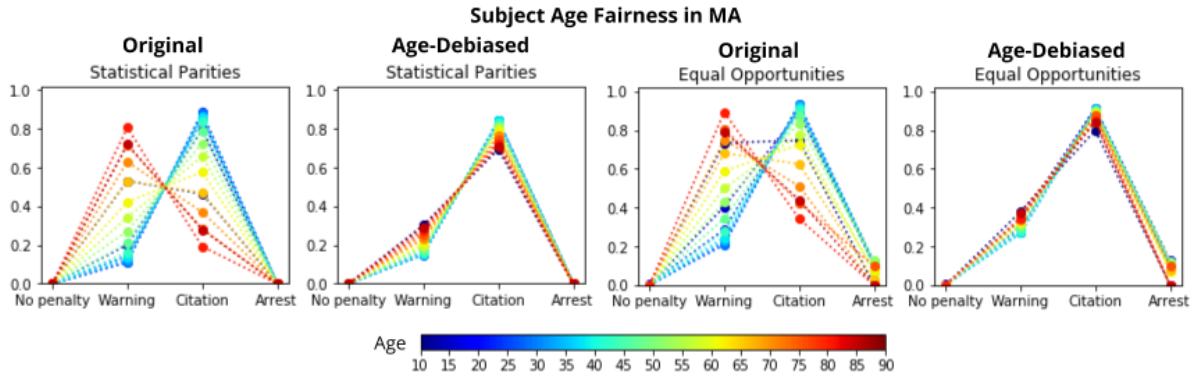

Figure 2: Subject age fairness of our model using the the Massachusetts's data set, before and after age debiasing.

are almost 4 times more likely than younger subjects to receive a warning, but are half as likely to be issued a citation. For a sanity check, we verified that, beyond these three states, the trends hold true in all 11 states in our data set.

Limits. We note that many possible confounders could explain the detected trends, e.g., younger subjects may have been caught at faster speeds, warranting harsher penalties. However, the data needed to control for such confounders is not available.

Interpretation. Two features found to be the most correlated with age are 1) the likelihood of a subject being searched after being stopped and 2) the likelihood of discovering contraband after a search is conducted. Younger subjects are more likely to be searched and are also more likely to be found with contraband. While this may indicate that police are not discriminating by age in their search practices, the ageism indicated by the discrepancy between the treatment of the young and old becomes even more pronounced after adjusting for whether or not contraband was found.

Modeling. We use XGBoost to model stop outcome prediction. We group subjects into age brackets of 5 years and generalize two fairness measures for multi-group analysis: statistical parity [1], which we define as Pr(Predicted Outcome|Age Group), and equal opportunity [2], which is the true positive rate of each age group. We debias ageism in our model by removing the subject age feature during pre-processing, and our age-debiased model effectively eliminates age bias and greatly improves statistical parity and equal opportunity fairness measures without impairing accuracy (cf., Figure 2).

References

[1] Benjamin Fish, Jeremy Kun, and Ádám D Lelkes. "A confidence-based approach for balancing fairness and accuracy". In: *Proceedings of the 2016 SIAM International Conference on Data Mining*. SIAM. 2016, pp. 144–152.

[2] Moritz Hardt, Eric Price, Nati Srebro, et al. "Equality of opportunity in supervised learning". In: *Advances in neural information processing systems*. 2016, pp. 3315–3323.

[3] Emma Pierson et al. "A large-scale analysis of racial disparities in police stops across the United States". In: (2019). Accessed: December 11, 2019. URL: `https://5harad.com/papers/100M-stops.pdf`.