Algorithmic Analysis: Law Schools & Bar Exams

April 2024

Prepared for:
Dr. Maria De-Arteaga
Assistant Professor
Department of Information, Risk & Operations Management
McCombs School of Business
The University of Texas at Austin
2110 Speedway, Austin, TX, 78705

Prepared By:
Ethan Wong, Ria Mittal, Nick DeLeon
Group 7

Model Development

Domain Selection

We chose to concentrate on education, with a specific focus on leveraging an algorithm to predict a student's likelihood of passing the bar examination, aiming for broader applicability in law school admissions. Our algorithm was trained using data from the 'LSAC National Longitudinal Bar Passage Study' by Linda Wightman in 1988, sourced from Kaggle. This dataset encompasses information gathered by the Law School Admissions Council (LSAC) from 1991 to 1997. It tracks approximately 22,400 law students from enrollment to graduation and eventual bar exam participation.

Data Preparation

Our dataset originally comprised 22,407 rows and 39 columns. Initially, we refrained from dropping rows with null values to prevent potential loss of insights during our empirical bias assessment, opting instead for missing value imputation methods. We mainly replaced null values with the mode of each column. Additionally, we employed a Gradient Boosting Regressor to predict null values in columns representing decile values. Subsequently, manual feature selection was conducted based on each column's context and correlation with the binary target variable, bar_passed. Many of our decisions during this stage were informed by an exploratory data analysis performed by another Kaggle user on the same dataset (Usylesses, 2022). Our finalized dataset for modeling contains 22,407 rows and 17 columns.

Algorithm Training

During model training, we independently conducted an algorithmic audit on gender and race as sensitive attributes. For gender, we split the data based on positive rates for men and women passing the bar and the group ratio for women in the dataset. For race, we utilized the *stratify* parameter in Scikit-Learn's *train_test_split* function to account for the five different race groups simultaneously. We then employed GridSearchCV to experiment with various model types (Logistic Regression, Bagging, Random Forest, K Nearest Neighbors, Adaptive Boosting, and Gradient Boosting) and hyperparameters to optimize balanced accuracy via 10-fold Stratified Cross Validation. The result was that a logistic regression model, standardized by removing the mean and scaling to unit variance, with a regularization strength between 0.01 and 1, and class weights adjusted to handle imbalance, performed optimally. Our model achieved 78% accuracy and an AUC of 88% on the testing data.

Characterization of Use

Historical Context

Law schools across the United States have long been accused of biased admissions. John Nussbaumer of the St. John's Law Review reports that making the LSAT the sole admissions criterion to improve their national rankings disadvantages minority students who average lower scores than the majority (Nussbaumer, Pg.170). John writes: "This unsubstantiated reliance on the LSAT reflects the elitist perspective of the schools that are represented on the Council. A majority of these schools have few African-American students, either in total numbers or as a percentage of their student populations. These exclusionary practices have a disparate impact on minority students in general, and on African-American students in particular." (Nussbaumer, Pg.179).

Algorithmic Reform

Law schools are shifting away from only shaping their admission decisions around LSAT scores to improve racial diversity. Some are using GRE scores instead of the LSAT, but this has proved futile in battling racially imbalanced student bodies. GRE scores alone have proven to increase selectivity among law schools without improving racial imbalance (Koons, 2022). As of 2024, there are seven factors when a student applies to law school: Transcripts, Test scores (e.g., LSAT), life experience, work/volunteer experience, recommendations, resilience, and goals (Kuris, 2024). Algorithms are adopted where each factor is assigned a weight and tested on the likelihood a student will pass the bar, as seen in our dataset (Whitman & Offer, 2022). Figuring out which attribute(s) hold the most predictive power requires some testing. Using the historical data, we can split the data into training and testing groups where our 'X' variable(s) are the different attributes, and our 'y' is the likelihood of passing the bar.

It is essential to consider algorithms are not inherently bias-free; the model is only as biased as the human behind the code. As a result, ethical considerations regarding which features to use for modeling and which should be left obscure arise. Will a model trained on demographic data have a disparate impact on what applicants are accepted? The Hechinger Report found fifteen flagship universities under-serving their respective state's Black and Hispanic populations by a figure of 10 points (Engler, 2021). It is paramount that admissions offices identify the implications of their models and the potential disparate impacts on minority demographics to ensure each student has a fair shot at attending their university.

Ethical Considerations

Sampling Bias

Sampling bias occurs when the data used to train an algorithm fails to represent the target population accurately. For instance, since our dataset was collected from 1991 to 1997, it likely reflects outdated sociocultural, economic, and educational conditions. Such historical data may not account for subsequent changes, potentially skewing predictions about current applicants' likelihood of passing the bar exam. Additionally, certain demographic groups or regions in the data could lead the algorithm to predict outcomes unfairly, disadvantaging underrepresented groups and reinforcing existing educational and professional inequities. This bias may result in fewer resources allocated to those deemed less likely to succeed, perpetuating a cycle of disadvantage (De-Arteaga, 2024).

Algorithmic Bias

Algorithmic bias arises when the criteria selected to predict outcomes are not ethically justified or reflect current values. In the context of law school admissions, the use of demographic attributes in algorithms risks perpetuating past injustices, such as the historical exclusion of minorities from legal education. Despite increased diversity following reforms, outdated models can still exacerbate racial disparities. For example, in 1994, only 7.5% of students in ABA-approved law schools were black, indicating potential bias if these figures influence current algorithms (Nussbaumer, Pg 167-168).

Unintended Consequences

Unintended consequences occur when outcomes not anticipated by developers affect system functionality. In law school admissions, an algorithm predicting a student’s success in passing the bar could reinforce existing gatekeeping in the legal profession. By prioritizing applicants based on outdated or biased data predictions, the algorithm might favor those from historically successful backgrounds, reducing diversity and inclusivity. Moreover, schools may alter their curricula to cater to the algorithm’s preferences rather than individual student needs, compromising the quality of education. There is also a risk of applicants manipulating their data, such as altering personal essays or demographics, to align with perceived algorithmic preferences - a tactic observed in 34% of college applicants (Spencer, 2021).

Empirical Bias Assessment

Gender

We analyzed the confusion matrix rates for men and women, noting only a 0.02 difference in false positives and true negatives (see Fig.1). For our model, a positive case indicates passing the bar, and a negative case indicates not passing. No demographic parity, equalized odds, or predictive parity violations were observed (see Fig.2). This data suggests equitable allocation of predictions across genders; however, it raises concerns about the potential disproportionate impact on individuals who might succeed but end up not being admitted due to model biases. In law school applications, minimizing false negatives is crucial, as missing out on admitting capable candidates could be detrimental. Our model development showed that selecting a different model optimized for metrics like ROC AUC without adjusting for class imbalance (95% of people passed the bar in the dataset) resulted in negligible false negatives and an 18% increase in test accuracy. Yet, this led to a significant rise in false positives and a sharp decrease in true negatives (see Fig.3). Despite attempts to optimize various performance metrics, we observed consistent results across confusion matrix elements and fairness metrics between genders. While false positives may seem less critical to law schools, prioritizing model fairness and utility over mere performance is essential, as aiming solely to reduce false negatives and boost accuracy can create a misleading perception of model effectiveness, primarily since it performs well in predicting bar passers but not in identifying those who are unlikely to pass.

Race

Our analysis of the confusion matrix rates between White, Black, Asian, Hispanic, and "Other" racial groups showed significant disparities (see Fig.4) and notable violations in demographic parity (see Fig.5) and equalized odds (see Fig.6). Despite this, predictive parity remained consistent across groups (see Fig.7). The most severe violations were between White and Black individuals - 0.6 for demographic parity and 0.57 for equalized odds (see Fig.8), with minor discrepancies (0.03) between Hispanic and "Other" persons. These issues likely stem from Whites comprising over 80% of our dataset and discrepancies in pass rates up to 19% among racial groups. This skewed outcome distribution disproportionately disadvantages minority groups. Given these findings, an intersectional analysis combining race and gender, similar to the work of Joy Buolamwini and Timnit Gebru, is recommended to enhance fairness metrics and better address any stakeholder concerns regarding the equity and inclusivity of our algorithm.

Appendix

Figure 1

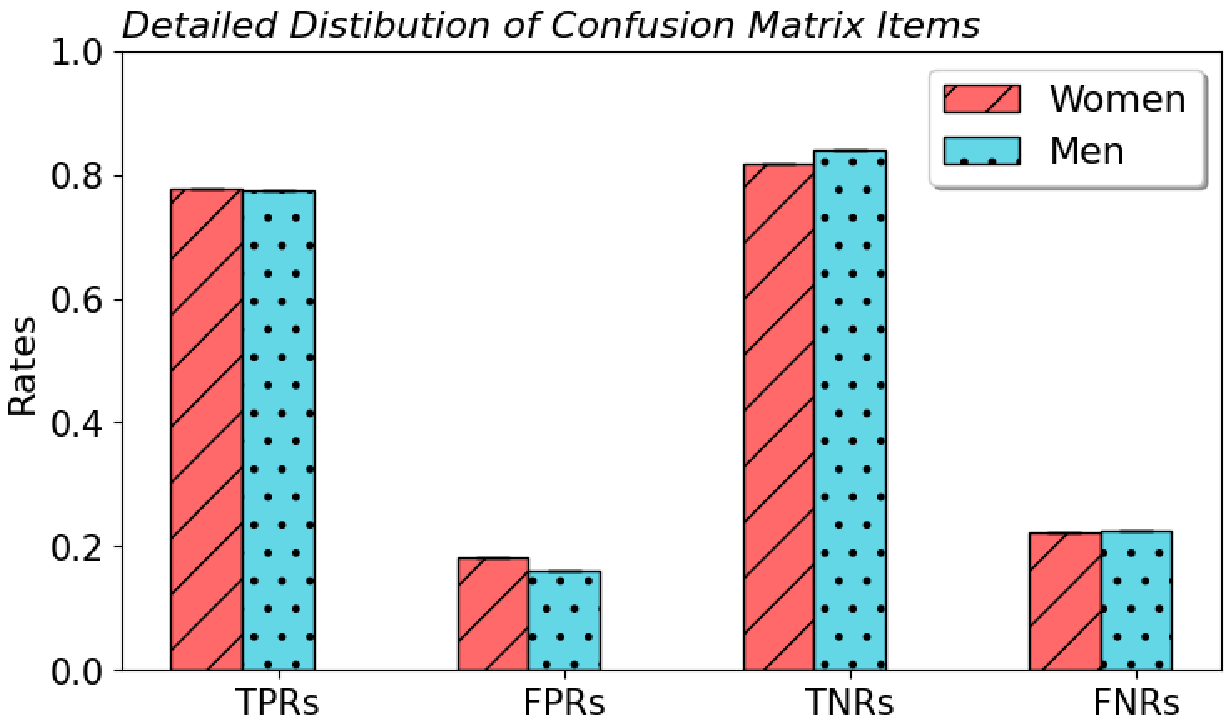


Figure 2

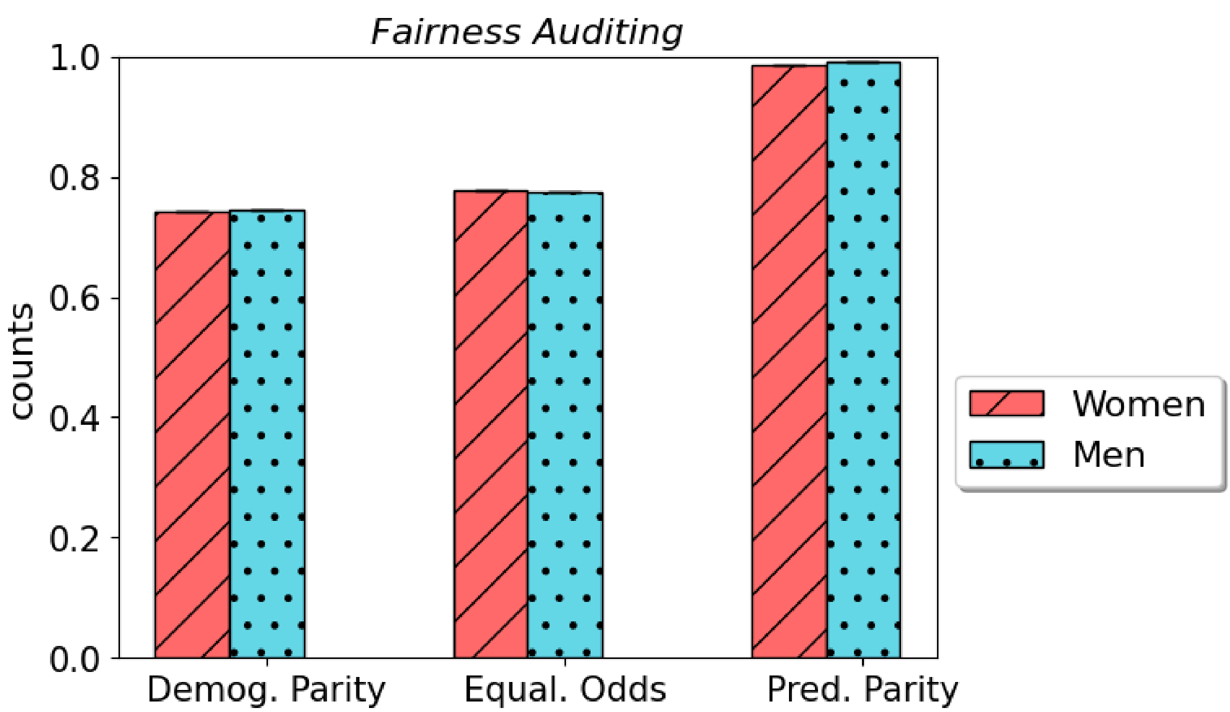


Figure 3

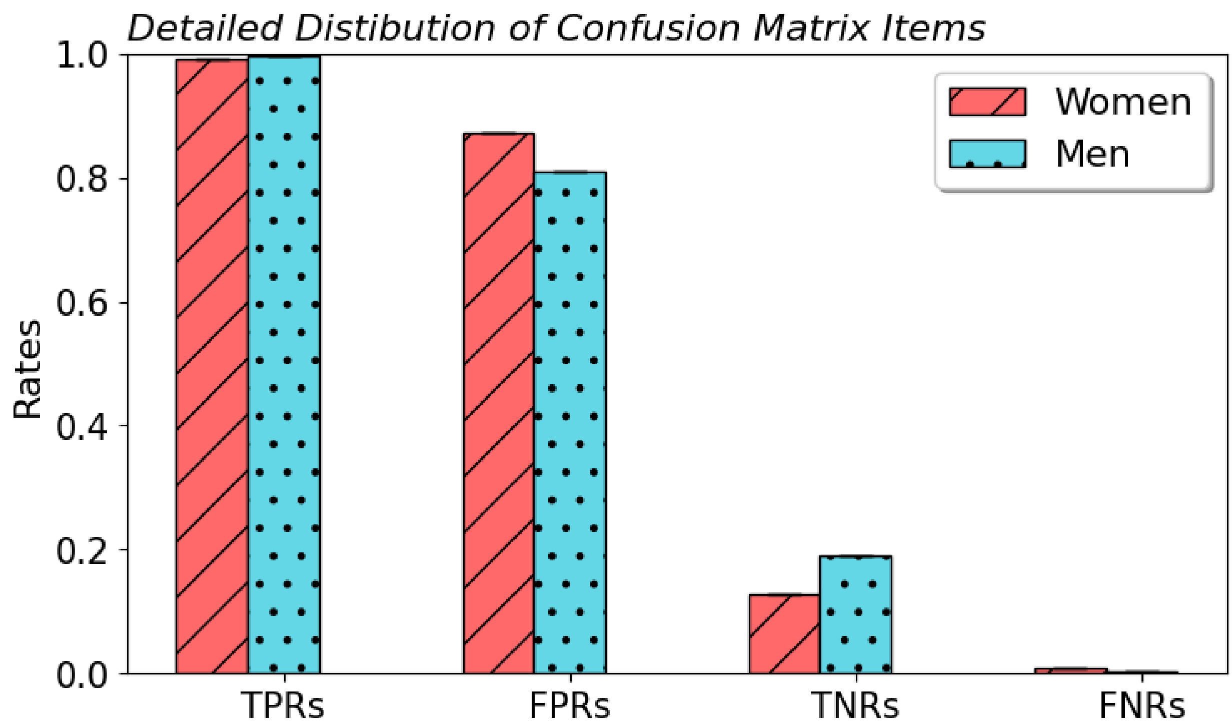


Figure 4

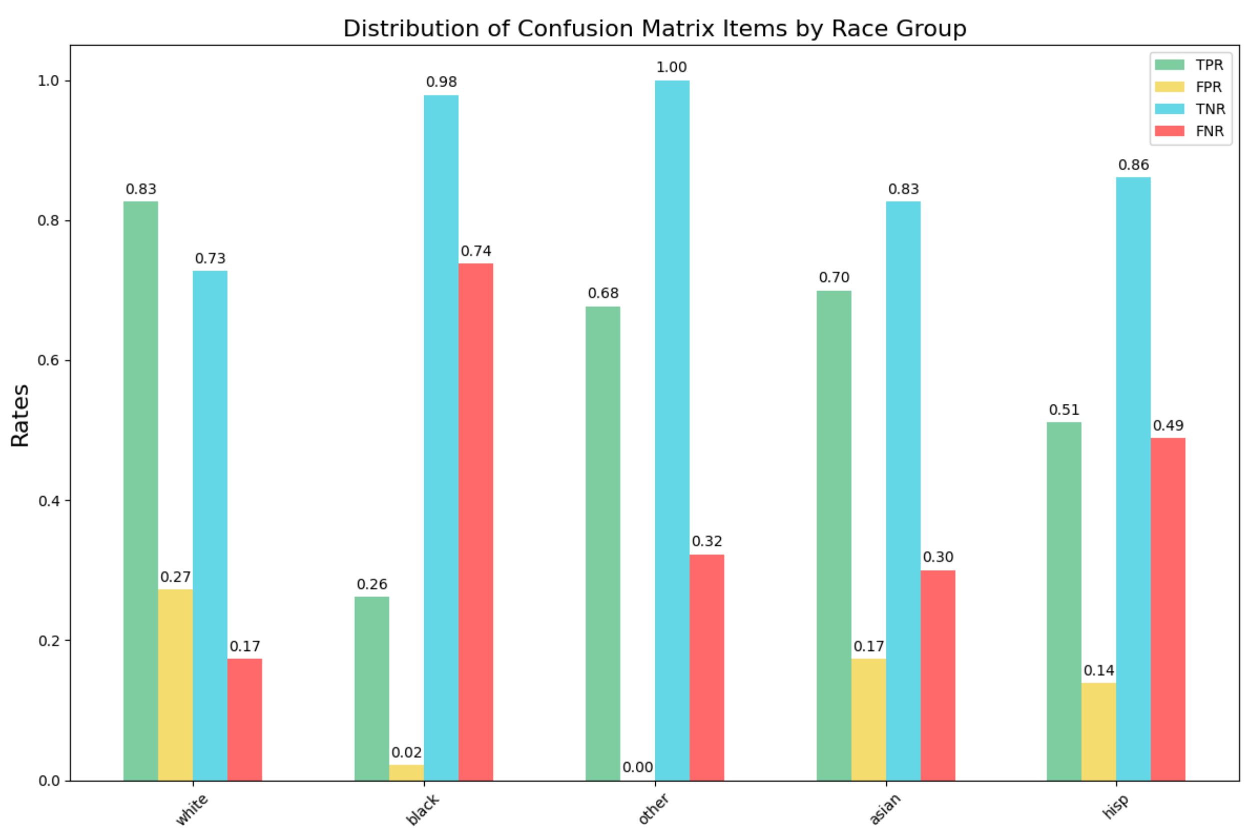


Figure 5

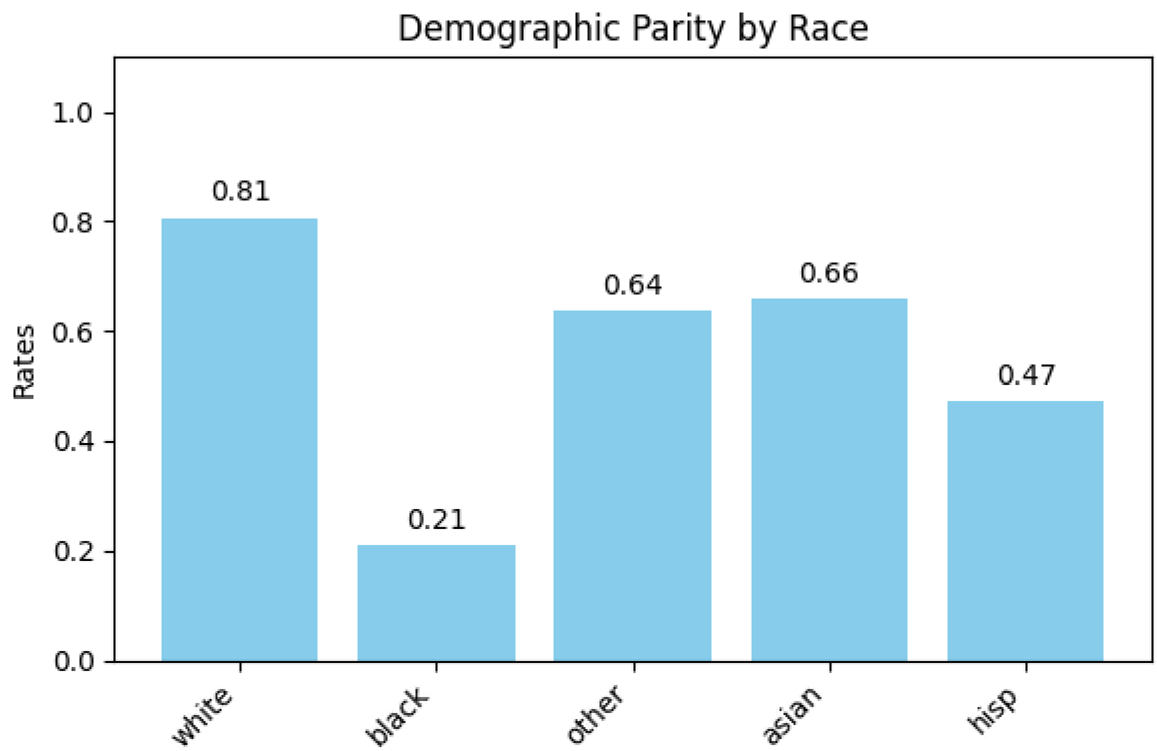


Figure 6

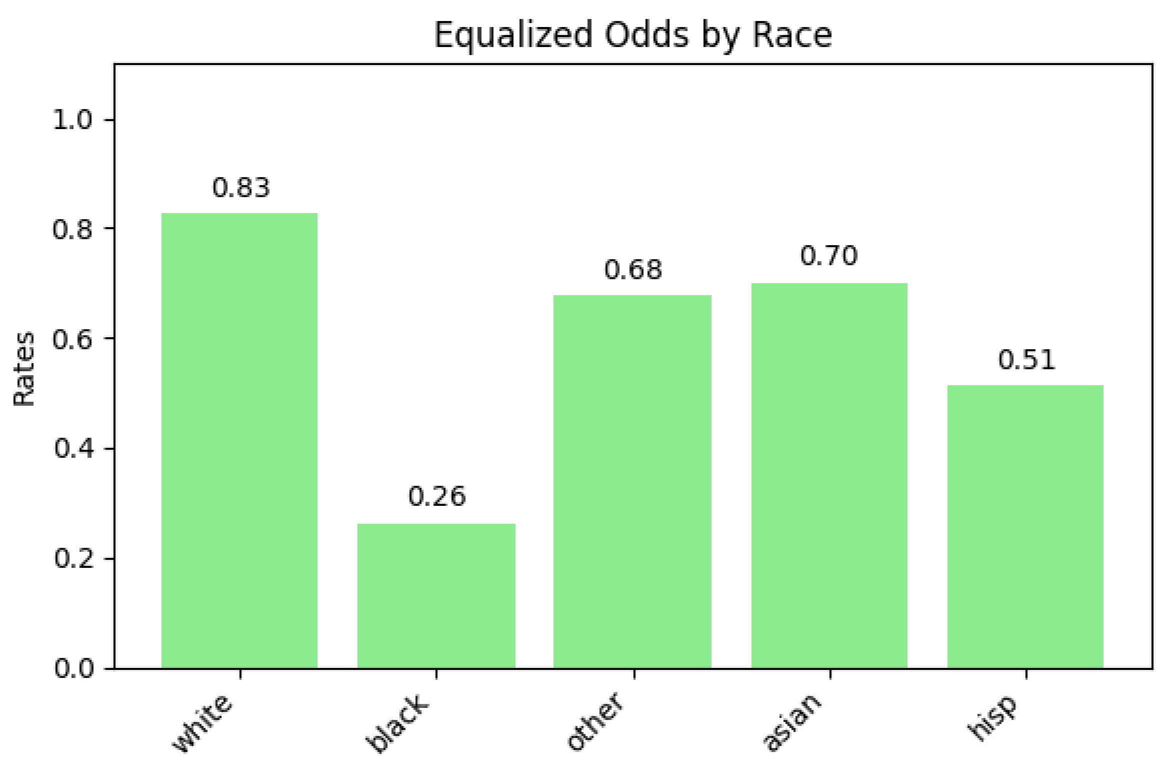


Figure 7

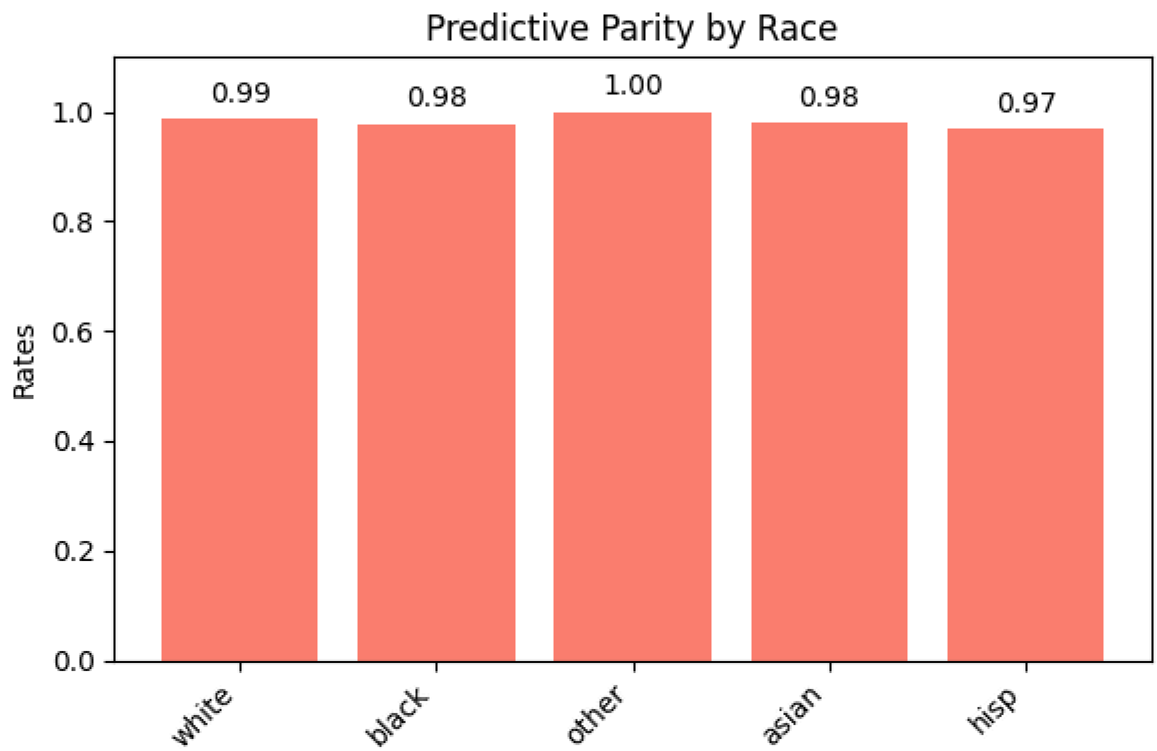
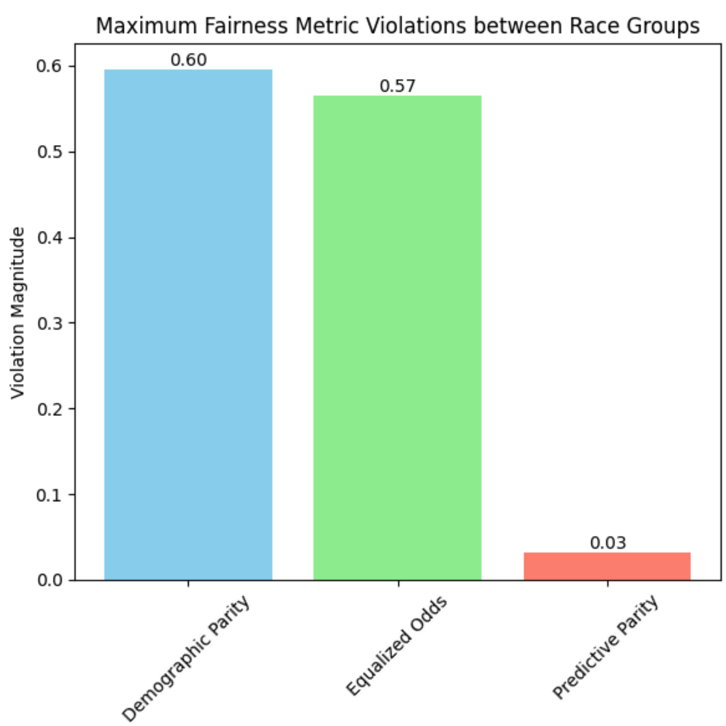


Figure 8



Works Cited

Buolamwini, J., & Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities In Commercial Gender Classification Proceedings of Machine Learning Research https://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf

De-Arteaga, M. (2024). S9-Class Ethics of AI: Measures of Algorithmic Bias. The University of Texas at Austin McCombs School of Business. *Canvas.* https://utexas.instructure.com/courses/1379214/files/76820760?module_item_id=13683842

Engler, A. (2021). Enrollment algorithms are contributing to the crises of higher education. *Research*. Brookings. https://www.brookings.edu/articles/enrollment-algorithms-are-contributing-to-the-crises-of-higher-education/

Koons, S. (2018). Researchers examine impact of law school admissions reform on universities. *Education*. https://www.psu.edu/news/education/story/researchers-examine-impact-law-school-admissions-reform-diversity/

Kuris, G. (2024). 7 Deciding Factors in Law School Admissions. Law Admissions Lowdown. U.S. News & World Report. https://www.usnews.com/education/blogs/law-admissions-lowdown/articles/7-deciding-factors-in-law-school-admissions#

Lowry, T. (2023). The Good, the Bad, and the Ugly of AI and Admissions. *Law:Fully*. LSAC. https://www.lsac.org/blog/good-bad-and-ugly-ai-and-admissions

Merritt, D. (2016). The Gender bias in Law School Admissions. Business & Practice. Bloomberg Law.

https://news.bloomberglaw.com/business-and-practice/the-gender-bias-in-law-school-admissions#

Nussbaumer, J. (2006). Misuse of the Law School Admissions Test, Racial Discrimination, and the De Facto Quota System for Restricting African-American Access to the Legal Profession. *St. John's Law Review*, *80* (Winter 2006), 162–182. https://scholarship.law.stjohns.edu/cgi/viewcontent.cgi?article=1180&context=lawreview

Spencer, Christian. (2021) "More than a Third of White Students Lie about Their Race on College Applications, Survey Finds." The Hill, The Hill, 1 Nov. 2021, thehill.com/changing-america/enrichment/education/577722-more-than-a-third-of-white-students-lie-about-their/.

Usylesses. (2022). LSAC dataset - EDA and predictions. Kaggle. https://www.kaggle.com/code/eds8531/lsac-dataset-eda-and-predictions

Wightman, L., Ofer, D (2022). Law School Admissions Bar Passage. Kaggle. https://www.kaggle.com/datasets/danofer/law-school-admissions-bar-passage/data.