

BIAS AUDIT OF MORGAN STANLEY'S USE OF EIGHTFOLD MODEL FOR SCORING APPLICANTS

July 31, 2024

Introduction

This report presents the results of a "Bias Audit" by BLDS, LLC, conducted in accordance with New York City's Automated Employment Decision Tools ordinance (Local Law 144, Int. 1894-2020) (herein, "New York City Ordinance"), of Morgan Stanley's use of the algorithmic model made available by Eightfold to screen applicants' resumes, thereby aiding in the user's evaluation of applicants to consider for employment.

Executive Summary

Morgan Stanley retained BLDS to conduct a bias audit of their use of the Eightfold model consistent with the New York City Ordinance. The Eightfold model scores applicants on how well the information on their resume matches the requirements for the job.

To conduct the bias audit, Morgan Stanley supplied BLDS with applicant records for the period 6/1/2023 through 6/1/2024 representing each applicant for a job. In total, BLDS analyzed the records of 254,539 applications.

In conducting the bias audit, BLDS computed the impact ratio for all applicants for each category required by the New York City Ordinance and the New York City Department of Consumer and Worker Protection's Notice of Adoption of Final Rule (Final Rule). BLDS followed the methodology set forth in the Final Rule, calculating the impact ratio based on a comparison of the scoring rates (defined by the law as the percent of applicants in each demographic group scoring above the median) by sex, race/ethnicity, and intersectional categories for all applicants.

Overall, our results show that Morgan Stanley's use of the Eightfold algorithmic model does not have a disparate impact on any protected group. The impact ratios for all categories ranges from 1 to .817, with the majority of the results being .9 or higher. An Impact Ratio of less than 0.80 is an initial indicator of adverse impact in selection under the federal government's Uniform Guidelines on Employee Selection Procedures. However, insofar as all of these ratios are above 0.80, they do not suggest a disparate impact.

BLDS is of the professional opinion that a more accurate bias audit would be conducted at the job level, i.e., considering the job for which the applicant applied. Thus, BLDS also computed the impact ratios following the same methodology required by the Final Rule, but we computed the median for each job rather than an overall median across all applicants for all jobs, and then summed overall by category the counts of those above the median score of the job for which they applied. The results are similar to those ignoring the job for which an applicant applied. The impact ratios for all categories, when the job applied for is considered, ranged from 1 to .847, which again, demonstrates that the Eightfold algorithmic model does not cause a disparate impact on any protected group.

Requirements of the New York City Ordinance

The New York City Ordinance specifies that a bias audit is required to be conducted of any automated employment decision tool (“AEDT”) used by an employer or employment agency within New York City before using the AEDT, and an employer “may not use or continue to use an AEDT if more than one year has passed since the most recent bias audit of the AEDT.” The New York City Ordinance defines an AEDT as “any computational process, derived from machine learning, statistical modeling, data analytics, or artificial intelligence, that issues simplified output, including a score, classification, or recommendation, that is used to substantially assist or replace discretionary decision making for making employment decisions that impact natural persons.” The Final Rule further defines the phrase “to substantially assist or replace discretionary decision making” to mean “(i) to rely solely on a simplified output (score, tag, classification, ranking, etc.), with no other factors considered, (ii) to use a simplified output as one of a set of criteria where the simplified output is weighted more than any other criterion in the set, or (iii) to use a simplified output to overrule conclusions derived from other factors including human decision making.”

BLDS has no independent opinion and has conducted no analysis to determine if Morgan Stanley’s use of the Eightfold model meets the definition of an AEDT under the New York City Ordinance or the Final Rule. Nevertheless, Morgan Stanley has retained BLDS to conduct a bias audit of the Eightfold model consistent with the requirements of the New York City Ordinance and Final Rule.

The Eightfold Model

The Eightfold algorithmic model produces a match score which represents how well a prospective job applicant’s resume matches the written job description for which they applied.

The Eightfold model has two inputs: (1) a job description for the relevant position, and (2) each applicant’s resume, which is then compared to the job description. Thus, at a high level, the model scores how good of a fit there is between the job description and a given resume. Morgan Stanley considers this objective output as one of the factors a recruiter will consider in deciding who to advance in the hiring process.

The model produces a match score which ranges from .5 to 5, with a higher match score indicating a better match between an applicant’s resume and the job requirements stated for the job.

Class Membership Information

The New York City Ordinance relies on the EEOC’s seven race/ethnicity groups¹ and requires that separate impact ratios be calculated for: (a) all race/ethnicity groups; (b) two sex groups (male/female); and (c) groups constructed inter-sectionally between sex² and the racial/ethnic groups. This results in fourteen groups. The highest scoring category, however, remains the reference group.

The Final Rule allows “an independent auditor to exclude a category that represents less than 2% of the data being used for the bias audit from the required calculation of the impact ratio.” However, the auditor “must include the justification for the exclusion and report the number of applicants and scoring rate or selection rate for the excluded category.”

¹ The EEOC’s seven race/ethnicity groups are: (1) American Indian or Alaska Native; (2) Asian; (3) Black or African American; (4) Hispanic or Latino; (5) Native Hawaiian or Other Pacific Islander; (6) Two or More Races; and (7) White.

² The only available options are Male and Female.

When conducting a bias audit, it is necessary to have demographic information (e.g., sex and race) for each observation. Consistent with legal requirements, applicants may choose whether or not to self-report race/ethnicity and sex information; some applicants choose to provide that information, and some do not. Our studies use the self-reported race and sex identifications. We excluded from our analysis any data for applicants for whom we do not have self-reported race/ethnicity or sex, but we report the number of such cases deleted from each analysis as required by the New York City Ordinance.

The Impact Ratio and Scoring Rate

Morgan Stanley supplied BLDS with applicant records from the period 6/1/2023 through 6/1/2024 for all individuals who applied for any job opening. Based on the available applicant records, we analyzed a total of 254,539 .

The Final Rule defines the impact ratio of an AEDT that an employer or employment agency does not use as a pass/fail selection device or assign the applicant to a classification as follows:

$$\text{impact ratio} = \frac{\text{scoring rate for each category}}{\text{scoring rate of the highest scoring category}}$$

The Final Rule defines the “scoring rate” as “the rate at which individuals in a category receive a score above the sample median score, where the score has been calculated by an AEDT.”

The New York City Ordinance and Final Rule suggest that a bias audit should report the overall results for all applicants scored by the AEDT.

Bias Audit Results for Morgan Stanley’s Use of the Eightfold Model

The impact ratio for the combined overall jobs for the various categories are presented in the following tables.

Sex Categories

	Number of Candidates	Scoring Rate	Impact Ratio
Female	86,561	37.9%	0.966
Male	159,121	39.2%	1.000
Unknown Sex	8,857		

Race/Ethnicity Categories

	Number of Candidates	Scoring Rate	Impact Ratio
Hispanic or Latino	25,712	42.8%	1.000
White (Not Hispanic or Latino)	75,139	40.0%	0.934
Black or African American (Not Hispanic or Latino)	25,719	40.3%	0.942
Asian (Not Hispanic or Latino)	98,764	36.1%	0.844
Two or More Races (Not Hispanic or Latino)	8,819	40.2%	0.939
Excluded AI/AN and NH/PI	1,102		
Unknown Race/Ethnicity	19,284		

Intersectional Categories

		Number of Candidates	Scoring Rate	Impact Ratio
Hispanic or Latino	Male	17,552	42.9%	1.000
	Female	8,075	42.6%	0.993
White (Not Hispanic or Latino)		53,875	40.2%	0.936
Black or African American (Not Hispanic or Latino)	Male	15,290	40.3%	0.939
Asian (Not Hispanic or Latino)		58,627	36.9%	0.858
Two or More Races (Not Hispanic or Latino)		5,819	39.8%	0.928
White (Not Hispanic or Latino)		21,169	39.5%	0.921
Black or African American (Not Hispanic or Latino)	Female	10,366	40.4%	0.942
Asian (Not Hispanic or Latino)		39,926	35.1%	0.817
Two or More Races (Not Hispanic or Latino)		2,857	40.9%	0.952
Excluded AI/AN and NH/PI		1,093		
Unknown Intersectional		19,890		

The above charts report the impact ratio for each sex category, race/ethnicity category, and the intersection of the sex and race/ethnicity categories based on the median being determined for overall applicants, regardless of for which job they applied. As noted above, the impact ratios for all categories ranges from 1 to .817. Again, all of these ratios are above 0.80, the level at which the federal government's Uniform Guidelines on Employee Selection Procedures state that, as a rule of thumb, disparate impact may exist.

As noted above as well, BLDS is of the professional opinion that a more accurate bias audit would be conducted at the job level, i.e., by counting the number in each category who applied for a certain job who scored above the median score of all applicants who applied for that job. Thus, BLDS also computed the impact ratios following the same methodology required by the Final Rule but computed the median for each job rather than an overall median across all applicants for all jobs and summed by category the counts of those above the median score of the job for which they applied. The results are similar to those ignoring the job applicants but the impact ratios more accurately ranged from 1 to .847.

The New York City Ordinance and Final Rule also allow the removal of categories that comprise less than 2% of the population from the calculation of the impact ratio. Thus, we did not include the Native American or Alaska Natives or the Native Hawaiian or Pacific Islanders in computing the impact ratios because both categories had less than 0.5% of the population. The inclusion of such small numbers would allow the race/ethnicity or intersectional categories of Native American or Alaska Natives or Native Hawaiian or Pacific Islanders to be the highest selection rate based on a trivial number of cases. Using such a small sample as the reference group to judge other categories would be questionable because the standard for judging the results of other categories for many jobs/sectors would be based on only a handful of cases.¹ The table below reports the data on the categories we did not use in computing the impact ratio.

Race/Ethnicity Categories

	Number of Candidates	Scoring Rate
Native Hawaiian or Pacific Islander (Not Hispanic or Latino)	448	44.0%
Native American or Alaska Native (Not Hispanic or Latino)	654	37.8%

Intersectional Categories

	Number of Candidates	Scoring Rate
Native Hawaiian or Pacific Islander (Not Hispanic or Latino)	266	41.4%
Male		
Native American or Alaska Native (Not Hispanic or Latino)	453	36.0%
Native Hawaiian or Pacific Islander (Not Hispanic or Latino)	179	48.0%
Female		
Native American or Alaska Native (Not Hispanic or Latino)	195	42.1%

¹ While, professionally, we believe they should not be used in computing the impact ratio, if they were incorporated into the analysis, the effect on the impact ratio of the other would be at worst a trivial decline and no category above 80% would become less than 80%.

Conclusion

In evaluating the impact ratios for the sex categories, the race/ethnicity categories, and the intersectional categories by combining the data over all jobs, as indicated by the New York City Ordinance and Final Rule, we do not find any instances in which the impact ratios fall below 0.80 (or 80%).

Appendix A – Background on BLDS, LLC

BLDS provides statistical and economic expert testifying and consulting services, primarily in the areas of compliance with anti-discrimination laws and regulations, with a major focus in labor and employment and consumer lending. BLDS has worked with numerous regulatory agencies, many Fortune-500 (and smaller) companies, and several state and local governments including New York City to apply statistical methods to identify and mitigate discrimination and bias in decision-making processes. BLDS's lead partner, Bernard Siskin Ph.D. was appointed by the U.S. Third Circuit as their statistical expert in assessing racial and gender fairness in the Court System. In our regulatory compliance work, BLDS assists clients in developing and implementing best practices for compliance with anti-discrimination laws, particularly as they relate to the use of statistical methods to make employment, credit, marketing, and insurance underwriting decisions. We have been retained by numerous agencies and regulators, including the EEOC, OFCCP, DOJ, and CFPB.