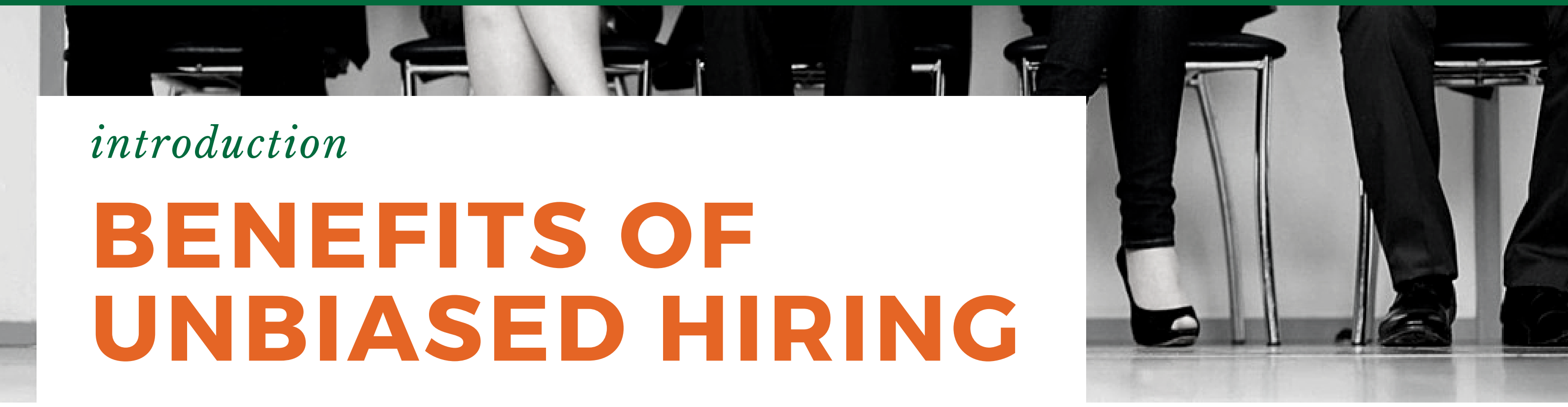
COMBATING ALGORITHMIC BIAS IN RECRUITING

THE PURSUIT OF EQUITABLE AND TRANSPARENT MACHINE LEARNING IN HUMAN RESOURCES

PREPARED BY

Heidi Fisher 'T'19



introduction

BENEFITS OF UNBIASED HIRING

Combatting algorithmic bias in hiring serves multiple purposes beyond the intrinsic value of pursuing fairness in the hiring process. First, **mitigating bias in algorithms used in talent acquisition protects the organization from unintended discrimination and related potential litigation.** While it may be difficult for an individual applicant to prove their resume was screened out due to a biased algorithm, the defensibility of algorithms has increasingly been tested in courts across the country (often at great expense).

Second, **algorithms can be refined to promote the organization's broader goals in recruiting,** like increasing the average tenure of employees. As these strategic priorities are often accomplished over the long term and difficult to incorporate in daily decision-making, the added input of data-driven tools encourages hiring managers to promote the organization's goals with every hire.

INCLUSIVE TEAMS MAKE BETTER BUSINESS DECISIONS UP TO **87%** OF THE TIME

AND MAKE THOSE DECISIONS **2X** AS FAST

DELIVERING **60%** BETTER RESULTS

Lastly, **unbiased hiring can promote better financial performance.** Evidence is mounting that diverse teams perform better:

25% **UPTICK IN GDP GROWTH PER CAPITA** over the past 50 years is attributed to increased labor force participation and contributions of women and people of color of both genders.

9.7% **MORE PROFITABLE VENTURE CAPITAL EXITS** were observed in firms that increased their female hires by 10%.

PHASES OF COMBATING BIAS

PREVENTION

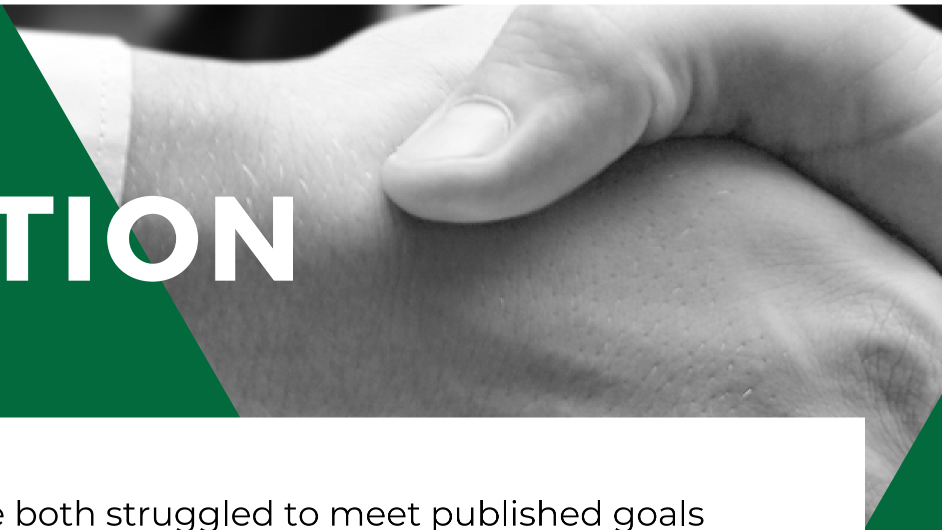
- Prevention of the inclusion of algorithmic bias should start with the team's formation
- Critical that teams evaluate their data for availability bias and represent a sufficiently wide range of experience to be able to do so comprehensively

DETECTION

- Budgets for every project should be required to conserve funds for constant monitoring
- Embracing a fire warden mentality is a cultural necessity
- Actively testing the efficacy of any project should become a procedural norm
- Organizations should consider a nested algorithm approach

DETERRENCE

- The nested algorithm approach needs to be implemented as a continuous learning mechanism, rather than a onetime exercise
- It is critical that if a candidate is resurrected, that feedback be incorporated back into the primary algorithm



PREVENTION

case studies

Amazon and Microsoft have both struggled to meet published goals around increasing the diversity of their respective workforces. For these and many other companies, relying on the resumes and performance of their existing employees as the training data for new recruiting tools will serve to reinforce the existing employee demographics.

amazon

Amazon recently scrapped an algorithm it was testing for recruiting after it taught itself that alumnae of women's colleges were subpar candidates. Given that Amazon's managers were 75% male in 2014, it's unsurprising the algorithm struggled to find a pattern of successful current employees who attended those schools.

Microsoft

While Microsoft has made considerable investments to recruit more diverse candidates, representation of black employees has hovered at 4%. An algorithm trained in their past ten years of employment data would likely de-prioritize or flatly reject candidates who graduated from HBCU's.

:) Affectiva

Boston-based AI company Affectiva learned the importance of diverse teams in their Cairo office: its database of four billion facial images enables Affectiva's algorithms to interpret emotional expressions but included no images of women wearing head coverings until a Hijabi employee asked why she was not represented in any data set.

If organizations intend to use algorithms to help meet their diversity goals or in any way shift the hiring trend away from historical patterns, THEY MUST NOT ASSUME THAT THE MOST ACCESSIBLE DATA IS EITHER REPRESENTATIVE OR EFFECTIVE IN BUILDING RECRUITING TOOLS.

STEPS OF DETECTION

BUDGET FOR MONITORING

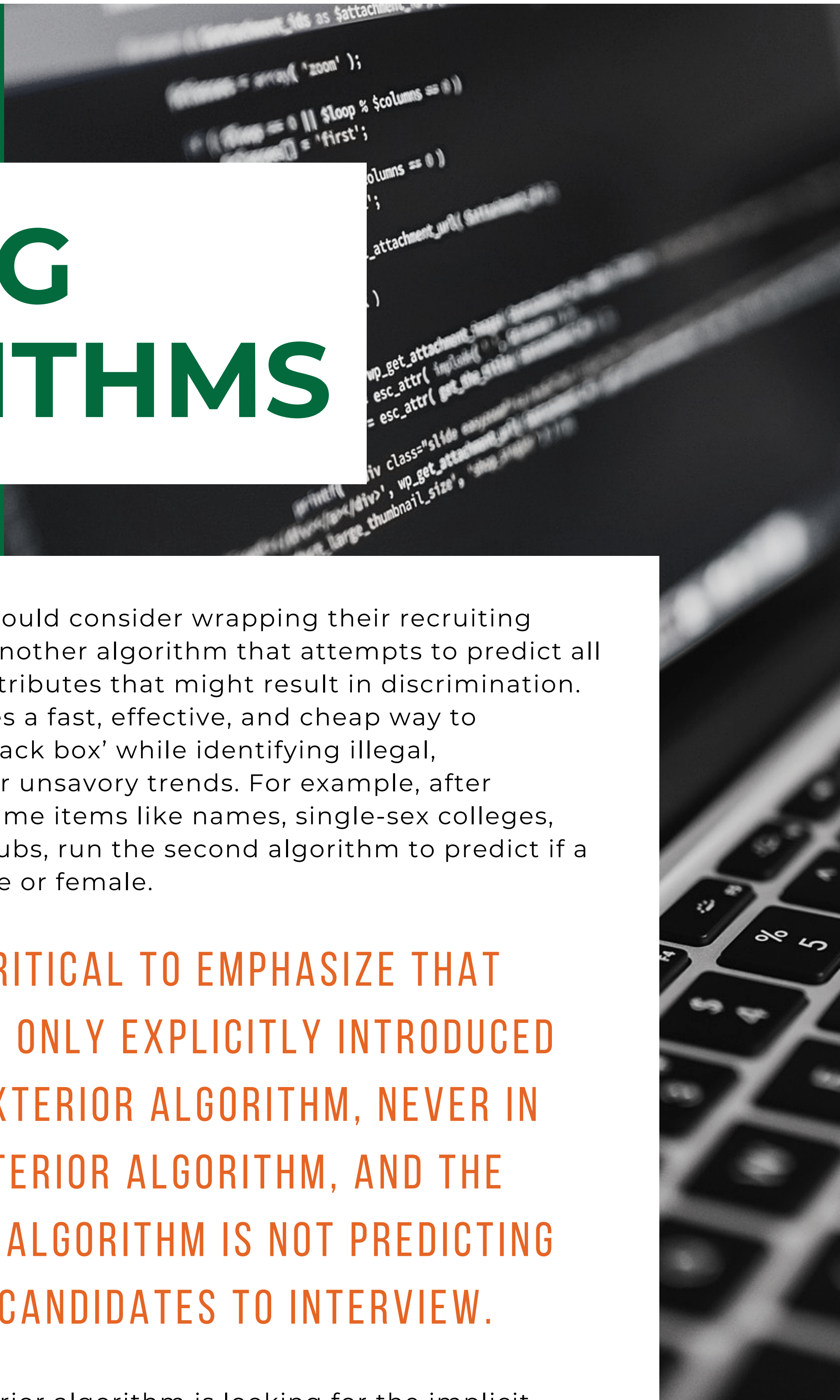
Budgets for every project should be required to conserve funds for constant monitoring. **Doing so not only encourages on-going attention to potential issues, but also signals a shift in priorities.** Projects that exhaust their budgets in development and implementation focus heavily on standing up a tool, risking a lack of attention to detecting undesirable or illegal outcomes.

FIRE WARDEN MENTALITY

Embracing a fire warden mentality is a cultural necessity. **The lack of effective government oversight in the use of machine learning programs necessitates that organizations self-regulate.** Developed by Accenture's AI lead Rumman Chowdhury, the Fire Warden model enables every team member to raise issues rather than relying on top-down police patrols. While it is not a fire warden's responsibility to solve the issue alone, it is their duty to follow the established alarm procedure to ensure each potential problem is evaluated and addressed appropriately. **Successful adoption of the Fire Warden model necessitates being comfortable with false alarms.** Much like the Toyota Production System where each worker is empowered to stop the line with one pull of the Andon cord, fire wardens need to be confident there is no penalty for raising issues and that the company culture would prefer to spend resources investigating false positives rather than risk perpetuating discriminatory biases.

ACTIVE EFFICACY TESTING

Actively testing the efficacy of any project should become a procedural norm. **Perhaps the easiest way to evaluate algorithms is the use of a control.** If the goal is to recruit high performers who will stay with the company for several years, comparing the average tenure of employees in the same position who were selected by the algorithm or by hand for interviews is an easy (albeit longer term) method. In the event that traditional recruiting methods outperform machine learning tools when measured against the organization's goals, **it is worth reevaluating the risk of inadvertently introducing algorithmic bias by using a homegrown solution.** If the updated recruiting process is not delivering substantive gains, an outside vendor might be a more pragmatic choice.

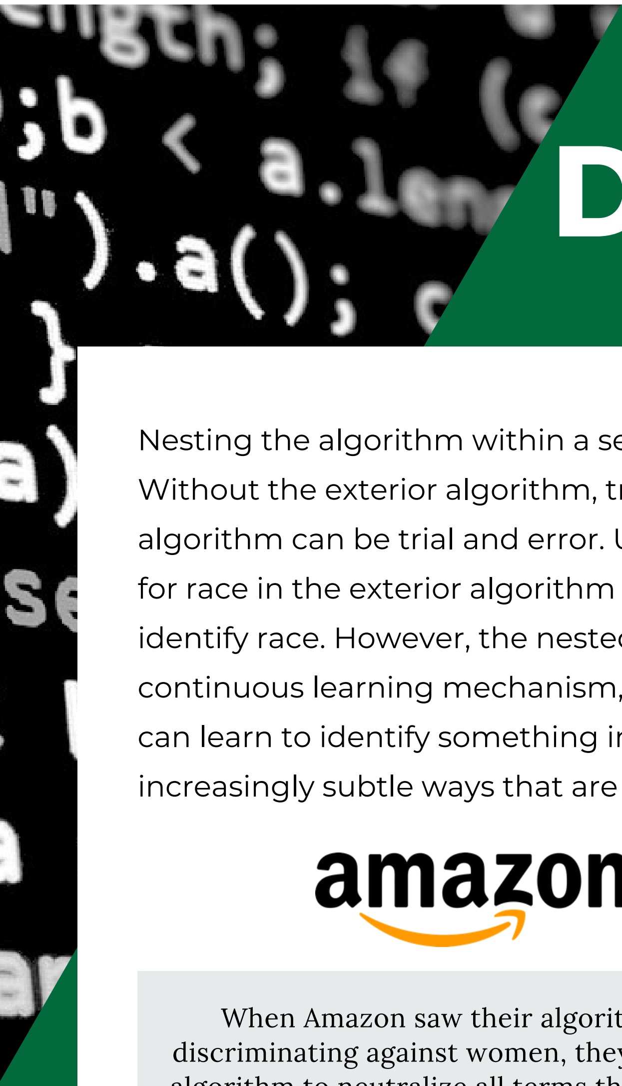


NESTING ALGORITHMS

Organizations should consider wrapping their recruiting algorithm with another algorithm that attempts to predict all the candidate attributes that might result in discrimination. Doing so provides a fast, effective, and cheap way to demystify the ‘black box’ while identifying illegal, discriminatory, or unsavory trends. For example, after neutralizing resume items like names, single-sex colleges, and gendered clubs, run the second algorithm to predict if a candidate is male or female.

> IT IS CRITICAL TO EMPHASIZE THAT GENDER IS ONLY EXPLICITLY INTRODUCED IN THE EXTERIOR ALGORITHM, NEVER IN THE INTERIOR ALGORITHM, AND THE EXTERIOR ALGORITHM IS NOT PREDICTING WHICH CANDIDATES TO INTERVIEW.

Instead, the exterior algorithm is looking for the implicit datapoints that may expose gender to understand if the interior algorithm is using gender as an input. Ideally, these predictions would be low-confidence results that match the candidate’s gender presentation at the interview with the fidelity of a coin toss. Realistically, some results may be uncomfortably accurate.



DETERRENCE

case studies

Nesting the algorithm within a second algorithm is also effective for deterring bias. Without the exterior algorithm, trying to curb something like racial bias in the primary algorithm can be trial and error. Using a nested algorithm, the most significant predictors for race in the exterior algorithm can help identify how the first algorithm taught itself to identify race. However, the nested algorithm approach needs to be implemented as a continuous learning mechanism, rather than a onetime exercise. If the primary algorithm can learn to identify something initially, it may re-learn that same identification in increasingly subtle ways that are harder for a manager to identify.

amazon

When Amazon saw their algorithm was discriminating against women, they edited the algorithm to neutralize all terms that explicitly identified gender, from sports teams to women's associations. After that adjustment, their algorithm relied on the hyper-masculine language (i.e. 'captured' and 'executed') more commonly used on male candidates' resumes to recommend candidates that most resembled the employee base used as training data. Resumes using neutral language that primarily belonged to female candidates were cast aside. Consequently, the importance of continuing to run the exterior algorithm after any adjustments to the primary algorithm are made cannot be overstated.

Microsoft

Google uses a secondary algorithm to review candidates to whom the primary algorithm declined to extend an interview invitation. Identifying candidates who struggled to differentiate themselves among a full field but have obvious potential on the second pass can help promote diverse hiring. It is critical that if a candidate is resurrected, that feedback be incorporated back into the primary algorithm. Doing so interrupts the feedback loops that could perpetuate biases against similar candidates.

It is critical that **IF A CANDIDATE IS RESURRECTED, THAT FEEDBACK BE INCORPORATED BACK INTO THE PRIMARY ALGORITHM.** Doing so interrupts the feedback loops that could perpetuate biases against similar candidates.

CONCLUSION

Talent acquisition is an imperfect art today. In fact, **simple equations have outperformed human decisions by 25% for hiring at all levels of organization.** In this context, it is clear why organizations are looking to recruiting algorithms as the next source of efficiency. Paramount to their success, however, is proper implementation and management.

While algorithmic bias is clearly a technical problem, **approaches to stymie bias cannot rely on technical solutions alone.** Adding controls, reconsidering availability bias, and nesting algorithms are strategies that need to be paired with commitments to change the culture of the organization around these projects. Successfully adopting a fire warden mentality, building diverse teams and promoting cross-functional collaboration may be just as important. What all of these tools have in common is the need for on-going attention; algorithmic bias cannot be quashed in a single action and combatting it requires constant vigilance at all phases of the project lifecycle.