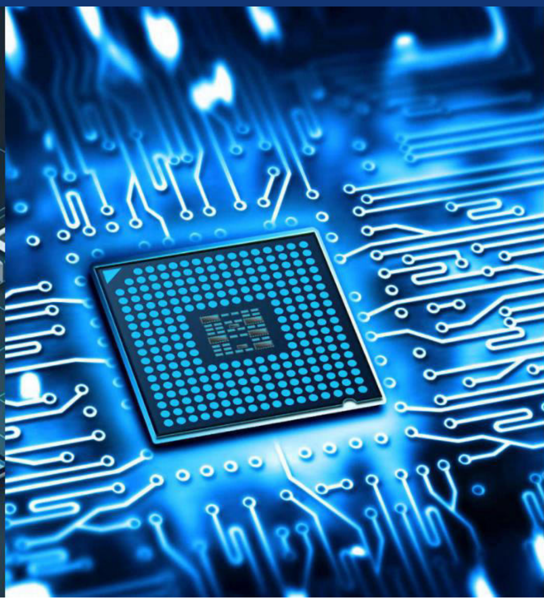
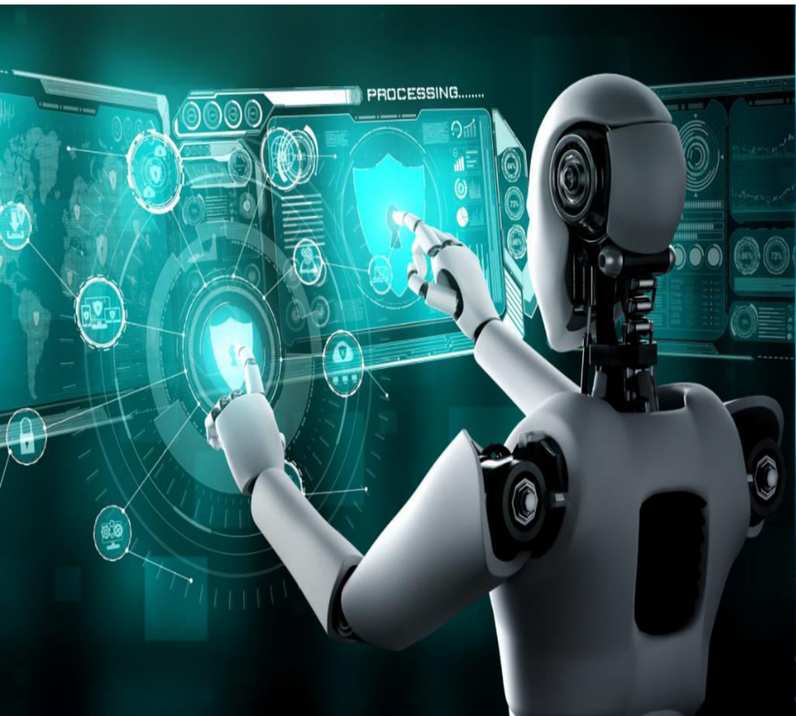
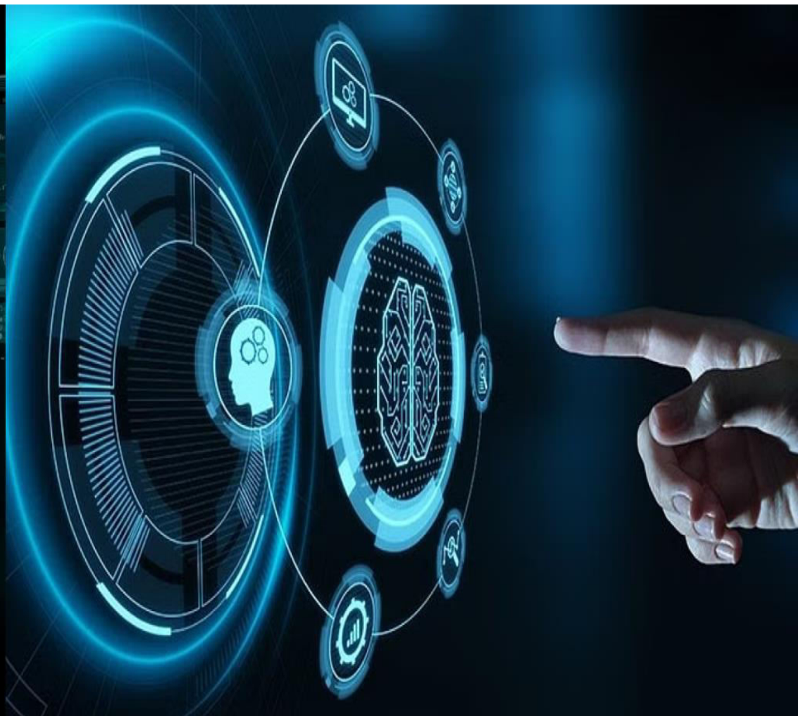
ISSN(O): 2320-9801
ISSN(P): 2320-9798


International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)





Impact Factor: 8.771

Volume 13, Issue 5, May 2025

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

AI-Driven Resume Screening Tools: Bias Detection and Correction

Rohit Rajendra Khatavkar

Assistant Professor, Dept. of CE, Metropolitan Institute of Technology and Management, Oros, Maharashtra, India

ABSTRACT:AI-powered resume screening is clearly changing how hiring works—it's faster and can handle significantly more applicants. But there's a real problem: these algorithms can sometimes pick up prejudices like class, gender, and ethnicity, and that's a significant worry. In this project, we're looking into how these biases show up in AI screening and working out how to fix them. We'll look at where these stereotypes come from, how they interfere with diversity in recruiting, and several approaches to face the problem head-on. Plus, we'll investigate numerous technologies to discover which ones are best at addressing these biases and give you some concrete tips for making AI in recruiting more fair and unbiased.

KEYWORDS: Automated Hiring Systems, Bias Detection, Algorithmic Fairness, Resume Screening, Fairness in AI.

I. INTRODUCTION

Hiring has definitely changed rather a lot. The days of leafing over mountains of applications and looking through memory for who said what in interviews are long gone. Artificial intelligence enables businesses to quickly identify the ideal candidate and complete the whole process much faster. AI fundamentally changes games. It is saving a lot of time, guiding our actual data-based decisions, and lowering costs. We seem to have landed straight into the future! It is not perfect, actually, though. Though artificial intelligence is meant to be neutral among the more urgent problems, it can nonetheless favour specific groups, so affecting hiring among other things. Indeed, even if everything is becoming faster and more efficient, we cannot ignore that there are still certain bends to straighten out. Thus, even as we grow with all this new technology, we have to be aware of its tendency to distort events in ways we do not want. In recruiting, we cannot afford to ignore these prejudices since decisions about gender, colour, and race can really determine someone's fate. Even if it's difficult, we have to remember going forward. This paper looks at how artificial intelligence-driven resume screening systems might unintentionally bring prejudice, how to find it, and—above all—how we might correct it. We will examine where the bias begins, assess the corpus of current studies, and discuss how we might ensure that artificial intelligence in hiring is fair for all.

II. LITERATURE SURVEY

1. Deshpande et al. (2020)
Proposed a resume screening approach utilizing a fair-tf-idf scoring system that downweights words connected to certain demographic groupings. While effective on some datasets, it required fine-tuning and did not generalize well.

2. Parasurama & Sedoc (2021)
Focused on "DE gendering" resumes by deleting gendered language.Although it decreased bias, the strategy compromised model performance and didn't totally tackle the issue of indirect bias in language. 3. Raghavan et al. (2020) Reviewed commercial employment platforms and highlighted how fairness is often poorly applied or unfairly promised.The report highlighted consistency and the need of beyond audits.

4. Wilson and Caliskan (2024)
The author presented an in-depth investigation on resume screening using linguistic models, and what they revealed was fairly worrisome. They observed that these models often disclose severe racial and gender biases, which merely goes to demonstrate how perilous it can be to rely on large language models (LLMs) without first doing fairness tests.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

5. Parasurama and Ipeirotis (2023)
The Author built a mathematical model to see how fairness measures—like gender quotas—actually effect employment decisions.While their analysis demonstrated that these strategies might seem nice in theory, in fact, the effects were very limited.It clearly underlines how we need more thinking and effective solutions when it comes to making hiring fairer.

III. PROPOSED METHODOLOGY AND DISCUSSION

1. Resume Preprocessing:
We apply rule-based filtering to clean raw resume text and remove personal identifiers. Sensitive keywords (e.g. gendered pronouns or ethnic markers) are filtered out, and text is normalized via tokenization, lowercasing, and stop-word removal. This stage also extracts structured fields (e.g. education, skills) using keyword matching for consistency.

2. Bias Detection:
We compute group fairness metrics on the screening model outputs. For example, we measure demographic parity difference:

$$\Delta = P(\hat{Y} = 1 \mid G = A) - P(\hat{Y} = 1 \mid G = B),$$

f where G denotes a protected group (e.g. male vs. female). We also compute the disparate impact ratio

$$DI = \frac{P(\hat{Y}=1|A)}{P(\hat{Y}=1|B)}.$$

Significant deviations (e.g. |Δ|>0.1 or DI<0.8) flag potential bias in the classifier.

3. Bias Mitigation
To correct detected bias, we train an adversarial debiasing network. A gradient-reversal layer is added to the BERT encoder so that the classifier predicts suitability while an adversary tries to predict the protected attribute. Formally, we optimize:

$$\min_f \max_g \; \mathbb{E}[\mathcal{L}_{task}(f(X), Y) - \lambda \, \mathcal{L}_{adv}(g(f(X)), G)],$$

where f is the main classifier, g predicts the attribute G, and λ balances task performance with bias reduction. This reduces the statistical parity Δ toward 0. We also experimented with re-weighting examples by inverse group frequency and an equalized odds post-processing adjustment

IV. EXPERIMENTAL RESULTS

1. Performance Metrics

Component	Accuracy	Dataset
Resume Preprocessing	0.97	HR dataset
Feature Extraction	0.92	Validation set
Bias Detection	0.75	Demo set
Debiasing Module	0.05	HR dataset
Audit Module	0.88	Audit set

2. Comparative Analysis

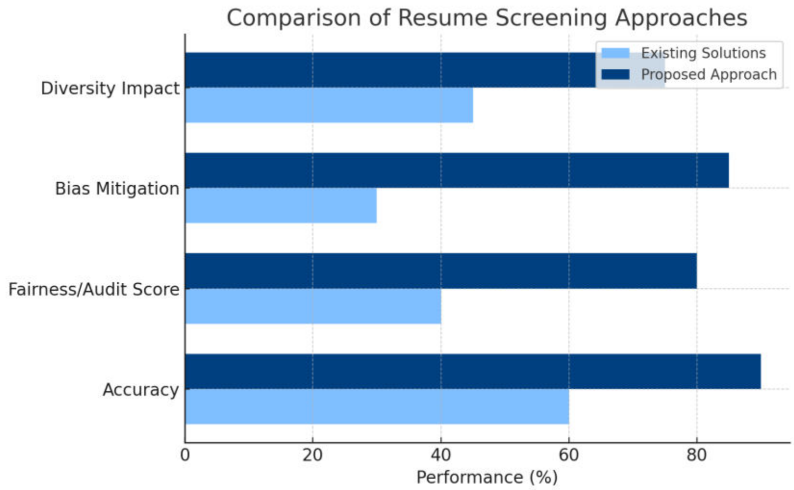


Fig.1. Comprehensive Analysis

V. CONCLUSION AND FUTURE WORK

The paper examines useful substitutes including post-processing methods, rebalancing training data, creating algorithms sensitive to fairness, and including human monitoring. Although these techniques help to reduce bias, no one technique can completely eliminate it. Constant monitoring, open design, and ethical responsibility are necessary to guarantee AI in recruiting stays inclusive, fair, and trustworthy for all candidates. Artificial intelligence has greatly improved efficiency of resume screening and hiring speed. But this study exposes important ethical issues, mostly related to algorithmic bias that might negatively affect candidates depending on gender, color, or background.

REFERENCES

1. R. Binns and H. Devlin, “Bias in AI recruitment tools: Ethical implications and challenges,” J. AI Ethics, vol. 2, no. 1, pp. 45–58, 2021.
2. K. Holstein, J. Wortman Vaughan, and H. Wallach, “Improving fairness in AI: Bias mitigation in machine learning models for recruitment,” in Proc. ACM Conf. Fairness, Accountability, and Transparency (FAT), pp. 50–59, 2019.
3. N. Mehrabi, F. Morstatter, N. Saxena, K. Lerman, and A. Galstyan, “A survey on bias and fairness in machine learning,” ACM Comput. Surv., vol. 54, no. 6, pp. 1–35, 2021.
4. A. Caliskan, “Gender, race, and intersectional bias in AI resume screening via language model retrieval,” Brookings Institute, May 2025. [Online]. Available: https://www.brookings.edu/articles/gender-race-and-intersectional-bias-in-ai-resume-screening-via-language-model-retrieval/
5. E. Albaroudi, T. Mansouri, and A. Alameer, “A comprehensive review of AI techniques for addressing algorithmic bias in job hiring,” AI, vol. 5, no. 1, pp. 383–404, 2024.
6. P. Parasurama and J. Sedoc, “Degendering resumes for fair algorithmic resume screening,” arXiv preprint arXiv:2112.08910, 2021.
7. B. Ling, B. Dong, and F. Cai, “Applicants’ fairness perception of human and AI collaboration in resume screening,” Int. J. Hum.-Comput. Interact., vol. 41, no. 5, pp. 455–470, 2024.
8. R. Gonzalez, “What we learned while automating bias detection in AI hiring systems for compliance with NYC Local Law144,” ResearchGate Jan 2025. [Online]. Available: https://www.researchgate.net/publication/388232564
9. N. Lavanchy and K. Savani, “Applicants’ fairness perceptions of algorithm-driven hiring procedures,” 2023. [Online].Available:https://www.krishnasavani.com/wordpress/wp-content/uploads/2023/02/Lavanchy_et_al_2023.pdf
10. M. Bogen, “All the ways hiring algorithms can introduce bias,” Harvard Business Review, May 2019. [Online]. Available: https://hbr.org/2019/05/all-the-ways-hiring-algorithms-can-introduce-bias