The effects of AI on underrepresented groups in recruitment processes

Mariella Inés Hemminga

Identification of subject

My research focuses on the intersection of technology and human resource management. Specifically, I examine how the implementation of AI-driven systems affects inclusive recruitment practices. By analyzing the origin and impact of algorithmic bias, I seek to identify strategies to mitigate its adverse effects and promote fairness in candidate selection.

Personal motivation and rationale

My previous work experience in human resources drives my interest in exploring the use of AI in recruitment. While AI is widely discussed within HR, the persistent issue of algorithmic bias and its impact on inclusive hiring remains underexplored. My research addresses this gap by analyzing how AI-driven processes influence hiring conditions. I aim to provide insights that can help businesses incorporate fairer recruitment practices to enhance organizational equity and performance. This research holds value not only for the academic community but also for businesses striving to create inclusive workplaces.

Research question

My central research question is whether bias in AI-driven recruitment algorithms impacts the inclusion of underrepresented groups in candidate selection processes. While these systems promise efficiency and objectivity, I hypothesize that they pose a significant risk of perpetuating existing biases, thus creating less diverse work environments.

Literature review

Artificial intelligence (AI) in recruitment refers to systems capable of interpreting external data and achieving specific goals through adaptive decision making (Kaplan and Haenlein, 2019b). Central to AI is Machine learning (ML), which enables these systems to improve from experience without explicit programming (Kaplan and Haenlein, 2019b). ML models, trained on labeled data to identify patterns and make predictions, are continuously refined for better performance (Cormen, 2009; Kumar, 2017; Zheng et al., 2013). Despite its potential for enhancing efficiency and data-driven hiring, AI encounters challenges such as label bias, where unrepresentative data embeds bias into algorithms (Mehlin et al., 2018; Lebovitz et al., 2021).

However, algorithmic bias manifests in various forms, including gender, racial, and systematic bias, where human prejudices transfer into AI (Ntoutsi et al., 2020). For instance, the Amazon AI recruiter exhibited bias against women due to training on biased datasets from predominantly male applicants (Kaplan and Haenlein, 2019a; Taniguchi et al., 2018). Consequently, the AI recruiter failed to rate candidates for technical jobs in a gender-neutral manner. This highlights the risk of HR tasks being automated by algorithms, which could compromise fairness and inclusivity (Callen, 2021).

Bias in AI-driven recruitment often stems from flawed algorithm design, biased training data, and a lack of diversity in development teams (Kaplan and Haenlein, 2019a; Shrestha et al., 2021; Ntoutsi et al., 2020). In fact, they arise from training with small or non-representative datasets that encode gender, ethnic, and cultural prejudices (Mehling et al., 2018). Moreover, Scholz (2017) argues that ambiguity in job expectations and requirements passed on to AI developers can compound these issues. However, according to Khandelwal (2018) and Gikopoulos (2019), AI has the capacity to speed up hiring processes, improve the quality of candidate selection and mitigate ingrained biases in recruitment. By leveraging algorithmic analysis, AI can investigate a larger applicant pool without being influenced by human heuristics or unconscious biases (Charlwood and Guenole, 2022). Nonetheless, the profit-driven focus of firms and the tendency to favor automation can lead to greater surveillance and monitoring technologies, raising concerns about the ethical and societal implications of widespread AI adoption in HR (Kellogg et al., 2020; Markoff, 2016). As a result, the challenge is not exclusively the inability to resolve issues of biased AI systems, but also the lack of interest among AI developers in addressing the ethical and societal consequences of AI deployment (Birhane, 2021; Crawford, 2021).

Hence, effective collaboration and knowledge sharing between HR managers and AI developers are crucial, as HR professionals provide insights into job requirements, while AI developers leverage this domain knowledge to create less biased systems (Scholz, 2017). This process encompasses understanding job requirements beforehand, harmonizing data during development, and correcting biases through retraining afterward (ibid.). Additionally, it is crucial to engage stakeholders affected by AI in the design and deployment stages (Leslie, 2019). Yet, identifying algorithmic bias remains complex due to the challenge of accurately representing minority groups in training datasets (Chouldechova and Roth, 2020). Despite the European Commission's emphasis on safeguards in AI for hiring, enforcement mechanisms are still lacking (De Stefano and Aloisi, 2021; Veale and Borgesius, 2021).

Methodology

My methodology encompasses three key stages. Firstly, I provide an overview of artificial intelligence (AI) in recruitment, delving into its definition, function, and relation to machine learning (ML). This sets the foundation for understanding the complexities and implications of AI-driven HR practices. Secondly, I examine the challenges and biases associated with AI in recruitment, analyzing various types of algorithmic bias and their impact on the hiring process. This involves reviewing relevant literature to illustrate a real-world example of bias in AI systems. Lastly, I explore mitigation strategies for addressing bias, focusing on cross-functional collaboration between HR professionals and AI developers, along with the importance of domain knowledge and stakeholder engagement. By applying this structure, I aim to provide a comprehensive analysis of the challenges presented by AI in recruitment while offering insights into potential pathways for mitigating bias and promoting ethical AI practices in HR.

References

Birhane, A. (2021) 'Algorithmic injustice: A relational ethics approach', *Patterns*, 2(2), 1–9.

Callen, A. (2021) 'When knowledge work and analytical technologies collide: The practices and consequences of black boxing algorithmic technologies', *Administrative Science Quarterly*, 66(4), 1173–1212.

Charlwood, A., & Guenole, N. (2022) 'Can HR adapt to the paradoxes of artificial intelligence?', *Human Resource Management Journal*, 32(4), 729–742. Available at: https://doi.org/10.1111/1748-8583.12433.

Chouldechova, A., & Roth, A. (2020) 'A Snapshot of the frontiers of fairness in machine learning', *Communications of the ACM*, 63(5), 82–89. Available at: https://doi.org/10.1145/3376898.

Cormen, T. (2009) ‘Introduction to algorithms’, *MIT Press.*

Crawford, K. (2021) ‘Atlas of AI’, *Yale University Press*.

De Stefano, V., & Aloisi, A. (2021) ‘Artificial intelligence and workers’ rights’, Available at: https:// socialeurope.eu/artificial-intelligence-and-workers-rights (Accessed: 10 May 2024).

Gikopoulos, J. (2019) ‘Alongside, not against: balancing man with machine in the HR function’, *Strategic HR Review*, Vol. 18 No. 2, pp. 56-61. Available at: https://doi.org/10.1108/SHR-12-2018-0103.

Kaplan, A., & Haenlein, M. (2019a) ‘Rulers of the world, unite! The challenges and opportunities of artificial intelligence’, *Business Horizons*, 63(1), 37–50. Available at: doi:10.1016/j.bushor.2019.09.003.

Kaplan, A., & Haenlein, M. (2019b) ‘Siri, Siri, in my hand: Who’s the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence’, *Business Horizons*, 62(1), 15–25. Available at: doi:10.1016/j. bushor.2018.08.004.

Kellogg, K., Valentine, M., & Christin, A. (2020) ‘Algorithms at work: The new contested terrain of control’, *Academy of Management Annals*, 14(1), 366–410.

Kumar, R. (2017) ‘Machine Learning and Cognition in Enterprises’, *Apress*. Available at: https://doi.org/10.1007/978-1-4842-3069-5.

Lebovitz, S., Levina, N., Lifshitz-Assaf, H., & Lifshitz-Assa, H. (2021) ‘Is AI-ground truth really true? The dangers of training and evaluating AI tools based on experts’ know-what’, *MIS Quarterly*, 45(3), 1501–1526.

Leslie, D. (2019) ‘Understanding artificial intelligence ethics and safety: A guide for the responsible design and implementation of AI systems in the public sector’, *The Alan Turing Institute.*

Markoff, J. (2016) ‘Machines of loving grace: The quest for common ground between humans and robots’, *Ecco Press*.

Mehling, M. A., Asselt, V., Das, K., & Droege, S. (2018) ‘Design AI so that it is fair’, *Springer Nature*, 559, 324–326.

Ntoutsi, E., Fafalios, P., Gadiraju, U., Iosifidis, V., Nejdl, W., Vidal, M. E., Ruggieri, S., Turini, F., Papadopoulos, S., Krasanakis, E., Kompatsiaris, I., Kinder-Kurlanda, K., Wagner, C., Karimi, F., Fernandez, M., Alani, H., Berendt, B., Kruegel, T., Heinze, C., & Staab, S. et al. (2020) ‘Bias in data-driven artificial intelligence systems—An introductory survey’, *Wiley Interdisciplinary Reviews. Data Mining and Knowledge Discovery*, 10(3), 1–14. Available at: doi:10.1002/widm.1356.

Scholz, T. M. (2017) ‘Big Data in Organizations and the Role of Human Resource Management’, *Peter Lang International Academic Publishers*. Available at: https://doi.org/10.3726/b10907.

Soleimani, M., Intezari, A., & Pauleen, D. J. (2022) ‘Mitigating Cognitive Biases in Developing AI-Assisted Recruitment Systems: A Knowledge- Sharing Approach’, *International Journal of Knowledge Management (IJKM),* 18(1), 1-18. Available at: http://doi.org/10.4018/IJKM.290022.

Taniguchi, H., Sato, H., & Shirakawa, T. (2018) ‘A machine learning model with human cognitive biases capable of learning from small and biased datasets’, *Scientific Reports*, 8(1), 7397. Available at: doi:10.1038/s41598-018-25679-z PMID:29743630.

Veale, M., & Borgesius, F. (2021) ‘Demystifying the draft EU Artificial Intelligence Act’, *Computer Law Review International*, 22(4), 97–112.

Zheng, H., Kulkarni, S., & Poor, V. (2013) ‘A sequential predictor retraining algorithm and its application to market prediction’, *Annals of Operations Research*, 208(1), 209–225. Available at: doi:10.1007/s10479-013-1396-2.

Keywords

Artificial intelligence: AI refers to the simulation of human intelligence in machines that are programmed to think and learn. These systems can interpret external data and use the acquired knowledge to achieve specific goals.

Machine learning: Machine learning enables systems to automatically learn and improve from experience without being explicitly programmed.

Algorithmic bias: Algorithmic bias is a systematic and repeatable error in decision-making that results from flawed assumptions or unintended preferences. This bias can lead to unfair outcomes.