

Assignment: Essay

Topic: AI application in Human resource management is creating biases

Name: Pinsi Wang

Course: Creative Computing

Word count: 2485

Introduction

Technological emergence significantly impacts several organizational operations, including Human resource management. Artificial Intelligence, or AI, is an emerging technology transforming business operations to a new level. As per a report, 80% of global organizations use AI technology within their Human Resource Management (HRM) system (Brin, 2019). However, with its several applicability, AI technology also brings the concern of biases. On the other hand, researchers suggested that AI is an impressive tool for removing bias. In this context, a study stated that using AI technology for hiring by HRM could effectively avoid unconscious bias (Raveendra, Satish and Singh, 2020).

Additionally, this study stated that bias-free hiring would ensure the quality and effectiveness of hiring and help any organisation gain more advantages. This essay will discuss the contemporary issues of AI application biasing in HRM. The essay will start by explaining the concept of AI in the HRM paradigm. Moreover, critical issues within AI applications will be briefly discussed, which will be expanded to the biasing problems caused by AI in HRM. Lastly, the future of AI in HRM will be outlined.

Discussion

Digital technologies have drastically altered the way of doing business across many industries. Automated machines and systems now perform repetitive and routine operations without the involvement of human interaction. Artificial Intelligence (AI) is a similar technology designed to reason and think like a human. This computational thinking technology possesses cognitive skills comparable to those of humans and consciousness, self-awareness, and the ability to think independently of human assistance (Zehir et al., 2020). Academicians have made many attempts to define Artificial Intelligence (AI). McCarthy (2004) defined AI as the engineering and science of creating intelligent machines, particularly computer programs. Simply put, AI integrates computer science and massive datasets to facilitate problem-solving.

Moreover, it includes the branches of AI known as deep learning and machine learning, commonly addressed together (IBM, 2023). These fields use AI algorithms to build expert systems that predict or categorise information based on incoming data. It can be said that AI is computational technology that utilises existing datasets to process tasks.

HRM is perceived as the major change agent in this digital transformation era. Historically, HR was conceived as a qualitative function distinct from objective evaluations since HRM is highly dependent upon subjective analysis and intuitions related to human resources (Zehir et al., 2020). However, as the technology progressed, HRM became more data-oriented with the support of computational technologies such as AI and data analytics. Therefore, AI has now become an integral part of HRM functioning such as candidate search, talent acquisition, recruitment, development and learning, planning and strategy, workforce planning, turnover prediction, succession planning, performance evaluation, payoff and compensation, motivation and engagement, performance evaluation (Falletta, 2014). Thus, AI has now penetrated key HR functions.

HRM utilises a large dataset available from within and outside the organisation. By processing and analysing data sets, HRM greatly benefits their department and overall business operations. AI in HRM has changed and transformed the way HRM functions and the requirement for HR executives. Today, HR competencies include skills and abilities that can manage digitisation (Zehir et al., 2020). Moreover, HR has begun to develop and practice digital HR strategies that complement digital business initiatives. The most important function of HR in the current digital era is successfully leading the organisation's digital transformation. One of the biggest impacts of AI in HRM is performance management. AI in HRM uses a data-driven approach that uses real-time performance data and provides feedback to the workers and organisation (Buck and Morrow, 2018). This way, employees can check their feedback and improvise their performance without losing any time. Another important impact of AI is eliminating repetitive tasks. For example, HRM uses AI-based chatbots to manage frequently asked queries of employees (Lengnick-Hall, 2018). Thus, AI has multiple roles in HRM and which is gradually expanding.

The main reason for introducing AI in HRM is to remove biases and human errors while doing HR functions. For example, Maity (2019) mentioned that AI helps HR professionals to select the appropriate candidate for the job role while avoiding cognitive biases based on gender or race. On the contrary, in recent years, AI is becoming a concern for algorithmic bias that limits its HRM uses for hiring applicants. As per a study, the retail giant Amazon utilised an AI-driven tool to review and sort applicant profiles and faced an issue of bias (Dastin, 2022). As per the study, the AI algorithm ignored applicant profiles that mentioned

the term 'women' or 'female' while sorting. The AI model was based on the data from the last 10 years. The book stated that the AI-integrated system of Amazon taught itself using the global data of IT profiles with more males compared to the number of females working in several organisations and prefers male candidates over female candidates while sorting applicant profiles (Mehan, 2022). Therefore, the AI model learned that men are more preferred candidates for job roles than women from the previous dataset. Dastin (2022) also mentioned that gender and race bias due to the AI algorithm happens again when Facebook allows advertisers to promote brands based on race and gender using AI technology. Therefore, AI is creating the problem it was designed to eliminate, biasing.

AI systems may become biased due to the engineers' preconceived notions, the technological constraints of their design, user input, and the data set the algorithms to employ, depending on what is overrepresented or absent from the training data (Cathey, 2018). Jessica Wade, a physicist who wrote more than 300 Wikipedia entries, said that Wikipedia is heavily biased towards women in science (Vincent, 2018). Despite being the fifth most visited website in the world, only 18% of biographies on Wikipedia are of women (Cathey, 2018). Therefore, AI bias is an issue that is persistent in the HRM field and general society. AI easily learn the unconscious bias of the users.

Moreover, feedback from the user further strengthens this unintentional bias (Cathey, 2018). For example, if the AI algorithm asks the user to share their experience by indicating the quality of the result to be 'good' or 'bad'. If the user is from a particular gender or race, he or she may end up rating results for other gender or ethnicities as 'poor'. Therefore, relying on such a review system to train and improvise AI is almost impossible.

Theoretically, AI may help firms go through large amounts of resumes more cost-effectively, strategically, and efficiently. However, in practice, it may encourage biased hiring due to its dependence on implicitly biased selection tendencies like gender, demographics, and language. Several data specialists contend that because predictive AI is frequently trained on incomplete and biased data sets, it supports the established order (Parikh, 2021). Therefore, another argument favouring AI in HRM can be interpreted from this: AI bias results from existing human biases. Leech (2018) mentioned that 'artificial' intelligence is not unnatural but based on a misguided understanding of human intelligence. AI learn from the data set

created by human brains. Therefore, AI learn biases as well from those data sets. Bias is an interesting feature of human nature. Humans can be trained to recognise, identify, and mitigate conscious or unconscious bias. However, since AI mimics and simulates human intelligence, it cannot be aware of the unconscious bias it learns from humans. AI bias results from the program's incompetency that is highly dependent upon mimicking human intelligence.

AI bias can be broadly categorised into historical, technical, emergent, and representation. If the AI is trained on historical recruitment data, it will learn biases favouring a particular gender or race. For example, in the past, an organisation hired more white men than Black men. AI will learn this pattern and shortlist candidates based on this. Barocas and Selbst (2016) argued that in such cases, the AI designer does not intend to marginalise candidates, yet it turns into a systematic disadvantage to an entire group. The second bias type is a technical bias resulting from the system's technical limitation. For example, limitation of the data set, hardware or peripherals.

Moreover, human intelligence, such as decision-making, is often contextual and difficult to quantify (Friedman and Nissenbaum, 1996). Therefore, it is extremely difficult for AI to simulate such complex human constructs in their entirety. AI can probably miss the specific context, which causes technical bias.

The third type is representation bias, which occurs when the data set underrepresents or overrepresents a certain group (Suresh and Guttag, 2019). For example, an AI model includes a data set that is dominated by data set for white males and females and fewer data on Hispanic females. Hispanic females are an underrepresented group in such cases. This will lead to discriminatory results of HRM function. The fourth type of bias is emergent bias. This occurs when there is a shift in the context, such as societal knowledge, cultural values, or population (Friedman and Nissenbaum, 1996). Since new knowledge is not incorporated into the AI system, it will judge based on existing knowledge. For example, an AI program was developed for a highly competitive and individualistic culture. However, the problem will occur when this AI program is implemented in another society which promotes collectivism and avoids competition. Therefore, emergent bias is a mismatch between the system design and user data. It can be summarised that AI biases have various forms and occur over time.

In HRM, fairness and justice are important for establishing positive relations between the organisation and employees. The perception of subjective and objective fairness of the employees and potential candidates regarding the application of AI is taken into account in HRM. According to Cropanzano et al. (2007), perceived justice or fairness is a personal opinion rather than an objective evaluation of reality. Studies have shown that employees treated fairly are likelier to show highly ethical and selfless behaviour (Cohen-Charash and Spector, 2001). On the other hand, unfair treatment of employees may result in employee dissatisfaction. For example, accepting job offers is highly dependent upon the fairness perception of the recruitment process (Bauer et al., 2001). If the candidate perceives the recruitment process fair and transparent, it is more likely that the candidate will accept the job offer or show interest in the job role. Hence, it is significant that organisations must know about the perception of the candidates or employees about AI in HRM since it has an immense impact on the employees' attitude, performance, behaviour, morale and intentions.

Furthermore, bad AI experiences or negative perceptions of AI decisions can damage the organisational reputation. A similar thing happened when Amazon's recruitment process revealed the AI biasing issue. It leads to distrust among potential candidates. Therefore, organisational performance heavily depends upon the fairness and justice AI decisions can provide.

Researchers and developers have suggested several norms to solve such issues related to AI-based hiring by the HRM. In this respect, a study stated that the Algorithmic de-biasing technique is the best suited for mitigating the limitation in AI-driven tools (Raghavan et al., 2019). This way, the organisation can identify biases in the data set and avoid AI bias. Furthermore, the HR department should avoid implementing AI for a crucial process that requires a multidimensional human perspective. (Panwar, 2021). Suppose the organisation is not having sufficient data set to design an AI model for a particular task. In that case, AI should be skipped for that task until the organisation generates enough high-quality data set. The organization's HR department should regularly check their database manually for any bias and ensure that it benefits the organisation and is inclusive. Moreover, organisations should unite and share the responsibility to manage AI bias within the industry (Panwar, 2021). Some actions that should be implemented at the industry level to avoid AI bias in

HRM include holding people responsible in the field, challenging power dynamics and institutions that may not be advantageous to everyone, and promoting diversity.

Human society is not based on society but also on technical or digital environments that decisions made by humans regulate. Costanza-Chock (2020) mentioned in her book that the digital environment is responsible for creating an infrastructure of affordances and constraints that impact human society's operation, task, and structure. As mentioned in this essay, a digital environment such as AI has now transformed HRM infrastructure in the business world. However, the social side of this environment is essential to integrate with the digital environment for effective decision-making. Social and technical aspects need to be considered in the common conversation so that the system, i.e., HRM, can manage the challenge of AI bias. Costanza-Chock (2020) suggest using the design justice method to address the challenge of digital environments such as AI and Machine Learning in a social context. Design Justice Network Principles are based on designing processes that empower, sustain, support voices, share, and provide a solution to the problem of the community while uplifting and honouring the indigenous, traditional, and local knowledge (Costanza-Chock, 2020). This way, the design justice approach ensures that the design computational process, such as AI, benefits the marginalised communities. AI must be autonomous, just, and sovereign, beyond fairness and inclusivity. Therefore, not only does AI need to be made more inclusive, it must be decolonised to support and amplify local and indigenous knowledge. This way, AI bias can be eliminated not only for gender and nationality but also from racial injustice.

Conclusion

This essay aimed to discuss the contemporary issues of AI application biasing in HRM. It was found that, theoretically, AI was designed to remove biasing in HRM processes such as recruitment, training, and performance management. However, AI encourages biasing in practice due to its dependence on implicitly biased selection tendencies like gender, demographics, and language. The findings revealed that AI mimics human intelligence. Human intelligence works on a contextual basis and is complex to copy.

Moreover, the existing data set used to program AI algorithms includes biased human behaviour. Therefore, AI adopts biasing unintentionally. Fairness and justice in AI decisions

are crucial in job satisfaction, employee performance, morale and behaviour and overall organisational performance. Therefore, organisations must manage bias in AI decisions in the HRM context. The organisation's HR department could implement some recommendations to address AI bias. For example, implementing the Algorithmic de-biasing technique, adopting Design Justice Network Principle, and sharing responsibility to manage AI within the industry. Lastly, it can be concluded that digitalisation cannot be prevented from influencing HRM, but it can be transformed into something which creates high value for the organisation and society.

References

- Barocas, S. and Selbst, AD (2016) "Big data's disparate impact," *California law review*, 104, pp.671-732.
- Bauer, T.N., Truxillo, D.M., Sanchez, R.J., Craig, J.M., Ferrara, P. and Campion, M.A. (2001) "Applicant reactions to selection: Development of the selection procedural justice scale (SPJS)," *Personnel psychology*, 54(2), pp.387-419.
- Brin, D. W. (2019) *Employers embrace artificial intelligence for HR*, SHRM. Available at: <https://www.shrm.org/resourcesandtools/hr-topics/global-hr/pages/employers-embrace-artificialintelligence-for-hr.aspx> (Accessed: February 1, 2023).
- Buck, B. and Morrow, J. (2018) "AI, performance management and engagement: keeping your best their best," *Strategic HR Review*, 17(5), pp.261-262.
- Cathey, G. (2018) *The risk of bias in AI-powered HR tech & what you can do about it*, Linkedin. Available at: <https://www.linkedin.com/pulse/risk-bias-ai-powered-hr-tech-what-you-can-do-glen-cathey> (Accessed: February 28, 2023).
- Cohen-Charash, Y. and Spector, P.E. (2001) "The role of justice in organizations: A meta-analysis," *Organizational behavior and human decision processes*, 86(2), pp.278-321.
- Costanza-Chock, S. (2020) *Design justice: Community-led practices to build the worlds we need*, The MIT Press.
- Cropanzano, R., Bowen, D.E. and Gilliland, S.W. (2007) "The management of organizational justice," *Academy of management perspectives*, 21(4), pp.34-48.
- Dastin, J. (2022) "Amazon scraps secret AI recruiting tool that showed bias against women," in *Ethics of Data and Analytics*. 1st Edition. Boca Raton: Auerbach Publications, pp. 296–299.
- Falletta, S. (2014) "In search of HR intelligence: evidence-based HR analytics practices in high performing companies," *People and Strategy*, 36(4), p.28.
- Friedman, B., Kahn, P.H., Borning, A. and Hultdtgren, A. (2013) "Value sensitive design and information systems," *Early engagement and new technologies: Opening up the laboratory*, pp.55-95.

IBM. (2023) *What is artificial intelligence?* Available at: <https://www.ibm.com/topics/artificial-intelligence> (Accessed: February 28, 2023).

Leach, N. (2018) “Design in the age of artificial intelligence”, *Landscape Architecture Frontiers*, 6(2), pp.8-20.

Lengnick-Hall, M.L., Neely, A.R. and Stone, C.B. (2018) “Human resource management in the digital age: Big data, HR analytics and artificial intelligence,” In *Management and technological challenges in the digital age*: CRC Press, pp. 1-30.

Maity, S. (2019) “Identifying opportunities for artificial intelligence in the evolution of training and development practices,” *Journal of Management Development*. 38(8), pp. 651-663.

McCarthy, J. (2004) *Roads to Human Level AI?*. Keynote Talk at Beijing University of Technology, Beijing, China.

Mehan, J. E. (2022). “Artificial Intelligence - Ethical, Social and Security Impacts for the Present and the Future”. United Kingdom: IT Governance Ltd.

Panwar, S. (2021) *AI bias challenges in hr and 6 ways companies can address them*, Spice Works. Available at: <https://www.spiceworks.com/hr/future-work/guest-article/ai-bias-challenges-hr/> (Accessed: February 28, 2023).

Parikh, N. (2021) *Understanding bias in AI-enabled hiring*, Forbes. Available at: <https://www.forbes.com/sites/forbeshumanresourcescouncil/2021/10/14/understanding-bias-in-ai-enabled-hiring/?sh=151b6f87b96e> (Accessed: February 28, 2023).

Raghavan, M. et al. (2019) “Mitigating bias in algorithmic hiring: Evaluating claims and practices,” *arXiv [cs. CY]*. Available at: <http://arxiv.org/abs/1906.09208>.

Raveendra, P. V., Satish, Y. M. and Singh, P. (2020) “Changing landscape of recruitment industry: A study on the impact of Artificial Intelligence on eliminating hiring bias from recruitment and selection process,” *Journal of computational and theoretical nanoscience*, 17(9), pp. 4404–4407. doi: 10.1166/jctn.2020.9086.

Suresh, H. and Gutttag, J. (2021) “A framework for understanding sources of harm throughout the machine learning life cycle”, In *Equity and access in algorithms, mechanisms, and optimization* (pp. 1-9).

Vincent, J. (2018) *AI spots 40,000 prominent scientists overlooked by Wikipedia*, The Verge. Available at: <https://www.theverge.com/2018/8/8/17663544/ai-scientists-wikipedia-primer> (Accessed: February 28, 2023).

Zehir, C., Karaboğa, T. and Başar, D. (2020) “The transformation of human resource management and its impact on overall business performance: Big Data analytics and ai technologies in strategic HRM,” *Digital Business Strategies in Blockchain Ecosystems: Transformational Design and Future of Global Business*, pp.265-279.