International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 12
Article Received: *25 October 2022* ***Revised:*** *12 November 2022* ***Accepted:*** *30 December 2022*

Ethical Considerations in Ai and Ml: Bias Detection and Mitigation Strategies

Jagbir Kaur
Program Manager, Independent Researcher, New Jersey, West Orange,USA.
jagbirkaurjk3@gmail.com

Ashok Choppadandi
Senior Data Architect, Independent Researcher, McKinney,Texas ,USA.
ashokch1@gmail.com.

Pradeep Kumar Chenchala
Software Development Engineer, Independent Researcher, Seattle, Washington,USA.
pradeepnita2015@gmail.com

Pandi Kirupa Kumari Gopalakrishna Pandian
Independent Researcher, AI ML Expert,USA.

Satyanarayan kanungo
Independent Researcher,Data Principal Engineer,USA.

Abstract

The ethical considerations in the artificial intelligence and machine learning can be discussed in this paper and create the proper opportunity to detect the biasness in the data that is provided in the system. There are teams that are associated with multi-disciplinary streams that help in providing proper algorithms to the machine using AI and machine learning systems. This report is essential to understand the process that can be adopted for the removal of the biases in the system of the data and the use AI and machine learning for the proper use of the technology.

Keywords: Artificial Intelligence, Machine Learning, Multi-disciplinary, Algorithms, Data, Bias detection

Introduction

The bias of Ai is used when the application of machine learning can expand out of its ethical use in the machine learning platform. The evaluation of these biases is important to discuss for the proper use of AI in different platforms. In this paper, the data that are provided to the AI is properly enhanced so that the bias is used for the detection of proper data and it is also noticed if the data that is provided in the system is properly trained to get better results out of the system. There are some proper strategies that are used for mitigating the AI in the machine learning platforms and delivering proper algorithms to the system. There are teams that are associated with multi-disciplinary streams that help in providing proper algorithms to the machine using AI and machine learning systems. The policies of ethical practices should be considered for applying the proper use of AI and machine learning to develop proper algorithms that could be provided to the system.

Literature Review

Bias in artificial intelligence algorithms and recommendations for mitigation

According to Nazer *et al.*, 2021, the use of AI and machine learning has greatly increased and created the opportunity to collaborate with this modern technology. There are different factors such as social norms and regulations that can help in implementing the use of AI and machine learning and can be integrated with modern technologies. The use of AI and machine learning have helped many people in improving their lifestyle and building

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 12
Article Received: *25 October 2022* ***Revised:*** *12 November 2022* ***Accepted:*** *30 December 2022*

a proper infrastructure for the society but on the other hand it has created some biases and discrimination in the health sector where the implementation of proper algorithms can be a challenging task (Nazer *et al.*, 2021). The importance of getting the source of different algorithms is necessary for implementing proper AI and machine learning processes in the system. This paper has helped in understanding the different sources of bias where different factors are considered as the stream of the application of the algorithm in the system of healthcare. The paper also discusses the biases and discrimination in providing proper algorithms and the way the Ai and machine learning system can comply with the ethical norms and consider the ethical rules to imply the algorithms properly in the system. The reduction of the bias and mitigating iot properly so that it can be neglected in the process of applying pepper algorithm. The developers and the users should work properly to mitigate the bias in AI and machine learning to exercise equity for using modern technologies in the healthcare system. The use of AI and machine learning have helped many people in improving their lifestyle and building a proper infrastructure for the society but on the other hand it has created some biases and discrimination in the health sector where the implementation of proper algorithms can be a challenging task. The proper utilization of AI and machine learning can be applied by mitigating the bias from the system of different algorithms of the system.

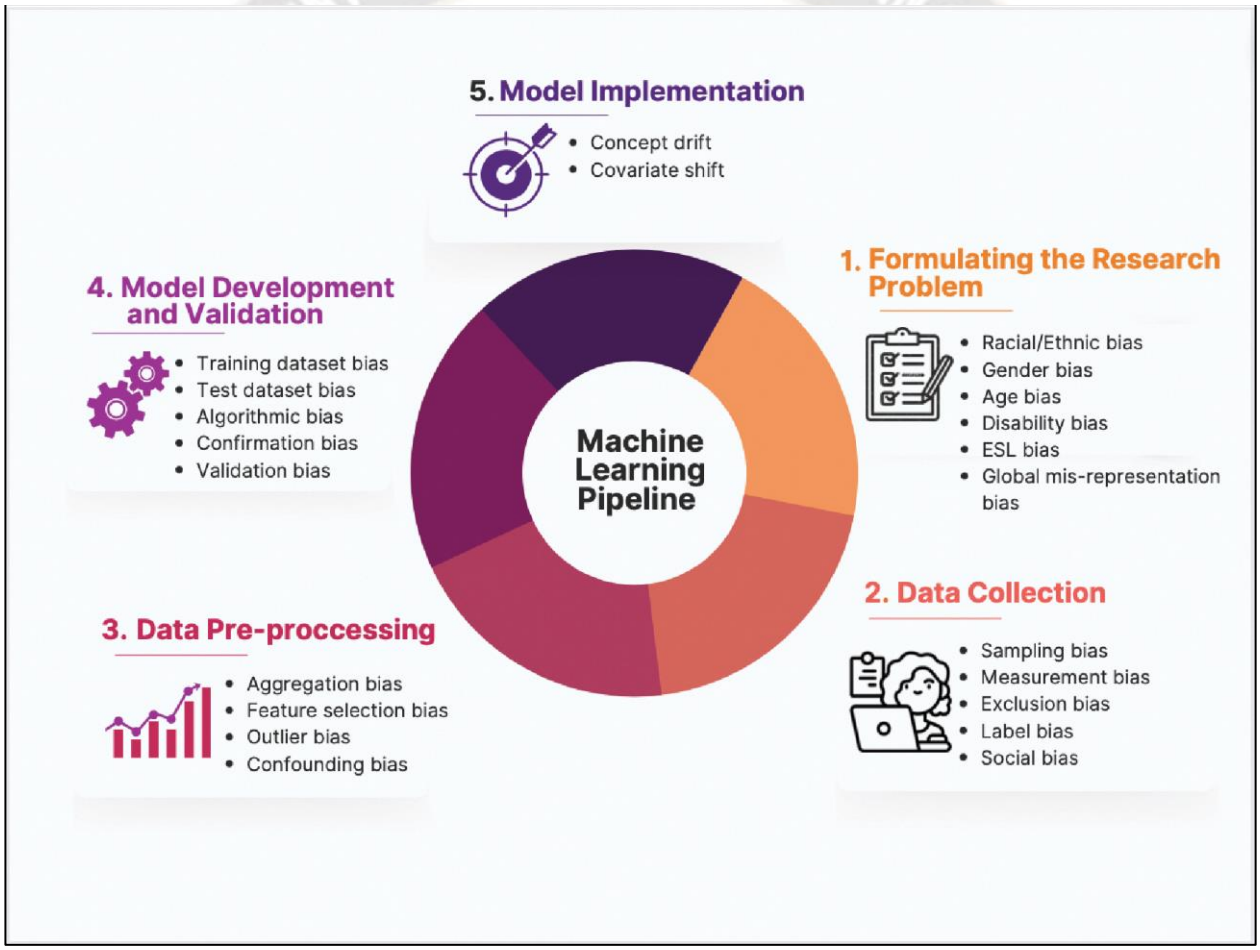

Figure 1: Bias in artificial intelligence
(Source: Nazer *et al.*, 2021)

A Survey on Bias and Fairness in Machine Learning

According to Mehrabi *et al.*, 2021, the fairness of getting proper use for each and every person has increased not only in the society but also in the application of Ai and machine learning in using proper algorithms in the system. The use of AI and machine learning can be improved by mitigating bias where it can take many important decisions in many difficult situations. The use of AI and machine learning can be thus used for delivering critical information and it should be free of discrimination so that it can provide proper decisions for developing effective algorithms for the system. The proper addressing of the biases are made in this paper to

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 12
Article Received: *25 October 2022* ***Revised:*** *12 November 2022* ***Accepted:*** *30 December 2022*

create the idea of the way it can be mitigated and further steps can be taken to avoid these biases in the algorithms that are needed to be done for greater use of modern technologies (Mehrabi *et al.*, 2021). The examination of different sub-parts and domains of AI should be done for addressing the bias in machine learning properly. Bias audits are used for specific algorithm techniques for the improvement of AI and machine learning in the system. The solutions are done in a systematic way and create the resolve of mitigating the bias to apply proper algorithms in the AI and machine learning of the system. This paper has helped in understanding the different sources of bias where different factors are considered as the stream of the application of the algorithm in the system of healthcare. The paper also discusses the biases and discrimination in providing proper algorithms and the way the AI and machine learning system can comply with the ethical norms and consider the ethical rules to imply the algorithms properly in the system. The survey that is conducted in this research paper will help in understanding the method of mitigating the bias of the AI and machine learning and apply proper algorithms in the system.

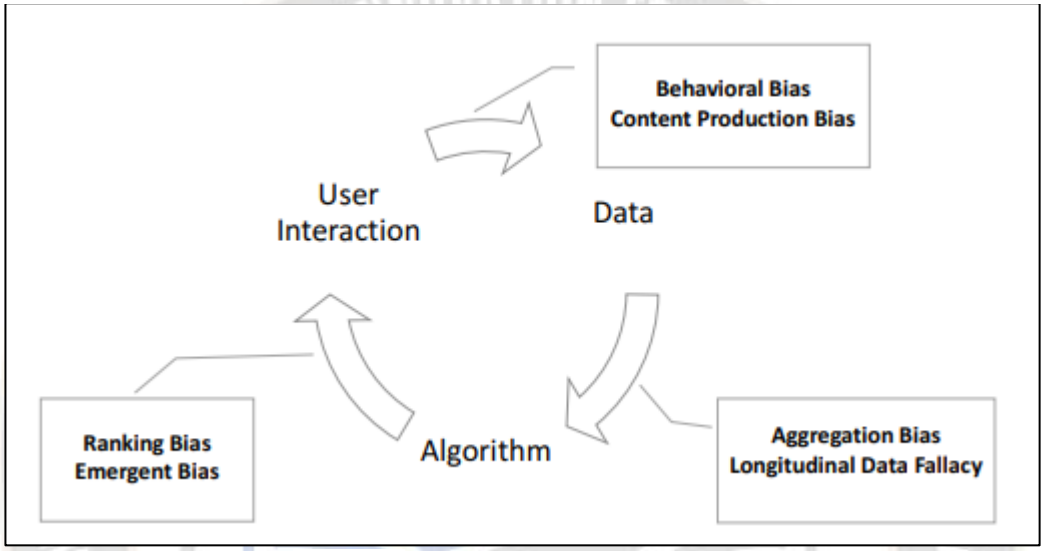

Figure 2: A Survey on Bias and Fairness
(Source: Mehrabi *et al.*, 2021)

Overcoming fairway: A classification of machine learning software to build fair with biases and mitigation methods

According to Chakraborty *et al*, 2020. ML software is becoming more and more used to make important decisions for people's life. However, the basic component of this program (the learnt model) might act in a biased manner, giving an unfair advantage to a certain group of individuals (where those categories are identified by gender, color, and so on). This "algorithmic discrimination" in AI software systems has raised severe concerns within the machine learning as well as software engineering communities. The application of AI and machine learning has increased greatly for which many streams of technologies have integrated with the different departments of the system. The examination of different sub-parts and domains of AI should be done for addressing the bias in machine learning properly. The solutions are done in a systematic way and create the resolve of mitigating the bias to apply proper algorithms in the AI and machine learning of the system. This paper has helped in understanding the different sources of bias where different factors are considered as the stream of the application of the algorithm in the system of healthcare. Fairness constraint algorithm can be used to resolve the biases in AI and machine leaning and improve the system. The dataset that is provided cannot be represented in a proper way or the models that are discussed can be in an improper manner which can create biases in the algorithms of the system. The use of AI and machine learning have helped many people in improving their lifestyle and building a proper infrastructure for the society but on the other hand it has created some biases and discrimination in the health sector where the implementation of proper algorithms can be a challenging task (Chakraborty *et al*, 2020). The proper utilization of AI and machine learning can be applied by mitigating the bias from the system of different algorithms of the system. The biases should be resolved in a way that can be applied in the real world and provide life changing decisions that could benefit the people and improve the system. The use of AI and machine learning can be improved by mitigating bias where it can take many important decisions in many difficult situations. The use of

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 12
Article Received: *25 October 2022* ***Revised:*** *12 November 2022* ***Accepted:*** *30 December 2022*

AI and machine learning can be thus used for delivering critical information and it should be free of discrimination so that it can provide proper decisions for developing effective algorithms for the system. The proper addressing of the biases are made in this paper to create the idea of the way it can be mitigated and further steps can be taken to avoid these biases in the algorithms that are needed to be done for greater use of modern technologies.

Methods

There are different methods that can be used for mitigating the bias and use ethical procedures for mitigating the bias in AI and machine learning and improving the algorithm in the system. The different methods that are discussed in this section can help in understanding the methods that can be adopted for mitigating the bias in AI and machine learning and provide proper algorithms in the system.

Analysis of the data

The data is analyzed to identify any biases in the algorithm such as improper distribution of the data (Nazer *et al.*, 2021). There are different methods such as testing in a statistical way and the visualization of data can be helpful in identifying different biases in the system.

Algorithms for mitigating bias

There are algorithms that can be applied during the training of the model and creating a proper system. There are techniques that can be applied such as techniques that are used for reweighting the data, and debiasing of the dataset for the system.

System for fairness

There are different metrics in python language that can be used such as opportunity that can be applied for equality, evaluation of the odds in equal manner, and equity in the demographic aspects of the system.

AI description

The application of different techniques and models that can enhance the decision-making process that can help in filtering the biases and the way that can be mitigated by applying different tools from the system.

These methods are useful for mitigating the bias in AI and machine learning and provide the proper algorithm of the system. There are other methods that can be implemented for better improvement of the mitigation process for the system.

Results

The data that are analyzed for this paper has helped in understanding the mitigation of the bias that is applicable for providing better algorithms in the system. The analysis of different groups which are related to the demographic aspects of the data can be evaluated for examining the different dataset that are provided to the system (Ntoutsi *et al.*, 2020). There are analyses available which are done in a statistical way such as analysis of the impact can evaluate the possible biases. There are various techniques that involve visualization processes that are done in python language such as scatter plots, histogram process, and other different plots that can help in segregating odd groups in the system. Once the biases are identified, the techniques involved for the mitigation process can be applied for the improvement of the algorithm in the system. The algorithms that can be applied for mitigation are debiasing techniques or applying optimization systems of algorithms for mitigation of the bias during the training of the model (Nadeem *et al.*, 2020). The examination of different sub-parts and domains of AI should be done for addressing the bias in machine learning properly. The solutions are done in a systematic way and create the resolve of mitigating the bias to apply proper algorithms in the AI and machine learning of the system.

Discussion

The discussion that is mentioned in this section can be elaborated by addressing the different sources that can cause the biases in AI and machine learning of the system. There are other causes that can be mentioned in this section such as labels of the proxy of the system, evaluation of the model, and other such methods that are the cause for the biases in the system. This can create the biases in the system and involve discriminatory remarks in the algorithm of the system (Safdar *et al.*, 2020). There are analyses available which are done in a statistical way such as analysis of the impact can evaluate the possible biases. There are various techniques that involve visualization processes such as scatter plots, histogram process, and other different plots that can help in segregating odd groups in the system (Lin *et al.*, 2020). The proper utilization of AI and machine learning can be applied by mitigating the bias from the system of different algorithms of the system. The biases should be resolved in a way that can be applied in the real world and provide life changing decisions that could benefit the people and improve the system (D'Antonoli, 2020). The use of AI and machine learning can be improved by mitigating bias where it can take many important decisions in many difficult situations. The use of AI and machine learning can be thus used for delivering critical information and it should be free of discrimination so that it can provide proper decisions for developing effective algorithms for the system.

Conclusion

In the conclusion it can be inferred that the mitigation process for biasing in AI and machine learning is discussed properly and that can help the programmers and other personnel to improve the algorithm in the system. The literature reviewed different journals that discussed different strategies that can be used for improving the algorithm of AI and machine learning of the system. The different fields in the world can also implement the improved AI and machine learning where the biases are mitigated and a proper algorithm is applied for the improvement of the system.

Reference List

Journals

1. D'Antonoli, T. A. (2020). Ethical considerations for artificial intelligence: an overview of the current radiology landscape. Diagnostic and Interventional Radiology, 26(5), 504. [Retrieved from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7490024/]
2. Hildebrandt, M. (2021). The Issue of Bias: The Framing Powers of Machine Learning. [Retrieved from: https://cris.vub.be/ws/portalfiles/portal/82585157/SSRN_id3497597.pdf]
3. Lin, Y. T., Hung, T. W., & Huang, L. T. L. (2021). Engineering equity: How AI can help reduce the harm of implicit bias. Philosophy & Technology, 34(Suppl 1), 65-90. [Retrieved from: https://link.springer.com/article/10.1007/s13347-020-00406-7] Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. ACM computing surveys (CSUR), 54(6), 1-35. [Retrieved from: https://arxiv.org/pdf/1908.09635.pdf?fbclid=IwAR3pBovhUWB420mkGiw5ehZ7dHOg4b02n5g2XbDQf13nEoaHwBbNt6KSfRQ] [Retrieved on:11.03.2024]
4. Nadeem, A., Abedin, B., & Marjanovic, O. (2020). Gender Bias in AI: A review of contributing factors and mitigating strategies. [Retrieved from:https://www.ijcst.com.pk/index.php/IJCST/article/download/376/336]
5. Nazer, L. H., Zatarah, R., Waldrip, S., Ke, J. X. C., Moukheiber, M., Khanna, A. K., ... & Mathur, P. (2021). Bias in artificial intelligence algorithms and recommendations for mitigation. PLOS digital health, 2(6), e0000278. [Retrieved from: https://www.dwt.com/-/media/files/blogs/artificial-intelligence-law-advisor/2022/03/nist-sp-1270--identifying-and-managing-bias-in-ai.pdf]
6. Ntoutsi, E., Fafalios, P., Gadiraju, U., Iosifidis, V., Nejdl, W., Vidal, M. E., ... & Staab, S. (2020). Bias in data-driven artificial intelligence systems—An introductory survey. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 10(3), e1356. [Retrieved from: https://journals.plos.org/digitalhealth/article?id=10.1371/journal.pdig.0000278&utm_source=miragenews&utm_medium=miragenews&utm_campaign=news]
7. Safdar, N. M., Banja, J. D., & Meltzer, C. C. (2020). Ethical considerations in artificial intelligence. European journal of radiology, 122, 108768. [Retrieved from: https://www.sciencedirect.com/science/article/pii/S0720048X19304188]]
8. Chakraborty, J., Majumder, S., Yu, Z., & Menzies, T. (2020, November). Fairway: a way to build fair ML software. In *Proceedings of the 28th ACM joint meeting on European software engineering conference and symposium on the foundations of software engineering* (pp. 654-665). [Retrieved from: https://dl.acm.org/doi/abs/10.1145/3368089.3409697]
9. Srivastav, P. Nguyen, M. McConnell, K. A. Loparo and S. Mandal, "A Highly Digital Multiantenna Ground-Penetrating Radar (GPR) System," in IEEE Transactions on Instrumentation and Measurement, vol. 69, no. 10, pp. 7422-7436, Oct. 2020, doi: 10.1109/TIM.2020.2984415.
10. Jhurani, Jayesh. "Revolutionizing Enterprise Resource Planning: The Impact Of Artificial Intelligence On Efficiency And Decision-making For Corporate Strategies." International Journal of Computer Engineering and Technology (IJCET) 13, no. 2 (2022): 156-165.
11. Kanungo, Satyanarayan. "Edge Computing: Enhancing Performance and Efficiency in IoT Applications." International Journal on Recent and Innovation Trends in Computing and Communication 10, no. 12 (December 2022): 242. Available at: http://www.ijritcc.org
12. Kanungo, Satyanarayan. "Hybrid Cloud Integration: Best Practices and Use Cases." International Journal on Recent and Innovation Trends in Computing and

Communication (IJRITCC), vol. 9, no. 5, May 2021, pp. 62-70. Available at: http://www.ijritcc.org

13. Kanungo, Satyanarayan, and Pradeep Kumar. "Machine Learning Fraud Detection System in the Financial Section." Webology, vol. 16, no. 2, 2019, p. 490-497. Available at: http://www.webology.org
14. Mohammad, Naseemuddin. "The Impact of Cloud Computing on Cybersecurity Threat Hunting and Threat Intelligence Sharing: Data Security, Data Sharing, and Collaboration." International Journal of Computer Applications (IJCA) 3, no. 1 (2022): 21-32. IAEME Publication.
15. Karuturi, S. R. V., Satish, Naseemuddin Mohammad. "Big Data Security and Data Encryption in Cloud Computing." International Journal of Engineering Trends and Applications (IJETA) 7, no. 4 (2020): 35-40. Eighth Sense Research Group.