



Threats and Ethical Considerations of Artificial Intelligence

N.B. Pushpa^{1*}, N.B. Prajwala², Kumar Satish Ravi³

Abstract

Intelligence when exhibited by machines is called as artificial intelligence. This technology enabled machines to reason, present knowledge, process language, solve problem etc. due to advanced researches and inventions, artificial intelligence has gained significant momentum in recent days. It is widely used in the of search engines, health, education, autonomous vehicles, personal identification and so on. Based on the previous actions, machines can analyse the current scenario and give predictions to the user. No doubt, because of its competency it has replaced manpower in various fields. When it comes of security issues there is always rising concern that How safe is Artificial intelligence? This article discusses about the ethical issues and threats resulting due to rampant use of AI.

Key Words: Artificial Intelligence, Labelled Data, Privacy and Security Risks.

DOI Number: 10.14704/nq.2022.20.8.NQ44382

NeuroQuantology 2022; 20(8): 3560-3562

Introduction

Artificial intelligence (AI) is an umbrella of technology which aims at making machine or computer system to learn, analyze, execute tasks for which they are trained for. For a system to reproduce human intelligence it needs contextual information, different situational solutions and the ability to overcome the challenge posed. Artificial intelligence is the result of man's obsession for technology which reduces the burden of various tasks [Ghahramani, 2015; Tirri et al, 2011]. There are various subsets of AI- Machine learning, Deep learning and cognitive computing to name. Machine learning mainly functions on available variables and thus providing possible accurate and predictive data for the future. This can be supervised or non-supervised by human. In supervised ML, the system uses the data labelled/ fed by the human; while in non-supervised ML, the system identifies the pattern from the hidden data which is not labelled by human. Supervised ML have role in the development of biological and artificial neuronal networks (ANN). The combination of supervised and non-supervised ML is used in reinforcement learning. Deep learning is highly evolved form of AI where the system is trained to filter, analyze and

learn from the data input so that it can mimic human brain (Raymond et al, 2018; Kulkarni et al, 2022). The applications of AI have been extensively used in day-to-day life which interferes one's personal space like Siri, google search engine, social media, AI chatbot etc.

3560

How Safe is Artificial Intelligence?

Multidimensional AI has become part of lifestyle. There are various apps which uses AI to monitor individuals using language translator, image recognition, location tracker etc. (Mayer-Schoenberger & Cukier, 2013). Although it aids to simplify and manage intricate data set and suggest most possible solutions, there are algorithms which can read through your google search and predict what has to come next (Flagella, 2020).

Corresponding author: N.B. Pushpa

Address: ^{1*}Assistant Professor, Department of Anatomy, JSS Medical College, Mysore, Karnataka; ²Software Engineer, Wipro Technologies, Bangalore, Karnataka; ³Professor (Additional), All India Institute of Medical Sciences, Rishikesh, Uttarakhand.



Now the question arises how safe is AI when it can access all personal data. Also, one of the greatest disadvantages of AI is inability to control it, as AI it includes various complicated algorithms. So, who is the ultimate beneficiary here, it is the organization or the company which can accumulate large consumer data and use for their benefit?(Jernigan,2009).

Machines usually fed with labelled data and trained to give accurate and predictable outcome. So, there are more chances of being biased while using labelled data(CDEI,2019).With the intension of making the system unbiased more complex algorithms are used which can result in inability to understand the AI analyzing process and also more intricate outcome(USCAM, 2017).This problem is called black box theory.With such an uninterpretable outcome it becomes increasingly difficult to rely on predictions of the system, which can neither be monitored nor controlled. Even in medical field there are threats where the AI application can predict wrong diagnosis and endanger life, as it mainly depends on enormously fed data set only (Topol 2019).

Ethical Issues and Threats of AI

Rapid global surge in AI and its applications has resulted in automation in various fields. This has directly or indirectly has increased the problem of unemployment(Boden 2018). Lack of thorough understanding of AI's functioning leads to disagreement with the result. Intrinsic bias could result in prejudiced decisions against particular population segment. This can result in potential exemption of beneficiaries to the citizen (Tao et al. 2005, Flick 2016).. Other threats related to AI are lack of accountability (USCAM, 2017), malicious psychological profiling, privacy and security risks(Ballings, 2015).

Regulations

Nations worldwide are working on principles and guidelines to monitor designing, development and applications of AI.High level expert group in the European union has released Ethics Guidelines for Trustworthy AI which suggests seven important requirements to label AI trust worthy (Hickman, 2021). In the similar way, Singapore is following Model AI Governance Framework(Singapore 2020) and the United States of America has guidelines for the Stewardship of AI Applications.General Data Protection Rules (EU,

2016), in the EU is a regulatory framework for protection of personal data and establishes the need for 'privacy by design' when developing automated solutions. In the USA, the Algorithmic Accountability Act of 2019 is a proposed bill that requires specified commercial entities to conduct assessments of high-risk systems that involve personal information or make automated decisions, such as systems that use artificial intelligence or machine learning(Senate, 2019).

Currently there is no overarching legislation specific to AI.Different authorities like the Ministry of Electronics and Information Technology, the Ministry of Commerce and Industry, the Department of Telecommunications, and the NITI Aayog have initiated to regulate the applications of AI so as to ensure non violation of human rights and privacy.NITI Aayog "National strategy for artificial intelligence" aims at helping the lagging behind sectors by developing more consistent, native pioneering solutions.NITI Aayog released a draft for discussion with stakeholders revolving the area around responsible AI.In this respect it advocates suggestions like: Setting up an IP regime for AI innovations and a task force that will comprise the Ministry Corporate Affairs and Department of Industrial Policy and Promotion to examine and issue a modification to intellectual property law;Developing a data privacy legal network to protect human rights and privacy and; and creating a sectoral regulatory guideline encompassing privacy, security, and ethics(Government of India,2018).

Ministry of Electronics and Information Technology constituted four committees for the development of the regulatory framework for artificial intelligence.The first committee for platforms and data on artificial intelligence. Second committee for leveraging AI for identifying national missions in critical sectors. Third committee for mapping technology capabilities, key policy enablers required across sectors, skilling and reskill; and the Forth committee for cybersecurity, safety, legal and ethical issues.

The four committees of MeitY, as mentioned earlier, lays down the following recommendations:

The development of an open National Artificial Intelligence Resource Platform (NAIRP) for knowledge integration and awareness for AI and ML.

Establishment of a committee of stakeholders to dissect the area of AI in a multidisciplinary way. The committee will review the existing laws to



make the amendments or modifications to align with AI development.

The stakeholders shall deliberate whether AI should be considered a legal person and establish a scheme or compensation fund to compensate for damages in civil liability claims.

Use of procurement contracts by the government to focus on the best practice relating to security and privacy issues.

AI frameworks should design broad principles, and the companies should be allowed to make their internal programs in compliance with the said framework. This will provide flexibility to adapt to the technology development.

The government should propose the development of safety parameters and safety thresholds to ensure that human interaction with AI does not harm people and property in any way.

Standards should be made to address the AI development cycle(NITI Ayog,2020).

Conclusion

Although AI is a milestone in the technology revolution there are opposite views regarding its wide spread use. Threats and ethical issues are major concerns in the widespread application of AI. Establishing such a framework would be crucial for providing guidance to various stakeholders in responsible management of Artificial Intelligence. So, regulations guiding AI strives to advocate Fairness, transparency, accountability and ethics.

References

- Ballings, M., Van Den Poel D, 2015. CRM in social media: Predicting increases in Facebook usage frequency. *EJOR* 244: 248–260.
- Boden MA. 2018. Artificial intelligence: a very short introduction. Reprint. Oxford, UK: Oxford University Press.
- CDEI (2019) Interim report: Review into bias in algorithmic decision-making. Centre for Data Ethics and Innovation, London.
<https://www.gov.uk/government/publications/interim-reports-from-the-centre-for-data-ethics-and-innovation/interim-report-review-into-bias-in-algorithmic-decision-making>. Accessed 30 Jan 2022
- EU (2016). 2016. (EU) 2016/679 GDPR. <https://eur-lex.europa.eu/eli/reg/2016/679/oj> Accessed on 30 Jan 2022
- Faggella D (2020) Everyday examples of artificial intelligence and machine learning. Emerj, Boston, MA.
<https://emerj.com/ai-sector-overviews/everyday-examples-of-ai/>. Accessed 30 Jan 2022
- Flick C (2016) Informed consent and the Facebook emotional manipulation study. *Res Ethics* 12. 10.1177/1747016115599568

- Ghahramani Z. 2015. Probabilistic machine learning and artificial intelligence. *Nature*, 521(7553): 452-9.
- Government of India. 2018. National strategy for artificial intelligence #AI for all.
<https://indiaai.gov.in/documents/pdf/NationalStrategy-for-AI-Discussion-Paper.pdf>. Accessed 30 Jan 2022
- Hickman, E., Petrin, M 2021. Trustworthy AI and Corporate Governance: The EU's Ethics Guidelines for Trustworthy Artificial Intelligence from a Company Law Perspective. *Eur Bus Org Law Rev.*, 22: 593–625.
- Jernigan C, Mistree BFT (2009) Gaydar: Facebook friendships expose sexual orientation. *First Monday* 14.
<https://firstmonday.org/ojs/index.php/fm/article/download/2611/2302>. Accessed 30 Jan 2022
- Kulkarni, P., Mahadevappa, M., & Chilakamarri, S. 2021. The emergence of artificial intelligence in cardiology: current and future applications. *Current cardiology reviews*, (in press, doi.org/10.2174/1573403X1766621119102220)
- Mayer-Schonberger V., Cukier K. Eamon Dolan/Houghton Mifflin Harcourt; Canada: 2013. Big data: A revolution that will transform how we live, work and think.
- Singh A. 2021. Regulation of artificial intelligence. Indian legal solution. <https://indianlegalsolution.com/regulation-of-artificial-intelligence/#:~:text=In%202020%2C%20NITI%20Aayog%20drafted,process%20to%20set%20particular%20standards>. Accessed 30 Jan 2022
- Raymond JL, Medina JF. 2018. Computational principles of supervised learning in the cerebellum. *Annu Rev Neurosci.*, 41: 233-53.
- Senate.Gov.USA 2019. Algorithmic Accountability Act of 2019. <https://www.wyden.senate.gov/imo/media/doc/Algorithmic%20Accountability%20Act%20of%202019%20Bill%20Text.pdf> Accessed 30 Jan 2022
- Singapore Government. 2020. Artificial intelligence governance framework. <https://www.pdpc.gov.sg/-/media/files/pdpc/pdf-files/resource-for-organisation/ai/sgmodelaigovframework2.pdf>. Accessed on 30 Jan 2022
- Tirri, Nokelainen. 2011. Measuring Multiple Intelligences and Moral Sensitivities in Education. *Moral Development and Citizenship Education*. Rotterdam, Netherlands. Sense Publishers.
- Tao J, Tan T, Picard R. Affective computing and intelligent interaction. Berlin: Springer; 2005. [Google Scholar]
- Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med.*, 2019;25:44–56. doi: 10.1038/s41591-018-0300-7.
- USACM (2017) Statement on algorithmic transparency and accountability. ACM US Public Policy Council, Washington DC. https://www.acm.org/binaries/content/assets/public-policy/2017_usacm_statement_algorithms.pdf. Accessed 30 Jan 2022

