King Fahd University of Petroleum and Minerals

ICS 309: Computing and Society (212)

Ethics of Incorporating AI/ Machine learning in Making Decisions in Human Life

Mohammed Alhassan

201831660

For

Dr. Samer Arafat

Abstract

This report discusses the effects of Artificial Intelligence and Machine Learning in the human life. It explains the ethical aspect of artificial intelligence such as privacy and human skill loss and elaborate on some possible solution to these problems.

Table of Contents

Introduction

Artificial Intelligence has gained incredible traction in all spheres of daily life. As it is ubiquitous, it impacts everyone, and its possibilities are rapidly evolving (Bostrom, 2014). In addition, AI can help us in many ways. It can help us with difficult, risky, or time-consuming tasks; it can help us save lives and deal with crises; and it can fascinate us and make our daily lives more pleasant (Nadikattu, 2016). Furthermore, AI systems are used to do complicated, data-intensive tasks, including detecting chip-based card networks for any fraudulent activity, permitting high-pitched stock trading, assisting with health diagnostics, and discovering cyber-security risks and threats. Incorporated as robots, AI will soon be used as robots and will be alongside us in the form of business, transit, healthcare, and defense robots. Nonetheless, current ideas and predictions about AI's potential are very different, and it's hard to come up with a consensus about how AI will affect society.

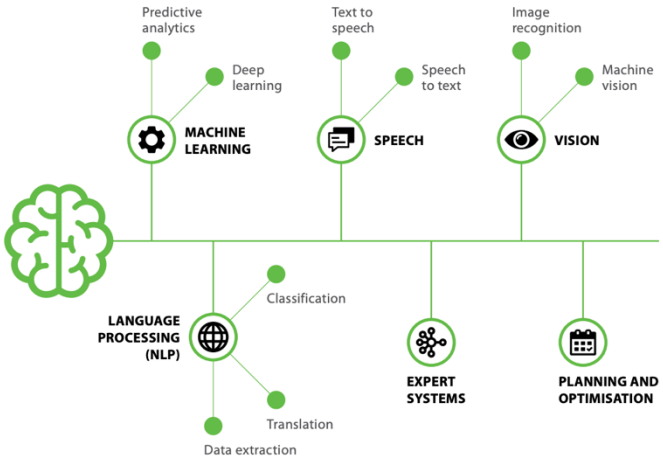

Figure 1- Components of AI (Walsh, 2019)

On the other hand, there is a growing realization that a rational approach to AI is important to ensure the safe, beneficial, and equitable usage of AI technologies. This involves the need to address the ethical consequences of machine decisions and to establish the legal standing of artificial intelligence. On the other hand, specific methods of ethical AI creation are largely absent (Riedl, 2019). For example, an appropriate design, implementation, and usage of AI systems are critical for AI technologies such as autonomous vehicles, companion, and nursing robots, as well as rating and profiling algorithms that are already impacting humanity or will do so in the coming years. In all of

these scenarios, AI understanding should be able to assess societal values, and moral and ethical concerns, prioritize the values held by individuals and in diverse multicultural situations, justify its rationale, and ensure transparency (Vakkuri, 2018).

Thus, it's not surprising that AI is becoming more involved in decision-making, whether as a resource, consultant, tool, or even management (Haesevoets, 2021). This means that today, smart technology is gradually gaining the ability to impact a wide variety of crucial societal outcomes. As is known, higher authority entails key responsibilities. As such, we must begin addressing the concern of whether AI is innately suited to act responsibly and, as such, in ways that we individuals - as the key end-users - review ethically. In order to address these concerns, a fundamental reconsideration of ethics is required, as is revising the concept of liberty in the context of transforming socio-technical reality. Additionally, putting ethical practices into AI-based practical applications will help us to gain a better understanding of ethics in general.

Moreover, this subject has drawn considerable attention as the adoption of Artificial Intelligence has raised ethical concerns regarding, among other things, privacy (compromise of personal information), biased decisions (based on incorrect statistical data), a lack of accountability, and the possibility of job loss due to automation. With such ethical implications, dread and even anxiety have emerged in society and industry around the employment and progress of AI. Curiously, the narrative that accompanies the debate over the moral implications of AI is defined by a proclivity for AI to exhibit human-like characteristics (De Cremer, 2021). As a result of this tendency, labelled anthropomorphism, we appear to have developed the idea that AI can be essentially malevolent or good. As we increasingly give magical and human-like abilities to AI, a trend is rising to view this smart technology as accountable for its acts and decisions. What are the implications of this trend?

This viewpoint recognizes the critical influence that human perceptions of machines play. Particularly, it appears as though our enthusiasm for AI's mystical capabilities has led us down a path in which we have downgraded ethics to a technological issue.

However, if we are being completely candid, AI is not miraculous. From this data, AI makes no new assertions or demonstrations that are not already present in the data from which it is learning. As such, AI cannot be regarded as a self-aware being capable of initiating and autonomously choosing to exhibit positive or negative behavior. While it is true that AI does have the potential to move between good and evil, it can only do so if the person or organization utilizing the technology has the appropriate intentions. As a result, the main point we are making is that artificial intelligence itself has no ability to make ethical decisions since it lacks ethics. Additionally, we believe that artificial intelligence (AI) cannot be made innately more ethical than other technologies (or even individuals) merely by being "smart" In this regard, it is important to note that artificial intelligence is a state-of-the-art technology created by humans and therefore functions as a mirror of our attitudes and biases. For example, the recent controversy in the United Kingdom, where machine learning algorithms were employed to forecast O/A level students' grades based on the past performance of secondary schools. This algorithm-driven strategy resulted in an unethical conclusion since many students' grades were dropped, especially those from lower-income institutions (Nast, 2020). It has been found that the utilization of algorithms, supposed to mitigate teachers' bias in forecasting students' results, actually adds to it. So, worrying about bias in AI is like complaining about one's reflection in a mirror. Consequently, since the "mirror" images reveal our own biases and weaknesses, we cannot predict whether artificial intelligence will magically become more ethical than humans.

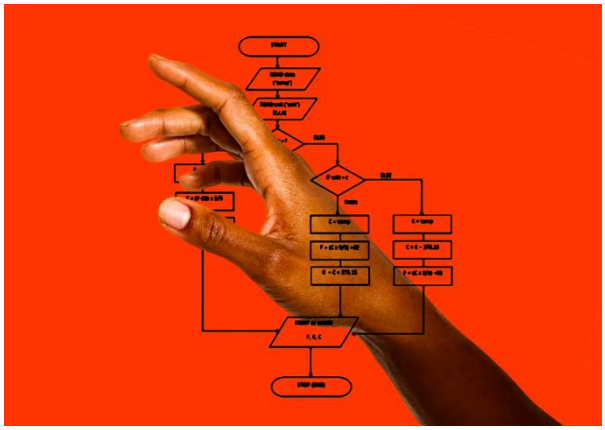

Figure 2 - Students grading algorithm (Nast, 2020)

Analysis

One of the factors contributing to this cynical perspective of mankind is that we have stopped viewing algorithms and new technologies as integral parts of our everyday lives. (Harari, 2016) cautions against dividing intelligence and awareness and highlights the risks associated with the use of subconscious but extremely intelligent algorithms. It is possible to argue that if organisms are algorithms and life is based on data processing, then we cannot play with a machine that is capable of assessing all present information, making decisions, and processing concerns without consciousness - or more precisely without it - and produce better results.

A straightforward illustration is the comparison of crashes involving driverless cars versus human-driven automobiles. Additionally, the widespread use of these kinds of cars would result in a massive surge in professional driver unemployment. Which would essentially establish the AI's dominance over humans. After all, dystopian perspectives imply that, inevitably, originally human-made judgments would be ruled by ever-improving algorithms, producing results superior to those achieved by humans subject to error. This would eventually alter the entire social stratification system. Exclusion or decrease of a person's position in critical decision-making processes related to the operation of society will inevitably result in their social inferiority.

How and for what reason AI is applied has a role in determining whether it generates benefits or downsides. For instance, (Braithwaite, 2020), a deployment setting in which an algorithm is independently tasked with deciding welfare payments, without substantial human supervision, and ultimately produces faulty computations, is one that can result in harm. Or artificial intelligence charged with evaluating individual performance in order to contribute, and possibly interact. The dismissal decision raises concerns about the transparency of data collection and the propriety of such deployment. These and other instances demonstrate that novel systems' incorporation poorly or in ways that marginalize individuals from the workflow can result in harm.

State-of-the-art advancements in AI are directed, in part by better access to substantially large datasets and novel CNN mapping (Walsh, 2019). Ethically, this sets an emphasis on the quality and type of the data used to develop AI outputs, as well as the degree to which such outputs are intelligible to people.

Accountability in Artificial Intelligence

When the result of an AI or machine learning model is questioned, transparency is necessary. Numerous possible fault causes can be recognized as the model is implemented, trained, and used. The current debate focuses largely on the question of whether training sets are sufficiently representative. Having an inadequate or biased sample population can lead to unfair results.

Figure 3 - Five safes framework in an AI algorithm (Walsh, 2019)

Additionally, flaws can occur at later phases, specifically if the source data put into the system is incorrect. There are two major ontological theories of accountability. The first is post-factum, which involves a culpable agent. The second is ethical, and society looks at how well-supervised learning systems work against established standards of fairness, justice, and so on.

The critical distinction between machine learning and human decision-making is that it requires individual flexibility, environment-relevant judgments, empathy, and sophisticated moral judgments, all of which are lacking in machine learning. In other circumstances (self-driving cars, aviation-monitoring systems), instant monitoring is not feasible. On the other hand, post-facto accountability is critical since it may be utilized and designed to detect a failure and develop better state-of-the-art systems.

Conclusion

Individuals' perceptions of automated decisions depend not only on the results but also on the process by which they are made. Additionally, the process of making ethical decisions may involve the capability, among other factors, to explicate a prediction or decision, which may not be possible if black-box frameworks are used. A significant shortcoming of machine learning is that it merely displays correlations, but we frequently treat its recommendations as evidence of causation. This is a recurring cause of contention.

The above also has greater ramifications for all of us. With the advent and deployment of artificial intelligence in our workplaces and society, governments worldwide have underlined the importance of everyone engaging in technological upskilling. A concern we observe is that in our justified quest for increased digital expertise, we appear to have forgotten the critical need to develop and even promote the distinctive human qualities that machines lack. As a consequence, we believe that by almost exclusively focusing on online upskilling, we are creating an environment where people will seek less attention to their strengths and, as a consequence, gradually lose their unique social skill abilities. Such a result would be detrimental to society in the future. And this will be true when it comes to the honest and moral usage of artificial intelligence. In terms of AI ethics, humans and their moral compass and consciousness can only make decisions about how smart technology is used.

As such, we recommend that researchers and professionals alike be urged to raise awareness about the fact that smart technology as it currently exists cannot serve as an alternative to a person's ethical compass. Rather than that, in addition to strengthening technology aspects that make data analysis more visible and hence easier to interpret, we need more ethical human decision-makers. In particular, people who make decisions will have to be more educated than ever to think about the ethical consequences of their actions and to be more aware of the ethical challenges that exist.

We will need to engage more in human development, particularly in the area of ethics. We must improve our ability to recognize our own positive and negative behaviors and employ those findings in treatments and practice sessions about how to use AI decision-making technologies responsibly. However, the substitute could lead to the devaluation of human rights and social norms, a lack of AI-based progress, and a missed opportunity to use AI to help people and the world.

Humanity experienced this lesson when it failed to control the technological revolution’s influence on labor workers and when it acknowledged the contextual consequences of widespread modernization. It has taken a long period, social disruption, and even regime change to secure people's rights and develop sustainable models.

References

Bostrom, N. a. Y. E., 2014. The ethics of artificial intelligence. The Cambridge handbook of artificial intelligence. In: The Cambridge handbook of artificial intelligence. pp. 316-334.

Braithwaite, V., 2020. Beyond the bubble that is Robodebt: How governments that lose integrity threaten democracy. Australian Journal of Social Issues, 55(3), pp. 242-259.

De Cremer, D. a. K. G., 2021. AI should augment human intelligence, not replace it. Harvard Business Review.

Haesevoets, T. D. C. D. D. K. a. V. H. A., 2021. Human-machine collaboration in managerial decision making.. Computers in Human Behavior, p. 106730.

Harari, Y., 2016. Homo Deus: A brief history of tomorrow... Random House.

Nadikattu, R., 2016. The emerging role of artificial intelligence in modern society.. International Journal of Creative Research Thoughts.

Nast, C., 2020. The lessons we all must learn from the A-levels algorithm debacle. https://www.wired.co.uk/article/gcse-results-alevels-algorithm-explained

Riedl, M. 2., 2019. Human-centered artificial intelligence and machine learning.. Human Behavior and Emerging Technologies, 1(1), pp. 33-36.

Vakkuri, V. a. A. P., 2018. The key concepts of ethics of artificial intelligence 2018 IEEE international conference on engineering, technology and innovation (ICE/ITMC).

Walsh, T. L. N. B. G. E. A. M. J. M. I. a. W. F., 2019. The effective and ethical development of artificial intelligence: an opportunity to improve our wellbeing.. Australian Council of Learned Acade.