AI as a Challenging Problem: Solvable without Data but Morally Intelligently-driven Insights

Mohammed Ehsan Ur Rahman
Kakatiya Institute of Technology and Science,Warangal
506015
+91 9848064062
ehsan81181@gmail.com

Md Sharfuddin Waseem
Assistant Prof., Kakatiya Institute of Technology and Science,Warangal
506015
+91 8019546945
waseem7602@gmail.com

Hari Dasyam Sri Saaketh Ram
Kakatiya Institute of Technology and Science,Warangal
506015
+91 8328532888
srisaakethram46@gmail.com

ABSTRACT

In this paper, we shall try to shed some light on the very pressing and challenging issue of ethics in technologies under the umbrella of AI. This concern has to be discussed extensively by philosophers, economists, and AI research scholars and this process of discussion is a continuous one. Autonomous services especially those involving the use of systems that have both aspects socio-technical [5] ones can increasingly become a major factor in making the environment hostile. The inclusion of human morals into human-level intelligence systems [8] is crucial.

The capacity of the human mind for formulating, ideating, conceptualizing, thinking, innovating and solving complex problems is minuscule compared with the mass of the problems, those of whom have solutions, which are essential for objectively rational and moral behavior in the real world or even for a novice reasonable approximation to such target sanity.

Equally important notions to be considered to avoid possible shortcomings in and havocs due to this technology are involvement of human ethics, consideration of social and moral implications of the same.

Concepts

• **Artificial Intelligence →Dangerous**

• **Human Ethics →Ethical Autonomous Intelligent Systems**

Keywords

AI-Artificial Intelligence, ML-Machine Learning, Intelligent agents, Ethics, Privacy, Big Data, Governance, Regulation, AGI-Artificial General Intelligence [2]

1. INTRODUCTION

The most interesting part when it comes to AI is that people while imagining about it take it to be a technology of such a kind wherein they will be finding solutions to almost all of their problems and feel fascinated by the tempting thoughts that it may transform the world into a utopia. But less of it is known that technologies like these may result in dire consequences if not regulated by its makers, and the intelligentsia. First, let's take a look at AI followed by ethics separately and then discuss their dependence on each other. The remaining paper is formatted to give an access to all concerned facts and figures about "AI-Ethics".

2. AI

The futuristic role that 'artificially intelligent' technology [9] – something depicted as more sophisticated and advanced than humans – could play in doing things accurately is not haunting until we humans realize the practical challenges it offers when is brought into widespread use by the public [7].

For the very first time in the history of human civilization, man is betting his life in building technology with the capacity to enhance itself further without human's interference.

The very idea of creating machines that can think , make predictions that can drive humans to take more advanced as well as informed decisions and taking decisions at instances when they are given the authority to do so , is seen very skeptically and raises a ton of most probable ethical concerns regarding them .These questions can be segregated into few categories like , those which are big deal of concern to the capitalists , those that are relevant to law-makers and advocates , those which are related to the governance and lastly those which have impactful results on the ultimate consumers of this technology i.e. the common man.

While dealing with AI and other related technologies, one can guess the very known and easily achievable solutions to the problems that may arise from them. Those who should take up active roles in making AI ethical have to be identified carefully and must be educated constantly about the issues coming up with the advent of new technologies and given clear path to tread for possible good solutions.

Scenarios: -

1. For capitalists it will get harder if they get lawsuits more often than not for the technology they are known for and plays a crucial role in their business model.
2. If state governance is done by an artificially intelligent agent, this may lead to overturning of existing relationships inside our delicate fabric of society as well as the current political and governance frameworks and having the perspective to enforce upon us some sort of autocratic inclinations.
3. Law-makers –

   Though a known fact but to be reiterated is that there is a very strong yet complex relationship between ethics and legislation.

The point to be made is that ethics and law go hand-in-hand and application of laws rigorously on budding technologies results in avoiding possible abnormal and un-deterministic performance by the technology in the future. A keen pursuit of ways to make law accommodate these technological advancements and search for concrete ways to embed ethics in technology. Diversity in thinking may make this task of compiling and putting forth laws a harder task.

4. Common man -This is the most vulnerable sector of people as they are in constant interaction and touch with various technologies that come under the umbrella of AI.

In fact, the short-term ethical concerns of AI stem from the fragility of it and existence of loopholes in the systems that are responsible for governing this automatic and unrestricted flow of valuable, private and vulnerable information, as well as due to their capacity to be co-opted in malicious ways in order to erode the collective genesis of truth. It is therefore utmost important to keep in mind the short-term dangers of AI-related methods through its impact that it has on the volatile labor markets, its tendency to reinforce the stereotypical conventions that have been around for long time and biases and its scope to stimulate authoritarian regime over humans.

The history has been a witness to the fact that ideas, methodologies, techniques , strategies and beyond have always been used by the people with power , wealth or so , in order to exploit others (resources and people alike) and this can serve as a great eye-opener for us and to all those who want to give this topic a thought ; that this technology can be used the same way , this question is important because people must get assurance about the safety and security of their resources (wealth , jobs , data , etc.) otherwise they may turn out to be the biggest hurdles in the advancement of the same technology .

The other crucial point that can be included in here, is that delay in inclusion of ethical knowledge [10] in such artificially intelligent systems [6] can lead to great havocs, time matters as decision making must never be slow especially in a world wherein everything is moving at a faster pace and unimaginable situations arise within the blink of an eye.

The third point that we want to put across is that we have got a lot of circumstances where the human morality plays a big role in dealing with a problem and also in getting to efficient solutions for the same. We can't confine our scope and reasoning for ethical AI on just the machine learning algorithms that we have currently, but this should be expanded to all the intelligent agents that are possible to be created [8]. In general, an agent's choice of action is always dependent on the entire perceptual inputs sequence fed into it.

3. HUMAN ETHICS

3.1 Ethics

Ethics is the science that deals with the normative and practical philosophy for governing and deciding on human decisions. It has three broad divisions [1]: -

- Consequentialist ethics: -When the agents tend to focus and prioritize consequences of choosing an option and base their working on the same strategy.
- Virtue ethics: -It can be imagined as a biased person but a good one who looks at the virtue or moral character of the person who is carrying out the action, not taking into consideration the aspects such as the ethical duties and rules, or the consequences of particular actions and choices made for prediction and decision making.
- Deontological ethics: It is an ethical theory that works on the principle that the moral correctness of a given action and/or choice must be solely based on whether it is right or wrong under given set of structured rules, rather than being based on the outcome of the action/choice made.

Humans sometimes fall into situations called as ethical dilemmas in which the available choices are so much co-related that it becomes next to impossible to avoid bad outcome by choosing any one option among the others. In such cases where ethical principles conflict with one another, decisions shall be made about how to prioritize them. Resolving these dilemmas should precisely reflect public opinion not just the programmers'.

3.2 Privacy and security

These terms bring to mind Isaac Asimov's "Three Laws of Robotics" [13] which can serve as a clear and simple, yet powerful strategies and impositions placed on AI that can help us take sufficient control over the technology [9].

In the world of data, wherein every enterprise seems to exploit for one's own benefits the loads of important but potentially private data of people across various fronts of life, these make them vulnerable to cyber attacks and other possible , regulation matters even in this areas.

3.3 Big Data and its flaws

Ethics is an abstract set of principles [3], they are hard to apply in a formal and deductive way; there exists no universal agreement on which ethical theory is best on to approach for and not much of ethics is connected to case-based study and results, it changes according to contexts which are hard for Big Data [4] to capture and analyze for.

3.4 Statistics

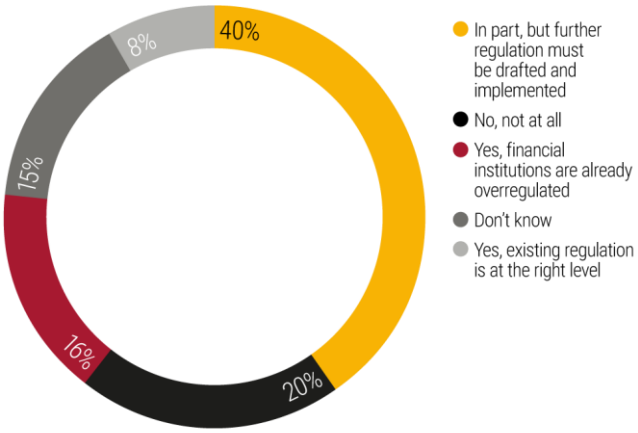

Figure 1 Responses to the question "Is existing regulation sufficient to address issues posed by AI?"

3.5 Algorithmic/Technical choices for now

The probable problems that can be solved by AI are real scheduling problems, which are inherently complex due to the presence of constraints on available resources. This gets us to the idea that if certain AI problems which are more vulnerable

to conditions like ethnic dilemmas shall be better modeled as real scheduling problems and solved through same procedures to obtain their solutions.

If machine learning algorithms involve complicated steps produced by directed evolution, which are harder for humans to analyze can land us up in situations where we don't understand the direct intuition as well as the working of the algorithm and thus making it difficult for us to understand and regulate them. Whereas a machine learning algorithm that is based on decision trees and alike is more feasible to inspection and analysis by the programmer and can enable people to trust it, because most of the skepticism in general public about AI-related technologies is that they are like a black box. Conclusively it is increasingly becoming a pressing issue to develop AI algorithms that are not just powerful in aspects of computation and performance but are also transparent to inspection/analyze [2] along with qualities such as fairness, privacy, and accountability embedded into them.

4. AI AS AN ETHICALLY DRIVEN TECH

The inducement of AI-driven technologies like computer vision , natural language processing , recommender systems , among many others , has led great minds in technology to worry about the regulations of these technologies , followed by lawmakers , governments and public to ponder over better governance for ethical decision-making AI as well as getting them into synchronization with the current trends in global culture , social interactions and by and large the establishment continually improving relation of ethically strong values which human beings has always revered and the best technology in making(AI).

When we can figure AI as a data-driven technology, in reality ,data must not only involve fundamental aspects that are more than sufficient to lead to better results but must also have those intricacies that build up their context, condition, morally-supportive or not, and answers to such kind of questions [12]. This is only possible through the use and firstly creation of very intensive and all-encompassing techniques for data collection. This also may/not ensure the acquiring of subtle patterns indicative of the potential involvement of human ethics and moral values being applied to solve the problem. And as always the data collected has noise, and if there are more than enough evidence of unethical behavior by humans in conditions and training examples on which the AI algorithms are trained upon , may definitely result in the algorithm taking up unethical behavior. This can make you picture AI as relief cum disaster at such instances.

Various taxonomies have been proposed that do the job of organizing the various ethical issues in order to regulate and generate the most productive solutions to them. The taxonomy proposed in [1] is the one which is better when compared to the rest in matters of simplicity and coverage of the ethical issues; there are four major chunks of this taxonomy: -

- Exploring Ethical Dilemmas: These are the advanced AI systems that use a feedback mechanism in order to let the programmers get to explore and comprehend the human preferences on a variety of ethical dilemmas.
- Individual Ethical Decision Frameworks: These are the generalizable decision-making and/or behavior prediction mechanisms that can enable individual agents to judge the ethics of its own choices and actions along with that of the other agents under a given context.
- Collective Ethical Decision Frameworks: Similar to the previous one, except that the decisions are collective.
- Ethics in Human-AI Interactions: very crucial from consumer point of view and also plays a dominant role in making the product earn fame and name for the business or erode its long-standing pride. These are AI frameworks that tend to incorporate ethical considerations into intelligent agents which are inherently designed to influence human behaviors or are in unmediated contact with them and/or their environments.

5. FIGURES

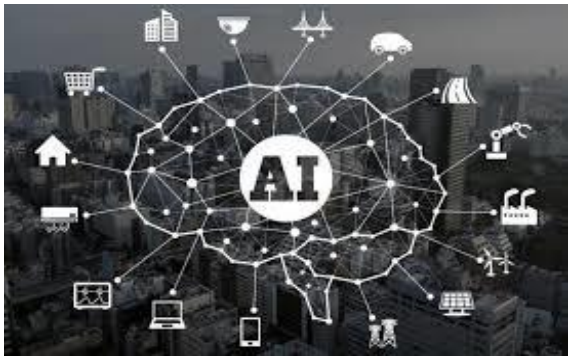

Figure 2 Illustrates possible sectors of industry and life that AI can have a great impact upon

Picturing AI as a relief to all the problems and difficulties human life on earth is facing is but a very optimistic truth, the harder part to imagine and being skeptical about is taking AI as not being driven by those human morals and resulting in disasters rather than alleviating us from our troubles.

Human ethics have always been debated upon, they are constantly evolving and one of the key factors of their betterment , advancement and branching is the invention, growth in technology and widespread problems that arise due to their existence and so on.

5.1 Technologies under the umbrella of AI

It is essential to note that AGI systems are yet to be a reality, and that all the AI applications come under the narrow AI branch.
AI technologies touch a lot of facets of our life, the most important ones being sectors like health, transport, financial services, education, etc.

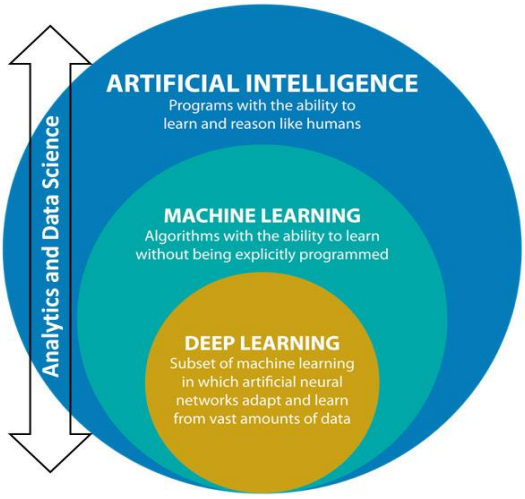

Figure 3 Umbrella of AI and scope of related technologies

List of technologies that can pose an ethical malfunction or threat:

1. Robots especially for military purposes [13]
2. Self-driving cars
3. Social media-Sentiment Analysis [7]
4. Financial systems – For example: - loan approval systems (as ML algorithms work by drawing distinguishing lines between group of populations, this may result in biased-driven performance and unfair chances for a particular segment of the population)

Emerging AI systems are serving as challenges for communities in most of the developing countries and hence this list is not a comprehensive collection of AI-related tech.

6. FUTURE SCOPE

With regards to information and communication technologies, there has to be a regulatory supervision that can lead to ethical and social assessment of it at a regular basis. Reducing the undesirable impacts of AI on human ethics and human rights shall be the main motto of governments worldwide. Dealing with the moral and emotional stigma connected with AI in effective ways and minimizing the incidents where humans are not aware they are interacting with a non-human and lead to potential heartbreaks.

Social values, political systems, religions and cultures differ across the globe and care must be taken to deploy AI systems [7] which don't take into consideration these distinctions and appropriate measures must be taken either to unify the laws governing technologies through formations of international organizations like Australian Computing Society or through the traditional institutions like UNO, etc.

Some AI disasters have the potential to lock down the research and investment going on in this technology and may also result in the wipeout of human life from the planet as per the predictions of a few noticeable personalities from the domains of AI and cognitive neuroscience and computing [8].

Other possible areas of research with regards to problems related to this new and powerful technology might be automation, unemployment, etc.

We also wish to enlist active centers of research in this domain whose work has played deterministic role in the shaping of this paper and has also served as inspiration to the authors: National Innovation Challenge on Active and Confident Ageing; US National Science Foundation Learning and Intelligent Systems; Human Brain Project; ETHICAA project by French National Agency for Research (ANR); Coordination, Organization, Institutions and Norms in Agent Systems (COIN) and Oxford University's AI Code of Ethics project.

7. REFERENCES

[1] Han Yu et al., Building Ethics into Artificial Intelligence

[2] Nick Bostrom Eliezer Yudkowsky ,
The ethics of Artificial Intelligence,
Draft for Cambridge Handbook of Artificial Intelligence. Wil liam Ramsey and Keith Frankish (Cambridge University Pre ss, 2011)

[3] Bruce M. McLaren,Carnegie Mellon University, Computational Models of Ethical Reasoning: Challenges, Initial Steps, and Future Directions.

[4] Bernd Carsten Stahl | De Montfort University David Wright | Trilateral Research Ethics and Privacy in AI and Big Data: Implementing Responsible Research and Innovation.

[5] Olivier Boifssier EMSE, France, Ethics by Reasoning in Socio-Technical and Cognitive Systems

[6] John C. Havens, IEEE Member, Creating the Human Standard for Ethical Autonomous and Intelligent Systems (A/IS)

[7] Nathan Hutchins, Zack Kirkendoll, and Dr. Loyd Hook, Social impacts of Ethical Artificial Intelligence and Autonomous System Design

[8] Nick Bostron, Superintelligence Paths, Dangers, Strategies

[9] Chapter –Artificial Intelligence, Stuart Russell, Peter Norvig, Artificial Intelligence- A Modern Approach-Prentice Hall (2002)-2nd edition

[10] Chapter – Uncertain knowledge and reasoning, Stuart Russell, Peter Norvig, Artificial Intelligence- A Modern Approach-Prentice Hall (2002)-2nd edition

[11] Chapter –Conclusions, Stuart Russell, Peter Norvig, Artificial Intelligence- A Modern Approach-Prentice Hall (2002)-2nd edition

[12] Jack Karsten, Research Analyst - Center for Technology Innovation, Who should answer the ethical questions surrounding artificial intelligence?

[13] Miguel González-Fierro, 10 Ethical Issues of Artificial Intelligence and Robotics