Artificial Intelligence: A Case Study on Risk Mitigation

Ahmed Al-Gindy

Chair, Department of Engineering AlDar University College Dubai, United Arab Emirates

Abstract

Artificial intelligence technologies are expanding at an extraordinary rate with wide range of applications from machine translation to medical image analysis. Many applications are being developed and there is no doubt that in the near future this will create substantial risk for humanity. In addition, Artificial intelligence is a dual use are of technology as it can be used toward useful of harmful ends, for example autonomous drones can be used to deliver packages faster and easier and on the other hand, they can be used to deliver explosives. This literature review research investigates different types of artificial intelligence risks and proposes ways to mitigate these risks.

Key words:
Artificial intelligence; Risk Mitigation, AI applications

Ⅰ. Introduction

The technology of artificial intelligence is growing very fast. Artificial intelligence applications are transforming our society to better decision making. It has an influence and significances in health community, our economics, security and governance. AI is software that improves and automates the knowledge-based work directed by humans. Artificial intelligence has several definitions based on the field where it comes from such as computer science, cognitive psychology, philosophy, neurology, Authors in [1], defines AI as technology that can think humanly, act humanly, think rationally, or act rationally. Other definitions, which fall into these categories, include: "The study of how to make computers do things at which, at the moment, people are better" and "The study of mental faculties through the use of computational models". Another AI textbook by Poole and Mackworth [2] defines AI as "the field that studies the synthesis and analysis of computational agents that act intelligently," where intelligently refers to acting appropriately, exhibiting flexibility, and learning from experience.

Artificial intelligence is the technology that can think humanly, it will learn, reason, and solve problems in complicated and uncertain environments as well as humans do. AI powerful and full competences do not yet exist and estimated to appear at 2030. Other estimations of artificial super-intelligence " a type of AI that far surpasses human intellect and abilities in nearly all areas" are forecasted to be developed over a century. Below is the diagram by Economist Intelligence Unit (EIU) that forecasts Artificial intelligence development over the next century.

Source: EIU (2016)
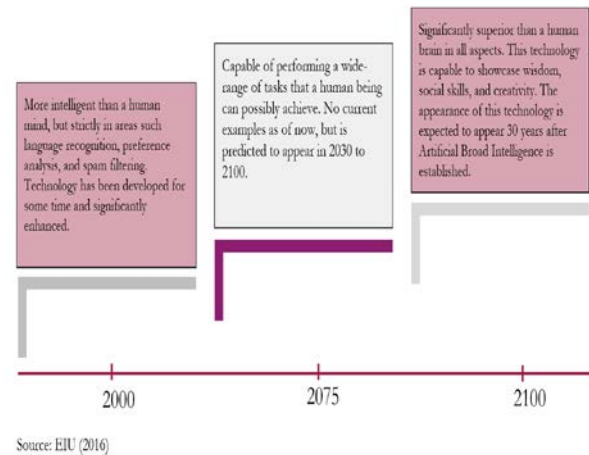
AI capabilities [3] include reasoning, machine learning, robotics, natural language processing, object perception, information storage and retrieval, and speech and handwriting recognition. These tools help power self-driving cars, automated translation, search engine results, and game-playing robots, among many other applications.

Reasoning system is a software system that produces decisions from available knowledge using logical techniques such as deduction and induction. Reasoning systems play an important role in the applications of artificial intelligence and knowledge-based systems [4].

Machine learning is an application of artificial intelligence (AI) that offers the ability to automatically learn and progress from experience without being obviously programmed. Machine learning focuses on the growth of computer programs that can access data and use it to learn for themselves [5].

Natural Language Processing (NLP) is a field of Artificial Intelligence which permits computers to interpret and recognise the human language. A user can have natural talks with computer. Natural language processing makes it possible for machines to process and understand audio and text data to automate tasks like translation, interactive dialogue and sentiment analysis [6].

Perception involves interpreting sights, sounds, smells and touch. Action includes the ability to negative through the world and manipulate objects. If we want to build robots that live in the world, we must understand these processes [7].

Recent researches concern about negative social impacts of AI. This has been growing in recent years as rapid technological developments bring the potentials and risks of AI closer to reality. "AI is more dangerous than North Korean nukes." Said Elon Musk, CEO and product architect of Tesla. On the other side, Mark Zuckerberg, - CEO and Co-founder of Facebook- and others have defended AI as essentially risk-free. There is a debate, as most people are wondering if AI is harmful or useful, dangerous or not. This research highlights the importance to answer the following two questions: What are the risks of AI? And what should be done about them?

II. Dimensions of AI Risk

Artificial Intelligence systems are changing our lives. As these systems become more talented, our world becomes more efficient and accordingly richer. However, with the strong competences of AI systems comes several dimensions for risk. When a robot-related accident or incident takes place, compound interrogations are essential: Is it an error from the robotics hardware by the manufacturer, or operating system. Perhaps it may be an artificial intelligence architecture problem. Is it a legal responsibility error? The opportunity for vagueness is clear. Thus, the need for a complete risk exposure and management solution [8].

A. Unemployment

The employment of labours are concerned mainly with automation. Several ways were developed to automate jobs. Taxi and truck driver companies currently employs millions of individuals in the United States or China alone. What will happen to them if the self-driving cars/trucks manufactured by Tesla's Elon Musk become widely available in the next decade?. Job losses due to the potential of automation technology need to restructure or requalify employees to keep them in jobs. Or simply to create a fair distribution mechanism of wealth created by those machines [9].

The same situation could happen to office workers, as well as to the majority of the workforce in developed countries. This is where we come to the question of how we are going to spend our time. Most people still rely on selling their time to have enough income to sustain themselves and their families. We can only hope that this opportunity will enable people to find meaning in non-labour activities,

such as caring for their families, engaging with their communities and learning new ways to contribute to human society [9].

While lower risk of accidents of the self-driving trucks seem like an ethical choice, if you ask an obedient intelligent car to take you to the airport as fast as possible, it might get you there chased by police helicopters by doing not what you wanted but literally what you asked for.

B. Inequality

The majority of companies are still dependent on hourly work when it comes to products and services. But by using artificial intelligence, a company can cut down on the human workforce, and this means that revenues will go to fewer people [10]. Consequently, individuals who have ownership in AI-driven companies will make all the money. Based on a study made by Deliotte University, staff resources could be freed up to do real work, with people having time to focus on creative projects and deal directly with clients and customers. When adopting artificial Intelligence applications, companies can free up to 1.2 billion of hours with a potential savings of 41.1 billion of Dollars.
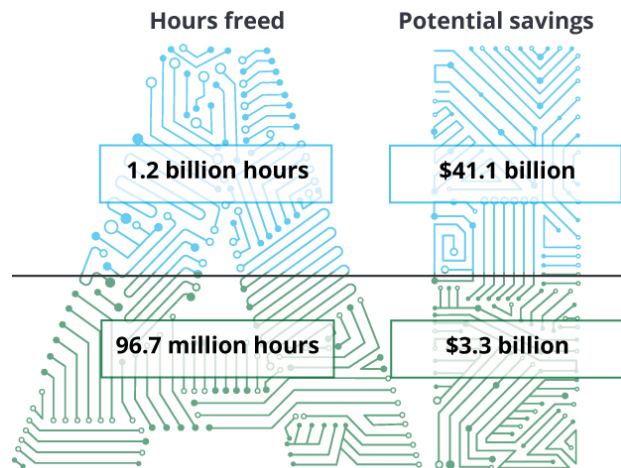
Source : Deliotte University Press

C. Humanity

Artificially intelligent bots are becoming better and better at mimic human interaction [11]. In 2015, a bot named "Eugene Goostman" won the Turing Challenge for the first time. In this challenge, human chat with an unknown entity, then guessed whether they had been chatting with a another human or a machine. Eugene Goostman fooled more than half of the human into thinking they had been talking to a another human being.

This breakthrough is only the start of a phase where we will normally interact with machines as if they are humans; whether in customer service or sales. While humans are limited in the attention and kindness that they can expend on another person, artificial bots can channel virtually unlimited resources into building relationships.

D. Artificial Stupidity

Intelligence comes from learning, whether you are human or machine [12]. Systems usually have a training segment in which they "learn" to make the right decision and act according to their input. Once a system is fully trained, it can then go into test segment, where it is hit with more examples and we see how it performs.

Obviously, the training segment cannot cover all imaginable examples that a system may deal with in the real world. These systems can be tricked in ways that humans wouldn't be. For example, random dot patterns can lead a machine to "see" things that aren't there. If we rely on AI to bring us into a new world of labour, security and efficiency, we need to ensure that the machine performs as planned, and that people can't overpower it to use it for their own ends [12].

E. Artificial Bias

Systems are created by humans, who can be biased and judgemental. Once again, if used right, or if used by those who attempt for social progress, artificial intelligence can become a tool for positive change.

F. Security

The more powerful a technology becomes, the more it can be used for harmful reasons as well as good ones. This applies to robots build to replace human soldiers, or autonomous weapons. AI systems can cause extreme damage if used unkindly. Cybersecurity fights will become even more important than the actual battleground fight. These systems are more fast and capable than humans.

G. Unintended Consequences

Imagine an AI system that is asked to eliminate cancer in the world. After a lot of computing, it spits out a formula that does, in fact, bring about the end of cancer – by killing everyone on the planet. The computer would have achieved its goal of "no more cancer" very efficiently, but not in the way humans intended it.

H. Singularity

Hhumans are on top of the food chain due to our intelligence. We can get the better of bigger, faster, stronger animals because we can create and use tools to control them: both physical tools such as cages and weapons, and cognitive tools like training and conditioning [13].

This brings a serious question about artificial intelligence: will it, one day, have the same advantage over us? We can't rely on just "pulling the plug" either, because a sufficiently advanced machine may anticipate this move and defend itself. This is what some call the "singularity": the point in time when human beings are no longer the most intelligent beings on earth [13].

III. Mitigating AI Risk

Based on the fears discussed above. Reducing risks associated with artificial intelligence is possible Those are as follows:

Strict use of AI-ethics which inject morals of how humans design, construct, use and treat robots and other artificially intelligent beings. It considers both how artificially intelligent beings may be used to harm humans and how they may be used to benefit humans. Researchers, universities and concerned institution such as IEEE who may look after and set the standard to ensure the AI-ethics issues. Additionally, Safety Standard organizations such as ISO, ANSI and other artificial intelligence industries must work to support safety standards.

Forming of a Specialized Risk Management Services that Support policy holders in navigating the new and shifting to artificial intelligence wide applications and helps to mitigate risks.

Utilization of Artificial intelligence such as intelligent robotics must have a regular workplace auditing to insure compatibility with laws mandated by authorities. Illness and Injury Prevention Programs can be adopted to overcome the danger of AI.

Safety robot can be built to control a destructive AI system like the case of weapon systems. It will be able to judge the capacity of the destructive system and make a program from its intelligence which can instantly take the decision and destroy its counterpart. But It should not destroy the human being and this decision should be taken from its artificial intelligence.

Additional risk mitigation steps might come from providing suitable safeguards to control and regulate the AI software itself with the needful of the development of testing protocols for the design of AI algorithms. Also, Improving cybersecurity protections, and validation standards and align them with each industry or individual application

IV. Conclusion

Artificial intelligence risk investigated by many scholars and researchers are real and worthy of immediate attention, even though artificial intelligence can potentially bring massive benefits to humanity. There is high level of uncertainty associated with both the speed of development and the potential risks of artificial intelligence. Thus, this will trigger the need for promoting a culture of responsibility among institutes, researchers and individuals when utilizing artificial intelligence applications. AI researchers and the organizations must deploy a unique position to shape the security dimensions of the AI-enabled world with highlights on ethical considerations and protection mechanisms, expectations and consequences when the artificial intelligence systems fail.

As artificial intelligent technology increases gradually toward the advancements it is necessary to facilitate widespread implementation. For example, In near future robots are going to be in situations that pose a number of courses of action. On the other hand, safety and preventative measures that are fail-safe must be given the priority and the moral responsibility of the researchers. However, it is not only the concern of the researcher and scholars rather of all. There is no doubt that the risk is realized by all. Now this is the time to take necessary action on it.

Destructive weapon and harmful side of artificial intelligence should not be allowed to the organizations or to countries where all developer of AI and researcher should come to an agreement that only the positive side of AI will be discovered. However, with human-being nature there will be a competition and one will try to bring destructive use to artificial intelligence.

Researchers need to advance both the researches on artificial intelligence as well as safety protocols and mechanisms when developing these systems. Industry leaders and policymakers should take the responsibilities to adopt the accurate strategies and regulations to allow the benefits of AI advances to be safely realized.

References

[1] Stuart Russell and Peter Norvig "Artificial Intelligence A Modern Approach" Prentice Hall, 2nd Edition 2002

[2] David Poole and Alan Mackworth "Artificial Intelligence foundation of Computational Agents" Cambridge University Press, 2nd Edition 2010

[3] Kutschenreiter and praszkiewicz, "Application of artificial neural network for determination of standard time in machining", J Intell Manuf, DOI10.1007/s.10845-008-0076-6,Springer Science Business Media, LLC 2008.

[4] Patrick j. Hayes , kenneth m. Ford & jack r. Adams-webber "Reasoning About Artificial Intelligence" in the Journal of Experimental & Theoretical Artificial Intelligence, Pages 247-263, 2007

[5] Jane E. Huggins, Gernot Müller-Putz & Jonathan R. Wolpaw, "Advances in Basic and Clinical Research" in the The sixth international brain–computer interface meeting, 2017.

[6] Kevin Crowston , Eileen E. Allen & Robert Heckman, "Using natural language processing technology for qualitative data analysis", International Journal of Social Research Methodology, Volume 15, Issue 6, 2011

[7] P V S Rao FIETE, "Fifth Generation Computers and Artificial Intelligence", in the IETE Journal of Research, Volume 34, 1988 - Issue 3, 2015

[8] Michael Grüning, "Artificial Intelligence Measurement of Disclosure (AIMD)", in the European Accounting Review, Volume 20, 2011 - Issue 3, 2011

[9] Michael A. Peters, "Technological unemployment: Educating for the fourth industrial revolution", in the Educational Philosophy and Theory, Volume 49-Issue 1, 2016.

[10] Ted Goertzel, The path to more general artificial intelligence, Journal of Experimental & Theoretical Artificial Intelligence, Volume 26-Issue 3, 2014

[11] Scott A. Wallace, Robert McCartney & Ingrid Russell, Games and machine learning: a powerful combination in an artificial intelligence course, Computer Science Education, Volume 20-Issue 1, 2010.

[12] Lee Mackinnon," Artificial Stupidity and the End of Men", Third Text, Volume 31, 2017 - Issue 5-6

[13] Jinhyun Kim , Giacomo Marani , Wan Kyun Chung & Junku Yuh, "Task reconstruction method for real-time singularity avoidance for robotic manipulators", Advanced Robotics, Volume 20 - Issue 4, 2006.