AI Ethics in Next Generation Wireless Networks: A Philosophical Outlook

Rahul Jaiswal, University of Agder, Norway

The Asian Conference on Ethics, Religion and Philosophy 2022
Official Conference Proceedings

Abstract
Artificial Intelligence (AI) algorithms are playing a great role in modern society nowadays. Developing AI-based algorithms more intelligent than humans, for example, it has beaten humans in many specific domains such as chess, and ensuring the use of their advanced intelligence for good rather than bad raises a lot of ethical issues including safety, security, privacy, human dignity, etc. Next-Generation Wireless Communication Networks (NGWNs) is one of the growing areas where new technologies are emerging, for example, 6G from 5G and deploys AI techniques such as Generative Adversarial Networks (GANs) for generating synthetic data in order to develop data-driven models, for example, real-time resource allocation, channel modelling, etc. Consequentialism decides that an action is good or bad depending on its outcome. To that end, this paper presents an overview of AI ethics in the context of wireless networks and investigates how AI ethics is related to AI from the philosophical perspective. It focuses on the ethical implications and moral questions that arise from the development and deployment of AI algorithms. Further, it verifies the claim that consequentialism drives AI, by addressing the possible impacts of deploying AI-based algorithms on society. It also verifies its ethical feasibility with the famous German philosopher Immanuel Kant, who defines ethics as "act as you would want all other people to act towards all other people". Several ethical initiatives taken by countries across the globe to address the ethical concerns and issues emerging in relation to AI-based algorithms are also reviewed.

Keywords: Artificial Intelligence (AI), Consequentialism, Ethics, Generative Adversarial Networks (Gans), Society

Introduction

The rapid developments in artificial intelligence (AI) including machine learning (ML) and deep learning (DL) bring huge potential benefits. For example, helping people to acquire new skills and training, improving voice over internet protocol (VoIP) applications such as Microsoft Skype, Google Meet, Apple FaceTime to connect people across the world, providing real-time quality monitoring of air pollution, etc. However, it is susceptible to errors and bias when developed with the malicious intent and/or trained with the adversarial data inputs. Therefore, it is necessary to explore the ethical, and social aspects of AI systems to avoid unintended, negative consequences and risks arising from the implementation of AI techniques in the society. Moreover, AI has enormous potential to be weaponized in the ways which can threaten public safety, security, and quality of life.

The European Commission's Communication on Artificial Intelligence (European Commission, 2018) defines artificial intelligence as: "Artificial Intelligence (AI) refers to the systems that display intelligent behavior by analyzing their environment and taking actions with some degree of autonomy to achieve specific goals. AI-based systems can be purely software-based, acting in the virtual world (e.g., voice assistant, image analysis software, search engine, speech and face recognition system, etc.) or AI can be embedded in the hardware devices (e.g., advanced robot, autonomous car, drone or internet of thing applications)."

A straightforward definition of "intelligence" is that intelligent behavior is "doing the right thing at the right time". Legg and Hunt (2007) identified three common features that defines intelligence: (a) property that an individual agent has, as it interacts with its environments; (b) agent's ability to succeed or profit with respect to some goal or objective; and (c) ability of that agent to adapt to different objectives and environments. They pointed out that intelligence involves adaptation, learning and understanding. In simplest form, intelligence is "the ability to acquire and apply knowledge and skills to manipulate one's environment".

In interpreting these definitions of intelligence, one needs to understand that for a *physical robot* its environment is the real world, which can be a human environment (for social robot), a city street (for an autonomous vehicle), a home care or hospital (for an assisted living robot), or a workplace (for a workmate robot). Similarly, the "environment" of a software AI is its context which might be clinical (for a medical diagnosis AI), or a public space for face recognition in airport, or virtual for face recognition in social media. But like the physical robots, software AIs almost always interact with humans, whether via question and answer interface (via text for chatbot) or via speech for digital assistants on mobile phones, that is, Siri or Alexa.

The long-term goal of AI is to explicitly develop AI systems that can learn incrementally, reason abstractly and act effectively over a wide range of domains just like the humans can, that is, AI systems should be comparable to the human intelligence. For example, a general-purpose "care robot" may be capable of preparing meals for an elderly person (and washing the dishes afterwards), helping them dress or undress, get into and out of bed or the bath, etc. In other words, AI systems may easily mimic human intelligence and may execute tasks from the most simplest to those that are even more complex.

The remainder of this paper is structured as follows: Section 2 introduces the usages of AI in wireless networks and other real-world applications. Section 3 defines AI ethics from the philosophical perspective. Section 4 describes the underlying impacts of AI on the society. Section 5 outlines the ethical initiatives governed for the ethical concerns in AI and its current progress. The concluding remarks are made in Section 6.

2. AI in Wireless Networks

In the era of exponentially increasing number of internet users and mobile devices, the fifth generation (5G) of wireless communication networks will be rolled out shortly. The Mobile Economy (ME, 2020) reported that by the end of year 2025, the growth in number of mobile subscribers and internet users is expected to be 5.8 billion (71% of global population). Therefore, the next generation wireless networks (NGWNs), starting with 5G, must embrace intelligent signal processing techniques such as AI including ML and DL to efficiently utilize available resources in the dimensions of time, frequency, code and space for achieving higher data-rates, low latency and high reliability in order to fulfil the desired human's quality of experience (QoE). Demand of end-to-end low latency, higher reliability, and higher per-user data-rate directs us to design more sophisticated software-oriented communication networks. The network should be able to support more flexible and on-demand network resources to meet user's expectations. The 5G (Fu et al., 2018), 6G technology (Letaief et al., 2019) and the next generation wireless networks are subjected to a range of non-linear factors due to which the classical theories and models lead to inaccurate results and sub-optimal performance. Therefore, novel intelligent techniques are necessary to design appropriate data-driven models which can perform real-time prediction of network resources and can make its efficient utilization to meet user's desired level of QoE. Using AI techniques, smart decision-making AI algorithms can be developed and introduced into the wireless infrastructure to improve the quality of service (QoS) and performance gains of the end-user device.

There are various problems which are associated in designing AI-based algorithms. For example, a data-driven AI-based communication model employs network data (Kibria et al., 2018) (e.g., channel gain maps, QoS-parameters such as delays, throughputs) to determine availability of resources based on user experience and environmental conditions. However, acquiring such real data in a cost-effective way for the fast-varying network conditions is a major challenge. Moreover, AI-based algorithms are data hungry and most real data are publicly unavailable. Therefore, it is important to address both the data scarcity (lack of data across some network domains and configurations) and the data sparsity (uneven distributions of data) challenges, by designing novel AI-based algorithms in order to generate rich and realistic data that can be used efficiently for building robust data-driven AI-based models in the next generation wireless networks.

In order to design AI-based algorithms, one approach is to obtain real data (e.g., channel gain maps, QoS-parameters such as delays, throughputs) from the network operators, which is very difficult due to data privacy, that is, general data protection regulations (GDPR) to protect citizen privacy. Another approach is to generate pseudo-real data from the testbed. However, getting such testbed (hardware device) is highly expensive. Besides this, it is also possible to generate rich synthetic data using various ML tools such as generative adversarial networks (GANs) (Goodfellow et al., 2014, Donahue et al., 2017).

2.1 Generative Adversarial Networks

Generative adversarial networks (GANs) (Goodfellow et al., 2014) is a new class of generative methods for data distribution learning, where the objective is to learn a model that can generate samples close to the target distribution. The intuition behind GANs is to exploit the potential of deep neural network (DNN) to model both nonlinear complex relationships as well as to classify complex signals. It is a two-player minimax game between a discriminator DNN and a generator DNN. The discriminator (D) learns to distinguish between the data generated by the generator (G) and the data from a real dataset, while the generator learns to generate synthetic samples. During the training, the generator maps an input noise (z) with prior distribution (p_Z(z)) to a synthetic data sample. Then, samples from the real data and those generated from the generator (e.g., channel gain maps, QoS-parameters such as delays, throughputs) are used to train the discriminator for maximizing its ability to distinguish between two categories, namely, real or fake/synthetic data. In theory, when the Nash equilibrium (Porter et al., 2008) is reached between the generator and the discriminator DNNs, then the pair of DNNs provide a generator (G) that can exactly duplicate or reproduce the distribution of real data so that the discriminator would not be able to identify whether a sample is synthetic, that is, whether it is generated by the generator DNN or whether it is from the real data. At this point, the synthetic data generated by the generator DNN are indistinguishable from the real data and thus are as realistic as possible. The performance of GANs, that is, the capability to distinguish between real samples and generated (synthetic) samples is usually evaluated by investigating whether the generated data actually follows the desired real data distribution. The structure of GANs is shown in Fig. 1.

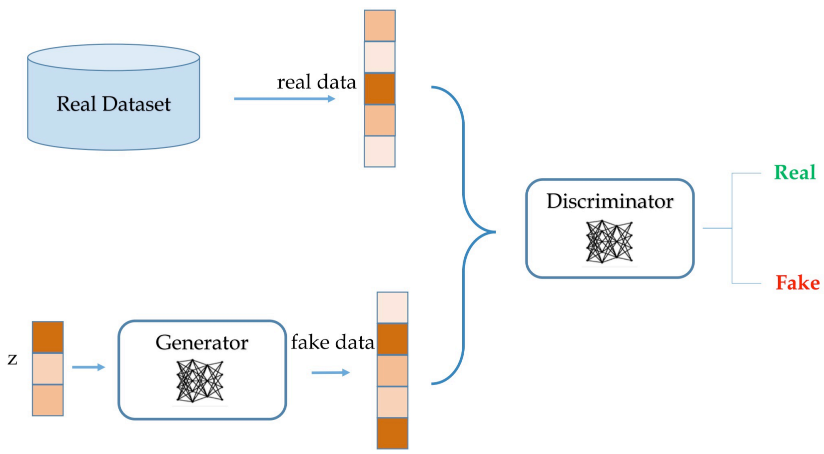

Figure 1: Structure of GANs (Ye et al., 2018).

Using the GAN-based data generation technique, one can meet the requirement of data hungry nature of AI-based algorithms in building robust AI-based framework. The expected impact will be enabling a paradigm shift from the traditional semi-manual sub-optimal operation of current wireless networks to the autonomous almost zero touch highly optimal operation in NGWNs, by providing new data generation AI-based algorithms for data-driven models.

Apart from the application of GANs in the field of wireless networks, there are other interesting applications of GANs in real-world. For example, generating image datasets, generating photographs of human faces, generating realistic photographs, generating cartoon characters, image-to-image translation, text-to-image translation, generating new human poses, photos-to-emojis, face aging, anomaly detection in medical field, 3D object generation, music generation, video prediction, etc. Moreover, in the coming year, one can probably see high-quality videos generating from the GANs.

3. AI Ethics and Philosophical Aspects

Ethics are moral principles that govern a person's behavior or the conduct of an activity. As a practical example, *one ethical principle is to treat everyone with respect*.

Philosophers have debated ethics for many centuries, and there are various well-known principles. One of the most famous German philosopher Immanuel Kant's (Kant, 2012) definition of ethics is: "act as you would want all other people to act towards all other people".

AI ethics is concerned with the important question of how human developers, manufacturers and operators should behave in order to minimize the ethical harms that can arise from AI in society, either arising from poor (unethical) design, inappropriate application or misuse. The scope of AI ethics spans immediate, medium-term and long-term concerns. For example, immediate concern includes data privacy and bias in current AI systems; near and medium-term concern includes impact of AI and robotics on jobs and workplaces; and long-term concern includes possibility of AI systems reaching or exceeding human-equivalent capabilities (so called super-intelligence).

In philosophy, "Consequentialism" is the view that normative properties depend only on the consequences. This embodies the basic intuition that what is best or right which makes the world best in the future. In other words, consequentialism suggests that an action is good or bad depending on its outcome. An action that brings more benefits than harms is good, while an action that causes more harms than benefits is bad. The most prominent example is "AI" and consequentialism drives AI. The next section discusses it by providing impacts of AI on the society.

4. Impacts of AI on Society

The deployment of AI systems brings positive as well as negative impacts on the society as these advanced technologies are used for the people and by the people. AI can also help in taking care of the planet by managing waste and pollution. For example, the adoption of autonomous vehicles can reduce greenhouse gas emissions as it can be programmed to follow the principle of eco-driving throughout a journey, reducing fuel consumptions and greenhouse gas emissions.

Since centuries, people concern about the displacement of workers due to the advancement in the technology. Automation, mechanization, and now more recently AI systems have been predicted to destroy jobs and create irreversible damage to the labor market, that is, people might be replaced by machines, just like the horses were made obsolete by the invention of internal combustion engines (Leontief, 1983). This may impact the economic growth and the productivity as well. The impact of these sizeable changes may be felt by all the members of society. Biavaschi et al., (2012) reported that the young people entering the labor market might be disproportionately affected, since they are at the beginning of their careers, and they may be the first generation to work alongside AI.

Since AI is created by humans that means it can be susceptible to *bias*. Systematic bias may arise as a result of the data used for training the AI systems. In the case of using GANs, there can be intended bias. For example, someone may generate fake data (e.g., channel gains) and may misuse the power spectrum for unethical usages. Similarly, someone may generate fake data e.g., images of an individual and may misuse it for face recognition or fingerprint verification. Unless model developers work to recognize and counteract these biases, AI applications and products may perpetuate unfairness and discrimination.

Data privacy is always important. AI applications need access to large amount of data, but data subjects do not know that how their data are used. There is a regulation, that is, GDPR to protect citizen privacy. However, the regulation only applies to the personal data, and not to the aggregated “anonymous” data that are usually used to train the AI models. This brings up a number of ethical issues. For example, what level of control subjects will have over the data that are collected about them? Should individuals have a right to use the model, or at least to know what it is used for, given their stake in training it? How to prevent the identity or personal information of an individual involved in training a model?

Further, AI machines and its uses in the society could have huge impact on the criminal law. The entire history of human laws has been built with the assumption that people make decisions, and not the AI machines. In a society in which complicated and important decisions are being handed over to the AI algorithms, there is the risk that the liable legal frameworks will be insufficient. The most important legal questions with AI machines are: *who or what should be liable for criminal, and contractual misconduct involving AI and under what conditions?* Pagallo (2018) questioned that what would happen, for example, if an AI program is chosen to predict successful investments and pick up on market trends, made a wrong evaluation that led to a lack of capital increase and hence, to the fraudulent bankruptcy of the corporation? Moreover, what should happen if someone commits banking fraud by forging a victim’s identity, including mimicking a person’s voice using AI models.

Our society relies on “*Trust*”. Therefore, for AI systems to take on tasks, such as surgery, people will need to trust the technology. Trust includes aspects such as fairness (i.e., AI will be impartial), transparency (i.e., our ability to understand how an AI system arrives at a particular decision), accountability (i.e., someone should be accountable for the mistakes made by AI system), and control (i.e., how one can shut down an AI system that becomes too powerful). Moreover, the increasing usages of AI systems come with large computational resources due to the training of huge amount of data, which necessitate substantial energy consumption and huge amount of carbon emissions (Strubell et al., 2020), affecting the natural environment.

With the definition of *ethics* by German philosopher Immanuel Kant’s (Kant, 2012) that “act as you would want all other people to act towards all other people”, it is feasible and important to the AI model developers that they should consider the various ethical aspects while designing and developing AI models such that the AI models should have minimal bias and very small chance of error. It should act as per the requirements (expectations) of the people for the benefit of the society.

5. Ethical Initiatives for AI and Current Progress

While AI technology may be used for good, potentially it may be misused. Hence, there is a need of ethical considerations accompanying the development and usages of AI. It can range from the fundamental human rights of citizens within a society to the security and utilization of gathered data, from the bias and understanding about the consequences and usage of any given AI system, leading to wrong decisions and subsequent harms. A wide range of initiatives have been sprung up in response to the ethical concerns and issues emerging in relation to AI. For example, The Institute for Ethics in Artificial Intelligence, Germany (IEAI, 2019); The Institute for Ethical AI & Machine Learning, United Kingdom (IEAML, 2018); AI4People, Belgium (AI4People, 2017); The Foundation for Responsible Robotics, Netherlands (FRR, 2017); Saidot: Enabling responsible AI ecosystems, Finland (Saidot, 2019); The Japanese Society for Artificial Intelligence, Japan (JSAI, 2017); The Future of Life Institute, United States (Fli, 2017); The Institute of Electrical and Electronics Engineers (IEEE), United States (IEEE, 2019), etc. These initiatives agree that the AI should be designed, developed, deployed, and monitored in ethical manner in their respective area. They aim to identify and form ethical frameworks and systems that establish human beneficence at the highest levels, prioritize benefits to both human society and the environment, and mitigate the risks and negative impacts associated with AI, with a focus on ensuring that AI is accountable and transparent.

Inspite of these initiatives which adhere to the view that AI must not impinge on the basic and fundamental human rights, such as human dignity, security, privacy, freedom of information, and protection of personal data, there are still ethical dilemmas which concern that “How do we ensure that AI upholds such fundamental human rights and prioritizes human well-being? or AI does not disproportionately affect vulnerable areas of society, such as children, disabilities, or the elderly, or reduce quality of life across society?” Moreover, one of the main concerns about AI is: its transparency, accountability, security, reproducibility, and interpretability. “Is it possible to discover why and how an AI system made a specific decision or acted in a different way?” as it has direct safety consequences about physical harm, for example, medical diagnosis systems. Without transparency, users may struggle to understand the AI systems and its associated consequences and then it will be difficult to hold the relevant person accountable and responsible. In other words, wherever AI is used to supplement or replace human decision-making, there is consensus that it must be safe, trustworthy, reliable, and act with integrity.

At present only one standard, namely, “British standard BS8611” (BS, 2016) addresses ethical design of AI systems (specifically for robots). BS8611 provides guidance on how designers can identify potential ethical harms, undertake an ethical risk assessment of their robot or AI, and mitigate any ethical risks identified. However, BS8611 documentation is not available publicly. The IEEE (IEEE, 2019) is progressing towards developing a number of standards and regulations that affect AI in a range of contexts. Currently, 14 IEEE standards working groups are working on drafting so-called “human” standards that have implications for AI.

6. Conclusion and Future Work

This paper presents a review on the ethical issues surrounding AI in the context of wireless networks and related areas that highlights a wide range of potential impacts, including in the personal, social, financial, legal and environmental domains. It, further, presents various examples that can impact the society in either ways, good or bad. For example, how AI can benefit the society and how AI can harm public privacy, etc. Inspite of several benefits, there are tremendous drawbacks of AI usages (good as well as bad consequences). It has been presented that how the current regulations are not sufficient to decide: who is responsible for the actions of AI or unethical practices, that is, either the programmers, or the manufacturers or, the end-users. Finally, these many examples conclude that AI ethics is driven by the consequentialism either in good or bad manners, that is, consequentialism drives AI. Further investigation in this context is the part of future direction.

References

AI4People, Belgium (2017). https://www.eismd.eu/ai4people/

Biavaschi, C., Eichhorst, W., Giulietti, C., Kendzia, M. J., Muravyev, A., Pieters, J. & Zimmermann, K. F. (2012). Youth unemployment and vocational training.

British Standard BS8611 (2016): Guide to the ethical design of robots and robotic systems. https://shop.bsigroup.com/ProductDetail/?pid=000000000030320089

Donahue, J., Krähenbühl, P., & Darrell, T. (2017). Adversarial feature learning. In 5th International Conference on Learning Representations (ICLR).

European Commission (2018). Communication from the Commission to the European Parliament, the European Council, the European Economic and Social Committee and the Committee of the Regions on Artificial Intelligence for Europe. https://ec.europa.eu/digital-singlemarket/en/news/communication-artificial-intelligence europe

Fu, Y., Wang, S., Wang, C. X., Hong, X., & McLaughlin, S. (2018). Artificial intelligence to manage network traffic of 5G wireless networks. IEEE Network, 32(6), 58-64.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., & Bengio, Y. (2014). Generative adversarial nets. Advances in neural information processing systems, 27.

Kant, I. (2012). Kant: Groundwork of the metaphysics of morals. Cambridge University Press.

Kibria, M. G., Nguyen, K., Villardi, G. P., Zhao, O., Ishizu, K., & Kojima, F. (2018). Big data analytics, machine learning, and artificial intelligence in next-generation wireless networks. IEEE access, 6, 32328-32338.

Legg, S., & Hutter, M. (2007). Universal intelligence: A definition of machine intelligence. Minds and machines, 17(4), 391-444.

Leontief, W. (1983). National perspective: The Definition of Problems and opportunities. In National Academy of Engineering Symposium (Vol. 30).

Letaief, K. B., Chen, W., Shi, Y., Zhang, J., & Zhang, Y. J. A. (2019). The roadmap to 6G: AI empowered wireless networks. IEEE Communications Magazine, 57(8), 84-90.

Mobile Economy (2020). https://www.gsma.com/r/mobileeconomy/

Pagallo, U. (2018). Apples, oranges, robots: four misunderstandings in today’s debate on the legal status of AI systems. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 376(2133), 20180168.

Porter, R., Nudelman, E., & Shoham, Y. (2008). Simple search methods for finding a Nash equilibrium. Games and Economic Behavior, 63(2), 642-662.

Saidot: Enabling responsible AI ecosystems, Finland (2019). https://www.saidot.ai/

Strubell, E., Ganesh, A., & McCallum, A. (2020). Energy and policy considerations for modern deep learning research. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 34, No. 09, pp. 13693-13696).

The Foundation for Responsible Robotics, The Netherlands (2017). https://responsiblerobotics.org/

The Future of Life Institute, United States (2017). https://futureoflife.org/

The Institute for Ethics in Artificial Intelligence, Germany (2019). https://ieai.mcts.tum.de/

The Institute for Ethical AI & Machine Learning, London (2018). https://ethical.institute/

The Institute of Electrical and Electronics Engineers, United States (2019). https://www.ieee.org/

The Japanese Society for Artificial Intelligence, Japan (2017). https://www.ai-gakkai.or.jp/en/

Ye, H., Li, G. Y., Juang, B. H. F., & Sivanesan, K. (2018). Channel agnostic end-to-end learning based communication systems with conditional GAN. In IEEE Globecom Workshop (pp. 1-5).

Contact email: rahul.jaiswal@uia.no