

# Ethics in Artificial Intelligence and Machine Learning

1<sup>st</sup> Mr.Soham Shripad Gumaste  
School Of Computer Science  
MIT World Peace University  
Pune, India

2<sup>nd</sup> Ms.Madhuri Pote  
School of Computer Science  
MIT World Peace University  
Pune, India

**Abstract**— Artificial Intelligence (AI), a field of computer science enabling computers to be better than humans at traditionally human tasks, is a developing field and hence it is of utmost importance to establish a guideline for ethical practices going forward. It can be used ethically to maximize quality of life in every aspect such as health-care, transport and city planning. However, unethically, it can be used to gather data and spy on individuals violating their personal space, as an enabler of an Orwellian state and as a means of war. Action needs to be taken now to ensure AI is used ethically. We describe some ethical and unethical uses cases and propose some laws and regulations to ensure ethical use of AI in the future.

**Keywords**—Ethics, Artificial Intelligence, Facial Recognition, Home Assistants,

## I. INTRODUCTION

Modern Artificial Intelligence allows computers to learn by themselves. It has been demonstrated that with this technology, computers can outperform humans in human-like tasks like identifying objects in an image. Hence, it is clear that this technology can a bane or a boon to humanity, both on indefinitely large scales. Consequently, it is important to discuss ethical practices in this developing field.

## II. INTRODUCTION TO MACHINE LEARNING AND ARTIFICIAL INTELLIGENCE

Historically, computers were best known for their ability to solve problems intellectually difficult to humans. Their ability to process numbers at a pace unmatched by any human was their greatest strength. Large products, factorials and integrals are suitable examples.

Nowadays, there is a greater need for computers to perform tasks trivial to humans, but previously impossible for computers. We as humans can intuitively identify cats, dogs, other humans and other near infinite number of objects with ease, however the same task was, until recently, near impossible for computers. This need for computers to perform human-like tasks comes from the trends of automation and mechanization in every industry.

AI, and by implication Machine Learning, fills this need and hence is in very high demand. In essence, Machine Learning involves teaching a computer system to make decisions without hard-coded instructions. It allows computers to learn from experience, from previously gathered data.

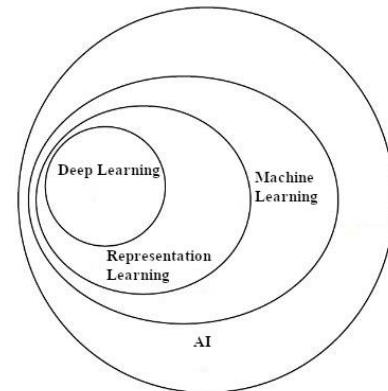


Figure 1: Venn diagram illustrating the relationship between AI, ML and Deep Learning [5]

### A. Deep Learning and Neural Networks

Deep Learning is a subset of Machine Learning which involves multi-layered graphs called "Neural Networks" or "Multilayer Perceptron" where every node is a "Neuron". A Deep Neural Network is a multi-layered mathematical model with multiple layers (of neurons or perceptron's) between the input layer and output layer. The model then "learns" the required mathematical functions to map the input to the output.

These can be used to perform simple linear regressions, but can scale up near indefinitely. A suitable example of a simple neural network would be the prediction of the chances of a given person having diabetes based on some demographic information including age, weight, height etc. Here, the information about the person (also called "features") would be the input layer and chances of diabetes would be the output. This model would need to be trained on actual data for it to learn the correlation between the information of a person (features, or the input) and their chances of having diabetes (the output).

There are different types of neural networks including Convolutional Neural Networks (CNNs) for image and video processing and Recurrent Neural Networks (RNNs) for natural language, which are beyond the scope of this paper.

### B. Use Cases

These days AI and specifically Deep Learning is used in a variety of fields. Self-driving cars, for example, use multiple Deep Learning technologies including Convolutional Neural Networks for detecting traffic signs. Various financial institutions including investment banks are deploying neural

networks to analyze and predict the movements and trends of the global markets.

A use case significant to the question of ethics is that of advertisements and recommendation engines. Advertising companies like Google and Facebook use deep learning technologies to analyze user behavior on their platforms to serve ads that would interest the user. Content based companies like YouTube and music providers make similar use of such technologies to predict which content the user would want to consume next.

### III. ETHICS OF THE FREE SERVICES ON THE INTERNET

The Internet is full of "free" services, with common examples being most Google services like Gmail, Maps, YouTube and others. These services are "free" in the sense that they do not require money to be paid by the user. However, almost all of these services require payment in the form of user privacy. The service providers entice new users with the free price tag to harvest their data, which is then used to fuel advertisement engines, generating revenue for the provider.

This practice poses an obvious ethical question: Do we want a future devoid of user privacy or one where every common service isn't free?

Unfortunately, there is no easy answer to this question. The free price tag enables access to technology to the destitute all the while being a clear violation of privacy. However, I believe the future should be one where the hunger of data of deep learning isn't satisfied by unethical collections, where a balance is struck between enabling access and user rights.

## IV. EFFECTS OF AI ON PERSONAL LIBERTIES

### A. Facial Recognition

The use of facial recognition technology has been in the news. Facial recognition uses Convolutional Neural Networks mentioned above. According to the American Civil Liberties Union [1], the greatest concern with this technology is the general, suspicion-less surveillance and spying that does not require the knowledge or consent of the subject. Such practices, especially by the State, are totally unethical as they intrude in the personal lives of the subjects. Hence, I propose extreme regulations if not a complete ban on facial recognition technologies.

Used carefully facial recognition technology can prove beneficial. It can be used to locate or identify criminals and missing persons. However, it is still debatable whether such a use case is ethical, as it would be trivial for such a system to transition to one of surveillance and spying.

### B. Biases in AI

Since AI relies on data provided by humans for training, it is extremely easy for human biases based on race, ethnicity, religion and the like to creep into AI models. This is what happened with Amazon's experimental recruitment engine, which was designed to sift through resumes to determine the best candidates [2]. The engine, apparently due to its training dataset, learned to associate feminine words with a lower chance of acceptance, exhibiting a clear case of gender bias.

Other incidents such as the case where a black child was ranked as a higher risk than a white serial thief [3] by an AI algorithm used by a US court show us the grave consequences

of letting human biases slip into AI models. Both the child and the man were accused of shoplifting worth a similar amount. Yet, the black child was deemed to be at a greater risk of committing a crime in the future, which we now know to be false.

Additionally, the rise in use of data hungry AI models in various forms of selections (interviews, schools etc.), poses another issue. Due to the nature of such models, it is impossible for its users-or developers-to explain why it made a particular decision. The model in essence turns into a black box, which no one can explain and hence hold accountable.

With all this in mind, we believe the only solution for this is open datasets and models. All such models should be open to the public for independent review to avoid any bias.

## V. ETHICS OF HOME ASSISTANTS

Home assistants are a relatively new group of products, pioneered by companies including Amazon, Google and Apple. They are in essence "smart" speakers that respond to commands as if they were human.

The issue with home assistants is clear when one realizes that these are always listening and recording, not only when they are summoned. The companies use this recorded data for whatever they please [6]. This practice clearly violates user privacy.

Used ethically, home assistants can improve quality of life. They can efficiently manage certain aspects of the home; for example a smart thermostat can save power compared to a manual one. However, it is critical that this ethical use is ensured and enforced.

## VI. EFFECTS OF AI ON POLITICS AND GOVERNANCE

AI can be used by Governments and leaders to both empower and progress the people, or to subjugate and impoverish them. Going forward, AI will play a critical role in the governance of States.

### A. Potential of AI as an enabler of fascism

AI can be used by fascist leaders to maintain their control. Such a State can deploy Natural Language Processing algorithms to monitor social media and identify opposition. These algorithms can also flag people as "potential" opposition, based on their online activities. The State is then free to oppress them by denying public services or worse, potentially using facial recognition technologies mentioned earlier. It can also identify potential areas of protest and impose curfews to silence public opinion.

### B. A Data Driven Democracy

In contrast, AI can do wonders in governance if used properly. City councils can use AI to predict traffic jams before they occur, and take necessary actions. AI plays a crucial role in the ever worsening problem of traffic in cities. Cities can also use AI to efficiently plan new roads in a way that minimizes traffic.

Health-care too can be augmented by AI. It can analyze patient data and predict chronic illnesses, which are best treated in early stages. It can also suggest ideal locations for building of hospitals so that no patient has to travel across the city to access health-care. Autonomous ambulances can aid in the

transportation of critical patients to hospitals with utmost efficiency.

In agriculture too AI shows its might. Meteorology departments can use AI to analyze rainfall and predict the ideal time for the sowing of seeds. Moreover, AI can be used to mechanize almost every aspect of the farm, reducing its laborious nature. Automated harvesters and processors are already surfacing in some parts of the world. The transportation of produce to its destination too can be optimized.

AI can be used in similar ways in various different fields to maximize quality of life, leading to a data-driven utopia

## VII. USE OF AI IN WAR

It is easy for one to conclude that AI should never be used in warfare. Weapons of mass destruction augmented by the power to learn by themselves can prove catastrophic for humanity as a whole. According to this open letter [4] signed by the likes of Stephen Hawking, Elon Musk and Steve Wozniak, as AI based weapons won't require any expensive components, they could become ideal for tasks such as assassinations, destabilizing nations, subduing populations and ethnic cleansings. Hence, such weapons should be banned outright through an international treaty.

## VIII. CONCLUSION: STEPS TO ENSURE THE RIGHT WAY FORWARD

It should now be evident that legislative steps need to be taken now to ensure that the field of AI heads in the right direction. We believe privacy should be declared a fundamental human right, putting an end to the privacy invading practices of the data giants.

Next, companies should be regulated so as to disallow intrusive tracking practices. Companies manufacturing home assistants should be required to keep all user data within the user's home network. Lastly, India should also pioneer an international treaty banning the use of AI weapons in warfare.

Apart from legislation, public education too plays a critical role in ensuring ethical AI. Schools should introduce students to AI at an appropriate age and imbibe in them the need for the ethical use of it. Workshops should also be conducted on the same for the public at large.

TABLE I. SUMMARY OF ETHICAL AND UNETHICAL USES OF AI

Use Case	Unethical Use	Ethical Use
Facial Recognition	Surveillance	Catching criminals, Finding missing persons
Court System	Unfair verdicts based on race	Fair, expedited justice
Media	Generate biased, sensationalized articles	Generate articles without any bias
Home Assistants	Listening to every conversation	Efficiently managing home appliances
Governance	Silencing public opinion	Efficiently managing welfare schemes
Warfare	No ethical use, must be banned	

## REFERENCES

- [1] Facial recognition technology." <https://www.aclu.org/issues/privacy-technology/surveillance-technologies/face-recognition-technology>.
- [2] "AI bias: It is the responsibility of humans to ensure fairness." <https://www.information-age.com/ai-bias-123479217/>.
- [3] J. Angwin, J. Larson, S. Mattu, and L. Kirchner, "Machine bias risk assessments in criminal sentencing," *ProPublica*, May, vol. 23, 2016.
- [4] "Autonomous weapons: An open letter from AI & robotics researchers." <https://www.futureoflife.org/open-letter-autonomous-weapons/>.
- [5] I. Goodfellow, Y. Bengio, and A. Courville, *DeepLearning*. MIT Press, 2016.
- [6] Matt Day, Giles Turner, Natalia Drozdiak, "Amazon Workers Are Listening to What You Tell Alexa", <https://www.bloomberg.com/news/articles/2019-04-10/is-anyone-listening-to-you-on-alexa-a-global-team-reviews-audio>