Reliability on Artificial Intelligence

Ziyad Maknojia, Rushank Rane

[4]*Student, NMIMS University, MPSTME, Mumbai, India.*

Abstract:

Artificial Intelligence is a software, in general terms, that is made with the concept of making it think like a human or more than a human. The fact that Artificial Intelligence can be made to think and do things beyond human imagination have raised a question of whether these machines should be trusted or not. This question can be figured out to some extent by looking closely to the development of ANN (artificial neural network) and deep learning of machines which is explained further and is explained by examples of self-driving cars and artificial intelligence in defense.

Keywords: *Artificial neural network (ANN), deep learning, frozen software, neocortex.*

1. Introduction:

With the rise in technology and the simplicity in work provided by them has increased the dependency of humans on them. The best major evolving example of such technology is Artificial Intelligence. Artificial intelligence is arguably the most exciting field in robotics. It's certainly the most controversial: Everybody agrees that a robot can work in an assembly line, but there's no consensus on whether a robot can ever be intelligent. Like the term robot itself, artificial intelligence is hard to define. Ultimate AI would be a recreation of the human thought process -- a man-made machine with our intellectual abilities. This would include the ability to learn just about anything, the ability to reason, the ability to use language and the ability to formulate original ideas.

Roboticists are nowhere near achieving this level of artificial intelligence, but they have made a lot of progress with more limited AI. Today's AI machines can replicate some specific elements of intellectual ability. Computers can already solve problems in limited realms. The basic idea of AI problem-solving is very simple, though its execution is complicated. First, the AI robot or computer gathers facts about a situation through sensors or human input.

The computer compares this information to stored data and decides what the information signifies. The computer runs through various possible actions and predicts which action will be most successful based on the collected information.

2. Trust:

Now the questions arise of whether it is reliable or safe to trust such an artificial intelligence because we are talking to include such a technology in doing our daily tasks which include many things which cannot go wrong at any cost. Trusting such an artificial intelligence creates an indirect link of trust between the user and the developer. As mentioned above artificial intelligence takes inputs from surrounding and react accordingly. So now is this reaction safe? How can anybody say that the decision taken by it is safe and accurate? This is the thing that differentiates a machine brain with human brain. The decisions taken by humans comprises of various steps like thinking, reasoning etc. The decision taken may not be accurate or right for a particular decision. Now developing such an artificial intelligence having reasoning power or thinking power via a code or algorithm is a difficult task.

3. Artificial Neural Networks (ANN):

The artificial neural networks form's the base for any artificial intelligence. ANN are similar to the neuron network in human brain which consists of millions of neurons connected to each other and share information via electric signals. The neural network in any artificial intelligence consists of logic that helps inn communicating with other artificial neurons and making a unanimous discussion regarding any situation. The ANN can be created by using a simple 9-line python code.

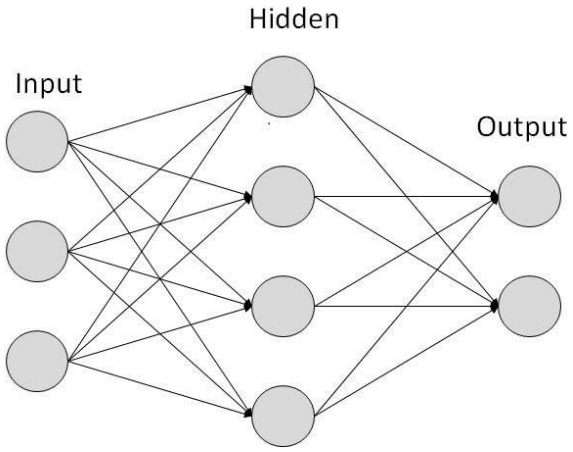

Fig 1: Structure of a Neural Network

Generally, the system of artificial neural network works on the concept of weight assigning. The paths indicate the flow of signals from one node to another. These paths are assigned weights i.e. an integer value and if the network produces a good desired output no alteration is done in these weights and if any error occurs or poor result takes pace as the output of the network the weights are adjusted accordingly.

The neural network works as follows:

Example: Consider a car standing in front of an AI. The AI has to identify the car.

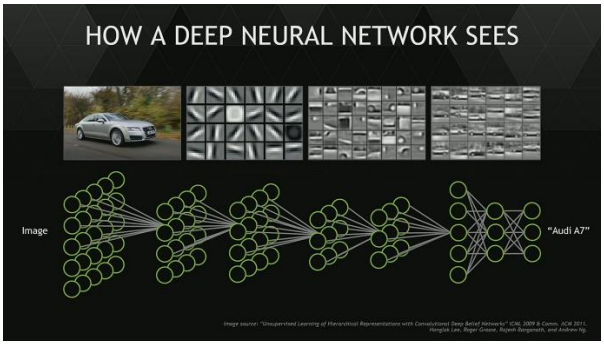

Fig 2: Working explained by using a car example

1. It takes input via a sensor, camera, etc. and divides the input taken in the form of image into smaller components. This smaller division created are larger in number. These divisions are analyzed and related to each other with the help of the logic introduced in the neural network. The AI carries out the analysis in every small part of the car.

2. The relation derived from the above step reduces the number of divisions. The AI will scan for car tires, door handles, headlight, taillight, window viper, etc.

3. Then the AI combines these results and forms the resulted image of a car. The result may consist of many images of a car. As it is difficult to get an accurate result at once.

4. The final step includes comparison between the original car and the AI derived image of a car. The AI searches for the resulted image that matches more accurately with the original one and then searches for the car name in its database or on the internet and then gives the name of the car.

5. This process may seem to be simple but every element in the neural network is responsible to carry out its calculations and give the result. So now creating such and advanced ANN is really a difficult job which determines whether the AI should be trusted or not.

4. Deep Learning:

Now let's move towards in learning in detail how the machines learn things.
As mentioned above an ANN is a machine brain that depicts the human brain. Now basically developers include that part of human brain or try to stimulate that part of brain that involves maximum thinking and learning i.e. neocortex. Neocortex is a part of mammalian brain that is involved in higher-order functioning but this idea is old and has led to many failures. Due to various improvements in mathematical formulas and logic building this old method has now proved to be useful and efficient. The software used recognizes patterns in digital form of sounds, images, etc.
The question mentioned above of whether an AI should be reliable or not can be cleared more by studying the concept of deep thinking and implementing the software's in correct organization with the hardware. Many AI's including deep thinking are developed out of which are driverless cars.

5. Automated Cars (AI):

Most of the artificial cars have been designed and tested. Many evolved techniques with the help of sensor fusion have helped researchers to develop a 3-D map of all activities that happen around the car. They have proved to be useful and dependable. Like a car developed by NVIDIA was tested and it responded very well. This car, like other artificial intelligence, didn't have machine learning concept but it worked on the algorithm to make its decision about handling the steering wheel and the brakes and other systems. Information from the car's sensor were provided as the input to the ANN that processed data and provided the outputs. Now the question arises that is this car capable of making all decisions that a human driver makes? There exist many situations where critical decisions are to be taken and in a fraction of time and if they are not perfect the consequences may be hazardous like it may hit the tree or stop on the green signal in the middle of the road. The systems are way too complicated that even the engineers who developed it struggle to fix errors in it. Such an AI cannot be implemented in real world as no one ones what may be the consequences. The algorithms work on the basis that they get their inputs from the environment. This complex algorithm can start developing their own logic because of the concept of machine learning embedded in their algorithm. This may lead to development of a new way of evaluating output from the input in a different way i.e. a new algorithm may even be generated by the machine itself which may be a good one or a bad one. This implies too much of risk to human life. Therefore, experiments conducted though being successful are not implemented in real life. Apart from the topic of AI in car's there many fields in which AI can be proved a boon for mankind like surgery procedures, decision making Multinational companies, military, etc. But imagine the consequences if the machine deployed in this field disrupts the work. It may kill someone in the field of surgery, can bring an entire multinational company to seize, can start a war between two nations and many other effects.

6. AI in Defense:

Some of the nations in today's world have proposed an idea of integrating the artificial machines in their defense system. Looking at the positive side of this scenario this artificial machines will replace the frozen software. The frozen software's are the software's that are designed to do a specific job like our basic computer software's. The artificial machine's work on complex algorithm's that may even consist of many frozen software's working together with co-ordination. As this machine learn and expand their selves starting from a small algorithm their role in defense increases as there are many different aspects that needs to be taken in defense as we are talking about a nations defense. Looking at the negative aspect of this integration of machine's in defense is that they can be corrupted and used which may have such a disastrous effect that no one can even imagine as this machine's get their input from surroundings. Their system may be provided with corrupted data or poisoned data or the entire program or algorithm may be changed and they can be used negatively. Adding to these consequences the concept of learning from the environment is the major aspect that can lead to destruction. The machines can use faulty ways in achieving their task and can cause even wars between nations or end of humankind.

Apart from physical defense lets have some glimpse on cyber defense or security. Many artificial robots are deployed to overview the activities going on the network. The hackers or attackers crack organizations precious data and sell it on black market. The machines designed deployed on the network take precautionary measures to resist such activities. At the negative side this robot can be manipulated and can be used against their own boss resulting in loss of precious data.

7. Literature Review:

From the paper (Pozna Claudiu, Antonya Csaba,2016), we observed a that cars get their input from sensors and is feed to the system (AI). The system consists of a program manager which reads the driving circumstances and produces the output in form of behavior of car i.e. what action it takes according to the situation that have appeared. Now assuming an ideal situation this AI works efficiently but in our daily work there may be many such situations that we are not aware and we may not have included in our AI which leads to exceptional situations which creates ambiguity for the AI which may create chaos or negative consequences.

From the paper (I. White,2002) A unique idea of insight came into picture. The human brain has various departments for thinking, reasoning, imaginations, etc. With the help of this factors the humans are able to make decisions which they feel right, but these decisions may not be accurate and can lead to circumstances. Invoking such factors of human brain in an artificial machine that to developed by a human having its own factors of thinking leads to an AI having the developer's brain. Now anyone can imagine the risk of implementing such an AI in defense. Considering all the factors that are involved in the decision process and also thinking of every exceptional case that may occur and developing an AI accordingly is somewhere near to impossible. So, involving such a machine in defense is not a very good approach for reducing human efforts.

8. Conclusion:

Nothing in this world is perfect which makes our question clear of whether to be dependent on the artificially created machines that may contribute to our daily work which includes small and very big tasks. This doesn't imply that we shouldn't include these machines for our task fulfillment but we should be aware of the limit to which they are included and be well prepared for the consequences if any things go upside down. The consequences we are talking about are not that small or simple which can be resolved by just shifting some things but these consequences are the one that may even can't be resolved and can even lead to extinction of human kind. Scientists fear that a day known as singularity may occur in future in which these machines may dominate over humans i.e., they have become super intelligent that they can do things without any human interference which can even lead to human extinction. The results of this event can't be even imagined. So, their restricted use and the desire of being less ambitious of humans may save human kind and we may gain the benefits of these machines in our daily life.

9. References:

9.1 Journal Articles

[1] Pozna Claudiu, Antonya Csaba, Issues about Autonomous Cars, IEEE 11th International Symposium on Applied Computational Intelligence and Informatics (SACI), 2016.

[2] I. White, Artificial Intelligence in defense: Wanted and Unwanted Research, IEE Colloquium on Strategic Industrial Issues in AI in Engineering, 2002.

9.2 Websites

[3] *https://phys.org/news/2016-10-reliability-artificial-intelligence.html.*

[4] *https://medium.com/technology-invention-and-more/how-to-build-a-simple-neural-network-in-9-lines-of-python-code-cc8f23647ca1.*

[5] *http://science.howstuffworks.com/robot6.html.*

[6] *https://www.livescience.com/49009-future-of-artificial-intelligence.html.*

[7] *https://www.rand.org/blog/2017/09/artificial-intelligence-and-the-military.html.*

[8] *https://www.google.co.in/amp/s/www.entrepreneur.com/amphtml/281040.*

[9] *https://phys.org/news/2016-10-reliability-artificial-intelligence.html.*