CSCI 470: Machine Learning Ethics Writing Assignment
Self-Driving Cars and Augmented Reality Game Trial

Nadima Dwihusna

The ethics session showcase a Black Mirror episode entitled "Playtest." The episode featured Cooper, an American who was travelling and got stuck in London with not enough money to get back home. To gain some income, Cooper decided to participate in a temporary experimental game testing job at a company called SaitoGemu. The newly developed game allowed Cooper to experience virtual reality. The game was cruel, taking Cooper's worst memory and nightmares to be used against him. The algorithm uses neural networks to learn the player's fears and apply the feared objects to the augmented reality haunted house game. The algorithm learns how the player response to his fears and increase the intensity level substantially. The ethical issues rise by the end of the game, when Cooper passed away from playing the horror game. Even though the SaitoGemu game company was doing trials on Cooper to make sure the game is safe and marketable for the public, the ethical issue still arises when testing newly potentially harmful algorithms to players. In the end, the death was determined to be the player breaking the rule of the game – turned on phone signal causing interference to the algorithm.

There are two stakeholders who are affected by the issue: Cooper and the SaitoGemu game company. The player Cooper was affected as he experiences his worst nightmares and faced his death in the augmented reality game. He became very dysfunctional and begs to be taken out of the game. The game manufacturer, SaitoGemu was responsible for developing the game and making sure the game is marketable. The SaitoGemu was not completely at fault for the player's death since the player's signed a voluntary agreement contract for the game trial. The player agrees to take the risk and SaitoGemu clearly defines the possible risk there are for playing the game at the contract. But with most issues in manufacturing, the list of risks and rules are not clear and emphasized enough to the player. As a result, the player's death was caused by not following the rules of the game (phone off).

Another similar ethical issue from the ethics session is also in self-driving cars. The stakeholders involve the driver of the car and the car manufacturer. Like the black mirror episode, the ethical issue arises when there are accidents and deaths involved. The manufacturers are responsible in developing the algorithm of the self-driving cars and stating the rules for drivers on how to use the cars. In most accidents, the drivers are at fault as they do not follow the rules when driving the cars. But same with the black mirror episode scenario, the manufacturer is also at fault for not listing the risks and rules on how to drive the car clearly.

Comparing this scenario to real life, it is difficult to impart ethics into a machine. People would hope that the machine will help ethical decisions. To understand the implications behind the algorithm. In this era, the scientists and engineers has too much criticism. There needs to be a model that connects between the public and the scientists and professional engineers. For example, these automotive cars have been expected to have a higher safety increase of these cars than traditional cars. This is psychological issue, no more than an ethical issue. The public expects higher expectations to self-driving cars than traditional cars even though traditional cars also has accidents. In fact, there have been studies showing self-driving cars to have lower accident rates than traditional cars.

The traditional responsibility of technology is changing. In general, most of these people who dies in the incidents of the self-driving car or game algorithm tries should not be blamed by the company's algorithm. For example, the first person who passed away from self-driving car did not follow the instruction – put hands on the wheel. Instead, the driver who passed away from the incident was watching a movie during the crash. Same with the black mirror episode, the player turned on his cell phone during the game which causes his death. In a sense, since the rules of using the algorithm is clearly stated by the manufacturer, it is not 100% the algorithms' fault.

Because of the new technology such as Artificial Intelligence, Machine Learning or Data Mining, the responsibility network of scientist and engineers have changed. Thus, scientists and engineers need to know ethics and to be able to defend themselves. As stated in the ethic's lecture, the professionals (scientists and engineers or algorithm developers) and public has a social contract. The society gives professions trust for to practice their expertise in which they are trained. Overall, professionals self-organized groups to share a common goal of using their expertise to contribute to the society. In a sense, professionals have a special training, and the society has expectation that professionals are ethical.

To decrease the amount of harm to the public, professionals must follow different kinds of responsibilities. For example, the Prohibitive Ethics, serves as a guideline for members of society to make decisions. This can be defined like the IEEE article showing certain actions one cannot do as a professional such as do not accept bribes, manipulate data, or do judgements on field outside ones' expertise. The Preventative Ethics explains how one can prevent the technology to stop the possibility of harming the public. The Aspirational Ethics shows how one can work to improve the life of the public which is a positive responsibility. Professionals must always take into consideration of the people that will be involve with the algorithms that they use. The problem with the manuals in technologies is that it is too long to read which is part of the contract limitation which is the main problem of these two ethical issues.

To prevent the self-driving car accidents and game trial deaths from rehappening, there are different actions the manufacturers can choose. As professionals, there are various ways to make decisions. As stated in Nature and Human Values, there are three main ethical theories which connects to the human ethical decision making. The theories are Deontology, Utilitarianism / Consequentialism, and Ritual Ethics. Deontology looks at the intent of the act itself and nothing else as consequences can distract oneself from looking at morally right things. Consequentialism looks at the different types of actions and look at the overall net value of everyone involve in the problem. Consequentialism are often not predictable, especially the long term and requires the fact that one treat others the way one would like to be treated. Contradictive to what most believes, it is not easy to make decisions.

Additional tests could also be used to determine if a specific course of action is ethical or not. For example, Michael Davis's eight moral tests include, harm test, publicity test, defensibility test, reversibility test, virtue test, professional test, colleague test, and organizational test. For example, from the reversibility test, one would imagine themselves as the user see how the decision impacts the user. Additionally, the consequence test, allows one to choose which actions to take in result of the consequence based on the law.

In the end, knowledge goes hand in hand with action. If the manufactures of the self-driving cars and game trails understand their values and benefits are the same as other peoples, they would feel more responsible and put more emphasis on the safety aspect and clarity of the rules in their product. There is a gap between moral knowledge and moral actions - even though people understand what is

morally correct, this does not mean they will go with the moral actions. There are many situations that could prevent someone from making good decisions even though one knows what the right course of actions. Moral Neutralization theory explains how someone can be vulnerable in making decisions. Denial responsibility shows that people do immoral things because it is out of their control. For instance, if person bikes, everyone else will drive, which will make no impact to the environment. These ethical issues made it clear that the responsibility must be taken seriously for all the stakeholders involved as small changes and course of actions can make a big impact.