Isabelle Yam

Professor Yvonne Su

Saturday, June 24th, 2023

HREQ4562

Layers of Violence/Photovoice Essay - How AI Can Manipulate People By Being Racist

AI is becoming a tool that has come into society in everyday life. It is becoming a trend that is used everywhere installed with the technological devices utilized in social social media to self-driving cars. The argument and message is that however, there is racism in AI which is a form of violence from manipulation that results from human error which comes from computer scientists, engineers, and other occupations that work in the field of AI. Racism can be seen in human life, when people are not treated fairly such as people treating you differently based on your skin colour, or making fun of others based on their skin colour. This type of racism can be seen within the examples of AI being racist. The fact that technology is now also incorporating the racism seen in society. How technology can be made to be more inclusive to all cultures, as someone growing up in Canada where it is multicultural might not want to be manipulated into sticking into one culture. The theme and topic that I am trying to explore is how humans program AI technology, and AI can be seen to be depicting violence in the form of being discrimination based on where you are from. This can be seen in everyday life from the technology that is being used anywhere from AI. Based on culture and skin colour, AI is excluding certain people or making people feel bad based on their appearance. This does not treat everyone fairly, and makes others feel violated with their technological tools. In this essay, I will discuss four different ways that show how AI can be racist which is seen from image algorithms, facial recognition software, AI chatbots, and analytical technological tools, to bring awareness to the topic of racism in AI.

Image algorithms are showing how AI is racist. The photo that I showed is typing They are programmed to a certain viewpoint which shows the racism in society. For instance, one photo on the google search that I took is on the question of: "What colour should my wedding dress be?" Ideally, it would show all kinds of wedding dresses to show the different types of wedding dresses and which types could look good. However, the results are mainly traditional US white dresses. "This stereotypes the terms of how the terms of "women", "dress", "wedding" and "bride" are traditionally models wearing the traditional US white wedding dress" (Schiebinger, Zou, 2018, pg. 1), when in reality there are many different types of wedding dresses from all different types of culture. It does not depict the range, and is racism by setting it to one type of wedding dress of how it looks like. "Whereas a traditional Indian dress would be considered with terms of "performance art" and "costume"." (Schiebinger, Zou, 2018, pg. 1). Even though that is a wedding dress too, it is being stereotyped as something else. The algorithm is showing the results leaning towards one type of result, when all cultures should be represented. When asking the color of a wedding dress, one would not want to wear the typical US white one, that is not the typical wedding dress as there are a variety of wedding dresses from all different cultures of what someone would wear depending on their culture and where they are from. By calling the North Indian dress as not a wedding dress is racist, and by saying that all wedding dresses are the white one is also racist. It is not an accurate representation of society. Excluding a certain type of race is considered racist because they are making others feel a certain way and making the norm for a certain type of photo of how people should behave and act. When really, they are able to be themselves, and be a part of their own culture.

Facial recognition exclusion is another way that AI is being racist. It is showing that technology recognition is not being able to recognize dark-skinned faces, and it is harder for darker- skinned people to be recognized with technology. In this photo, I depicted myself in the facial recognition. As an Asian person, it is notably harder to recognize then a person of white colour using facial recognition on my phone. It is displaying the feature of how to use

facial recognition and people with lighter shades of skin colour are depicted as the technology working. "These technologies are failing to see blackness in many ways" (Boccio, 2022, pg. 1)..., some examples include a computer not being able to follow the black-skinned person, but being able to follow a white-skinned person. Another is a soap dispenser, that is able to dispense soap under the white-skinned hand, but not the dark-skinned hand. "By not being able to see these people, they make the bodies and experiences of those who are dark-skinned as invisible, because they feel like they cannot use the technology, and have the same experiences as everyone else, or white-skinned people." (Boccio, 2022, pg. 1). This is racist because they are not being included in a product that is supposed to be designed for anybody to use. Everyone uses the same product, but does not have similar results according to the technology. It makes people feel left out, based on their skin tone, like they are not considered worthy of using the technology like everyone else. This is exploitative of people's skin colour as a form of violence in AI development. "It has a default to whiteness underneath all this technology." (Boccio, 2022, pg. 1. The way that technology is coded by people, showing that they are not being able to be receptive to all skin tones and colours, which shows the racism in the programmer and technology.

AI can also be racist seen through AI chatbots. The photo I chose shows the racism in the AI chatbot asking what a common Chinese phrase means, but then not answering it correctly, showing how the chatbot is not acceptable to all cultures. They do not understand that all languages are typically set to the white default.

> "This can be seen when asking a chatbot what the meaning of the phrase, "Have you eaten yet?" means. This phrase is a Chinese phrase that means "How's it going?" in the western equivalent of it. However, the chatbot usually fails to recognize and answer, making the bot default set to whiteness. If it answers with "No, I have not eaten yet" this displays that is racist because the phrase should be responded with

"Yes, I have eaten", which means "I'm fine", in the Western equivalent. When asking the snapchat AI bot, the phrase, it answers that the phrase only means if you have eaten something or not, and this is wrong because it fails to detect other meanings of it." (Cheuk, 2021, pg. 5).

This means that the bot is default to Western style, and many bots can be seen to be defaulted this way, failing to recognize other types of phrases in other cultures. This is racist because the AI bot has been programmed to Western culture, and it assumes that you are a Western person right away. It does not detect languages of all cultures, even though it can be inputted to different cultures to not be racist. This is racist because not all cultures are being represented, and the Western is being shown as the norm, or the right way of communicating in society. The bot interacts solely in Western communication, oftentimes, ruling out other types of cultures in terms of communication. This is unfair and biased. This relates to how chatbots can show violence with the harsh terms that they show such as mentioned, when people are affected by this, if they come across this would also feel violated and biased against. Chatbots can be violent and cause violence for humans, showing preference for certain races.

AI is also seen to be racist in its analyzing software. The photo I chose is of a courtroom to show how the technology used for analyzing crime for black defendants is unfair for white defendants.

"A technology in this type of protection software is prejudiced against black reportedly. The formula is likely to flag that black defendants in the future will commit crime, wrongly labeling them twice as much as white defendants. Also, white defendants were mislabeled as low risk often more times than black defendants. it states, that in the, "Fairness in Criminal Justice Risk Assessments: The State of the Art, in their analysis noted for black people it is a higher false positive rate and it

seems like black offenders are more likely to be classified as violent mistakenly than white people. Both of these error rates are showing the representation of blacks being predicted as violent." (Cain, Watkins, 2022, pg. 1)

These differences can support the claims of racial injustice as black people are being stereotyped as committing crime. When in reality, they have not done anything, but because of their skin color they are being prejudiced against and accused of committing crime based on how they look. White people are able to get away with the violent acts they committed as the technology does not scan them as guilty but rather innocent, due to looking more white. This shows that the technology is racist because black people are automatically being affected by the algorithm to be considered criminals based on the error of the technology. In the law court, two sides should always be treated and then it is determined whether or not they are guilty with their points, rebuttals, counter arguments, but automatically setting someone to be higher chance of being detected as a criminal is a violation of their human rights.

Overall, racism in humans shows the racism in AI because it is humans who program and create AI afterall. The argument that I am trying to make with all these photos is that AI has seen to be racist from the human error of people who program the AI software, and racism from humans can be seen within the technology. As a result of this, shows how AI can be racist. Not all AI is bad, as some can be seen to be used for good such as in healthcare, and for helping with studying. It will definitely continue to be used in the future as it is growing within our society widely. The argument that can be made with all these photos, is that In all the examples, it shows how AI can be racist because of the way the technology is programmed to function, not showing equal treatment to all regardless of their background or ethnicity. For instance, in the photos they all show an example of how that specific technology in AI was being racist to the user. In these examples AI is showing how they are treating others poorly, and not fairly and injustice to equality of all the races. In one photo it is

the depiction of the racist chatbot phrases, the other is the racist google image algorithms, the other is the racist racial recognition software, and lastly the racist analyzing software. These are all types of AI that are being developed, and show some racism that are indirectly or directly affecting others by their race.

I learned from looking at the different photos, of ways AI could be racist that I never knew of. From the examples I was shocked to hear that the ways that AI are programmed can exclude certain groups or make other groups seem more superior. The ways that races are not being fairly treated is definitely showing violence in the AI community. I learned that it may not always be implicitly shown the bias, but it can still be there. I felt that the community is definitely showing how AI can be racist and how it needs to change to inherently avoid the racism in humans showing up in AI. I learned about myself that the technology I use is not all acceptable to the ways that they are not treating everyone fairly in compared to how human rights I learned in this course and how violence should not be tolerated even though it can show up in something that is so widely used such as AI. In the future, to bring more awareness on the topic so that humans can see change in their society, towards being a diverse multiracial world, where people can express their own cultures. AI should be programmed to be less racist, and inclusive to all to avoid the manipulation of racism that is affecting others.

References

Boccio, R. (2022). [Review of the book *Race After Technology: Abolitionist Tools for the New Jim Code*, by Ruha Benjamin]. *Configurations 30*(2), 236-238. doi:10.1353/con.2022.0013.

Cheuk. (2021). Can AI be racist? Color‑evasiveness in the application of machine learning to science assessments. Science Education (Salem, Mass.), 105(5), 825–836. https://doi.org/10.1002/sce.21671

Eastwood. (2017). Why racist AI bots are our fault. CIO.

Hong, J.-W., & Williams, D. (2019). Racism, responsibility and autonomy in HCI: Testing perceptions of an AI agent. *Computers in Human Behavior, 100*(Complete), 79–84. https://doi.org/10.1016/j.chb.2019.06.012

Howard, A., & Borenstein, J. (2018). The Ugly Truth About Ourselves and Our Robot Creations: The Problem of Bias and Social Inequity. *Science and Engineering Ethics*, *24*(5), 1521–1536. https://doi.org/10.1007/s11948-017-9975-2

Levendowski. (2018). HOW COPYRIGHT LAW CAN FIX ARTIFICIAL INTELLIGENCE'S IMPLICIT BIAS PROBLEM. Washington Law Review, 93(2), 579–630.

Lu, & Qiu, J. L. (2022). Empowerment or warfare? dark skin, AI camera, and Transsion's patent narratives. Information, Communication & Society, 25(6), 768–784. https://doi.org/10.1080/1369118X.2022.2056500

Tian, J., Xie, H., Hu, S., & Liu, J. (2021). Multidimensional Face Representation in a Deep Convolutional Neural Network Reveals the Mechanism Underlying AI Racism. *Frontiers in Computational Neuroscience*, *15*. https://doi.org/10.3389/fncom.2021.620281

Watkins, T., & Cain, J. O. (2022). Systemic racism in data practices. *IASSIST Quarterly*, *46*(4).

Zou, & Schiebinger, L. (2018). AI can be sexist and racist - it's time to make it fair. Nature (London), 559(7714), 324–326. https://doi.org/10.1038/d41586-018-05707-8

Instagram Account For Layers of Violence/Photovoice Presentation:

Username: ai_racism
Password: schoolproject123