BRAC UNIVERSITY
CSE422 : Artificial Intelligence
Assignment 5

Consider that your assigned **case number = (ID % 7) + 1**. Answer the following with respect to your assigned case.

1. Evaluate the sustainability impacts of deploying the AI system. Consider both positive contributions and negative impacts.

2. Analyze the ethical challenges of the AI solution. In the response, you may discuss issues such as bias in predictions, transparency regarding the model, and accountability if the system's errors cause harm to communities.

3. Examine the societal and legal implications of deploying such an AI system. In addition, responsibilities of the engineers and organizations to ensure safety, fairness, and compliance with regulatory frameworks across different jurisdictions.

Cases

Case 1: In 2014, Amazon started to develop and use AI programs to mechanise highly time-intensive human resources (HR) work, namely the shortlisting of applicants for jobs. Amazon "literally wanted it to be an engine where I'm going to give you 100 resumes, it will spit out the top five, and we'll hire those" (Reuters 2018). The AI tool was trained on CVs submitted over an earlier ten-year period and the related staff appointments. Following this training, the AI tool discarded the applications of female applicants, even where no direct references to applicants' gender were provided. Given the predominance of successful male applicants in the training sample, Amazon found that the system penalised language such as "women's chess club captain" for not matching closely enough the successful male job applicants of the past. While developers tried to modify the system to avoid gender bias, Amazon abandoned its use in the recruitment process in 2015 as a company "committed to workplace diversity and equality" (ibid).

Case 2: China is one of the world's leading nations in AI development. It embraces the use of large amounts of data that it collects on its citizens, for instance in its social credit scoring system. This system uses a large number of data points, including social media data, local government data and citizens' activities, to calculate a trustworthiness score for every citizen. Several data platforms are used to integrate data into "a state surveillance infrastructure" (Liang et al. 2018). High scores lead to the allocation of benefits, such as lower utility rates and favourable booking conditions, whereas low scores can lead to the withdrawal of services (Raso et al. 2018). Within China, the system benefits from high levels of approval because Chinese citizens "interpret it through frames of benefit-generation and promoting honest dealings in society and the economy instead of privacy-violation" (Kostka 2019: 1565).

Case 3: Clearview AI is a software company headquartered in New York which specialises in facial recognition software, including for law enforcement. The company, which holds ten billion facial images, aims to obtain a further 90 billion, which would amount to 14 photos of each person on the planet (Harwell 2022). In May 2021, legal complaints were filed against Clearview AI in France, Austria, Greece, Italy and the United Kingdom. It was argued that photos were harvested from services such as Instagram, LinkedIn and YouTube in contravention of what users of these services were likely to expect or have agreed to (Campbell 2021). On 16 December 2021, the French Commission Nationale de l'Informatique et des Libertés announced that it had "ordered the company to cease this illegal processing and to delete the data within two months" (CNIL 2021).

Case 4: The 2008 US presidential election has been described as the first that "relied on large-scale analysis of social media data, which was used to improve fundraising efforts and to coordinate volunteers" (Polonski 2017). The increasing availability of large data sets and AI-enabled algorithms led to the recognition of new possibilities of technology use in elections. In the early 2010s, Cambridge Analytica, a voter-profiling company, wanted to become active in the 2014 US midterm election (Rosenberg et al. 2018). The company attracted a $15 million investment from Robert Mercer, a Republican donor, and engaged Stephen Bannon, who later played a key role in President Trump's 2016 campaign and was an important early member of the Trump cabinet. Cambridge Analytica lacked the data required for voter profiling, so it solved this problem with Facebook data (Cadwalladr and Graham-Harrison 2018). Using a permission to harvest data for academic research purposes that Facebook had granted to Aleksandr Kogan, a researcher with links to Cambridge University, the company harvested not just the data of people who had been paid to take a personality quiz, but also that of their friends. This allowed Cambridge Analytica to harvest in total 50 million Facebook profiles, which allowed the delivery of personalised messages to the profile holders and also – importantly – a wider analysis of voter behaviour.

Case 5: In May 2016, a Tesla car was the first known self-driving car to be involved in a fatal crash. The 42-year-old passenger/driver died instantly after colliding with a tractor-trailer. The tractor driver was not injured. "According to Tesla's account of the crash, the car's sensor system, against a bright spring sky, failed to distinguish a large white 18-wheel truck and trailer crossing the highway" (Levin and Woolf 2016). An examination by the Florida Highway Patrol concluded that the Tesla driver had not been attentive and had failed to take evasive action. At the same time, the tractor driver had failed, during a left turn, to give right of way, according to the report (Golson 2017).

Case 6: "Automation can be an asset to a company, but there needs to be a way for humans to take over if the machine makes a mistake," says Ibrahim Diallo (Wakefield 2018). Diallo was jobless for three weeks in 2017 (Diallo 2018) after being dismissed by an automated system for no reason his line manager could ascertain. It started with an
inoperable access card, which no longer worked for his Los Angeles office, and led to him being escorted from the building "like a thief" (Wakefield 2018) by security staff following a barrage of system-generated messages. Diallo said the message that made him jobless was "soulless and written in red as it gave orders that dictated my fate. Disable this, disable that, revoke access here, revoke access there, escort out of premises ... The system was out for blood and I was its very first victim" (ibid). After three weeks, his line manager identified the problem (an employee who had left the company had omitted to approve an action) and reinstated Mr Diallo's contractual rights.

Case 7: Seasonal climate forecasting (SCF) is used to predict severe weather, such as droughts and floods, in order to provide policymakers and farmers with the means to address problems in an anticipatory rather than a reactive manner (Klemm and McPherson 2017). Lemos and Dilling (2007) have argued that the benefits of SCF mostly reach those "that are already more resilient, or more resource-rich ... in terms of ... ability to cope with hazards and disasters". By contrast, those who are most at risk of being pushed below the poverty line by severe weather have been harmed in cases in Zimbabwe, Brazil and Peru. In Zimbabwe and Brazil, poor farmers were denied credit after SCF results predicted a drought (ibid). In Zimbabwe, "misinterpretation of the probabilistic nature of the forecast by the banking sector" might have played a role in decision-making about credits (Hammer et al. 2001). SCF forecasting in Peru also led to accelerated layoffs of workers in the fishing industry due to "a forecast of El Niño and the prospect of a weak season." (Lemos and Dilling 2007)