Week 4 Quiz

TOTAL POINTS 10

1. What are the current limitations of AI technology? (Select all that apply) | 1 point |

   ☐ There are no limitations to AI technology

   ☑ AI technology can discriminate

   ☑ AI technology can be biased

   ☑ Explainability is hard

   ☑ AI technology is susceptible to adversarial attacks

2. What is the Goldilocks Rule of AI? | 1 point |

   ⊙ One shouldn't be too optimistic or too pessimistic about AI technology

   ○ One should allocate many resources to defend the world from giant killer robots

   ○ AI's technology will continue to grow and can only benefit society

   ○ An AI winter is coming

3. Say you are building an AI system to help make diagnoses from X-ray scans. Which of the following statements about explainability of AI do you agree with?

○ Explainability is usually achieved through building a chatbot to talk to the user to explain its outputs.

○ AI systems are intrinsically "black box" and cannot give any explanation for their outputs.

○ Most AI systems are highly explainable, meaning that it's easy for a doctor to figure out why an AI system gave a particular diagnosis.

⦿ Lack of explainability can hamper users' willingness to trust and adopt an AI system.

4. Using current AI technology, if a machine learning system learns from text that reflects unhealthy biases/stereotypes, then the resulting AI software may also exhibit similarly unhealthy biases/stereotypes.

⦿ True

○ False

5. Using current AI technology, if a machine learning system learns only from text that is completely neutral and does not reflect any gender biases, then we would expect it to exhibit no, or at most minimal, gender bias.

⦿ True

○ False

6. Which of these are good practices for addressing bias in AI? (Select all that apply)   1 point

   ☐ Using an adversarial attack on the AI system to change its outputs to be less biased

   ☑ Using more inclusive/less biased data

   ☑ Technical solution such as "zeroing out" bias

   ☑ Systematic auditing processes to check for bias

7. Which of these are examples of adversarial attacks on an AI system? (Select all that apply)   1 point

   ☑ Subtly changing an image to make an AI system mistakenly recognize a dog as a cat.

   ☐ Using AI to synthesize a fake video of a politician saying something they never actually said.

   ☑ Subtly modifying an audio clip to make a speech recognition system think someone said "Yes, authorized" when they actually said "No, reject."

   ☑ Adding a sticker to a stop sign to make an AI system fail to detect it.

8. If a developing economy has a strong and thriving coffee bean manufacturing industry (or some other vertical industry), then it has an advantage in applying AI to coffee bean manufacturing (or other vertical industry).   1 point

○ True

○ False

9. What are the jobs that AI is most likely to displace over the next several years? 1 point

○ All jobs will be displaced

● Jobs that comprise primarily of routine, repetitive work

○ Most jobs involving office work (white collar jobs)

○ Jobs that comprise primarily of non-routine, non-repetitive work

10. Congratulations! You deserve a pat on the back for finishing this course. 1 point

● True

○ False