

Week 4

✓ **Congratulations! You passed!**

Grade
received 90%

Latest Submission
Grade 90%

To pass 80% or
higher

[Go to next item](#)

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

1 / 1 point

☒ AI technology can discriminate

✓ Correct

☒ AI technology is susceptible to adversarial attacks

✓ Correct

☒ AI technology can be biased

✓ Correct

☒ Explainability is hard

✓ Correct

☐ There are no limitations to AI technology

2. What is the Goldilocks Rule of AI?

1 / 1 point

☐ An AI winter is coming

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

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

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

✓ Correct

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?

1 / 1 point

☐ 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.

☐ AI systems are intrinsically "black box" and cannot offer any explanation for their outputs.

-
- ☒ Lack of explainability can hamper users' willingness to trust and adopt an AI system.
 - ☐ Explainability is usually achieved through building a chatbot to talk to the user to explain its outputs.

✓ Correct

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.

0 / 1 point

- ☐ True
- ☐ False

✗ Incorrect
You didn't select an answer.

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.

1 / 1 point

- ☒ True
- ☐ False

✓ Correct

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

1 / 1 point

☒ Using more inclusive/less biased data

✓ Correct

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

☒ Systematic auditing processes to check for bias

✓ Correct

☒ Technical solution such as "zeroing out" bias

✓ Correct

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

1 / 1 point

☐ A user providing false information to an AI system to cause it to make a wrong decision

- ☐ 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."

✓ Correct

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

✓ Correct

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

✓ Correct

- 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 / 1 point

- ☒ True
- ☐ False

✓ Correct

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

1 / 1 point

- ☒ Jobs that comprise primarily of routine, repetitive work
- ☐ All jobs will be displaced
- ☐ Most jobs involving office work (white collar jobs)
- ☐ Jobs that comprise primarily of non-routine, non-repetitive work

✓ Correct

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

1 / 1 point

- ☒ True
- ☐ False

✓ Correct