

## Understanding Machine Learning Projects

LATEST SUBMISSION GRADE  
100%

1. Which of the following statements about the Machine Learning Process Lifecycle are true? (click all that apply)

1 / 1 point

- ☐ Data wrangling is only a small part of an ML project
- ☐ Looking at whether you can answer your questions based on the data you have is part of the final phase.
- ☒ Testing the QuAM is an important part of the Machine Learning and Model Evaluation phase

✓ **Correct**  
Correct! Well done!

- ☐ Lifecycle re-sets are easy to anticipate
- ☐ Steps can easily be skipped if they're taking too long
- ☒ A clearly defined process helps with communication

✓ **Correct**  
Correct! Well done!

- ☐ Documentation isn't important if the QuAM is being built and used internally
- ☒ The first phase considers exactly what features are going to be used

✓ **Correct**  
Correct! Well done!

- ☒ Having a clear, specific question for the QuAM to answer is an important part of the first phase

✓ **Correct**  
Correct! Well done!

- ☒ Iteration is almost always a given; an ML project is rarely straight-forward and linear

✓ **Correct**  
Correct! Well done!

2. What does it mean to work backwards through the ML Process in order to define a question?

1 / 1 point

- ☐ Start with the required action, identify impact on the business goals, define and refine a question
- ☐ Start with the answer you want your QuAM to give, decide on the algorithm, then define operational data.
- ☒ Start from the broadest business outcome, identify actions that accomplish that outcome, identify questions that support those actions, refine the question.
- ☐ Start with the learning data you have, define the question you want your QuAM to answer, choose an algorithm.

✓ **Correct**  
Correct! Well done!

3. Why might a business choose to use a QuAM that has lower accuracy than the current system?

1 / 1 point

- ☐ Because the cost of developing the QuAM was high, and the business doesn't want to waste that time and money
- ☒ Because the QuAM can do the task much faster than the humans can, and a faster process is highly advantageous for the business

✓ **Correct**  
Correct! The time savings from using ML can be more important than perfect accuracy.

- ☐ Because your business needs to have machine learning to survive in the modern age
- ☒ Because the employees hate doing the task, and handing it off to a machine frees them to do other useful things

✓ **Correct**  
Correct! Best use of human time is an important consideration when analyzing use of ML.

- ☒ Because the QuAM does the job adequately and provides other benefits

✓ **Correct**  
Correct! Well done!

4. Why is it important to have a hold-out test set?

1 / 1 point

- ☐ A) It is important to test your QuAM's ability to generalize before deploying it
- ☐ B) It's important to protect the integrity of the test data so that you have confidence in the performance of your QuAM
- ☐ C) Your QuAM is going to have to answer questions about unseen data, so you need to test it on unseen data
- ☐ A and B
- ☒ A, B and C

✓ **Correct**  
Correct! Great job!

5. Which of the following statements are true? (click all that apply)

1 / 1 point

- ☐ A true positive means someone is pregnant
- ☐ False negatives are always a bigger problem than false positives
- ☒ Telling a patient that they have cancer when they don't is a false positive

✓ **Correct**  
Correct! Well done!

- ☐ Telling a patient that they have cancer when they don't is a false negative
- ☒ True positives and true negatives are correct predictions

✓ **Correct**  
Correct! Well done!