ARTIFICIAL INTELLIGENCE ALGORITHMS

ARTIFICIAL INTELLIGENCE DESIGN AND IMPLEMENTATION

TEST 1

YOUR NAME: <Type here>

COLLEGE_ID: <Type here>

DURHAM COLLEGE
SUCCESS MATTERS

ARTIFICIAL INTELLIGENCE ALGORITHMS

ARTIFICIAL INTELLIGENCE DESIGN AND IMPLEMENTATION

Read the question and USE **X** , in all boxes that are correct in your perspective.

X

1 What is Artificial Intelligence?

X ☐ **Getting computers to behave better without explicitly programming every little thing**

X ☐ **A sub-field of computer science**

X ☐ **Expert systems or search, for example**

☐ **Computers with human-level or superhuman intelligence**

☐ **Anything intelligent that isn't human**

☐ **An algorithm**

2. **What is Machine Learning?**

- ◯ **Machines that learn like humans do**
- ◯ **Computers learning to do tasks perfectly**
- ◯ Magic
- X ◯ **Computers learning from existing examples to guess well on new examples**

3. **The difference between Data Science and Machine Learning is:**

- Data science is always about providing insights into data patterns; machine learning is never about providing insights
- Data science never builds models; machine learning always builds models
- X Data science is about extracting insights; machine learning is about generalizing from data

4. **What type of machine learning is the spam filter on your email platform that automatically sends suspicious emails to your spam folder?**

- Reinforcement Learning
- Unsupervised Learning
- X Supervised Learning

5. **What type of machine learning is the Optical Character Recognition system that scans envelopes to determine their destination?**

X ○ Supervised Learning

○ Reinforcement Learning

○ Unsupervised Learning

6. **What type of machine learning is a system that classifies an image as ‘cat’ or ‘not cat’?**

- Unsupervised Learning
- Reinforcement Learning
- X Supervised Learning

7. **What type of machine learning is a system that predicts the retail value of a new house on the market?**

X
- ○ Supervised Learning
- ○ Reinforcement Learning
- ○ Unsupervised Learning

8. **When you're shopping online, a system recommends similar items to what is in your cart. What type of machine learning is this?**

- Reinforcement Learning
- X Unsupervised Learning
- Supervised Learning

9 **What type of machine learning is a system that clusters data in order to detect anomalies that could indicate instances of fraud?**

X
- ○ Unsupervised Learning
- ○ Reinforcement Learning
- ○ Supervised Learning

10 **What type of machine learning is a system that learns to play Atari games to get a high score?**

X ○ Reinforcement Learning

○ Unsupervised Learning

○ Supervised Learning

11 **What type of machine learning is a system that explores a maze until it finds an optimal solution to get to the finish line?**

X
- ○ Reinforcement Learning
- ○ Unsupervised Learning
- ○ Supervised Learning
- ○ Planning
- ○ Search

QuAM = Machine Learning Model.

12 Machine learning systems do things like mistake a woman's hair for a cat, or think that the name of a band from Montreal could be part of a familiar saying because:

- [ ] The question the QuAM was answering was too specific
- [ ] QuAMs are able to generalize
- X [ ] They might have had limited training data
- X [ ] Because of the learning algorithm being used

QuAM = Machine Learning Model.

13 Why are features important?

- They describe important characteristics of the data.
- They are used by a QuAM to distinguish between categories
- They can filter out irrelevant data
- X All of the above

QuAM = Machine Learning Model.

14 What are some of the considerations you need to keep in mind about your QuAM for when you're using it on operational data? (click all that apply)

X ☐ For some applications like legal rulings, explainability is important

X ☐ For some applications, the operations of the QuAM can cause the operating environment to change, making the QuAM inaccurate or ineffective.

☐ For some applications like medical diagnosis, the bedside manner of the QuAM is more important than accuracy.

☐ For all applications, the effects of the QuAM are far less relevant than what the system costs

X ☐ For some applications there is a minimum standard of accuracy needed, and less than that standard makes the QuAM useless

X ☐ For some applications, the QuAM's performance must be verified, which may mean setting up a parallel system

QuAM = Machine Learning Model.

15 What are some of the important questions when it comes to QuAM performance and business evaluation? (check all that apply)

X ☐ Are you measuring what you think you're measuring?

X ☐ What is acceptable performance?

X ☐ What does success look like?

X ☐ What is ideal performance?

X ☐ How are you performing by your current metrics?

☐ Are you using the latest technology?

X ☐ What is the cost of making a mistake?

☐ How do we get perfect performance?

16 What should you prefer a simple proxy?

- It is easier to measure.
- Improvements in a simple metric are likely to be real improvements to start with.
- A simple proxy is easier to validate as measuring what you think it's measuring.
- X All of the above

17 Which of the following are sources of data that can be used for machine learning? (click all that apply)

X ☐ Readings from sensors such as temperature, pressure, pH monitors, etc.

X ☐ Government data such as census results.

☐ Personal data collected without permission

X ☐ Data collected by a business about their own operations

X ☐ Data collected by a business about their customers

X ☐ Text data from the Internet, such as Amazon reviews or Wikipedia

☐ Data handwritten in a notebook

X ☐ Government archives

X ☐ Data purchased from third party data "brokers"

18 Which of the following are issues of ethics and responsibility in machine learning? (click all that apply)

X ☐ The fair treatment of the people collecting and processing the data

X ☐ The representativeness of the data

X ☐ The anonymization of the data, as much as is possible

X ☐ The proper consent of the original owners of the data

X ☐ The security of the data, so that it isn't easily lost or stolen

19 How can data be biased? (click all that apply)

- [ ] It can't; data is data and it reflects the real world
- [X] It might include data collected under different conditions, and so not reflect operational data
- [X] It might not include enough training data on a range of gender and ethnic groups, and so not reflect operational data
- [X] It might not include data from underrepresented socioeconomic groups, and so not reflect real-world data

QuAM = Machine Learning Model.

20 What is the batch effect?

X
- When data from different sources have variations that aren't meaningful, but the algorithm takes as meaningful
- When data from different times have included measurements of different things
- When you train your QuAM several times in different batches
- When hospitals don't have the same scan results

QuAM = Machine Learning Model.

21 Which of the following statements are true about data and data pipelines?

X ☐ Transformed data will need to be accessible to your QuAM

☐ Automating data retrieval is a straight-forward process

☐ Long term data storage is never a concern

X ☐ Machine learning is an ongoing process, so new, incoming data is important

X ☐ Integrating data from multiple sources can cause formatting issues

X ☐ Learning data and operational data need to be in the same format

X ☐ Features that were used in the learning data must be present in operational data

QuAM = Machine Learning Model.

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

- [ ] The first phase considers exactly what features are going to be used
- [ ] Documentation isn't important if the QuAM is being built and used internally
- [X] A clearly defined process helps with communication
- [ ] Lifecycle re-sets are easy to anticipate
- [ ] Data wrangling is only a small part of an ML project
- [X] Having a clear, specific question for the QuAM to answer is an important part of the first phase
- [ ] Steps can easily be skipped if they're taking too long
- [X] Iteration is almost always a given; an ML project is rarely straight-forward and linear
- [X] Testing the QuAM is an important part of the Machine Learning and Model Evaluation phase
- [ ] Looking at whether you can answer your questions based on the data you have is part of the final phase.

QuAM = Machine Learning Model.

23 What does it mean to work backwards through the ML Process in oder to define a question?

- ○ Start with the learning data you have, define the question you want your QuAM to answer, choose an algorithm.
- ○ 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.
- X ○ Start from the broadest business outcome, identify actions that accomplish that outcome, identify questions that support those actions, refine the question.

QuAM = Machine Learning Model.

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

X ☐ Because the QuAM does the job adequately and provides other benefits

X ☐ Because the QuAM can do the task much faster than the humans can, and a faster process is highly advantageous for the business

☐ Because the cost of developing the QuAM was high, and the business doesn't want to waste that time and money

☐ Because your business needs to have machine learning to survive in the modern age

X ☐ Because the employees hate doing the task, and handing it off to a machine frees them to do other useful things

QuAM = Machine Learning Model.

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

- 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
- X A, B and C

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

- [ ] A true positive means someone is pregnant
- [ ] False negatives are always a bigger problem than false positives
- X [ ] Telling a patient that they have cancer when they don't is a false positive
- [ ] Telling a patient that they have cancer when they don't is a false negative
- X [ ] True positives and true negatives are correct predictions