AiU™ Certified Machine Learning Engineer (CMLE)

Sample exam — 15 Questions

Released Version

2022 Syllabus

Artificial Intelligence United

Purpose of this document

This document contains 15 sample exam questions for AiU Certified Machine Learning Engineer (CMLE) in the English language.

The sample questions, answer sets and associated justifications in this document have been created by a team of subject matter experts and experienced question writers with the aim of assisting people who are planning to take the AiU Certified Machine Learning Engineer (CMLE) examination.

None of these questions are used in the official AiU Certified Machine Learning Engineer (CMLE) examination, but they are written to the same level of difficulty as the official certification exam and considered to be a sampling, as there are 50 questions in the real exam.

Instructions

The question-and-answer sets are organized in the following way:

- Chapters
- Question - including any scenario followed by the question stem
- Answer Set with explanations

General Information on the sample exam paper:

- Number of Questions: 15
- Number of points: 26 (1-3 per question)
- Please only choose one answer per question

List of Chapters

Question 1

(correct answer is worth 2 points)

You are working in a project where there is a large variation in the values of a continuous input variable.

Which **ONE** of the following options is the **MOST SUITABLE** transformation that you can apply?

- ☐ (a) One hot encoding
- ☐ (b) PCA
- ☐ (c) Binning
- ☐ (d) Normalization

Question 2

(correct answer is worth 2 points)

Which **ONE** of the following options would you be performing if you were to propose a movie for your friend to watch based on your likes?

- ☐ (a) Regression
- ☐ (b) Classification
- ☐ (c) Collaborative filtering
- ☐ (d) Clustering

Question 3

(correct answer is worth 1 point)

Jennifer is trying to guess the age of an oak tree by looking at various factors, i.e., its spread, number of roots, etc.

Which **ONE** of the following problem types is Jennifer performing in this above scenario?

- ☐ (a) Regression
- ☐ (b) Classification
- ☐ (c) Clustering
- ☐ (d) Association

Question 4

(correct answer is worth 1 point)

"Bullseye" is a sales support agency that aims to group supermarket customers who buy similar items together, so that they can be targeted better in advertisements in the future.

Which **ONE** of the following unsupervised learning methods is **BEST** suited for this task?

- ☐ (a) Clustering
- ☐ (b) Association Analysis
- ☐ (c) Dimensionality Reduction
- ☐ (d) Anomaly Detection

Question 5

(correct answer is worth 2 points)

In a regression problem the R-square was observed as being very close to the value of 1.

Which **ONE** of the following reasons would you expect to be the cause of this above situation?

- ☐ (a) There is a large range in the input variables for X.
- ☐ (b) There are (in general) a large number of input variables.
- ☐ (c) There is a high correlation between the output and an input variable for X.
- ☐ (d) The output has a very high range of variation.

Question 6

(correct answer is worth 1 point)

Which **ONE** of the following options is **NOT** a challenge when developing a linear regression?

- ☐ (a) Collinearity
- ☐ (b) Continuous output variable
- ☐ (c) A high number of dimensions
- ☐ (d) A lack of correlation between input and output variables

Question 7

(correct answer is worth 2 points)

When training a neural network, it has been noticed that the training and validation losses keep decreasing until a certain point where the validation loss starts increasing.

Which **ONE** of the following changes would be **LEAST HELPFUL** to solve this above-mentioned issue.

- ☐ (a) Removing some of the layers from the neural network.
- ☐ (b) Adding more neurons to specific layers of the neural network.
- ☐ (c) Changing the optimizer of the neural network.
- ☐ (d) Adding more data to the neural network.

Question 8

(correct answer is worth 3 points)

A deep learning classifier has been built to diagnose cancer. When using this classifier, a patient with a positive prediction is required to perform additional tests, while a patient with a negative prediction is directly free to go.

Which **ONE** of the following metrics should be focused on the **MOST** when evaluating such a classifier?

- ☐ (a) Recall
- ☐ (b) Precision
- ☐ (c) Accuracy
- ☐ (d) F1-Score

Question 9

(correct answer is worth 2 points)

For a given model, it is required to pad the sequences using the following line of code:

```
padded_docs = pad_sequences (encoded_docs, maxlen=5, padding='post')
```

Which of the following (I – IV) could be a representation for a random sentence?

I. [1 1 15 1 0 0]
II. [0 10 12 3 4]
III. [1 2 32 4 25]
IV. [1 2 3 19 0]

☐ (a) Options I and II could be a representation for a random sentence.

☐ (b) Options I and IV could be a representation for a random sentence.

☐ (c) Options II and III could be a representation for a random sentence.

☐ (d) Options III and IV could be a representation for a random sentence.

Question 10

(correct answer is worth 2 points)

Assume that you have a feature map of 6X6 elements.

While max pooling 2X2, which **ONE** of the following would you expect the resultant matrix to be?

☐ (a) 2X2

☐ (b) 2X3

☐ (c) 3X2

☐ (d) 3X3

Question 11

(correct answer is worth 1 point)

In which **ONE** of the following networks is it **NOT** possible to predict long range dependencies?

☐ (a) GRU

☐ (b) RNN

☐ (c) LSTM

☐ (d) Bidirectional LSTM

Question 12

(correct answer is worth 2 points)

In a production run for forecasting stock market data, the forecasts suddenly started deviating significantly from the actual data.

Which **ONE** of the following options could be a **VIABLE** reason for this sudden change?

- ☐ (a) A change in the distribution of data
- ☐ (b) Improper execution of a pipeline
- ☐ (c) Wrong configuration of an inference pipeline
- ☐ (d) An overloaded pipeline

Question 13

(correct answer is worth 2 points)

You deployed a model that differentiates between the different types of animals. Your model's predictions are accurate, but one time, your model received two visually identical images of a giraffe and it classified one of them as a giraffe and the other one as a horse.

Which **ONE** of the following options would you suspect is the reason for the misclassification?

- ☐ (a) Data Drift
- ☐ (b) Concept Drift
- ☐ (c) An adversarial attack
- ☐ (d) Potential difference in image size

Question 14

(correct answer is worth 2 points)

John has deployed a model on the edge, and he needs to decrease the associated inference time of the given model.

Which **ONE** of the following options would **NOT** help John to solve this issue?

- ☐ (a) Caching
- ☐ (b) Cloud Deployment
- ☐ (c) Model Quantization
- ☐ (d) Model Pruning

Question 15

(correct answer is worth 1 point)

Which **ONE** of the following options is used to store a deployable model after several trials and tuning?

- ☐ (a) Model Experiment
- ☐ (b) Model Execution
- ☐ (c) Model Hyperparameter
- ☐ (d) Model Registry

Answer Key:

Question 1

a) Incorrect – It would be more suitable for discrete values.
b) Incorrect – It would be better for reducing dimensions.
c) Incorrect – It would be more suitable for uniform behavior for ranges of data.
d) Correct – It would bring the range of variable to 0, 1 hence manageable.

Question 2

a) Incorrect – No prediction of output is involved.
b) Incorrect – No predicting of class of output is involved.
c) Correct – Using likes of friend to recommend based on user based collaborative filtering.
d) Incorrect – No grouping of items is involved.

Question 3

a) Correct – The prediction of output continuous variable is involved.
b) Incorrect – The output variable is not discrete.
c) Incorrect – No grouping of items is involved.
d) Incorrect – No notion of finding is occurring together.

Question 4

a) Correct – Clustering groups of similar entities together.
b) Incorrect – Association analysis generates rules about different products.
c) Incorrect – Dimensionality reduction would reduce the number of features only.
d) Incorrect – Anomaly detection would learn from normal behavior to detect abnormal behavior.

Question 5

a) Incorrect – This does not cause low error.
b) Incorrect – This does not cause low error.
c) Correct – A high correlation would imply a direct linear equation between the input variable and output variable hence very less error.
d) Incorrect – This does not cause low error.

Question 6

a) Incorrect – It is a challenge as it causes unnecessary additional variables to be part of the input.
b) Correct – It is expected, as regression works for continuous variables.
c) Incorrect – Due to this it causes the possibility for collinearity among some of the variables.
d) Incorrect – This may cause a high error model.

Question 7

a) Incorrect – This would decrease the complexity of the network and help overcome overfitting.
b) Incorrect – This would increase the complexity and does not prevent overfitting.
c) Correct – This would least affect overfitting.
d) Incorrect – Adding more data would help reduce overfitting.

Question 8

We want the model not to miss positive predictions (low FN), and there is no harm in having FP (because of the additional tests the patients will perform).

a) Correct – Recall evaluates how much of the real positives was the model able to correctly predict.
b) Incorrect – Precision evaluates how much of the predicted positive were actually positive.
c) Incorrect – Accuracy evaluates how much the model predictions were right regardless of what these predictions were (positive or negative).
d) Incorrect – F1-score is good when precision and recall are equally important.

Question 9

I. Incorrect – This sentence is of `length=6`.
II. Incorrect – Here the padding has been performed as 'pre' and not 'post'.
III. Correct – The sentence has a length of 5 and no "0" in the representation.
IV. Correct – The length is 5 and the "0" is appended at the end.

Therefore, D is the only possible correct answer.

Question 10

a) Incorrect – Each side gets divided by 2.
b) Incorrect – Each side gets divided by 2.
c) Incorrect – Each side gets divided by 2.
d) Correct – Each side gets divided by 2.

Question 11

a) Incorrect – These architectures are built basically to address this issue of long-range predictions.
b) Correct – Due to vanishing gradients issue RNN s are unable to do long-range predictions.
c) Incorrect – These architectures are built basically to address this issue of long-range predictions.
d) Incorrect – These architectures are built basically to address this issue of long-range predictions.

Question 12

a) Correct – Change in distribution of data will cause deviation of the new distribution from trained model.
b) Incorrect – This does not cause distribution drift.
c) Incorrect – This cannot be a cause for distribution drift.
d) Incorrect – This cannot lead to distribution drift.

Question 13

a) Incorrect – A data drift would cause a decreased performance on all images.
b) Incorrect – A concept drift would cause a decreased performance on all images.
c) Correct – The incorrectly classified image might have been carefully manipulated with the goal of "fooling" the classifier.
d) Incorrect – Images would be resized before they are fed to the model.

Question 14

a) Incorrect – This would save the outputs corresponding to frequent inputs and it would speed up the inference time.
b) Correct – This would add a delay caused by the network during inference.
c) Incorrect – This would make the models faster and result in a reduction of the inference time.
d) Incorrect – This would make the models smaller and result in a reduction of the inference time.

Question 15

a) Incorrect – An experiment is an individual run.
b) Incorrect – This is the act of running the model.
c) Incorrect – This is the tunable parameter for models.
d) Correct – This is where the final deployable version can be stored, after hyperparameter tuning.