Sample Machine Learning quiz.

A typical quiz includes 9 questions similar to these, involving (1) multiple choice questions, (2) open questions with text answer, (3) open questions with calculations required.

For (2), you are expected to be clear and concise (including a readable handwriting if answering on paper). For (3), you need to show all steps of the calculations.

Typical allotted time is 30 minutes.

Question **1**

Generalisation will be bad (much higher probability of error on test sets than on the training set) when:

Select one:

- ☐ Data in the training set are many but the learning machine is not powerful enough (underfitting)
- ☐ The hypothesis space is wrong
- ☐ Data in the training set are too few and the learning machine too powerful (overfitting)

Question **2**

Given a linear classifier with sigmoid output and square error loss, is the resulting objective function quadratic?

Select one:

- ☐ True
- ☐ False

Question **3**

What is the **major limitation** of earlier neural networks, that multi-layer neural networks can **overcome**?

Select one:

- ☐ Inability to learn perceptual problems even under scenario 1 (known population)
- ☐ Inability to learn classification problems even under scenario 2 (known probabilities)
- ☐ Inability to learn non-linearly-separable problems even under scenario 1 (known population)

Question **4**

Does PCA compute transformations of input values?

Select one:

- ☐ True
- ☐ False

Question **5**

Write down all the possible bootstrap samples for the following data set: {A, B, C}

Question **6**

A sigmoid activation in neural networks can be used:

Select one:

- ☐ Only in the first (input) layer
- ☐ In all layers, especially the output layer in regression problems
- ☐ In all layers, including the output layer in 2-class classification problems

Question **7**

The frequency of disease X is one over 1000 people (Pr(X) = 1/1000). The probability of having symptom Y is Pr(Y|X) = 1/100 when I have disease X and Pr(Y|not X) = 1/1000 when I don't have disease X. Observing only symptom Y and according to Bayes' theory, is it more probable that I have or that I don't have disease X? Write down the steps you followed to reach the conclusion.