

[Dashboard](#) / [My courses](#) / [COMS3007A-ML-2021](#) / [Quizzes](#) / [Quiz 1: Intro, Decision Trees, Naive Bayes](#)

Started on Saturday, 22 May 2021, 1:15 AM

State Finished

Completed on Saturday, 22 May 2021, 2:43 AM

Time taken 1 hour 27 mins

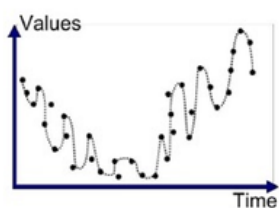
Grade **15.50** out of 16.00 (**97%**)

Question 1

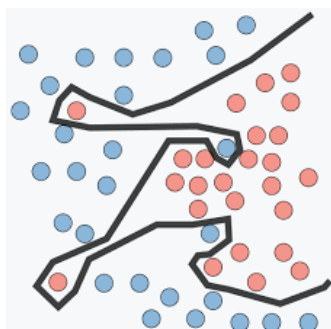
Correct

Mark 2.00 out of 2.00

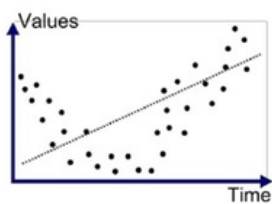
For each of the following cases, say whether the model has overfit, underfit, or neither.



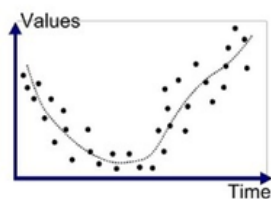
Overfit



Overfit



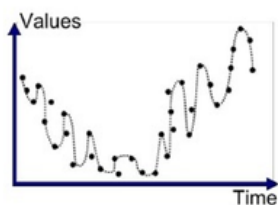
Underfit



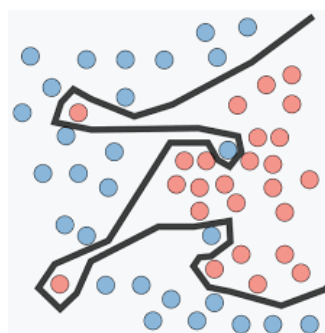
Neither



Your answer is correct.

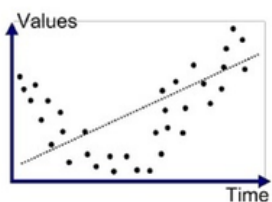


→ Overfit,

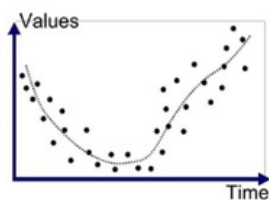


→ Overfit,

The correct answer is:



→ Underfit,



→ Neither

Question **2**

Partially correct

Mark 2.50 out of 3.00

For each of the following scenarios, state whether the problem is a supervised learning problem, unsupervised learning problem, or reinforcement learning problem.

An autonomous car learning to drive.

Reinforcement ✓

Segmenting drivers into groups based on their driving styles and behaviours.

Unsupervised ✓

Predicting the outcome of an election.

Supervised ✓

Dividing patients arriving at a hospital into 5 different categories.

Supervised ✗

Translating documents from Zulu to English.

Supervised ✓

Estimating tomorrow's Bitcoin price.

Supervised ✓

Your answer is partially correct.

You have correctly selected 5.

The correct answer is: An autonomous car learning to drive. → Reinforcement, Segmenting drivers into groups based on their driving styles and behaviours. → Unsupervised, Predicting the outcome of an election. → Supervised, Dividing patients arriving at a hospital into 5 different categories. → Unsupervised, Translating documents from Zulu to English. → Supervised, Estimating tomorrow's Bitcoin price. → Supervised

Question 3

Correct

Mark 4.00 out of 4.00

Consider the following dataset:

TABLE 1. Classification dataset

class	X	Y	Y	X	X	X	Y	Y
feature 1	A	A	B	B	A	A	B	A
feature 2	C	D	D	C	C	D	C	C

Using ID3, which feature should be placed at the root of the tree, and what is the gain?

Select one:

- ☐ a. Feature 1. Gain = 0.9710
- ☐ b. Feature 1. Gain = 0.1261
- ☐ c. Either feature. Gain = 0.9710
- ☐ d. Feature 2. Gain = 0.9183
- ☒ e. Either feature. Gain = 0.0488
- ☐ f. Feature 1. Gain = 0.0488
- ☐ g. Feature 2. Gain = 0.9710
- ☐ h. Feature 2. Gain = 0.0488



Your answer is correct.

The correct answer is: Either feature. Gain = 0.0488

Question 4

Correct

Mark 3.00 out of 3.00

Which of the following statements are TRUE about decision trees?

Select one or more:

- ☒ a. When using ID3, the average entropy of the [data](#) after a split is always less than the average entropy of the [data](#) before a split. ✓
- ☒ b. The entropy of {0.5, 0.5} < the entropy of {0.25, 0.25, 0.25, 0.25} ✓
- ☐ c. The same feature can be used multiple times down the same branch of the tree.
- ☐ d. Decision trees cannot overfit.
- ☐ e. Sum of squares error is used instead of entropy when you have continuous features.
- ☒ f. Decision trees are discriminative models. ✓

Your answer is correct.

The correct answers are: Decision trees are discriminative models., The entropy of {0.5, 0.5} < the entropy of {0.25, 0.25, 0.25, 0.25}, When using ID3, the average entropy of the [data](#) after a split is always less than the average entropy of the [data](#) before a split.

Question **5**

Correct

Mark 4.00 out of 4.00

Consider the training [data](#) in Table 1. We now want to classify a new datapoint (B, C)

TABLE 1. Classification dataset

class	X	Y	Y	X	X	X	Y	Y
feature 1	A	A	B	B	A	A	B	A
feature 2	C	D	D	C	C	D	C	C

Select the most accurate answer describing the true class and the probabilities.

Select one:

- ☐ a. (B, C) is in class X. The probability is given by 0.75
- ☐ b. (B, C) is in class X. The probability is given by 0.43
- ☒ c. (B, C) is in class Y. The probability is given by 0.57
- ☐ d. (B, C) is in class X. The probability is given by 0.19
- ☐ e. (B, C) is in class Y. The probability is given by 0.25
- ☐ f. (B, C) is in class Y. The probability is given by 0.81
- ☐ g. (B, C) is in class X. The probability is given by 0.57
- ☐ h. (B, C) is in class Y. The probability is given by 0.43



Your answer is correct.

The correct answer is: (B, C) is in class Y. The probability is given by 0.57

[◀ Announcements](#)

Jump to...

[Quiz 2: Linear and logistic regression, and neural networks ▶](#)