X

yashraj.devrat.aids.2020@vpkbiet.org ⌄

NPTEL (https://swayam.gov.in/explorer?ncCode=NPTEL) » Introduction To Machine Learning (course)

≡

Click to register
for Certification
exam
(https://examform.nptel.ac.in/2023_10/exam_form/dashboard)

If already
registered, click
to check your
payment status

Course outline

**How does an
NPTEL
online
course
work? ()**

Week 0 ()

Week 1 ()

Week 2 ()

Week 3 ()

Week 4 ()

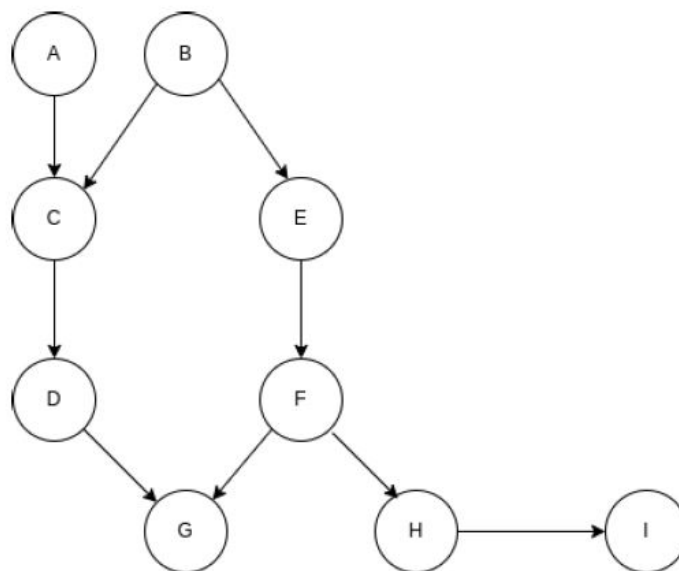**Week 5 ()**

Week 6 ()

Week 8: Assignment 8

The due date for submitting this assignment has passed.

Due on 2023-09-20, 23:59 IST.

Assignment submitted on 2023-09-20, 14:26 IST

1) The figure below shows a Bayesian Network with 9 variables, all of which are   **1 point**
binary.

Which of the following is/are always true for the above Bayesian Network?

◯
$$P(A, B|G) = P(A|G)P(B|G)$$

◉
$$P(A, I) = P(A)P(I)$$

◯

$$P(B, H|E, G) = P(B|E, G)P(H|E, G)$$

○

$$P(C|B, F) = P(C|F)$$

Yes, the answer is correct.
Score: 1

Accepted Answers:
$$P(A, I) = P(A)P(I)$$

2) Consider the following data for 20 budget phones, 30 mid-range phones, and 20 **1 point**
high-end phones:

Type	Single SIM	5G Comaptability	NFC	Total
Budget	15	5	0	20
Mid-Range	20	20	15	30
High End	15	15	15	20

Consider a phone with 2 SIM card slots and NFC but no 5G compatibility. Calculate the
probabilities of this phone being a budget phone, a mid-range phone, and a high-end phone
using the Naive Bayes method. The correct ordering of the phone type from the highest to the
lowest probability is?

○ Budget, Mid-Range, High End

○ Budget, High End, Mid-Range

◉ Mid-Range, High End, Budget

○ High End, Mid-Range, Budget
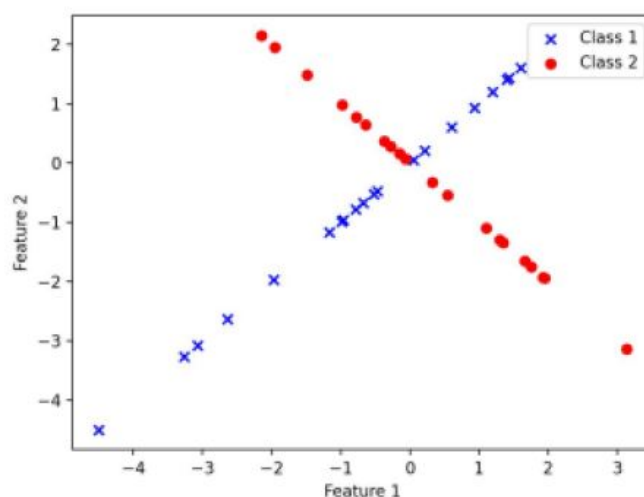
Yes, the answer is correct.
Score: 1

Accepted Answers:
Mid-Range, High End, Budget

3) A dataset with two classes is plotted below.   **1 point**

Does the data satisfy the Naive Bayes assumption?

○ Yes

○ No

◉ The given data is insufficient

○ None of these

No, the answer is incorrect.
Score: 0

Accepted Answers:
No

4) A company hires you to look at their classification system for whether a given **1 point**
customer would potentially buy their product. When you check the existing classifier on different
folds of the training set, you find that it manages a low accuracy of usually around 60%.
Sometimes, it's barely above 50%.
With this information in mind, and without using additional classifiers, which of the following
ensemble methods would you use to increase the classification accuracy effectively?

- ○ Committee Machine
- ◉ AdaBoost
- ○ Bagging
- ○ Stacking

Yes, the answer is correct.
Score: 1

Accepted Answers:
AdaBoost

5) Which of the following algorithms don't use learning rate as a hyperparameter? **1 point**

- ☑ Random Forests
- ☐ Adaboost
- ☑ KNN
- ☑ PCA

Yes, the answer is correct.
Score: 1

Accepted Answers:
Random Forests
KNN
PCA

6) Consider the two statements: **1 point**
Statement 1: Bayesian Networks need not always be Directed Acyclic Graphs (DAGs)
Statement 2: Each node in a bayesian network represents a random variable, and each edge
represents conditional dependence.
Which of these are true?

- ◉ Both the statements are True.
- ○ Statement 1 is true, and statement 2 is false.
- ○ Statement 1 is false, and statement 2 is true.
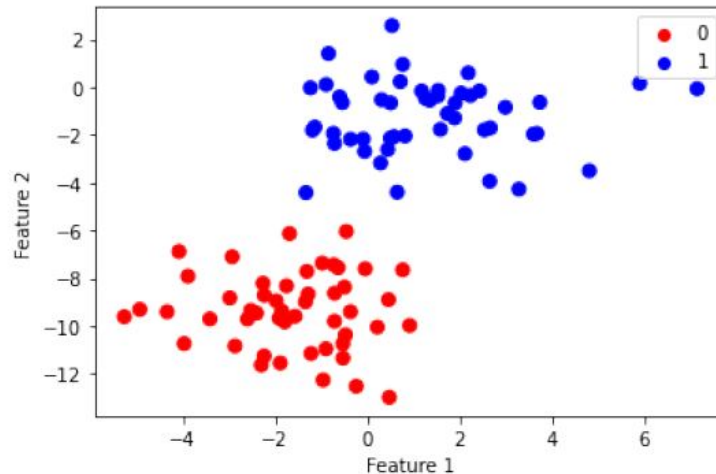- ○ Both the statements are false.

No, the answer is incorrect.
Score: 0

Accepted Answers:
Statement 1 is false, and statement 2 is true.

7) A dataset with two classes is plotted below.   **1 point**

Does the data satisfy the Naive Bayes assumption?

○ Yes

○ No

○ The given data is insufficient

○ None of these

Yes, the answer is correct.
Score: 1

Accepted Answers:
Yes

8) Consider the below dataset: **1 point**

x	y
India won the match.	Cricket
The Mercedes car was driven by Lewis Hamilton.	Formula 1
The ball was driven through the covers for a boundary	Cricket
Max Verstappen has a fast car.	Formula 1
Bumrah is a fast bowler.	Cricket
Max Verstappen won the race	Formula 1

Suppose you have to classify a test example "The ball won the race to the boundary" and are asked to compute P(Cricket |"The ball won the race to the boundary"), what is an issue that you will face if you are using Naive Bayes Classifier, and how will you work around it? Assume you are using word frequencies to estimate all the probabilities.

○ There won't be a problem, and the probability of P(Cricket |"The ball won the race to the boundary") will be equal to 1.

● Problem: A few words that appear at test time do not appear in the dataset.
Solution: Smoothing.

○ Problem: A few words that appear at test time appear more than once in the dataset.
Solution: Remove those words from the dataset.

○ None of these

Yes, the answer is correct.

Score: 1

Accepted Answers:

Problem: A few words that appear at test time do not appear in the dataset.

Solution: Smoothing.