Introduction to Machine Learning -IITKGP

Assignment - 4

TYPE OF QUESTION: MCQ/MSQ

Number of questions: 15 Total mark: 2 * 15 = 30

Q1.

Answer Questions 1-4 with the data provided below:

A spam filtering system has a probability of 0.95 to classify correctly a mail as spam and 0.10 probability of giving false positives. It is estimated that 0.5% of the mails are actual spam mails.

Suppose that the system is now given a new mail to be classified as spam/ not-spam, what is the probability that the mail will be classified as spam?

a. 0.89575
b. 0.10425
c. 0.00475
d. 0.09950

Correct Answer: b

Detailed Solution:

Let S = 'Mails correctly marked spam by the system', T= 'Mails misclassified by the system'(Marked as spam when not spam or Marked as not spam when it is a spam) , M = 'Spam mails'.

P(S|M) = 0.95 , P(S|M′) = 0.10 , P(M) = 0.005

We are to find the probability of mail being classified as spam which can either be if a spam mail is correctly classified as spam or if a mail is misclassified as spam.

$$\therefore P(S) = P(S|M) * P(M) + P(S|M') * P(M') = 0.95 * 0.005 + 0.10 * 0.995$$
$$= 0.10425$$

Q2. Find the probability that, given a mail classified as spam by the system, the mail actually being spam.

a. 0.04556
b. 0.95444
c. 0.00475
d. 0.99525

Correct Answer: a

Detailed Solution: We are to find P(M|S),

$$P(M|S) = \frac{P(S|M)*P(M)}{P(S)} = \frac{0.95 * 0.005}{0.10425} = 0.0455$$

Q3. Given that a mail is classified as not spam, the probability of the mail actually being not spam

a. 0.10425
b. 0.89575
c. 0.003
d. 0.997

Correct Answer: d

Detailed Solution: We are to find P(M′ | S′):

$$P(M'|S') = \frac{P(S'|M')* P(M')}{P(S')} = \frac{(1-P(S|M'))* P(M')}{1-P(S)} = \frac{0.9*0.995}{(1-0.10425)} = 0.997$$

P(S|M′) calculated in question 2 and P(S) in question 1.

Q4. Find the probability that the mail is misclassified:

a. 0.90025
b. 0.09975
c. 0.8955
d. 0.1045

Correct Answer: b

Detailed Solution: We are to find P(T) now:

$$P(T) = P(S \cap M') + P(S' \cap M) = P(S|M') * P(M') + P(S'|M) * P(M)$$
$$= 0.10 \; * \; 0.995 + 0.05 * 0.005$$

= 0.09975

Q5. What is the naive assumption in a Naive Bayes Classifier?

a. All the classes are independent of each other
b. All the features of a class are independent of each other
c. The most probable feature for a class is the most important feature to be considered for classification
d. All the features of a class are conditionally dependent on each other.

Correct Answer: b

Detailed Solution:

Naive Bayes Assumption is that all the features of a class are independent of each other which is not the case in real life. Because of this assumption, the classifier is called Naive Bayes Classifier.

Q6.

Answer Questions 6-7 with the data provided below:

Consider the following dataset. a,b,c are the features and K is the class(1/0):

a	b	c	K
1	0	1	1
1	1	1	1
0	1	1	0
1	1	0	0
1	0	1	0
0	0	0	1

Classify the test instance given below into class 1/0 using a Naive Bayes Classifier.

a	b	c	K
0	0	1	?

a. 0
b. 1

Correct Answer: b

Detailed Solution:

P(K=1|a=0, b=0, c=1) = $\frac{P(K=1)*P(a=0|K=1)*P(b=0|K=1)*P(c=1|K=1)}{P(a=0,b=0,c=1)}$[By Naive Bayes' Assumption]

$= \frac{3}{6} * \frac{1}{3} * \frac{2}{3} * \frac{2}{3}$[The denominator can be ignored since all the features have the same probability for each class]

= 0.07407

P(K=0| a=0, b=0, c=1) = $\frac{3}{6} * \frac{1}{3} * \frac{1}{3} * \frac{2}{3} = 0.03703 < 0.07407$

Hence, the test example will have class 1.

Q7. Find P (K=0| a=1, b=1).

a. $\frac{1}{3}$
b. $\frac{2}{3}$
c. $\frac{1}{9}$
d. $\frac{8}{9}$

Correct Answer: b

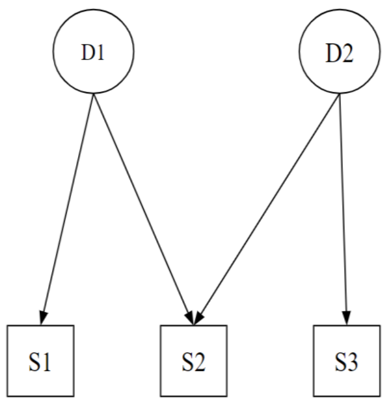
Detailed Solution:

$$P(K=0|a=1,b=1) = \frac{P(K=0) * P(a=1|K=0) * P(b=1|K=0)}{P(K=0) * P(a=1|K=0) * P(b=1|K=0) + P(K=1) * P(a=1|K=1) * P(b=1|K=1)} = \frac{2}{3}$$

Q8.

Answer Questions 8-10 with the data given below:

A patient goes to a doctor with symptoms S1, S2 and S3. The doctor suspects disease D1and D2 and constructs a Bayesian network for the relation among the disease and symptoms as the following:



What is the joint probability distribution in terms of conditional probabilities?

a. $P(D1) * P(D2|D1) * P(S1|D1) * P(S2|D1) * P(S3|D2)$
b. $P(D1) * P(D2) * P(S1|D1) * P(S2|D1) * P(S3|D1,D2)$
c. $P(D1) * P(D2) * P(S1|D2) * P(S2|D2) * P(S3|D2)$
d. $P(D1) * P(D2) * P(S1|D1) * P(S2|D1,D2) * P(S3|D2)$

Correct Answer: d

Detailed Solution:

From the figure, we can see that D1 and D2 are not dependent on any variable as they don't have any incoming directed edges. S1 has an incoming edge from D1, hence S1 depends on D1. S2 has 2 incoming edges from D1 and D2, hence S2 depends on D1 and D2. S3 has an incoming edge from D2, S3 depends on D2. Hence, (d) is the answer.

Q9. Suppose P(D1) = 0.4, P(D2) = 0.7 , P(S1|D1)=0.3 and P(S1| D1′)= 0.6. Find P(S1)

a. 0.12
b. 0.48
c. 0.36
d. 0.60

Correct Answer: b

Detailed Solution:

$$P(S1) \ = P(S1|D1) * P(D1) + P(S1|D1')P(D1') = \mathbf{0.3 * 0.4 + 0.6 * 0.6 = 0.48}$$

Q10. What is the Markov blanket of variable, S3

a. D1
b. D2
c. D1 and D2
d. None

Correct Answer: b

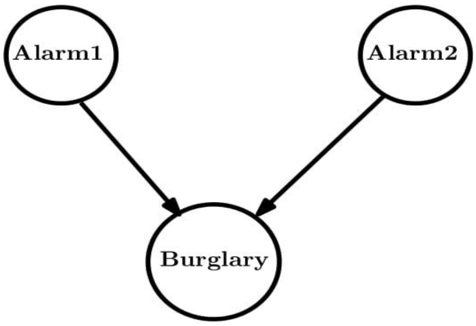
Detailed Solution:

In a Bayesian Network, the Markov blanket of node, X is the set consisting of

- X's parents
- X's children
- Parents of X's children

In the given diagram, variable, S2 has a parent D2 and no children. Hence, the correct answer is (b).

Q11. Consider the following Bayes' network:



Alarm1 means that the first alarm system rings, Alarm2 means that the second alarm system rings, and Burglary means that a burglary is in progress. Now assume that:

P(Alarm1) = 0.1

P(Alarm2) = 0.2

P (Burglary | Alarm1, Alarm2) = 0.8

P (Burglary | Alarm1, ¬Alarm2) = 0.7

P (Burglary | ¬Alarm1, Alarm2) = 0.6

P (Burglary | ¬Alarm1, ¬Alarm2) = 0.5

Calculate P (Alarm2 | Burglary, Alarm1).

a. 0.78
b. 0.22
c. 0.50
d. 0.10

Correct Answer: b

Detailed Solution:

P (Alarm2 | Burglary, Alarm1) = P (Alarm1, Alarm2, Burglary) / P (Burglary, Alarm1)

= 0.016 / 0.072 ≈ 0.22 with

P (Alarm1, Alarm2, Burglary) = P(Alarm1) * P(Alarm2) * P (Burglary | Alarm1, Alarm2)

= 0.1 * 0.2 * 0.8 = 0.016

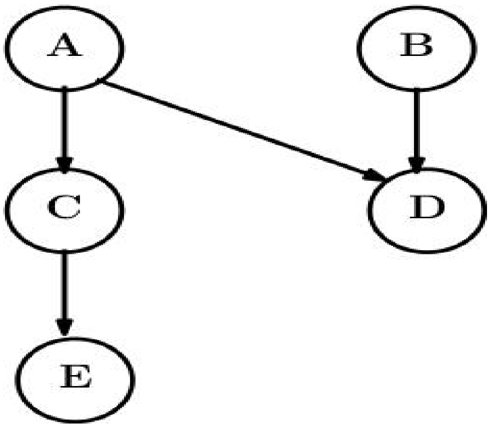
P (Alarm1, ¬ Alarm2, Burglary) = P(Alarm1) * P (¬ Alarm2) * P (Burglary | Alarm1, ¬ Alarm2)

= 0.1 * 0.8 * 0.7 = 0.056

P (Burglary, Alarm1) = P (Alarm1, Alarm2, Burglary) + P (Alarm1, ¬ Alarm2, Burglary)

= 0.016 + 0.056 = 0.072

Q12. Consider the following Bayesian network.



The values of the conditional probabilities are given below. Find $P(D)$.

Assume,

$$P(A) = 0.3$$

$$P(B) = 0.6$$

$$P(C|A) = 0.8$$

$$P(C|\underline{A}) = 0.4$$

$$P(D|A, B) = 0.7$$

$$P(D|A, \underline{B}) = 0.8$$

$$P(D|\underline{A}, B) = 0.1$$

$$P(D|\underline{A}, \underline{B}) = 0.2$$

$$P(E|C) = 0.7$$

$$P(E|\underline{C}) = 0.2$$

a. 0.68
b. 0.32
c. 0.50
d. 0.70

Correct Answer: b

Detailed Solution:

$P(C) = P(C|A) * P(A) + P(C|\underline{A}) * P(\underline{A}) = 0.8 * 0.3 + 0.4 * 0.7 = 0.24 + 0.28 = 0.52$

$\underline{C} = 1 - P(C) = 1 - 0.52 = 0.48$

Next, let's calculate the probability of event D using the law of total probability and the given conditional probabilities:

$P(D) = P(D|A, B) * P(A) * P(B) + P(D|A, \underline{B}) * P(A) * P(\underline{B}) + P(D|\underline{A}, B) * P(\underline{A}) * P(B) + P(D|\underline{A}, \underline{B}) * P(\underline{A}) * P(\underline{B})$

$P(D) = 0.7 * 0.3 * 0.6 + 0.8 * 0.3 * 0.4 + 0.1 * 0.7 * 0.6 + 0.2 * 0.7 * 0.4 = 0.126 + 0.096 + 0.042 + 0.056 = 0.32$

So, the probability P(D) is 0.32.

Q13.

<u>Answer Questions 13-14 with the data given below:</u>

In an oral exam you have to solve exactly one problem, which might be one of three types, A, B, or C, which will come up with probabilities 30%, 20%, and 50%, respectively. During your preparation you have solved 9 of 10 problems of type A, 2 of 10 problems of type B, and 6 of 10 problems of type C.

What is the probability that you will solve the problem of the exam?

a. 0.61
b. 0.39
c. 0.50
d. 0.20

Correct Answer: a

Detailed Solution:

A: Problem of type A.

B: Problem of type B.

C: Problem of type C.

S: You solve the problem

P(A) = 0.30

P(B) = 0.20

P(C) = 0.50

P(S|A) = 9/10

P(S|B) = 2/10

P(S|C) = 6/10

P(S) = P(S|A) P(A) + P(S|B) P(B) + P (S|C) P(C)

= (9/10) * (0.30) + (2/10) * (0.20) + (6/10) * (0.50)

= 0.61

Q14. Given you have solved the problem, what is the probability that it was of type A?

a. 0.35
b. 0.50
c. 0.56
d. 0.44

Correct Answer: d

Detailed Solution:

A: Problem of type A.

S: You solve the problem

P(A) = 0.30

P(S|A) = 9/10

P(S) = 0.61

P(A|S) = $\frac{\boldsymbol{P(S|A) * P(A)}}{\boldsymbol{P(S)}} = = \frac{\boldsymbol{\left(\frac{9}{10}\right)(0.30)}}{\boldsymbol{(0.61)}}$ = 0.4426

Q15. Naive Bayes is a popular classification algorithm in machine learning. Which of the following statements is/are true about Naive Bayes?

a. Naive Bayes assumes that all features are independent of each other, given the class.
b. It is particularly well-suited for text classification tasks, like spam detection.
c. Naive Bayes can handle missing values in the dataset without any special treatment.
d. It is a complex algorithm that requires a large amount of training data.

Correct Answers: a, b

Explanation:

a. Correct. Naive Bayes assumes that features are conditionally independent given the class. This simplifying assumption allows the algorithm to estimate probabilities efficiently even with a limited amount of data.

b. Correct. Naive Bayes is commonly used for text classification tasks, such as spam detection and sentiment analysis, due to its ability to handle high-dimensional feature spaces.

c. Incorrect. Naive Bayes does not handle missing values naturally. Missing values need to be handled before applying the algorithm.

d. Incorrect. Naive Bayes is actually known for its simplicity and ability to work well with small amounts of training data. It is not considered complex and often provides good results with relatively simple computations.

*********END*********