QUESTIONS & ANSWERS
DEMO VERSION
(LIMITED CONTENT)

Question 1

Suppose A, B , and C are events. The probability of A given B , relative to P(|C), is the same as the probability of A given B and C (relative to P). That is,

Options:

A- P(A,B|C) P(B|C) =P(A|B,C)
B- P(A,B|C) P(B|C) =P(B|A,C)
C- P(A,B|C) P(B|C) =P(C|B,C)
D- P(A,B|C) P(B|C) =P(A|C,B)

Answer:

A

Explanation:

From the definition, P(A,B|C) P(B|C) =P(A,B.C)/P(C) P(B.C)/P(C) =P(A,B.C)

P(B,C) =P(A|BC)

This follows from the definition of conditional probability, applied twice: P(A,B)=(PA|B)P(B)

Question 2

Select the correct objectives of principal component analysis

Options:

A- To reduce the dimensionality of the data set
B- To identify new meaningful underlying variables
C- To discover the dimensionality of the data set
D- Only 1 and 2
E- All 1, 2 and 3

Principal component analysis (PCA) involves a mathematical procedure that transforms a number of (possibly) correlated variables into a (smaller) number of uncorrelated variables called principal components. The first principal component accounts for as much of the variability in the data as possible: and each succeeding component accounts for as much of the remaining variability as possible.

Objectives of principal component analysis

1. To discover or to reduce the dimensionality of the data set.
2. To identify new meaningful underlying variables.

Answer:

E

Question 3

Question Type: MultipleChoice

You are working on a problem where you have to predict whether the claim is done valid or not. And you find that most of the claims which are having spelling errors as well as corrections in the manually filled claim forms compare to the honest claims. Which of the following technique is suitable to find out whether the claim is valid or not?

Options:

A- Naive Bayes
B- Logistic Regression
C- Random Decision Forests
D- Any one of the above

In this problem you have been given high-dimensional independent variables like texts, corrections, test results etc. and you have to predict either valid or not valid (One of two). So all of the below technique can be applied to this problem.

Support vector machines Naive Bayes Logistic regression Random decision forests

Answer:

D

Question 4

Question Type: MultipleChoice

Suppose there are three events then which formula must always be equal to P(E1|E2,E3)?

Options:

A- P(E1,E2,E3)P(E1)/P(E2:E3)

B- P(E1,E2;E3)/P(E2,E3)

C- P(E1,E2|E3)P(E2|E3)P(E3)

D- P(E1,E2|E3)P(E3)

E- P(E1,E2,E3)P(E2)P(E3)

This is an application of conditional probability: P(E1,E2)=P(E1|E2)P(E2). so

P(E1|E2) = P(E1.E2)/P(E2)

P(E1,E2,E3)/P(E2,E3)

If the events are A and B respectively, this is said to be 'the probability of A given B'

It is commonly denoted by P(A|B): or sometimes PB(A). In case that both 'A' and 'B' are categorical variables, conditional probability table is typically used to represent the conditional probability.

Answer:

B

Question 5

Question Type: MultipleChoice

You are designing a recommendation engine for a website where the ability to generate more personalized recommendations by analyzing information from the past activity of a specific user, or the history of other users deemed to be of similar taste to a given user. These resources are used as user profiling and helps the site recommend content on a user-by-user basis. The more a given user makes use of the system, the better the recommendations become, as the system gains data to improve its model of that user. What kind of this recommendation engine is ?

Options:

A- Naive Bayes classifier

B- Collaborative filtering

C- Logistic Regression

D- Content-based filtering

Another aspect of collaborative filtering systems is the ability to generate more personalized recommendations by analyzing information from the past activity of a specific user, or the history of other users deemed to be of similar taste to a given user. These resources are used as user profiling and help the site recommend content on a user-by-user basis. The more a given user makes use of the system, the better the recommendations become, as the system gains data to improve its model of

that user

Thank You for trying Databricks-Certified-Professional-Data-Scientist PDF Demo

To try our Databricks-Certified-Professional-Data-Scientist practice exam software visit link below

https://prepbolt.com/Databricks-Certified-Professional-Data-Scientist.html

Start Your Databricks-Certified-Professional-Data-Scientist Preparation

Use Coupon "SAVE50" for extra 50% discount on the purchase of Practice Test Software. Test your Databricks-Certified-Professional-Data-Scientist preparation with actual exam questions.