Pass2Lead
https://Pass2Lead.com

E20-007^Q&As

Data Science and Big Data Analytics

Pass EMC E20-007 Exam with 100% Guarantee

Free Download Real Questions & Answers **PDF** and **VCE** file from:

https://www.pass2lead.com/e20-007.html

100% Passing Guarantee
100% Money Back Assurance

Following Questions and Answers are all new published by EMC
Official Exam Center

⚙ **Instant Download** After Purchase

⚙ **100% Money Back** Guarantee

⚙ **365 Days** Free Update

⚙ **800,000+** Satisfied Customers

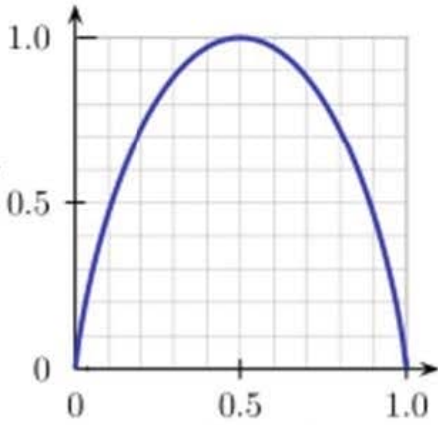![Pass2Lead](https://Pass2Lead.com)
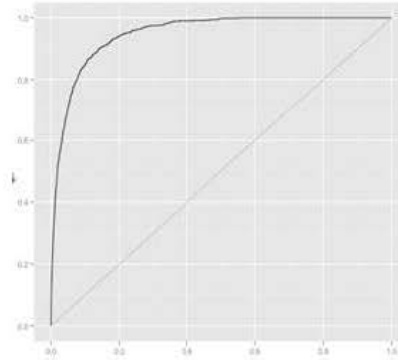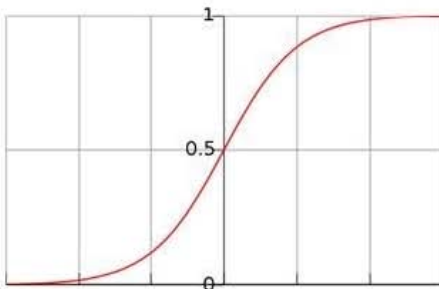**QUESTION 1**
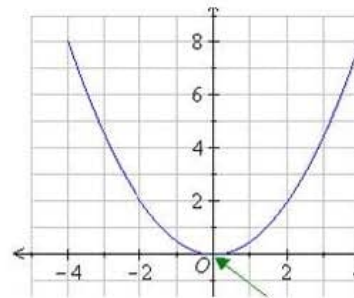
Refer to the exhibit.

Fig-A

Fig-B

Fig-C

Fig-D

$$H = -\sum_{c} p(c) \log_2 p(c)$$

p (c) is the probability of a given class

The exhibit shows four graphs labeled as Fig A thorough Fig D. Which figure represents the entropy function relative to a Boolean classification and is represented by the formula shown in Exhibit?

A. Fig-A

B. Fig-B

C. Fig-C

D. Fig-D

Correct Answer: A

QUESTION 2

Data visualization is used in the final presentation of an analytics project. For what else is this technique commonly

used?

A. Assessing data quality

B. Descriptive statistics

C. ETLT

D. Model selection

Correct Answer: A

QUESTION 3

Consider these itemsets:

(hat, scarf, coat)

(hat, scarf, coat, gloves)

(hat, scarf, gloves)

(hat, gloves)

(scarf, coat, gloves)

What is the confidence of the rule (hat, scarf) -> gloves?

A. 66%

B. 40%

C. 50%

D. 60%

Correct Answer: A

QUESTION 4

You have been assigned to run a linear regression model for each of 5, 000 distinct districts, and all the data is currently stored in a PostgreSQL database. Which tool/library would you use to produce these models with the least effort?

A. MADlib

B. Mahout

C. R

D. HBase

Correct Answer: A

QUESTION 5

Which R data structure allows elements to have different data types?

A. List

B. Vector

C. Matrix

D. Array

Correct Answer: A