

DATABRICKS-CERTIFIED-PR OFESSIONAL-DATA-SCIENTIST^{Q&As}

Databricks Certified Professional Data Scientist Exam

Pass Databricks DATABRICKS-CERTIFIED-PROFESSIONAL-DATA-SCIENTIST Exam with 100% Guarantee

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

https://www.pass4itsure.com/databricks-certified-professional-data-scientist.html

100% Passing Guarantee 100% Money Back Assurance

Following Questions and Answers are all new published by Databricks Official Exam Center https://www.pass4itsure.com/databricks-certified-professional-data-scientis 2024 Latest pass4itsure DATABRICKS-CERTIFIED-PROFESSIONAL-DATA-SCIENTIST PDF and VCE dumps Download

- Instant Download After Purchase
- 100% Money Back Guarantee
- 😳 365 Days Free Update

VCE & PDF

Pass4itSure.com

800,000+ Satisfied Customers





QUESTION 1

Consider flipping a coin for which the probability of heads is p, where p is unknown, and our goa is to estimate p. The obvious approach is to count how many times the coin came up heads and divide by the total number of coin flips. If we flip the coin 1000 times and it comes up heads 367 times, it is very reasonable to estimate p as approximately 0.367. However, suppose we flip the coin only twice and we get heads both times. Is it reasonable to estimate p as 1.0? Intuitively, given that we only flipped the coin twice, it seems a bit rash to conclude that the coin will always come up heads, and _______ is a way of avoiding such rash conclusions.

- A. Naive Bayes
- B. Laplace Smoothing
- C. Logistic Regression
- D. Linear Regression
- Correct Answer: B

Explanation: Smooth the estimates: consider flipping a coin for which the probability of heads is p, where p is unknown, and our goal is to estimate p. The obvious approach is to count how many times the coin came up heads and divide by the total number of coin flips. If we flip the coin 1000 times and it comes up heads 367 times, it is very reasonable to estimate p as approximately 0.367. However, suppose we flip the coin only twice and we get heads both times. Is it reasonable to estimate p as 1.0? Intuitively, given that we only flipped the coin twice, it seems a bit rash to conclude that the coin will always come up heads, and smoothing is a way of avoiding such rash conclusions. A simple smoothing method, called Laplace smoothing (or Laplace\\'s law of succession or add-one smoothing in RandN), is to estimate p by (one plus the number of heads) / (two plus the total number of flips). Said differently, if we are keeping count of the number of heads and the number of tails, this rule is equivalent to starting each of our counts at one, rather than zero. Another advantage of Laplace smoothing is that it avoids estimating any probabilities to be zero, even for events never observed in the data. Laplace add-one smoothing now assigns too much probability to unseen words

QUESTION 2

In which lifecycle stage are test and training data sets created?

- A. Model planning
- B. Discovery
- C. Model building
- D. Data preparation
- Correct Answer: C

Explanation: In Phase 1, the team learns the business domain, including relevant history such as whether the organization or business unit has attempted similar projects in the past from which they can learn. The team assesses the resources available to support the project in terms of people, technology time, and data. Important activities in this phase include framing the business problem as an analytics challenge that can be addressed in subsequent phases and formulating initial hypotheses (IHs) to test and begin learning the data. Data preparation: Phase 2 requires the presence of an analytic sandbox, in which the team can work with data and perform analytics for the duration of the project. The team needs to execute extract, load, and transform (ELT) or extract, transform and load (ETL) to get data into the sandbox. The ELT and ETL are sometimes abbreviated as ETLT Data should be transformed in the ETLT process so



the team can work with it and analyze it. In this phase, the team also needs to familiarize itself with the data thoroughly and take steps to condition the data Model planning: Phase 3 is model planning, where the team determines the methods, techniques, and workflow it intends to follow for the subsequent model building phase. The team explores the data to learn about the relationships between variables and subsequently selects key variables and the most suitable models. Model building: In Phase 4, the team develops datasets for testing, training, and production purposes. In addition, in this phase the team builds and executes models based on the work done in the model planning phase. The team also considers whether its existing tools will suffice for running the models, or if it will need a more robust environment for executing models and workflows (for example, fast hardware and parallel processing, if applicable). Communicate results: In Phase 5, the team, in collaboration with major stakeholders, determines if the results of the project are a success or a failure based on the criteria developed in Phase 1. The team should identify key findings, quantify the business value, and develop a narrative to summarize and convey findings to stakeholders. Operationalize: In Phase 6, the team delivers final reports, briefings, code, and technical documents. In addition, the team may run a pilot project to implement the models in a production environment.

QUESTION 3

Logistic regression is a model used for prediction of the probability of occurrence of an event. It makes use of several variables that may be.....

- A. Numerical
- B. Categorical
- C. Both 1 and 2 are correct
- D. None of the 1 and 2 are correct

Correct Answer: C

Explanation: Logistic regression is a model used for prediction of the probability of occurrence of an event. It makes use of several predictor variables that may be either numerical or categories.

QUESTION 4

What are the key outcomes of the successful analytical projects?

- A. Code of the model
- B. Technical specifications
- C. Presentations for the Analysts
- D. Presentation for Project Sponsors
- Correct Answer: ABCD

Explanation: When your analytical project successfully completed they come up with the following at the end of the projects. Presentations- You will be having presentations like for the all the stakeholders, generally these presentation will

help seniors executives to make better decisions. Similarly you would be creating presentations for the other teams like analysts various visuals you would be creating like ROC Curves, Heat Maps, and Bar Charts etc.



Whatever tools you have used like SAS, R, or Python then accordingly code was developed and you will get that code as one of the outcome. Also you would have created a technical specifications for implementing the codes.

QUESTION 5

Select the choice where Regression algorithms are not best fit

- A. When the dimension of the object given
- B. Weight of the person is given
- C. Temperature in the atmosphere
- D. Employee status
- Correct Answer: D

Explanation: Regression algorithms are usually employed when the data points are inherently numerical variables (such as the dimensions of an object the weight of a person, or the temperature in the atmosphere) but unlike Bayesian algorithms, they\\'re not very good for categorical data (such as employee status or credit score description).

Latest DATABRICKS-CERT DATABRICKS-CERTIFIED- DATABRICKS-CERTIFIED-IFIED-PROFESSIONAL-DATA-SCIENTIST Dumps

PROFESSIONAL-DATA-SCIENTIST PDF Dumps

PROFESSIONAL-DATA-SCIENTIST Practice Test