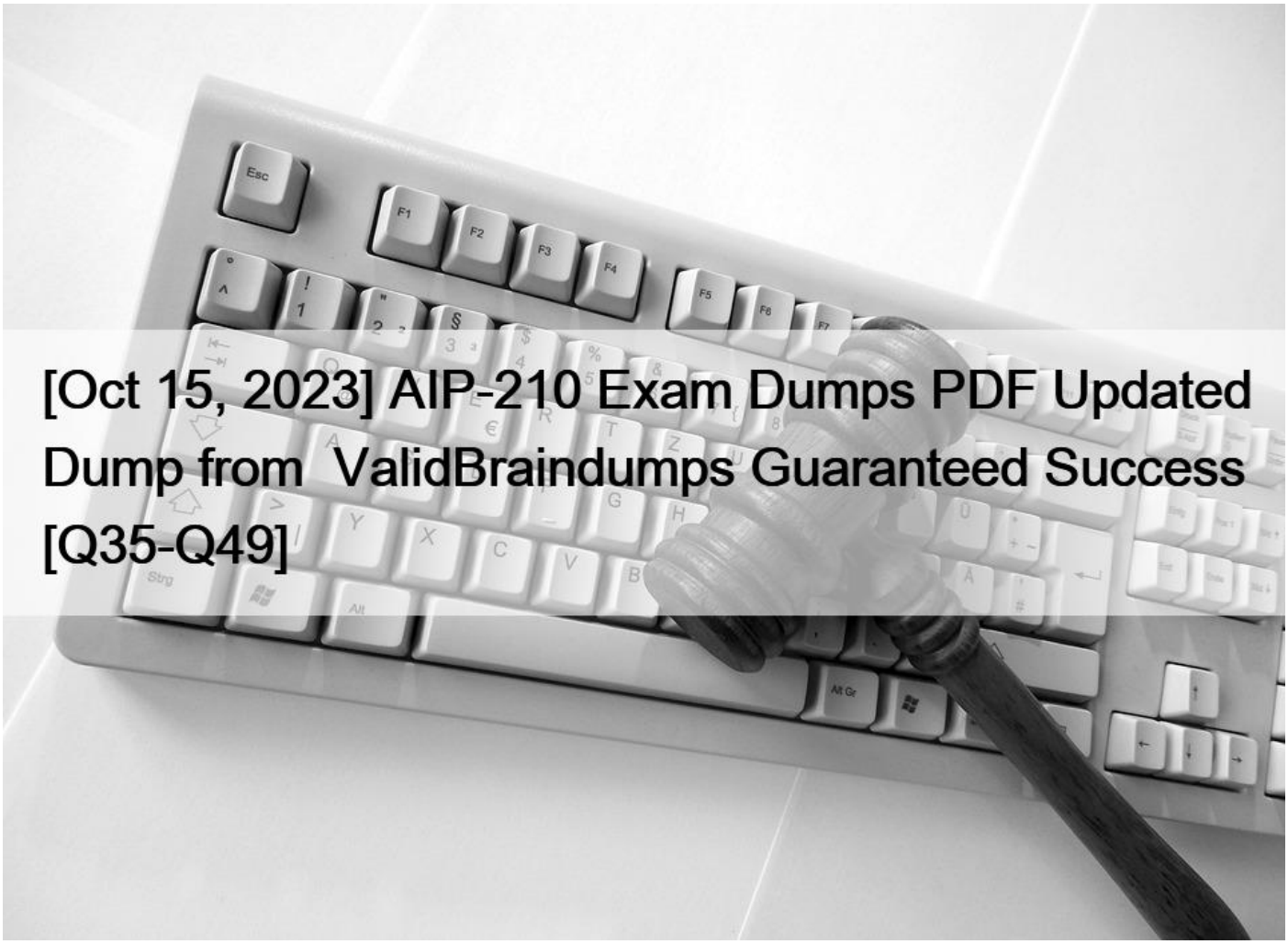
[Oct 15, 2023 AIP-210 Exam Dumps PDF Updated Dump from ValidBraindumps Guaranteed Success [Q35-Q49



[Oct 15, 2023] AIP-210 Exam Dumps PDF Updated Dump from ValidBraindumps Guaranteed Success
Pass Your CertNexus Exam with AIP-210 Exam Dumps

CertNexus AIP-210 Exam Syllabus Topics:

TopicDetailsTopic 1- Design machine and deep learning models- Explain data collection- transformation process in ML workflow
Topic 2- Recognize relative impact of data quality and size to algorithms- Engineering Features for Machine LearningTopic 3- Understanding the Artificial Intelligence Problem- Analyze the use cases of ML algorithms to rank them by their success probabilityTopic 4- Address business risks, ethical concerns, and related concepts in training and tuning- Work with textual, numerical, audio, or video data formats

QUESTION 35

Which of the following are true about the transform-design pattern for a machine learning pipeline? (Select three.) It aims to

separate inputs from features.
* It encapsulates the processing steps of ML pipelines.
* It ensures reproducibility.
* It represents steps in the pipeline with a directed acyclic graph (DAG).
* It seeks to isolate individual steps of ML pipelines.
* It transforms the output data after production.

Explanation

The transform-design pattern for ML pipelines aims to separate inputs from features, encapsulate the processing steps of ML pipelines, and represent steps in the pipeline with a DAG. These goals help to make the pipeline modular, reusable, and easy to understand. The transform-design pattern does not seek to isolate individual steps of ML pipelines, as this would create entanglement and dependency issues. It also does not transform the output data after production, as this would violate the principle of separation of concerns.

QUESTION 36

A data scientist is tasked to extract business intelligence from primary data captured from the public. Which of the following is the most important aspect that the scientist cannot forget to include?
* Cyberprotection
* Cybersecurity
* Data privacy
* Data security

Explanation

Data privacy is the right of individuals to control how their personal data is collected, used, shared, and protected. It also involves complying with relevant laws and regulations that govern the handling of personal data. Data privacy is especially important when extracting business intelligence from primary data captured from the public, as it may contain sensitive or confidential information that could harm the individuals if misused or breached .

QUESTION 37

R-squared is a statistical measure that:
* Combines precision and recall of a classifier into a single metric by taking their harmonic mean.
* Expresses the extent to which two variables are linearly related.
* Is the proportion of the variance for a dependent variable thaf’ s explained by independent variables.
* Represents the extent to which two random variables vary together.

Explanation

R-squared is a statistical measure that indicates how well a regression model fits the data. R-squared is calculated by dividing the explained variance by the total variance. The explained variance is the amount of variation in the dependent variable that can be attributed to the independent variables. The total variance is the amount of variation in the dependent variable that can be observed in the data. R-squared ranges from 0 to 1, where 0 means no fit and 1 means perfect fit.

QUESTION 38

Which of the following unsupervised learning models can a bank use for fraud detection?
* Anomaly detection
* DB5CAN
* Hierarchical clustering
* k-means

Explanation

Anomaly detection is an unsupervised learning technique that identifies outliers or abnormal patterns in data, which can be useful for fraud detection. Anomaly detection algorithms can learn the normal behavior of transactions and flag the ones that deviate significantly from the norm, indicating possible fraud.

QUESTION 39

Below are three tables: Employees, Departments, and Directors.

Employee_Table

ID	Firstname	Lastname	Age	Salary	Dept
1256	Jake	Burton	33	$ 68,230	IT
3489	Zoe	Clark	22	$ 65,000	IT
4577	Ciara	Ortiz	47	$ 75,625	Analyst
8900	Josh	Green	28	$ 43,670	HR

free.validbraindumps.com

Department_Table

ID	Name	Director ID
1	Accountants	4566
2	HR	1230
3	Analyst	9077
4	IT	1346
5	Legal	2088

free.validbraindumps.com

Director_Table

ID

Firstname

Lastname

Age

Salary

DeptJD

4566

Joey

Morin

62

$ 122,000

1

1230

Sam

Clarck

43

$ 95,670

2

9077

Lola

Russell

54

$ 165,700

3

1346

Lily

Cotton

46

$ 156,000

4

2088

Beckett

Good

52

$ 165,000

5

Which SQL query provides the Directors’ Firstname, Lastname, the name of their departments, and the average employee’s salary?
* SELECT m.Firstname, m.Lastname, d.Name, AVG(e.Saiary) as Dept_avg_Saiary FROM Employee_Table as e LEFT JOIN Department_Table as d on e.Dept = d.Name LEFT JOIN Directorjable as m on d.ID = m.DeptJD GROUP BY m.Firstname, m.Lastname, d.Name
* SELECT m.Firstname, m.Lastname, d.Name, AVG(e.Salary) as Dept_avg_Salary FROM Employee_Table as e RIGHT JOIN Departmentjable as d on e.Dept = d.Name INNER JOIN Directorjable as m on d.ID = m.DeptJD GROUP BY d.Name
* SELECT m.Firstname, m.Lastname, d.Name, AVG(e.Salary) as Dept_avg_Salary FROM Employee_Table as e RIGHT JOIN Department_Table as d on e.Dept = d.Name INNER JOIN Directorjable as m on d.ID = m.DeptJD GROUP BY e.Salary
* SELECT m.Firstname, m.Lastname, d.Name, AVG(e.Salary) as Dept_avg_Salary FROM Employee_Table as e RIGHT JOIN Department_Table as d on e.Dept = d.Name INNER JOIN Directorjable as m on d.ID = m.DeptID GROUP BY m.Firstname, m.Lastname, d.Name
Explanation

This SQL query provides the Directors’ Firstname, Lastname, the name of their departments, and the average employee’s salary by joining the three tables using the appropriate join types and conditions. The RIGHT JOIN between Employee_Table and Department_Table ensures that all departments are included in the result, even if they have no employees. The INNER JOIN between Department_Table and Directorjable ensures that only departments with directors are included in the result. The GROUP BY clause groups the result by the directors’ names and departments’ names, and calculates the average salary for each group using the AVG function. References: SQL Joins – W3Schools, SQL GROUP BY Statement – W3Schools

QUESTION 40

Which two of the following criteria are essential for machine learning models to achieve before deployment?

(Select two.)
* Complexity
* Data size
* Explainability
* Portability
* Scalability
Explanation

Scalability and explainability are two criteria that are essential for ML models to achieve before deployment.

Scalability is the ability of an ML model to handle increasing amounts of data or requests without compromising its performance or quality. Scalability can help ensure that the model can meet the demand and expectations of users or customers, as well as adapt to changing conditions or environments. Explainability is the ability of an ML model to provide clear and intuitive explanations for its predictions or decisions.

Explainability can help increase trust and confidence among users or stakeholders, as well as enable accountability and responsibility for the model’s actions and outcomes.

QUESTION 41

Which of the following equations best represent an LI norm?
* |x| + |y|
* |x|+|y|2
* |x|-|y|
* |x|2+|y|2

Explanation

An L1 norm is a measure of distance or magnitude that is defined as the sum of the absolute values of the components of a vector. For example, if x and y are two components of a vector, then the L1 norm of that vector is |x| + |y|. The L1 norm is also known as the Manhattan distance or the taxicab distance, as it represents the shortest path between two points in a grid-like city.

QUESTION 42

Which of the following items should be included in a handover to the end user to enable them to use and run a trained model on their own system? (Select three.)
* Information on the folder structure in your local machine
* Intermediate data files
* Link to a GitHub repository of the codebase
* README document
* Sample input and output data files

Explanation

A handover is the process of transferring the ownership and responsibility of an ML system from one party to another, such as from the developers to the end users. A handover should include all the necessary information and resources that enable the end users to use and run a trained model on their own system. Some of the items that should be included in a handover are:

Link to a GitHub repository of the codebase: A GitHub repository is an online platform that hosts the source code and version control of an ML system. A link to a GitHub repository can provide the end users with access to the latest and most updated version of the codebase, as well as the history and documentation of the changes made to the code.

README document: A README document is a text file that provides an overview and instructions for an ML system. A README document can include information such as the purpose, features, requirements, installation, usage, testing, troubleshooting, and license of the system.

Sample input and output data files: Sample input and output data files are data files that contain examples of valid inputs and expected outputs for an ML system. Sample input and output data files can help the end users understand how to use and run the system, as well as verify its functionality and performance.

QUESTION 43

Which of the following is the primary purpose of hyperparameter optimization?
* Controls the learning process of a given algorithm
* Makes models easier to explain to business stakeholders
* Improves model interpretability
* Increases recall over precision
Explanation

Hyperparameter optimization is the process of finding the optimal values for hyperparameters that control the learning process of a given algorithm. Hyperparameters are parameters that are not learned by the algorithm but are set by the user before training. Hyperparameters can affect the performance and behavior of the algorithm, such as its speed, accuracy, complexity, or generalization. Hyperparameter optimization can help improve the efficiency and effectiveness of the algorithm by tuning its hyperparameters to achieve the best results.

QUESTION 44

Which of the following sentences is TRUE about the definition of cloud models for machine learning pipelines?
* Data as a Service (DaaS) can host the databases providing backups, clustering, and high availability.
* Infrastructure as a Service (IaaS) can provide CPU, memory, disk, network and GPU.
* Platform as a Service (PaaS) can provide some services within an application such as payment applications to create efficient results.
* Software as a Service (SaaS) can provide AI practitioner data science services such as Jupyter notebooks.
Explanation

Cloud models are service models that provide different levels of abstraction and control over computing resources in a cloud environment. Some of the common cloud models for machine learning pipelines are:

Software as a Service (SaaS): SaaS provides ready-to-use applications that run on the cloud provider’s infrastructure and are accessible through a web browser or an API. SaaS can provide AI practitioner data science services such as Jupyter notebooks, which are web-based interactive environments that allow users to create and share documents that contain code, text, visualizations, and more.

Platform as a Service (PaaS): PaaS provides a platform that allows users to develop, run, and manage applications without worrying about the underlying infrastructure. PaaS can provide some services within an application such as payment applications to create efficient results.

Infrastructure as a Service (IaaS): IaaS provides access to fundamental computing resources such as servers, storage, networks, and operating systems. IaaS can provide CPU, memory, disk, network and GPU resources that can be used to run machine learning models and applications.

Data as a Service (DaaS): DaaS provides access to data sources that can be consumed by applications or users on demand. DaaS can host the databases providing backups, clustering, and high availability.

QUESTION 45

You and your team need to process large datasets of images as fast as possible for a machine learning task.

The project will also use a modular framework with extensible code and an active developer community.

Which of the following would BEST meet your needs?

* Caffe
* Keras
* Microsoft Cognitive Services
* TensorBoard

Explanation

Caffe is a deep learning framework that is designed for speed and modularity. It can process large datasets of images efficiently and supports various types of neural networks. It also has a large and active developer community that contributes to its code base and documentation. Caffe is suitable for image processing tasks such as classification, segmentation, detection, and recognition

QUESTION 46

Which database is designed to better anticipate and avoid risks of AI systems causing safety, fairness, or other ethical problems?

* Asset
* Code Repository
* Configuration Management
* Incident

Explanation

An incident database is a database that is designed to better anticipate and avoid risks of AI systems causing safety, fairness, or other ethical problems. An incident database collects and stores information about incidents or events where AI systems have caused or contributed to negative outcomes or harms, such as accidents, errors, biases, discriminations, or violations. An incident database can help identify patterns, trends, causes, impacts, and solutions for AI-related incidents, as well as provide guidance and best practices for preventing or mitigating future incidents.

QUESTION 47

An AI practitioner incorporates risk considerations into a deployment plan and decides to log and store historical predictions for potential, future access requests.

Which ethical principle is this an example of?

* Fairness
* Privacy
* Safety
* Transparency

Explanation

Transparency is an ethical principle that describes the degree to which an AI system can provide clear and understandable information about its inputs, outputs, processes, and decisions. Transparency can help increase trust and confidence among users and stakeholders, as well as enable accountability and responsibility for the system’s actions and outcomes. Logging and storing historical predictions for potential, future access requests is an example of transparency, as it can help provide evidence and explanation for the system’s recommendations, as well as facilitate auditing and feedback.

QUESTION 48

Which of the following methods can be used to rebalance a dataset using the rebalance design pattern?

* Bagging
* Boosting
* Stacking
* Weighted class

Explanation

Weighted class is a technique to rebalance a dataset by assigning different weights to each class, according to their frequency in the dataset. The weights are inversely proportional to the class frequency, meaning that rare classes have higher weights and common classes have lower weights. This helps to reduce the bias towards the majority class and improve the model performance on the minority class. References: 4. Data Validation – Building Machine Learning Pipelines, A guide to React design patterns – LogRocket Blog

QUESTION 49

Which of the following metrics is being captured when performing principal component analysis?
* Kurtosis
* Missingness
* Skewness
* Variance

Explanation

Principal component analysis (PCA) is a technique that reduces the dimensionality of a dataset by transforming it into a set of new variables called principal components. The principal components are linear combinations of the original variables that capture the maximum amount of variance in the data. The first principal component explains the most variance, the second principal component explains the second most variance, and so on. The goal of PCA is to retain as much variance as possible while reducing the number of variables.

New Real AIP-210 Exam Dumps Questions: https://www.validbraindumps.com/AIP-210-exam-prep.html]