

Boost up Your Certification Score

DASCA

SDS

Senior Data Scientist



For More Information – Visit link below:

<https://www.examsboost.com/>

Product Version

- ✓ Up to Date products, reliable and verified.
- ✓ Questions and Answers in PDF Format.

Latest Version: 6.0

Question: 1

SpamAssassin has been developed to detect:

- A. Spam emails
- B. Email with big attachments
- C. Email with virus
- D. None of the above

Answer: A

Explanation:

Apache SpamAssassin is one of the most widely used open-source tools for spam email detection. It applies a rule-based system combined with Bayesian filtering, heuristics, and collaborative filtering methods to classify incoming emails as spam or legitimate.

Option A (Spam emails): Correct, this is the main function.

Option B (Big attachments): Incorrect. Large attachment filtering is not its primary purpose.

Option C (Email with virus): Incorrect. That falls under antivirus or malware detection tools, not SpamAssassin.

Option D: Incorrect since A is valid.

Thus, the correct answer is Option A (Spam emails).

[Reference:, DASCA Data Scientist Knowledge Framework (DSKF) – Business Applications of Data Science: Email Filtering and Text Mining.,]

Question: 2

Machine learning can be categorized as:

- A. Supervised learning
- B. Unsupervised learning
- C. Reinforcement learning
- D. All of the above

Answer: D

Explanation:

Machine learning (ML) can be broadly divided into three main paradigms:

Supervised Learning (Option A):

Data includes labeled outputs (e.g., classification, regression).

Goal: Learn a mapping from input to output.

Unsupervised Learning (Option B):

Data has no labels.

Goal: Discover hidden patterns (e.g., clustering, dimensionality reduction).

Reinforcement Learning (Option C):

Agent interacts with an environment and learns by maximizing cumulative rewards through trial and error.

Used in robotics, game AI, and autonomous systems.

Since all three categories are valid, the correct answer is Option D (All of the above).

[Reference:, DASCA Data Scientist Knowledge Framework (DSKF) – Machine Learning Paradigms: Supervised, Unsupervised, Reinforcement.,]

Question: 3

The DevOps movement is an outgrowth of which of the following software development methodologies?

- A. Agile
- B. Waterfall
- C. Promise-based algorithms
- D. Test-driven development and model-driven development

Answer: A

Explanation:

The DevOps movement evolved as a natural extension of the Agile methodology.

Agile (Option A): Agile emphasizes iterative development, collaboration, and flexibility. While Agile improved software development speed, it created challenges in integrating development with IT operations. DevOps emerged to address this by bringing operations into the Agile cycle — enabling continuous integration, delivery, and deployment.

Waterfall (Option B): Incorrect. Waterfall is a rigid, sequential methodology, fundamentally opposite to the DevOps philosophy.

Promise-based algorithms (Option C): Not a methodology — irrelevant here.

Test-driven development and model-driven development (Option D): While these practices support DevOps, they are not the origin of the movement.

Thus, the DevOps movement is an outgrowth of Agile methodology.

[Reference:, DASCA Data Scientist Knowledge Framework (DSKF) – Agile and DevOps in Data Science Projects.,]

Question: 4

Which of the following standardizes scores similar to a percentile rank but preserves equal interval properties of a Z-score?

- A. Trend analysis
- B. Normal Curve Equivalent (NCE)
- C. High Curve Equivalent (HCE)
- D. Medium Curve Equivalent (MCE)
- E. None of the above

Answer: B

Explanation:

Normal Curve Equivalent (NCE) scores are standardized scores designed to:

Range between 1 and 99.

Be comparable to percentile ranks but with the advantage of equal-interval properties like Z-scores.

This makes NCE scores useful in educational assessments, survey analysis, and statistical modeling.

Option A (Trend analysis): Incorrect. Not related to score standardization.

Option B (Correct): NCE fits the definition perfectly.

Option C (HCE) & D (MCE): Not recognized standard measures in statistics.

Option E: Incorrect, since Option B is valid.

Thus, the correct answer is Option B: Normal Curve Equivalent (NCE).

[Reference:, DASCA Data Scientist Knowledge Framework (DSKF) – Statistical Methods in Data Science: Z-scores, Percentiles, and NCE.,]

Question: 5

What is DevOps?

- A. Software Development
- B. Software Operations
- C. Quality Assurance
- D. All

Answer: D

Explanation:

DevOps is not just about coding (development) or system administration (operations). It is a holistic cultural and technical practice that unifies:

Software Development (Option A): Writing and building applications.

Software Operations (Option B): Deploying, monitoring, and maintaining systems in production.

Quality Assurance (Option C): Ensuring the reliability, security, and performance of applications through testing and automation.

Thus, DevOps encompasses all three dimensions, making the correct answer Option D (All).

[Reference:, DASCA Data Scientist Knowledge Framework (DSKF) – Business Applications of Data Science: DevOps & Continuous Delivery., , ,]

Question: 6

Which of the following is a DevOps Practice?

- A. Continuous build
- B. Continuous integration
- C. Continuous delivery
- D. All of the above

Answer: D

Explanation:

DevOps is a collaborative practice that integrates software development (Dev) and IT operations (Ops) to shorten development cycles and deliver applications reliably. Common DevOps practices include: Continuous Build (Option A): Automating compilation and packaging of source code to ensure consistent builds.

Continuous Integration (Option B): Developers frequently merge code into a shared repository, which is automatically tested to catch integration issues early.

Continuous Delivery (Option C): Automating software release pipelines so applications can be deployed to production quickly and reliably.

Since all of these are essential DevOps practices, the correct answer is Option D (All of the above).

[Reference:, DASCA Data Scientist Knowledge Framework (DSKF) – Business Applications of Data Science: DevOps Practices in Data Science Projects.,]

Question: 7

Which of the following is an example of graphical model?

- A. Markov Random Fields
- B. Bayesian Networks
- C. Geographical Networks
- D. Both A and B
- E. Both A and C

Answer: D

Explanation:

Graphical models are probabilistic models that represent variables and dependencies using graphs:

Markov Random Fields (Option A): Undirected graphical models that capture joint distributions over variables with neighborhood dependencies.

Bayesian Networks (Option B): Directed acyclic graphical models that encode conditional dependencies between random variables.

Geographical Networks (Option C): While they are graphs, they are not probabilistic graphical models used in statistics/ML.

Thus, the correct answer is Option D (Both A and B).

[Reference:, DASCA Data Scientist Knowledge Framework (DSKF) – Analytics: Graphical Models (Bayesian Networks & Markov Random Fields).,]

Question: 8

A workflow refers to a:

- A. Directed cyclic graph
- B. Indirected cyclic graph

- C. Directed acyclic graph
- D. Indirected acyclic graph

Answer: C

Explanation:

In data pipelines and process orchestration, a workflow is represented as a Directed Acyclic Graph (DAG):

Directed: Each edge has a direction, representing task dependencies.

Acyclic: No cycles exist; tasks must follow a sequence without looping back.

Graph: Represents tasks as nodes and dependencies as edges.

This structure is common in tools like Apache Airflow, Spark DAGs, and Hadoop MapReduce job schedulers.

Option A & B: Incorrect, as workflows cannot have cycles (would cause infinite loops).

Option D: Incorrect, because workflows are directed, not indirected.

Thus, the correct answer is Option C (Directed Acyclic Graph).

[Reference:, DASCA Data Scientist Knowledge Framework (DSKF) – Data Engineering Architectures: Workflow Management with DAGs.]

Question: 9

Self-driving car is an example of:

- A. Supervised learning
- B. Unsupervised learning
- C. Reinforcement learning
- D. All of the above

Answer: C

Explanation:

Self-driving cars (autonomous vehicles) are an application of Reinforcement Learning (RL) in machine learning:

In RL, an agent (car) interacts with an environment (roads, obstacles, traffic) and learns to maximize rewards (e.g., safe driving, efficient navigation).

The system improves performance through trial-and-error learning, guided by reward signals such as staying in a lane or avoiding collisions.

Supervised learning (A): Used in some supporting tasks like image recognition (e.g., identifying stop signs), but not the core paradigm for self-driving.

Unsupervised learning (B): Useful for clustering sensor data, but again not the main paradigm.

Reinforcement learning (C): Correct, since self-driving fundamentally depends on RL decision-making.

Thus, the correct answer is Option C (Reinforcement Learning).

[Reference:, DASCA Data Scientist Knowledge Framework (DSKF) – Machine Learning Paradigms: Reinforcement Learning and Autonomous Systems.,]

Question: 10

Which classification steps are performed in inductive techniques?

- i. Training Step
- ii. Test Step
- iii. Validation Step
- iv. Application Step

- A. i, ii
- B. ii, iii
- C. i, ii, iv
- D. i, ii, iii, iv

Answer: D

Explanation:

Inductive learning techniques in machine learning (such as decision trees, neural networks, or SVMs) follow a systematic sequence of steps for classification:

Training Step (i): A model is built using training data, where the system learns relationships between features and target labels.

Test Step (ii): The trained model is evaluated on unseen test data to measure its performance and generalizability.

Validation Step (iii): Often, a validation set is used to fine-tune model parameters, avoid overfitting, and choose the best model configuration.

Application Step (iv): The final validated model is applied to classify new, real-world data.

Since all four steps (i, ii, iii, iv) are essential to inductive classification, the correct answer is Option D (i, ii, iii, iv).

[Reference:, DASCA Data Scientist Knowledge Framework (DSKF) – Analytics & Machine Learning: Classification and Inductive Learning Techniques.,]

Thank You for Trying Our Product

For More Information – **Visit link below:**

<https://www.examsboost.com/>

15 USD Discount Coupon Code:

G74JA8UF

FEATURES

- ✓ **90 Days Free Updates**
- ✓ **Money Back Pass Guarantee**
- ✓ **Instant Download or Email Attachment**
- ✓ **24/7 Live Chat Support**
- ✓ **PDF file could be used at any Platform**
- ✓ **50,000 Happy Customer**



Visit us at: <https://www.examsboost.com/test/sds>