Isaca Data-Science-Fundamentals Dumps

Shared by Maddox on 26-11-2026

For More Free Questions and Preparation Resources

Check the Links on Last Page

Question 1

Question Type: MultipleChoice

Which of the following does the deep learning branch of machine learning mimic?

Options:

A- the brain's neural network.
B- a network system's wiring.
C- hierarchies within complex supercomputers.

Answer:

A

Explanation:

Deep learning is a subset of machine learning built on artificial neural networks, a design inspired by the way biological neurons are interconnected in the brain. An artificial network is composed of layers of nodes, each of which receives weighted inputs, applies an activation function, and passes a signal onward, loosely

paralleling how neurons transmit impulses across synapses. The word 'deep' refers to the presence of many hidden layers, which allow the model to learn increasingly abstract representations: early layers might detect edges in an image, later layers shapes, and the final layers whole objects. Training adjusts the connection weights through backpropagation and gradient descent. Option B confuses the idea with computer networking; the topology of cables and switches has no bearing on how a neural model learns. Option C is also wrong, because the layered structure of a network is a computational abstraction and does not copy the organization of supercomputer hardware, although such hardware, particularly GPUs, is what makes training large networks practical. The biological analogy is imperfect, yet it is the conceptual foundation the study guide uses to explain deep learning. Reference: Data Science Fundamentals Study Guide -- Data Science Concepts: Machine Learning and Deep Learning (Neural Networks).

Question 2

Question Type: MultipleChoice

Which type of database is BEST suited to storing flexible, semi-structured records such as JSON documents whose fields vary from one record to the next?

Options:

A- Relational database
B- Document-oriented NoSQL database
C- Data warehouse built on a star schema

Answer:

B

Explanation:

Choosing a storage technology begins with the structure of the data. Relational systems require a predefined schema in which every row in a table has the same columns, a model that excels for consistent, transactional records but becomes awkward when attributes differ between records or change often. Document-oriented NoSQL databases, such as MongoDB and Couchbase, store each record as a self-contained document, commonly in JSON or a similar format, and permit nested structures and differing fields without schema changes. This flexibility suits product catalogs, user profiles , event payloads, and content management, where requirements evolve rapidly, and it scales horizontally across servers, which explains the pairing of NoSQL with big data workloads. Option B is therefore correct. Option A can store JSON in some implementations, but doing so departs from its core design and forfeits the strengths that normalized tables provide. Option C is a modeling approach for analytical warehouses, in which a central fact table is surrounded by dimension tables; it is intended for structured, historical reporting rather than variable documents. Other NoSQL categories include key-value, column-family, and graph stores, each tuned to a particular access pattern. Reference: Data Science Fundamentals Study Guide -- Data Management: Database Types (Relational and NoSQL) and Data Storage.

Question 3

Question Type: MultipleChoice

Which of the following best describes what a data lake architecture typically involves?

Options:

A- a cloud-based repository with parallel processing.
B- a central repository with batch processing.
C- a relational database management system (RDBMS).

Answer:

A

Explanation:

A data lake is a large-scale repository that holds data in its native form, including structured tables, semi-structured logs and JSON, and unstructured files such as documents, images, and audio. Data are stored without a predefined schema, and structure is applied only when the data are read, an approach known as schema-on-read. Modern implementations are overwhelmingly built on cloud object storage, which offers elastic capacity and low cost per terabyte, and they are paired with distributed engines such as Apache Spark, Hadoop, and cloud query services that spread the work across many nodes in parallel. This combination of scalable cloud storage and parallel processing is the characteristic described in option A. Option B is less precise: a lake is indeed a central repository, but it is not limited to batch processing, since the same platform supports streaming ingestion, interactive queries, and machine learning workloads. Option C describes an entirely different technology. An RDBMS enforces a rigid schema, stores structured rows and columns, and is optimized for transactions, so it cannot economically accommodate the variety and volume that lakes target.

Without governance and metadata management a lake can degrade into a so-called data swamp, which is why cataloging and access controls are integral to the architecture. Reference: Data Science Fundamentals Study Guide -- Data Management: Data Storage Architectures (Data Lakes and Data Warehouses).

Question 4

Question Type: MultipleChoice

Which of the following tasks is an example of a regression problem?

Options:

A- Determining whether an incoming email is spam
B- Predicting the sale price of a house from its size and location
C- Grouping customers into segments based on purchasing behavior

Answer:

B

Explanation:

Supervised learning problems divide according to the type of output being predicted. Classification assigns an observation to a discrete category, whereas regression estimates a continuous numeric quantity. Predicting the sale price of a house from attributes such as floor area, number of rooms, and neighborhood yields a value that can fall anywhere along a numeric scale, so it is a regression task; linear regression, regression trees, and gradient boosting are common approaches, and performance is judged with error measures such as mean absolute error or root mean squared error. Option A, identifying spam, produces one of two labels and is therefore binary classification, assessed through metrics like precision and recall. Option C, grouping customers by behavior, has no predefined target variable at all; it is a clustering exercise belonging to unsupervised learning, in which the algorithm discovers structure on its own. A reliable shortcut is to ask what the model must output: a number on a continuum points to regression, a label from a fixed set points to classification, and the absence of any target points to unsupervised methods. Selecting the right problem type is the first step in choosing algorithms and evaluation criteria. Reference: Data Science Fundamentals Study Guide -- Data Science Concepts: Machine Learning Approaches (Regression, Classification and Clustering).

Question 5

Question Type: MultipleChoice

What is the correct definition of a hypothesis?

Options:

A- knowledge based on evidence and facts collected via repeatable tests to support a theory.
B- a proposed Explanation: made on the basis of limited evidence as a starting point for further investigation.
C- observational collection of signals and responses to stimuli in order to define a theory.

Answer:

B

Explanation:

A hypothesis is a tentative, testable statement that proposes an Explanation: for an observed pattern or relationship. It is formed from limited evidence, prior knowledge, or reasoning, and it

serves as the starting point for further investigation rather than a conclusion. In data science, a hypothesis might state that customers who receive a discount code are more likely to renew. It is expressed so that data can support or contradict it, usually as a null and an alternative hypothesis, and it is evaluated through experiments or statistical tests. Option B captures this definition. Option A describes a theory: an Explanation: supported by substantial evidence and repeatable testing, which has survived repeated attempts at refutation and therefore carries far greater confidence than a hypothesis. Option C describes observation, the systematic gathering of information about phenomena, which supplies the evidence from which hypotheses are drawn but is not itself a proposed Explanation:. The scientific method links the three: observation leads to a question, the question yields a hypothesis, testing generates evidence, and accumulated, consistent evidence may eventually support a theory. Distinguishing these terms is important because treating an untested hypothesis as established fact is a frequent source of analytical error. Reference: Data Science Fundamentals Study Guide -- Data Science Concepts: The Scientific Method and Hypothesis Formulation.

P2P exams

Question 6

Question Type: MultipleChoice

What does Bayes' Theorem state?

Options:

A- There are four classes of scales that are unchanging.
B- Processing power doubles in size every two years.
C- If something is true, something related to it is potentially true.

P2P exams

Answer:

C

Explanation:

Bayes' Theorem is a result in probability theory describing how to revise the probability of a hypothesis once new evidence is observed. Expressed mathematically, the probability of A given B equals the probability of B given A multiplied by the prior probability of A, divided by the probability of B. The theorem formalizes conditional reasoning: knowledge that one event has occurred or is true changes the likelihood of a related event. This is why it appears in spam filtering, medical diagnostic testing, and the naive Bayes classifier,

where prior beliefs are combined with observed features to yield a posterior probability. Option C

expresses that relationship in general terms, namely that if something is true, a related event becomes potentially true. Option A confuses the theorem with the four scales of measurement (nominal, ordinal, interval, and ratio) associated with S. S. Stevens, which concern how variables are classified rather than how probabilities are updated. Option B restates Moore's law, an observation about transistor density and computing power with no connection to probability. Candidates should be able to distinguish these three named principles, since each belongs to a different strand of the data science foundation. Reference: Data Science Fundamentals Study Guide -- Data Science Concepts: Probability and Statistics Fundamentals (Bayes' Theorem).

Question 7

Question Type: MultipleChoice

Which of the following is an example of exploratory analysis?

Options:

A- Confirming a preconceived assumption
B- Noticing patterns in visualization charts
C- Disproving a preconceived assumption

Answer:

B

Explanation:

Exploratory analysis is an open-ended investigation of a data set carried out without a fixed hypothesis. The analyst examines distributions, plots relationships, and scans summaries to learn what the data contain, detect anomalies, and generate questions worth pursuing. Visualization is a core technique, since patterns such as clusters, trends, seasonality, and outliers are often visible in a chart long before they can be described numerically. Noticing patterns in visualization charts is therefore a characteristic exploratory activity, and option B is correct. Options A and C both begin from a preconceived assumption and use the data to test it, either supporting or refuting it. This is confirmatory, hypothesis-driven analysis, in which a claim is formulated in advance and evaluated through statistical tests with defined criteria. The two approaches are complementary. Exploration produces hypotheses and informs data preparation, and confirmation then tests them rigorously. A recognized caution is that patterns uncovered through exploration must not be validated on the same data that suggested them, because doing so overstates their significance; fresh data or a hold-out sample should be used. Knowing which mode of analysis is under way clarifies the standard of evidence expected and shapes how

findings should be reported to stakeholders. Reference: Data Science Fundamentals Study Guide -- Data Science Process: Exploratory and Confirmatory Data Analysis.

Question 8

Question Type: MultipleChoice

How would you define a key performance indicator (KPI)?

Options:

A- a measure of progress in achieving an intended result.
B- an output that establishes what must be accomplished.
C- one of the focus areas of a balanced scorecard (BSC).

Answer:

A

Explanation:

A key performance indicator is a quantifiable measure used to evaluate how effectively an organization, team, or process is progressing toward a defined objective. It turns a strategic goal into something observable by pairing a metric, such as customer churn rate, average order value, or on-time delivery percentage, with a target and a time frame, so that results can be tracked and compared against expectations. Well-chosen KPIs are relevant to the objective, measurable, and tied to decisions that someone can influence. Option A defines the concept accurately as a measure of progress in achieving an intended result. Option B describes a goal or objective, which states what must be accomplished; the KPI is the yardstick used to judge whether the goal is being met, not the goal itself. Option C refers to the perspectives of the balanced scorecard, namely financial, customer, internal process, and learning and growth. Those perspectives are categories used to organize strategy and group measures, whereas KPIs are the individual indicators placed within them. KPIs are commonly distinguished as leading indicators, which anticipate future performance, and lagging indicators, which report outcomes after the fact. Data science teams use them to link analytical work to business value.

Question 9

Question Type: MultipleChoice

Which business analytics technique are business stakeholders using when they try to forecast holiday sales for the month of December?

Options:

A- Prescriptive analytics
B- Descriptive analytics
C- Predictive analytics

Answer:

C

Explanation:

Business analytics is commonly organized into descriptive, diagnostic, predictive, and prescriptive techniques, and each answers a different question. Forecasting December holiday sales asks what is likely to happen in the future, which is the defining purpose of predictive analytics. The technique applies statistical models, time series methods, or machine learning to historical sales, seasonalit y, promotions, and economic indicators in order to estimate a future value together with its uncertainty. Descriptive analytics, option B, summarizes what has already occurred, such as last December's revenue by region, and produces reports and dashboards rather than projections. Prescriptive analytics, option A, goes a step further by recommending a course of action, for example how much stock to order or what discount to offer, typically using optimization or simulation built on top of a forecast. The scenario stops at estimating the sales figure itself, so no recommendation is involved. Stakeholders frequently move through the three in sequence: they review past results, forecast the coming period, and then decide on actions. Identifying which stage a described activity belongs to is a routine skill expected of data science practitioners working with business teams. Reference: Data Science Fundamentals Study Guide -- Data Science Process: Types of Analytics (Descriptive, Predictive, Prescriptive).

To Get Premium Files for Data-Science-Fundamentals Visit
https://www.p2pexams.com/products/data-science-fundamentals

For More Free Questions Visit
https://www.p2pexams.com/isaca/pdf/data-science-fundamentals

