



Up-to-date Questions and Answers from authentic resources to improve knowledge and pass the exam at very first attempt. ---- Guaranteed.



C1000-059 Dumps  
C1000-059 Braindumps  
C1000-059 Real Questions  
C1000-059 Practice Test  
C1000-059 Actual Questions



[killexams.com](http://killexams.com)

**IBM**

**C1000-059**

*IBM AI Enterprise Workflow V1 Data Science Specialist*

ORDER FULL VERSION

<https://killexams.com/pass4sure/exam-detail/C1000-059>



## Question: 1

Which of the following is a popular framework used for distributed data processing and machine learning tasks in the IBM AI Enterprise Workflow?

- A. TensorFlow
- B. PyTorch
- C. Apache Spark
- D. Scikit-learn

Answer: C

Explanation: The popular framework used for distributed data processing and machine learning tasks in the IBM AI Enterprise Workflow is Apache Spark. Apache Spark provides a unified analytics engine for big data processing and supports various programming languages, including Python, Scala, and Java. It offers distributed computing capabilities, allowing efficient processing of large-scale datasets and enabling scalable machine learning workflows.

## Question: 2

Which of the following is an unsupervised learning algorithm used for dimensionality reduction?

- A. K-means clustering
- B. Decision tree
- C. Support Vector Machine (SVM)
- D. Principal Component Analysis (PCA)

Answer: D

Explanation: Principal Component Analysis (PCA) is an unsupervised learning algorithm commonly used for dimensionality reduction. PCA transforms a

high-dimensional dataset into a lower-dimensional space while preserving the most important patterns or variations in the data. It achieves this by identifying the principal components, which are linear combinations of the original features that capture the maximum variance in the data. By reducing the dimensionality of the data, PCA can simplify complex datasets, remove noise, and improve computational efficiency in subsequent analyses.

### **Question: 3**

Which of the following techniques is used to address the problem of overfitting in machine learning models?

- A. Regularization
- B. Feature scaling
- C. Cross-validation
- D. Ensemble learning

Answer: A

Explanation: Regularization is a technique used to address the problem of overfitting in machine learning models. Overfitting occurs when a model becomes too complex and starts to capture noise or random fluctuations in the training data, leading to poor generalization to unseen data. Regularization introduces a penalty term to the model's objective function, discouraging overly complex or extreme parameter values. This helps to control the model's complexity and prevent overfitting by finding a balance between fitting the training data well and generalizing to new data.

### **Question: 4**

In the context of data science, what does the term "feature engineering" refer to?

- A. Creating artificial intelligence models
- B. Extracting relevant features from raw data
- C. Developing data visualization techniques
- D. Implementing data cleaning algorithms

Answer: B

Explanation: In data science, "feature engineering" refers to the process of extracting relevant features from raw data. It involves transforming the raw data into a format that is suitable for machine learning algorithms to process and analyze. Feature engineering includes tasks such as selecting important variables, combining or transforming features, handling missing data, and encoding categorical variables. The goal of feature engineering is to enhance the predictive power of machine learning models by providing them with meaningful and informative input features.

### Question: 5

What is the technique called for vectorizing text data which matches the words in different sentences to determine if the sentences are similar?

- A. Cup of Vectors
- B. Box of Lexicon
- C. Sack of Sentences
- D. Bag of Words

Answer: D

Explanation: The correct technique for vectorizing text data to determine sentence similarity is called "Bag of Words" (BoW). In this technique, the text data is represented as a collection or "bag" of individual words, disregarding grammar and word order. Each word is assigned a numerical value, and the presence or absence of words in a sentence is used to create a vector

representation. By comparing the vectors of different sentences, similarity or dissimilarity between them can be measured.

**Question: 6**

Which of the following is a popular algorithm used for natural language processing tasks, such as text classification and sentiment analysis?

- A. Random Forest
- B. K-nearest neighbors (KNN)
- C. Long Short-Term Memory (LSTM)
- D. AdaBoost

Answer: C

Explanation: Long Short-Term Memory (LSTM) is a popular algorithm used for natural language processing (NLP) tasks, particularly text classification and sentiment analysis. LSTM is a type of recurrent neural network (RNN) that can effectively capture long-range dependencies in sequential data, such as sentences or documents. It is well-suited for handling and analyzing text data due to its ability to model context and sequential relationships. LSTM has been widely applied in various NLP applications, including machine translation, speech recognition, and text generation.

**Question: 7**

Which of the following evaluation metrics is commonly used for classification tasks to measure the performance of a machine learning model?

- A. Mean Squared Error (MSE)
- B. R-squared ( $R^2$ )
- C. Precision and Recall

## D. Area Under the Receiver Operating Characteristic Curve (AUC-ROC)

Answer: C

Explanation: Precision and Recall are commonly used evaluation metrics for classification tasks. Precision measures the proportion of correctly predicted positive instances out of all instances predicted as positive, while Recall measures the proportion of correctly predicted positive instances out of all true positive instances. These metrics are particularly useful when dealing with imbalanced datasets or when the cost of false positives and false negatives is different. They provide insights into the model's ability to make accurate positive predictions and identify relevant instances from the dataset.

### Question: 8

Which of the following is an example of a supervised learning algorithm?

- A. K-means clustering
- B. Apriori algorithm
- C. Linear regression
- D. Principal Component Analysis (PCA)

Answer: C

Explanation: Linear regression is an example of a supervised learning algorithm. In supervised learning, the algorithm learns from labeled training data, where each instance is associated with a corresponding target or output value. In the case of linear regression, the algorithm aims to find the best-fitting linear relationship between the input features and the continuous target variable. It learns a set of coefficients that minimize the difference between the predicted values and the actual target values. Once trained, the linear regression model can be used to make predictions on new, unseen data.

### Question: 9

Which of the following is the goal of the backpropagation algorithm in neural networks?

- A. to randomize the trajectory of the neural network parameters during training
- B. to smooth the gradient of the loss function in order to avoid getting trapped in small local minima
- C. to scale the gradient descent step in proportion to the gradient magnitude
- D. to compute the gradient of the loss function with respect to the neural network parameters

Answer: B

Explanation: The goal of the backpropagation algorithm is to smoothly propagate the gradient of the loss function back through the neural network in order to update the network's parameters during training. By smoothing the gradient, it helps to avoid getting trapped in small local minima and allows the optimization process to converge towards a better global minimum.

Randomizing the trajectory, scaling the gradient descent step, or computing the gradient alone are not the specific goals of the backpropagation algorithm.

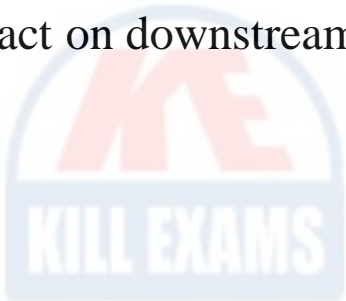
### Question: 10

Which of the following techniques is commonly used for imputing missing values in a dataset?

- A. Random sampling
- B. Median imputation
- C. One-hot encoding
- D. Principal Component Analysis (PCA)

Answer: B

Explanation: Median imputation is a commonly used technique for imputing missing values in a dataset. In this approach, the missing values are replaced with the median value of the corresponding feature. Median imputation is particularly useful for handling missing values in numerical variables, as it preserves the central tendency of the data. By filling in missing values with the median, the overall distribution and statistical properties of the variable are preserved to a certain extent. However, it is important to note that imputation techniques should be chosen carefully, considering the nature of the data and potential impact on downstream analyses.





# SAMPLE QUESTIONS



*These questions are for demo purpose only. **Full version** is up to date and contains actual questions and answers.*

*Killexams.com is an online platform that offers a wide range of services related to certification exam preparation. The platform provides actual questions, exam dumps, and practice tests to help individuals prepare for various certification exams with confidence. Here are some key features and services offered by Killexams.com:*



**Actual Exam Questions:** *Killexams.com provides actual exam questions that are experienced in test centers. These questions are updated regularly to ensure they are up-to-date and relevant to the latest exam syllabus. By studying these actual questions, candidates can familiarize themselves with the content and format of the real exam.*

**Exam Dumps:** *Killexams.com offers exam dumps in PDF format. These dumps contain a comprehensive collection of questions and answers that cover the exam topics. By using these dumps, candidates can enhance their knowledge and improve their chances of success in the certification exam.*

**Practice Tests:** *Killexams.com provides practice tests through their desktop VCE exam simulator and online test engine. These practice tests simulate the real exam environment and help candidates assess their readiness for the actual exam. The practice tests cover a wide range of questions and enable candidates to identify their strengths and weaknesses.*

**Guaranteed Success:** *Killexams.com offers a success guarantee with their exam dumps. They claim that by using their materials, candidates will pass their exams on the first attempt or they will refund the purchase price. This guarantee provides assurance and confidence to individuals preparing for certification exams.*

**Updated Content:** *Killexams.com regularly updates its question bank and exam dumps to ensure that they are current and reflect the latest changes in the exam syllabus. This helps candidates stay up-to-date with the exam content and increases their chances of success.*

**Technical Support:** *Killexams.com provides free 24x7 technical support to assist candidates with any queries or issues they may encounter while using their services. Their certified experts are available to provide guidance and help candidates throughout their exam preparation journey.*

For More exams visit <https://killexams.com/vendors-exam-list>  
*Kill your exam at First Attempt....Guaranteed!*