

DOMANDE SELEZIONE PUBBLICA PER L'ASSUNZIONE A TEMPO INDETERMINATO DI 100 UNITÀ PER LA TERZA AREA FUNZIONALE, FASCIA RETRIBUTIVA F1, PROFILO PROFESSIONALE FUNZIONARIO INFORMATICO - DS50 - PROVA SCRITTA - BUSTA 2

-
- 1) Nel linguaggio di programmazione "R", quale costrutto può essere usato per memorizzare un insieme di elementi di tipo diverso?
- A Lista.
 - B Vettore.
 - C Non esiste in R un costrutto con queste caratteristiche.
-
- 2) Nel linguaggio di programmazione "R" un data frame viene utilizzato per memorizzare...
- A dati in formato tabellare.
 - B dati non strutturati.
 - C dati esclusivamente numerici.
-
- 3) Con riferimento al linguaggio "R", qual è una sintassi corretta per assegnare ad una variabile x un vettore di 3 elementi interi
- A `x=[1,2,3]`
 - B `int[] x=new int[3];`
 - C `x <- c(1,2,3)`
-
- 4) Qual è l'errore di sintassi contenuto nel seguente frammento di codice Python?
- ```
for x in vettore {
 somma = somma + x
}
```
- A La mancanza di parentesi tonde nel costrutto for.
  - B Non c'è nessun errore.
  - C Le parentesi graffe usate per delimitare un blocco di codice.
- 
- 5) In Python, quale tra i seguenti nomi di variabili NON è corretto?
- A `global`
  - B `_pippo`
  - C `Pippo1`
- 
- 6) In quale tra i linguaggi di programmazione R e Python il seguente frammento di codice è sintatticamente corretto:
- ```
x=5  
print(x)
```
- A In nessuno dei due linguaggi.
 - B Solo in Python.
 - C Sia in R che in Python.
-
- 7) In quali strutture sono archiviati i dati in un database relazionale?
- A Oggetti.
 - B Gerarchie.
 - C Tabelle.
-
- 8) In SQL, considerando l'operatore UNION tra due result set A e B, quale delle seguenti affermazioni è vera?

- A A e B devono avere stesso numero di colonne.
 - B A e B possono avere colonne di tipo diverso in posizione coincidente.
 - C A e B possono avere numero di colonne diverso.
-

9) **Data una tabella IVA avente le colonne codice_fiscale, codice_attivita, volume_affari, quale espressione SQL restituisce il totale del volume d'affari per codice attivita?**

- A `SELECT SUM(volume_affari)
FROM IVA
GROUP BY codice_fiscale, codice_attivita`
 - B `SELECT ADD(volume_affari)
FROM IVA
GROUP BY codice_attivita`
 - C `SELECT SUM(volume_affari)
FROM IVA
GROUP BY codice_attivita`
-

10) **La seguente espressione SQL**

```
SELECT *  
FROM Impiegati RIGHT JOIN Dipartimenti  
ON Impiegati.ID_dipartimento = Dipartimenti.ID_dipartimento
```

restituisce...

- A solo le righe corrispondenti a dipartimenti a cui sono assegnati degli impiegati.
 - B anche le righe corrispondenti a dipartimenti senza impiegati.
 - C le colonne più a destra della tabella Dipartimenti.
-

11) **Quale dei seguenti comandi deve essere usato, in SQL, per cancellare una tabella (T) da un database?**

- A `delete table T;`
 - B `drop table T;`
 - C `drop relation T;`
-

12) In SQL, con riferimento alla tabella mostrata nella figura, che effetto avrà questa espressione?

```
START TRANSACTION;  
INSERT INTO Indirizzo  
VALUES (  
    'Parodi',  
    'Franco',  
    'Tevere',  
    '22',  
    'Firenze',  
);  
COMMIT;
```

Indirizzo				
Cognome	Nome	Via	Civico	Citta
Rossi	Paolo	Adige	33	Genova
Verdi	Antonio	Po	55	Torino
Bianchi	Luigi	Piave	44	Roma
Viola	Dario	Isonzo	66	Roma

- A La sostituzione dei valori contenuti nell'ultima riga con i valori descritti nell'istruzione.
- B L'inserimento di una riga con i valori descritti nell'istruzione.
- C Nessuno, l'ultima istruzione (COMMIT) annulla le precedenti.

13) In un database relazionale, come viene rappresentato un attributo?

- A Con una relazione.
- B Con una colonna.
- C Con una riga.

14) Nell'ambito dell'inferenza statistica come si chiamano le due ipotesi messe a test?

- A Ipotesi primaria e ipotesi secondaria.
- B Ipotesi esatta e ipotesi errata.
- C Ipotesi nulla e ipotesi alternativa.

15) Nell'ambito dell'inferenza statistica quali sono le specie di errore che si possono commettere nella verifica delle ipotesi?

- A Di prima e seconda specie.
- B Di misura e di copertura.
- C Campionario e non campionario.

16) Nell'ambito dell'inferenza statistica quale delle seguenti NON è una proprietà degli stimatori?

- A Consistenza.
 - B Nullità.
 - C Efficienza.
-

17) In statistica quale dei seguenti NON è un metodo di verifica delle ipotesi?

- A Metodo del chi-quadro.
 - B Test t di Student.
 - C Metodo di Gauss.
-

18) Quale caratteristica di una distribuzione denota la seguente espressione:

$$\mathbb{E} \left[(X - \mathbb{E}[X]) (Y - \mathbb{E}[Y]) \right].$$

- A Indice di simmetria.
 - B Media.
 - C Covarianza.
-

19) Quali dei seguenti metodi nell'ambito del machine learning necessita di dati etichettati?

- A Machine learning supervisionato.
 - B Cluster analysis.
 - C Machine learning non supervisionato.
-

20) Quale dei seguenti è un aspetto di criticità della cluster analysis?

- A Aumenta il rumore dei dati.
 - B I dati su cui opera necessitano di etichettamento.
 - C La possibile difficoltà di interpretazione dei cluster individuati.
-

21) Nell'ambito dell'intelligenza artificiale, quale delle seguenti è una tecnica di apprendimento?

- A Apprendimento induttivo.
 - B Apprendimento per analisi.
 - C Apprendimento con rinforzo.
-

22) Quale tipo di machine learning viene solitamente utilizzato per effettuare il clustering dei dati?

- A Apprendimento con rinforzo.
 - B Apprendimento non supervisionato.
 - C Apprendimento supervisionato.
-

23) Quale tipo di machine learning viene solitamente utilizzato per effettuare la riduzione dimensionale dei dati?

- A Apprendimento supervisionato.
 - B Apprendimento con rinforzo.
 - C Apprendimento non supervisionato.
-

24) Quale dei seguenti è un tipo di modello di apprendimento supervisionato?

- A Compensazione.
- B Etichettamento.
- C Regressione.

-
- 25) **Nel metodo di regressione lineare...**
- A si stimano i coefficienti delle variabili dipendenti che forniscono la migliore approssimazione della relazione lineare con la variabile indipendente.
 - B si stimano i coefficienti delle variabili indipendenti che forniscono la migliore approssimazione della relazione lineare con la variabile dipendente.
 - C si stimano i coefficienti delle variabili indipendenti e dipendenti che massimizzano una relazione lineare.
-
- 26) **Quale informazione sulla bontà di un modello di regressione lineare è fornita dalla somma dei quadrati dei residui?**
- A Minore è questo valore, minore è l'errore che il modello di predizione commette.
 - B Maggiore è questo valore, minore è l'errore che il modello di predizione commette.
 - C Minore è questo valore, maggiore è l'errore che il modello di predizione commette.
-
- 27) **Il bootstrap è...**
- A un metodo di individuazione degli outlier.
 - B un metodo di ricampionamento.
 - C un algoritmo per la cluster analysis.
-
- 28) **Nell'ambito dell'analisi dei dati cosa si intende per outlier?**
- A Un dato sentinella per verificare se il modello di stima funziona correttamente.
 - B Un dato chiave del set di dati da cui far partire l'algoritmo di stima.
 - C Un dato anomalo e distante dai dati presenti nel dataset.
-
- 29) **L'accuratezza di un modello predittivo dovrebbe essere valutata...**
- A sul test set.
 - B sull'insieme di tutti i dati a disposizione.
 - C sul training set.
-
- 30) **L'algoritmo K-Means si pone come obiettivo...**
- A minimizzare la varianza tra i diversi cluster.
 - B minimizzare la varianza all'interno di uno stesso cluster.
 - C minimizzare il numero di cluster.
-
- 31) **Per cosa viene utilizzato l'algoritmo di backpropagation nell'ambito delle reti neurali?**
- A A calcolare i pesi di una rete neurale a partire dai dati.
 - B A disegnare la struttura della rete neurale.
 - C A calcolare a ritroso gli input a partire dagli output.
-
- 32) **Quale di queste operazioni è tipicamente parte di un processo di analisi del linguaggio naturale?**
- A Il conteggio della punteggiatura.
 - B La compressione dei file di testo da analizzare.
 - C Il conteggio delle occorrenze delle parole presenti nel testo.
-
- 33) **Quale tra i seguenti concetti non è legato alla teoria dei grafi?**
- A Apice.
 - B Collegamento.
 - C Nodo.
-
- 34) **Nell'ambito del text mining, le stopwords indicano...**
- A dei termini che separano diverse frasi.
 - B dei termini che non devono essere considerati nell'analisi perché non rilevanti.
 - C dei termini che fanno interrompere l'analisi del testo.

35) **Quale delle seguenti NON è una tipica applicazione del text-mining?**

- A Supporto clienti.
 - B Predizione dell'andamento delle vendite
 - C Monitoraggio dei social media.
-

36) **In un grafo, il grado di un nodo è definito come...**

- A il numero di connessioni con altri nodi.
 - B il numero di attributi connessi al nodo.
 - C la distanza media da tutti gli altri nodi.
-

37) **Per determinare il livello di qualità dei dati in un'organizzazione...**

- A è sufficiente fornire delle valutazioni qualitative.
 - B è necessario definire delle metriche.
 - C esiste un insieme standard di metriche adatto a qualsiasi contesto.
-

38) **Nell'ambito della qualità dei dati, la dimensione di accuratezza di un dataset indica...**

- A il livello in cui i dati descrivono correttamente il corrispondente oggetto del mondo reale a cui si riferiscono.
 - B la percentuale di dati valorizzati rispetto alla popolazione di riferimento.
 - C la precisione dei formati numerici.
-

39) **Quali delle seguenti caratteristiche attiene alla qualità dei dati?**

- A Completezza.
 - B Distorsione.
 - C Variabilità.
-

40) **In un dataset anagrafico, la presenza nell'attributo "nome di persona" del valore "Giuseppe" incide sulla dimensione di qualità...**

- A completezza.
- B accuratezza.
- C consistenza.