

<b>Iniziato</b>	martedì, 17 dicembre 2024, 12:11
<b>Stato</b>	Completato
<b>Terminato</b>	martedì, 17 dicembre 2024, 12:52
<b>Tempo impiegato</b>	41 min. 18 secondi
<b>Valutazione</b>	27,00 su un massimo di 30,00 (90%)

**Domanda 1**

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Selezionare la sentenza ERRONEA riguardo agli alberi di decisione

- a. Il costo computazionale della predizione è molto basso
- b. Possono esprimere qualunque funzione di classificazione
- c. Possono essere utilizzati solo con features discrete ✓
- d. Hanno una forte tendenza all'overfitting

Risposta corretta.

La risposta corretta è: Possono essere utilizzati solo con features discrete

**Domanda 2**

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Selezionare la sentenza ERRATA riguardo all'apprendimento supervisionato

- a. Può comprendere sia problemi di regressione che di classificazione
- b. La definizione della ground truth può richiedere l'intervento umano ed essere onerosa
- c. Richiede la costante supervisione di un esperto durante il training ✓
- d. Si riferisce all'apprendimento di funzioni basato su esempi di training composti da coppie input-output

Risposta corretta.

La risposta corretta è:

Richiede la costante supervisione di un esperto durante il training

**Domanda 3**

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Una variabile aleatoria discreta con valori  $a, b$  e  $c$  has la seguente distribuzione di probabilità:  $P(a)=1/4$ ,  $P(b)=1/2$ ,  $P(c)=1/4$ .

Qual'è la sua entropia?

- a.  $\log(3)$
- b.  $4/5$
- c.  $5/4$
- d.  $3/2$  ✓

Risposta corretta.

La risposta corretta è:

$3/2$

**Domanda 4**

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Se un modello calcola una distribuzione di probabilità, aggiungere alla funzione obiettivo una componente tesa a diminuire l'entropia avrà l'effetto di:

- a. focalizzare le scelte sui casi più probabili ✓
- b. favorire l'uscita da minimi locali
- c. ridistribuire le probabilità in modo più bilanciato tra tutti i casi
- d. nessun effetto concreto

Risposta corretta.

La risposta corretta è:

focalizzare le scelte sui casi più probabili

**Domanda 5**

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Selezionare la sentenza corretta relativa alla distribuzione congiunta di probabilità

- a. Non permette di fare nessun tipo di predizione
- b. Non consente una visione distinta delle singole features
- c. Il suo calcolo presenta problemi di scalabilità all'aumentare delle features ✓
- d. Non permette il calcolo di eventi condizionali

Risposta corretta.

La risposta corretta è:

Il suo calcolo presenta problemi di scalabilità all'aumentare delle features

**Domanda 6**

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Qual è lo scopo del metodo di cross-validation K-Fold?

- a. Aumentare la dimensione del training set.
- b. Selezionare automaticamente il miglior modello.
- c. Ridurre il numero di dati necessari per l'addestramento.
- d. Valutare il modello su più set di test, migliorando la generalizzazione. ✓

Risposta corretta.

La risposta corretta è:

Valutare il modello su più set di test, migliorando la generalizzazione.

**Domanda 7**

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Quale è l'obiettivo principale dell'algoritmo di clustering K-means?

- a. minimizzare la distanza tra punti in cluster differenti
- b. trovare i K clusters che minimizzano la somma delle distanze quadratiche tra i punti del cluster e il rispettivo centroide. ✓
- c. massimizzare la distanza tra i punti all'interno di uno stesso cluster
- d. massimizzare il numero di clusters del dataset

Risposta corretta.

La risposta corretta è: trovare i K clusters che minimizzano la somma delle distanze quadratiche tra i punti del cluster e il rispettivo centroide.

**Domanda 8**

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Ci sono due dadi, uno normale e uno truccato che restituisce un 6 con probabilità .5, e gli altri valori con probabilità 0.1

Faccio tre lanci con lo stesso dado e osservo un 3, un 6 e un 2. Cosa posso concludere?

- a. La probabilità di usare uno o l'altro dei dadi è esattamente la stessa
- b. E' più probabile che il dato sia truccato ✓
- c. E' più probabile che il dato sia normale
- d. Nulla

Risposta corretta.

La risposta corretta è:

E' più probabile che il dato sia truccato

**Domanda 9**

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Perchè la tecnica Naive Bayes è detta "naive" (ingenua)?

- a. Perchè fornisce un modo semplice ma preciso di calcolare la distribuzione congiunta di probabilità delle features
- b. Perchè suppone ingenuamente che la teoria possa avere applicazioni pratiche
- c. Perchè suppone ingenuamente che i dati di training rispecchino i dati reali
- d. Perchè suppone ingenuamente che le features siano indipendenti tra loro, date le classi ✓

Risposta corretta.

La risposta corretta è:

Perchè suppone ingenuamente che le features siano indipendenti tra loro, date le classi

**Domanda 10**

Risposta errata

Punteggio ottenuto 0,00 su 1,00

Quale distribuzione di probabilità si assume abitualmente per una variabile obiettivo continua nel caso della tecnica Naive Bayes?

- a. Una distribuzione Normale, ovvero Gaussiana(0,1)
- b. Una distribuzione binomiale
- c. Una distribuzione esponenziale
- d. Una distribuzione uniforme in [0,1] ✗

Risposta errata.

La risposta corretta è:

Una distribuzione Normale, ovvero Gaussiana(0,1)

**Domanda 11**

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Qual è il principale scopo della normalizzazione dei dati prima dell'addestramento di un modello di machine learning?

- a. Garantire che tutte le feature abbiano la stessa scala, migliorando la convergenza dell'ottimizzazione. ✓
- b. Evitare che i dati contengano valori mancanti.
- c. Ridurre il numero di parametri del modello.
- d. Aumentare la precisione delle previsioni del modello.

Risposta corretta.

La risposta corretta è:

Garantire che tutte le feature abbiano la stessa scala, migliorando la convergenza dell'ottimizzazione.

**Domanda 12**

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Quale delle seguenti situazioni NON è particolarmente problematica dal punto di vista dell'overfitting?

- a. Effettuare un training molto prolungato
- b. Disporre di un modello molto espressivo
- c. Avere pochi dati di training
- d. Avere dati molto rumorosi ✓

Risposta corretta.

La risposta corretta è:

Avere dati molto rumorosi

**Domanda 13**

Risposta errata

Punteggio ottenuto 0,00 su 1,00

Selezionare la sentenza ERRATA relativa alla Principal Component Analysis (PCA)

- a. E' una tecnica per selezionare le features migliori tra quelle date, senza modificarle.
- b. E' una tecnica per ridurre il numero delle features ✖
- c. Vengono create nuove features come combinazioni lineari delle features esistenti
- d. I dati sono convertiti in un nuovo sistema di coordinate in modo da preservare la varianza massima tra di essi

Risposta errata.

La risposta corretta è:

E' una tecnica per selezionare le features migliori tra quelle date, senza modificarle.

**Domanda 14**

Risposta errata

Punteggio ottenuto 0,00 su 1,00

Un dataset contiene  $\frac{2}{3}$  di positivi e  $\frac{1}{3}$  di negativi. La precisione del modello è  $\frac{9}{10}$ . Che percentuale dei dati totali sono Falsi Positivi?

- a.  $\frac{1}{9}$
- b. Non può essere stabilito
- c.  $\frac{2}{27}$  ✖
- d.  $\frac{1}{10}$

Risposta errata.

La risposta corretta è:

Non può essere stabilito

**Domanda 15**

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Selezionare la sentenza corretta relativa alla probabilità condizionata  $P(A|B)$  tra due eventi A e B

- a.  $P(A|B)$  è sicuramente maggiore o uguale di  $P(A)$
- b.  $P(A|B)$  è sicuramente minore o uguale di  $P(A \text{ and } B)$
- c.  $P(A|B)$  è sicuramente maggiore o uguale di  $P(A \text{ and } B)$  ✓
- d.  $P(A|B)$  è sicuramente minore o uguale a  $P(A)$

Risposta corretta.

La risposta corretta è:

$P(A|B)$  è sicuramente maggiore o uguale di  $P(A \text{ and } B)$

**Domanda 16**

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Selezionare la sentenza **SCORRETTA** riguardo alla regressione logistica

- a. Permette di associare una probabilità alla predizione della classe
- b. La predizione dipende dal bilanciamento dei dati di training rispetto alle classi
- c. I parametri del modello possono essere tipicamente calcolati in forma chiusa, mediante una formula esplicita ✓
- d. Il calcolo della predizione si basa sulla loglikelihood dei dati di training

Risposta corretta.

La risposta corretta è: I parametri del modello possono essere tipicamente calcolati in forma chiusa, mediante una formula esplicita

**Domanda 17**

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Qual è il ruolo principale della funzione di loss in una rete neurale?

- a. Fornire un criterio quantitativo per valutare le prestazioni del modello.
- b. a) Misurare la discrepanza tra le previsioni del modello e i valori reali, guidando l'ottimizzazione dei parametri. ✓
- c. Regolare il tasso di apprendimento durante l'addestramento.
- d. Penalizzare modelli troppo complessi per migliorare la generalizzazione.

Risposta corretta.

La risposta corretta è:

a) Misurare la discrepanza tra le previsioni del modello e i valori reali, guidando l'ottimizzazione dei parametri.

**Domanda 18**

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Selezionare la sentenza SCORRETTA riguardo ai modelli generativi

- a. Generative Adversarial Networks, Variational Autoencoders e Diffusion models sono esempi di tecniche generative profonde
- b. Sono modelli che cercano di apprendere la distribuzione di probabilità dei dati
- c. Sono modelli meta-teorici rivolti alla automatizzazione della generazione di reti neurali ✓
- d. Un tipico esempio di tecnica generativa è Naive Bayes

Risposta corretta.

La risposta corretta è:

Sono modelli meta-teorici rivolti alla automatizzazione della generazione di reti neurali

**Domanda 19**

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Selezionare la sentenza corretta

- a. Un neurone artificiale può apprendere solo funzioni lineari
- b. Un neurone artificiale tipicamente calcola una combinazione lineare dei suoi input, seguita dalla applicazione di una funzione di attivazione non lineare ✓
- c. Il numero dei parametri di un neurone artificiale è quadratico nella dimensione dei suoi inputs
- d. Un neurone artificiale può apprendere qualunque funzione dei suoi input

Risposta corretta.

La risposta corretta è:

Un neurone artificiale tipicamente calcola una combinazione lineare dei suoi input, seguita dalla applicazione di una funzione di attivazione non lineare

**Domanda 20**

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Quale è lo scopo dell'optimizer in Tensorflow/Keras?

- a. aggiungere una penalità ai pesi del layer su cui viene istanziato
- b. salvare i migliori pesi del modello durante il processo di training
- c. contrastare l'overfitting
- d. definire l'algoritmo che calcola i gradienti della loss e aggiorna i pesi del modello ✓

Risposta corretta.

La risposta corretta è:

definire l'algoritmo che calcola i gradienti della loss e aggiorna i pesi del modello

**Domanda 21**

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Quale delle seguenti tecniche NON può essere utilizzata per contrastare l'overfitting?

- a. aggiunta di una loss di regolarizzazione
- b. early stopping
- c. aggiunta di skip connections ✓
- d. data augmentation

Risposta corretta.

La risposta corretta è:

aggiunta di skip connections

**Domanda 22**

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Selezionare la sentenza SCORRETTA relativa al problema della scomparsa del gradiente (vanishing gradient)

- a. Il problema è mitigato dall'uso di link residuali all'interno della rete
- b. Se il gradiente tende a zero i parametri non sono più aggiornati e la rete smette di apprendere
- c. Il problema è fortemente attenuato dall'uso di ReLU (o sue varianti) come funzione di attivazione per i livelli nascosti della rete
- d. Se il gradiente tende a zero anche i parametri e le attivazioni dei neuroni tendono a zero ✓

La risposta corretta è: Se il gradiente tende a zero anche i parametri e le attivazioni dei neuroni tendono a zero

**Domanda 23**

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Selezione al sentenza ERRATA relativa alle Transposed Convolutions

- a. Richiedono la trasposizione dell'input prima di calcolare la convoluzione del Kernel ✓
- b. Possono essere interpretate come convoluzioni normali con stride sub-unitario
- c. Sono prevalentemente utilizzate in architetture per Image-to-Image processing, come autoencoders o U-Nets
- d. Sono essenzialmente equivalenti alla applicazione di un livello di upsampling seguito da una convoluzione normale

La risposta corretta è: Richiedono la trasposizione dell'input prima di calcolare la convoluzione del Kernel

**Domanda 24**

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Selezionare la sentenza SCORRETTA relativa alla U-Net

- a. E' un componente tipico dei modelli generativi a diffusione
- b. Può essere usata per la rimozione del rumore (denoising) di immagini
- c. Viene spesso utilizzata nell'ambito della classificazione dei generi musicali ✓
- d. E' spesso impiegata per problemi di segmentazione semantica di immagini

La risposta corretta è: Viene spesso utilizzata nell'ambito della classificazione dei generi musicali

**Domanda 25**

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Quale delle seguenti sentenze relative agli autoencoders è SCORRETTA?

- a. La rappresentazione interna prodotta dall'encoder abitualmente ha una dimensione ridotta rispetto a quella di partenza
- b. Gli Autoencoders richiedono l'uso di livelli densi ✓
- c. Possono essere utilizzate per la rimozione di rumore (denoising)
- d. L'encoder e il decoder non devono essere necessariamente simmetrici

La risposta corretta è: Gli Autoencoders richiedono l'uso di livelli densi

**Domanda 26**

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Il tensore di input di un layer convolutivo 2D ha dimensione (16,16,32). Sintetizzo 8 kernel con dimensione spaziale (3,3), stride 2, nessun padding (valid mode). Quale sarà la dimensione dell'output?

- a. (7,7,15)
- b. (8,8,32)
- c. (8,8,8)
- d. (7,7,8) ✓

Risposta corretta.

La risposta corretta è:

(7,7,8)

**Domanda 27**

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Quale delle seguenti affermazioni relative alla backpropagation è corretta?

- a. Si effettua solo durante il "training" ✓
- b. E' molto più costosa, in termini di tempo, del calcolo "in avanti" (inference) lungo la rete
- c. Viene fatta sia durante la fase di "inference" (calcolo in avanti) che in quella di "training"
- d. Viene effettuate unicamente lungo le skip connections delle reti residuali, per evitare perdita del gradiente

La risposta corretta è: Si effettua solo durante il "training"

**Domanda 28**

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Quale è l'effetto tipico della riduzione della dimensione del minibatch durante il training?

- a. La Backpropagation è effettuata meno frequentemente ma l'aggiornamento dei parametri è più accurato
- b. La Backpropagation è effettuata più frequentemente e l'aggiornamento dei parametri è più accurato
- c. La Backpropagation è effettuata più frequentemente ma l'aggiornamento dei parametri è meno accurato ✓
- d. La Backpropagation è effettuata meno frequentemente e l'aggiornamento dei parametri è meno accurato

La risposta corretta è: La Backpropagation è effettuata più frequentemente ma l'aggiornamento dei parametri è meno accurato

**Domanda 29**

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Selezionare la sentenza corretta:

- a. Una GAN è una rete che permette di generare attacchi per un qualunque modello predittivo
- b. Le GAN basano il loro training su una funzione di loglikelihood relativa ai dati generati
- c. Le GAN hanno una struttura encoder-decoder, simile a quella di un autoencoder.
- d. le GAN possono soffrire del fenomeno di "mode collapse", cioè la tendenza a focalizzare la generazione su un unico o pochi esempi ✓

La risposta corretta è: le GAN possono soffrire del fenomeno di "mode collapse", cioè la tendenza a focalizzare la generazione su un unico o pochi esempi

**Domanda 30**

Risposta corretta

Punteggio ottenuto 1,00 su 1,00

Selezionare la sentenza SCORRETTA relativa alla tecnica a discesa del gradiente.

- a. Può essere applicata solo se la funzione da minimizzare ha una superficie concava ✓
- b. Il risultato può dipendere dalla inizializzazione dei parametri del modello
- c. E' opportuno decrementare il learning rate verso la fine dell'apprendimento
- d. Potrebbe convergere a un minimo locale

La risposta corretta è: Può essere applicata solo se la funzione da minimizzare ha una superficie concava

[◀ Quiz 16 dicembre 2024](#)[Ricostruzione della temperatura marina superficiale da occlusione per nuvole ▶](#)