

[DASHBOARD](#) / [I MIEI CORSI](#) / [INTRODUZIONE ALL'APPRENDIMENTO AUTOMATICO](#) / [SEZIONI](#) / [QUIZ FEBBRAIO 2023](#) / [QUIZ FEBBRAIO 2023](#)

Iniziato lunedì, 13 febbraio 2023, 14:11

Stato Completato

Terminato lunedì, 13 febbraio 2023, 15:22

Tempo impiegato 1 ora 11 min.

Punteggio 15,00/24,00

Valutazione **6,25** su un massimo di 10,00 (**63%**)



Domanda **1**Risposta
correttaPunteggio
ottenuto 1,00 su
1,00

Nel caso di un albero di decisione con features discrete, cosa si può dire della profondità dell'albero?

- a. E' minore o uguale al numero delle features
- b. Non si può dire nulla
- c. E' minore o uguale del numero delle classi
- d. E' sicuramente maggiore del logaritmo in base due del numero dei dati



Risposta corretta.

La risposta corretta è:

E' minore o uguale al numero delle features



Domanda **2**

Risposta errata

Punteggio
ottenuto 0,00 su
1,00

Selezionare la sentenza errata relativa alle Random Forest (foreste di alberi decisionali)

- a. Richiedono tecniche opportune per la creazione di alberi di decisione diversi relativi a uno stesso dataset ✘
- b. Tendono a migliorare l'explainability (spiegabilità) degli alberi di decisione riducendo l'instabilità nella selezione degli attributi
- c. Tentano di mitigare il fenomeno dell'overfitting tipico degli alberi di decisione
- d. E' una tecnica di apprendimento ad "ensemble" basata su di una combinazione di alberi di decisione

Risposta errata.

La risposta corretta è:

Tendono a migliorare l'explainability (spiegabilità) degli alberi di decisione riducendo l'instabilità nella selezione degli attributi



Domanda **3**

Risposta errata

Punteggio
ottenuto 0,00 su
1,00

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

Qual'è la sua entropia?

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



Risposta errata.

La risposta corretta è:

$3/2$



Domanda **4**Risposta
correttaPunteggio
ottenuto 1,00 su
1,00

In che situazioni si parla di apprendimento auto-supervisionato (self-supervised)

- a. Quando il modello è in grado di riconfigurare in modo automatico la propria architettura
- b. Quando il modello è supposto contribuire alla creazione di nuovi dati di training
- c. Qualora i dati di input possano essere considerati come annotazioni (labels) per guidare l'apprendimento, come nel caso degli autoencoders
- d. Quando l'apprendimento perde una sinergia tra l'uomo e la macchina



Risposta corretta.

La risposta corretta è:

Qualora i dati di input possano essere considerati come annotazioni (labels) per guidare l'apprendimento, come nel caso degli autoencoders



Domanda **5**

Risposta errata

Punteggio
ottenuto 0,00 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 $P(A \text{ and } B)$
- b. $P(A|B)$ è sicuramente maggiore o uguale di $P(A)$
- c. $P(A|B)$ è sicuramente minore o uguale di $P(A \text{ and } B)$
- d. $P(A|B)$ è sicuramente minore o uguale a $P(A)$



Risposta errata.

La risposta corretta è:

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



Domanda **6**Risposta
correttaPunteggio
ottenuto 1,00 su
1,00

Selezionare la sentenza ERRATA relativa alla distribuzione congiunta di probabilità per N variabili aleatorie discrete

- a. E' la distribuzione di probabilità di tutte le possibili tuple di valori per le variabili
- b. Richiede il calcolo di un numero esponenziale di parametri
- c. Non permette il calcolo di probabilità condizionali tra le features
- d. Consente il calcolo delle probabilità marginali delle singole features



Risposta corretta.

Le risposte corrette sono:

Richiede il calcolo di un numero esponenziale di parametri,

Non permette il calcolo di probabilità condizionali tra le features



Domanda **7**

Risposta errata

Punteggio
ottenuto 0,00 su
1,00

Avendo 5 categorie di dati e 3 features di input booleane, quanti parametri indipendenti devono essere stimati secondo la tecnica Naive Bayes (compresi i priors).

- a. 15
- b. 16
- c. 19
- d. 20



Risposta errata.

La risposta corretta è:
19



Domanda **8**

Risposta errata

Punteggio
ottenuto 0,00 su
1,00

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

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



Risposta errata.

La risposta corretta è:

Avere dati molto rumorosi



Domanda **9**Risposta
correttaPunteggio
ottenuto 1,00 su
1,00

Selezionare la sentenza ERRATA relativa al learning rate

- a. E' una metrica che misura la capacità di apprendimento del modello
- b. Un learning rate alto tipicamente velocizza il training ma potrebbe saltare sopra al minimo
- c. E' un iper-parametro che definisce la lunghezza del passo durante la discesa del gradiente
- d. Il learning rate può variare durante il training



Risposta corretta.

La risposta corretta è:

E' una metrica che misura la capacità di apprendimento del modello



Domanda **10**Risposta
correttaPunteggio
ottenuto 1,00 su
1,00

Ci sono due monete, una normale e una che restituisce testa con probabilità $3/4$ e croce con probabilità $1/4$.

Faccio due lanci con la stessa moneta e osservo una testa e una croce. Che cosa posso concludere?

- a. Nulla
- b. E' più probabile che la moneta sia normale
- c. E' più probabile che la moneta sia truccata
- d. La probabilità di usare uno o l'altra moneta è esattamente la stessa



Risposta corretta.

La risposta corretta è:

E' più probabile che la moneta sia normale



Domanda **11**Risposta
correttaPunteggio
ottenuto 1,00 su
1,00

Selezionare la sentenza corretta riguardo alla regressione logistica

- a. I parametri del modello sono tipicamente calcolati mediante discesa del gradiente ✔
- b. La predizione non 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 non si basa sulla loglikelihood dei dati di training, in quanto si tratta di una tecnica discriminativa

Risposta corretta.

La risposta corretta è: I parametri del modello sono tipicamente calcolati mediante discesa del gradiente



Domanda **12**Risposta
correttaPunteggio
ottenuto 1,00 su
1,00La derivata della funzione logistica $\sigma(x)$ è:

- a. $\sigma(x)/\sigma(1-x)$
- b. $\sigma(x) \cdot (1 - \sigma(x))$
- c. $\sigma(x)/(1 - \sigma(x))$
- d. $\sigma(x) \cdot \sigma(1-x)$



Risposta corretta.

La risposta corretta è:

$$\sigma(x) \cdot (1 - \sigma(x))$$



Domanda **13**

Risposta errata

Punteggio
ottenuto 0,00 su
1,00

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

- a. $1/9$
- b. Non può essere stabilito
- c. $2/27$
- d. $1/10$



Risposta errata.

La risposta corretta è:
Non può essere stabilito



Domanda **14**

Risposta errata

Punteggio
ottenuto 0,00 su
1,00

Siano date le seguenti distribuzioni di probabilità P e Q: $P(0)=3/8$, $P(1)=1/2$, $P(2)=1/8$, and $Q(0)=1/2$, $Q(1)=1/4$, $Q(2)=1/4$.

Quanto vale la crossentropy $H(P|Q)$ tra P e Q?

- a. $13/8$
- b. $3/2 + \log(3)$
- c. $5/2 - \log(3)/2$
- d. 2



Risposta errata.

La risposta corretta è:

$13/8$



Domanda **15**Risposta
correttaPunteggio
ottenuto 1,00 su
1,00

Selezionare la sentenza ERRATA relativa alla regressione lineare

- a. Cerca di determinare un iperpiano di separazione tra due categorie di dati ✔
- b. Il problema di ottimizzazione ammette una soluzione in forma chiusa
- c. La funzione di loss è tipicamente una distanza quadratiche tra i valori predetti e quelli osservati
- d. Cerca di stabilire una relazione tra i valori di una variabile di output e i valori di una o più features di input

Risposta corretta.

La risposta corretta è:

Cerca di determinare un iperpiano di separazione tra due categorie di dati



Domanda **16**Risposta
correttaPunteggio
ottenuto 1,00 su
1,00

Quale delle seguenti non è una applicazione tipica degli autoencoders

- a. Segmentazione di immagini (semantic segmentation)
- b. Rimozione del rumore (denoising)
- c. Riduzione delle dimensioni (dimensionality reduction)
- d. Rilevamento di anomalie (anomaly detection)




La risposta corretta è: Segmentazione di immagini (semantic segmentation)



Domanda **17**Risposta
correttaPunteggio
ottenuto 1,00 su
1,00

Selezionare la sentenza SCORRETTA relativa alla backpropagation per reti neurali

- a. E' l' algoritmo per il calcolo della derivata parziale della loss rispetto a ogni parametro della rete
- b. Tipicamente, il gradiente viene artificialmente rinforzato ad ogni layer attraversato per contrastare il fenomeno della sua scomparsa (vanishing) 
- c. Si riduce a semplici calcoli algebrici facilmente parallelizzabili in strutture di calcolo tipo GPU
- d. L'algoritmo calcola il gradiente un layer alla volta, sfruttando la regola matematica per la derivazione di funzioni composte

La risposta corretta è: Tipicamente, il gradiente viene artificialmente rinforzato ad ogni layer attraversato per contrastare il fenomeno della sua scomparsa (vanishing)

[◀ Polar Express](#)[Progetto Febbraio 2023: Inpainting ▶](#)

Domanda **18**Risposta
correttaPunteggio
ottenuto 1,00 su
1,00

Componendo due layer Conv2D con stride 1, il primo con kernel 5x5 e il secondo con kernel 3x3 quale sarà il campo ricettivo dei neuroni finali?

- a. 7
- b. 8
- c. dipende dal padding
- d. 3



La risposta corretta è: 7



Domanda **19**Risposta
correttaPunteggio
ottenuto 1,00 su
1,00

Il tensore di input di un layer convolutivo 2D ha dimensione (16,16,8). Sintetizzo 4 kernel con dimensione spaziale (5,5), stride 2, nessun padding (valid mode). Quale sarà la dimensione dell'output?

- a. (6,6,4)
- b. (8,8,8)
- c. (7,7,4)
- d. (7,7,8)



Risposta corretta.

La risposta corretta è:
(6,6,4)

Domanda **20**Risposta
correttaPunteggio
ottenuto 1,00 su
1,00

Selezionare la sentenza errata relativa all' "inception module"

- a. Sfutta kernel di dimensione diversa
- b. Tende a ridurre il costo computazionale sfruttando convoluzioni unarie per diminuire il numero dei canali
- c. Utilizza al proprio interno delle skip-connections per bypassare l'applicazione di parte dei kernel ✓
- d. E' un componente tipico della rete Inception-v3

La risposta corretta è: Utilizza al proprio interno delle skip-connections per bypassare l'applicazione di parte dei kernel



Domanda **21**

Risposta errata

Punteggio
ottenuto 0,00 su
1,00

Quale delle seguenti tecniche non può aiutare ad uscire da minimi locali durante la fase di training.

- a. Ridurre la dimensione del minibatch ✘
- b. Fare clipping del gradiente in una range prefissato
- c. Aumentare il learning rate
- d. Aggiungere un "momento" al gradiente, cioè parte del gradiente del passo precedente

Risposta errata.

La risposta corretta è:

Fare clipping del gradiente in una range prefissato



Domanda **22**Risposta
correttaPunteggio
ottenuto 1,00 su
1,00

Un layer convolutivo 2D con stride 1, kernel size 1x1, e senza padding prende in input un layer con dimensioni (32,32,16) e restituisce un layer di dimensione (32,32,4). Quanti sono i suoi parametri?

- a. 2
- b. 68
- c. 8
- d. 64



Risposta corretta.

La risposta corretta è:
68



Domanda **23**

Risposta errata

Punteggio
ottenuto 0,00 su
1,00

Selezionare la sentenza SCORRETTA relativa alla Intersection over Union (IoU)

- a. E' frequentemente utilizzata come misura di similitudine tra bounding boxes
- b. Restituisce un valore nel range [0,1]
- c. Non è una funzione simmetrica dei suoi input
- d. E' una metrica principalmente utilizzata nel campo della Object Detection



Risposta errata.

La risposta corretta è:

Non è una funzione simmetrica dei suoi input



Domanda **24**Risposta
correttaPunteggio
ottenuto 1,00 su
1,00

Selezionare la sentenza errata relativa ai transformes.

- a. Hanno una tipica struttura encoder-decoder, ognuno formato da uno stack di sotto-componenti modulari
- b. Sono alla base delle reti della famiglia BERT e GPT
- c. Utilizzano pesantemente il meccanismo di attenzione
- d. Aggiungono ad ogni livello della rete un encoding posizionale per enfatizzare la posizione realtiva dei tokens ✓

Risposta corretta.

La risposta corretta è:

Aggiungono ad ogni livello della rete un encoding posizionale per enfatizzare la posizione realtiva dei tokens

