

### Domanda 1

Risposta  
corretta

Punteggio  
ottenuto 1,00  
su 1,00



[Contrassegna  
domanda](#)

Selezionare la sentenza ERRONEA riguardo agli alberi di decisione

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

Risposta corretta.

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

### Domanda 2

Risposta  
corretta

Punteggio  
ottenuto 1,00  
su 1,00



[Contrassegna  
domanda](#)

Il range della entropia per la distribuzione di probabilità di una variabile aleatoria discreta è:

- a. tra 0 e 1
- b. tra -1 e 1
- c. tra 0 e  $\log n$  dove  $n$  sono i possibili valori di  $X$  ✓
- d. tra 0 e infinito

Risposta corretta.

La risposta corretta è:  
tra 0 e  $\log n$  dove  $n$  sono i possibili valori di  $X$

**Domanda 3**

Risposta errata

Punteggio  
ottenuto 0,00  
su 1,00[Contrassegna  
domanda](#)

Selezionare la sentenza corretta relativa alla distribuzione congiunta di probabilità

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

Risposta errata.

La risposta corretta è:

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

**Domanda 4**Risposta  
correttaPunteggio  
ottenuto 1,00  
su 1,00[Contrassegna  
domanda](#)

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

- a. focalizzare le scelte sui casi più probabili
- b. nessun effetto concreto
- c. ridistribuire le probabilità in modo più bilanciato tra tutti i casi **✔**
- d. contrastare il fenomeno della perdita di gradiente

Risposta corretta.

La risposta corretta è:

**Domanda 5**

Risposta  
corretta

Punteggio  
ottenuto 1,00  
su 1,00



Contrassegna  
domanda

Un dataset contiene  $\frac{1}{3}$  di positivi e  $\frac{2}{3}$  di negativi. La recall del modello è di  $\frac{2}{3}$ . Che percentuale dei dati sono Falsi Negativi?

- a.  $\frac{1}{3}$
- b.  $\frac{2}{9}$
- c.  $\frac{1}{9}$  ✓
- d. Non può essere stabilito

Risposta corretta.

La risposta corretta è:  
 $\frac{1}{9}$

**Domanda 6**

Risposta  
corretta

Punteggio  
ottenuto 1,00  
su 1,00



Contrassegna  
domanda

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 due lanci con lo stesso dado e osservo un 3 e un 6. Cosa posso concludere?

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

Risposta corretta.

La risposta corretta è:

**Domanda 7**

Risposta  
corretta

Punteggio  
ottenuto 1,00  
su 1,00



[Contrassegna  
domanda](#)

Selezionare la sentenza erronea relativa alla crossentropy  $H(P,Q)$  tra  $P$  e  $Q$

- a. Misura la loglikelihood di  $Q$  data la distribuzione  $P$
- b. Ha un valore minimo quando  $P = Q$
- c. E' una funzione simmetrica:  $H(P,Q) = H(Q,P)$  ✓
- d. E' uguale alla divergenza di Kullback-Leibler  $KL(P,Q)$  più l'entropia  $H(P)$  di  $P$

Risposta corretta.

La risposta corretta è:

E' una funzione simmetrica:  $H(P,Q) = H(Q,P)$

**Domanda 8**

Risposta  
corretta

Punteggio  
ottenuto 1,00  
su 1,00



[Contrassegna  
domanda](#)

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

- a. Può essere applicata solo se la funzione da minimizzare ha una superficie concava
- b. il risultato non dipende dalla inizializzazione dei parametri del modello
- c. Potrebbe convergere a un minimo locale ✓
- d. permette sempre di individuare il minimo globale, se questo esiste

La risposta corretta è: Potrebbe convergere a un minimo locale



**Domanda 9**

Risposta  
corretta

Punteggio  
ottenuto 1,00  
su 1,00



[Contrassegna  
domanda](#)

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

- a. La Backpropagation è effettuata più frequentemente e l'aggiornamento dei parametri è più accurato
- b. La Backpropagation è effettuata meno frequentemente e l'aggiornamento dei parametri è meno accurato
- c. La Backpropagation è effettuata meno frequentemente ma l'aggiornamento dei parametri è più accurato
- d. La Backpropagation è effettuata più frequentemente ma 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 10**

Risposta  
corretta

Punteggio  
ottenuto 1,00  
su 1,00



[Contrassegna  
domanda](#)

Selezionare la sentenza ERRATA riguardo alla regressione logistica

- a. Si basa su una combinazione lineare delle features di input
- b. Non dipende dal bilanciamento dei dati di training rispetto alle classi ✓
- c. Nel caso di classificazione binaria, la superficie di confine tra le classi è un iperpiano
- d. La probabilità della predizione cresce se ci si allontana dalla superficie di confine tra le classi

Risposta corretta.

La risposta corretta è:

**Domanda 11**

Risposta  
corretta

Punteggio  
ottenuto 1,00  
su 1,00



Contrassegna  
domanda

In quale di questi casi la regressione logistica potrebbe essere in difficoltà:

- a. Quando non tutte le features di input sono rilevanti ai fini della classificazione
- b. Quando le features sono indipendenti tra loro, data la classe.
- c. Quando esiste una elevata correlazione tra le features
- d. Quando la classificazione dipende da un confronto tra features ✓

Risposta corretta.

La risposta corretta è:

Quando la classificazione dipende da un confronto tra features

**Domanda 12**

Risposta  
corretta

Punteggio  
ottenuto 1,00  
su 1,00



Contrassegna  
domanda

Riguardo alla regressione multinomiale, selezionare la sentenza ERRATA tra le seguenti

- a. il peso delle features indica la loro importanza ai fini della classificazione
- b. Il peso con cui è valutata ogni feature è tipicamente diverso per ogni classe
- c. Per ogni input, esiste almeno una classe con probabilità  $> 0.5$  ✓
- d. Per  $n$  features di input e  $m$  classi, il numero dei parametri del modello è  $n \times m + m$

Risposta corretta.

La risposta corretta è: Per ogni input, esiste almeno una classe con probabilità  $> 0.5$

### Domanda 13

Risposta  
corretta

Punteggio  
ottenuto 1,00  
su 1,00



[Contrassegna  
domanda](#)

Selezionare la sentenza ERRATA relativa alla funzione softmax

- a. Produce valori compresi nell'intervallo  $[-1,1]$  ✓
- b. Permette di calcolare una distribuzione di probabilità sulle classi
- c. Generalizza la funzione logistica al caso multiclasse
- d. Per un dato input, la somma dei suoi valori su tutte le classi è sempre 1

Risposta corretta.

La risposta corretta è:

Produce valori compresi nell'intervallo  $[-1,1]$

### Domanda 14

Risposta  
corretta

Punteggio  
ottenuto 1,00  
su 1,00



[Contrassegna  
domanda](#)

Quale funzione di loss è tipicamente utilizzata per una rete neurale per classificazione a categorie multiple che utilizza softmax come attivazione finale?

- a. absolute error
- b. binary crossentropy
- c. categorical crossentropy ✓
- d. mean squared error

La risposta corretta è: categorical crossentropy

**Domanda 15**


Risposta  
corretta

Punteggio  
ottenuto 1,00  
su 1,00



[Contrassegna  
domanda](#)

Cosa si intende con tecniche discriminative?

- a. Tecniche che cercano di discriminare i dati in base alle diverse distribuzioni di probabilità delle varie classi
- b. Tecniche di classificazione che si focalizzano sulla definizione delle frontiere di decisione (decision boundaries) 
- c. Tecniche tipiche di unsupervised learning che tentano di separare i dati in clusters distinti
- d. Tecniche che cercano di identificare gli outliers all'interno del data set

Risposta corretta.

La risposta corretta è:

Tecniche di classificazione che si focalizzano sulla definizione delle frontiere di decisione (decision boundaries)

**Domanda 16**


Risposta  
corretta

Punteggio  
ottenuto 1,00  
su 1,00



[Contrassegna  
domanda](#)

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

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

Risposta corretta.



**Domanda 17**

Risposta  
corretta

Punteggio  
ottenuto 1,00  
su 1,00



[Contrassegna  
domanda](#)

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

- a. Ridurre il numero di clusters al minimo
- b. Trovare il punto medio del dataset
- c. Raggruppare i punti di un cluster attorno al loro centroide ✓
- d. ottimizzare il numero dei clusters basandosi sulla distribuzione Gaussiana dei dati

Risposta corretta.

La risposta corretta è:

Raggruppare i punti di un cluster attorno al loro centroide

**Domanda 18**

Risposta  
corretta

Punteggio  
ottenuto 1,00  
su 1,00



[Contrassegna  
domanda](#)

Cosa si intende con "deep" features?

- a. Features ottenute mediante utilizzo di sensori ottici di profondità
- b. Features soggette a una approfondita supervisione da parte umana
- c. Features sintetizzate in modo automatico a partire da altre features ✓
- d. Features relative a dati in 2 o più dimensioni

La risposta corretta è: Features sintetizzate in modo automatico a partire da altre features

### Domanda 19

Risposta  
corretta

Punteggio  
ottenuto 1,00  
su 1,00



[Contrassegna  
domanda](#)

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

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

Risposta corretta.

La risposta corretta è:

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

### Domanda 20

Risposta  
corretta

Punteggio  
ottenuto 1,00  
su 1,00



[Contrassegna  
domanda](#)

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

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

**Domanda 21**

Risposta  
corretta

Punteggio  
ottenuto 1,00  
su 1,00



Contrassegna  
domanda

Il tensore di input di un layer convolutivo 2D ha dimensione (32,32,8). Sintetizzo un unico kernel con dimensione spaziale (4,4), stride 2, nessun padding (valid mode). Quale sarà la dimensione dell'output?

- a. (15,15,8)
- b. (16,16,8)
- c. (15,15,1) ✓
- d. (16,16,1)

Risposta corretta.

La risposta corretta è:  
(15,15,1)

**Domanda 22**

Risposta  
corretta

Punteggio  
ottenuto 1,00  
su 1,00



Contrassegna  
domanda

Qual'è l'effetto di uno stride non unitario ( $>1$ ) in un layer convolutivo?

- a. Nessun effetto spaziale, il numero dei canali decresce
- b. La dimensione spaziale aumenta
- c. La dimensione spaziale diminuisce ✓
- d. Nessun effetto spaziale, il numero dei canali aumenta

La risposta corretta è: La dimensione spaziale diminuisce

**Domanda 23**

Risposta  
corretta

Punteggio  
ottenuto 1,00  
su 1,00



Contrassegna  
domanda

Il numero dei parametri di un layer convolutivo dipende da:

- a. unicamente dalle dimensioni dei layers di input e di output
- b. Lo stride del kernel e tutte le dimensioni di input e output, compresi i canali
- c. Lo stride del kernel e le dimensioni spaziali di input e output
- d. La dimensione spaziale del kernel e il numero dei canali di input e output ✓

Risposta corretta.

La risposta corretta è:

La dimensione spaziale del kernel e il numero dei canali di input e output

**Domanda 24**

Risposta  
corretta

Punteggio  
ottenuto 1,00  
su 1,00



Contrassegna  
domanda

il campo ricettivo (receptive field) di un neurone di una CNN dipende da:

- a. La dimensione del kernel e il numero dei canali del layer in cui si trova il neurone
- b. La profondità del layer in cui si trova il neurone e le dimensioni dei kernel dei layers precedenti, ma non dai loro strides.
- c. La profondità del layer in cui si trova il neurone e le dimensioni e gli strides dei kernel dei layers precedenti ✓
- d. Unicamente dalla profondità del layer a cui si trova il neurone

La risposta corretta è: La profondità del layer in cui si trova il neurone e le dimensioni e gli strides dei kernel dei layers precedenti



**Domanda 25**

Risposta  
corretta

Punteggio  
ottenuto 1,00  
su 1,00



[Contrassegna  
domanda](#)

Quale è la derivata della funzione di MaxPooling?

- a. Non è una funzione derivabile
- b. 1 ovunque
- c. l'identità
- d. 1 in corrispondenza del massimo e 0 altrove ✓

Risposta corretta.

La risposta corretta è:

1 in corrispondenza del massimo e 0 altrove

**Domanda 26**

Risposta  
corretta

Punteggio  
ottenuto 1,00  
su 1,00



[Contrassegna  
domanda](#)

Quale è la tipica struttura di una rete neurale per classificazione di immagini?

- a. Solo livelli densi
- b. Una sequenza alternata di convoluzioni e downsampling, seguita da flattening e pochi livelli densi finali ✓
- c. Un encoder, seguito da un decoder
- d. Una sequenza di convoluzioni che preservano la dimensione spaziale dell'input

La risposta corretta è: Una sequenza alternata di convoluzioni e downsampling, seguita da flattening e pochi livelli densi finali

**Domanda 27**

Risposta  
corretta

Punteggio  
ottenuto 1,00  
su 1,00



[Contrassegna  
domanda](#)

Quale delle seguenti sentenze relative agli autoencoders è corretta?

- 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. E' una rete neurale che codifica se stessa
- d. L'encoder e il decoder devono essere strettamente simmetrici

La risposta corretta è: La rappresentazione interna prodotta dall'encoder abitualmente ha una dimensione ridotta rispetto a quella di partenza

**Domanda 28**

Risposta  
corretta

Punteggio  
ottenuto 1,00  
su 1,00



[Contrassegna  
domanda](#)

Con modelli generativi si intende:

- a. modelli che cercano di apprendere la distribuzione di probabilità dei dati ✓
- b. l'applicazione di tecniche genetiche al deep learning
- c. il processo di automatizzazione della generazione di reti neurali
- d. l'uso di attacchi avversariali allo scopo di aumentare la robustezza dei modelli

Risposta corretta.

La risposta corretta è:  
modelli che cercano di apprendere la distribuzione di probabilità dei dati

**Domanda 29**

Risposta  
corretta

Punteggio  
ottenuto 1,00  
su 1,00



[Contrassegna  
domanda](#)

Selezionare la sentenza corretta:

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

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



[Contrassegna  
domanda](#)

Qualte tra i seguenti è un tipico campo di applicazione della U-Net?

- a. Object Detection
- b. Generazione musicale
- c. Segmentazione semantica ✓
- d. Natural Language Processing

La risposta corretta è: Segmentazione semantica