Fondamenti di Data Science e Machine Learning - Prof. G. Polese - Anno Accademico 2021/22 Prova Scritta (ore 15:00) 04/07/2022

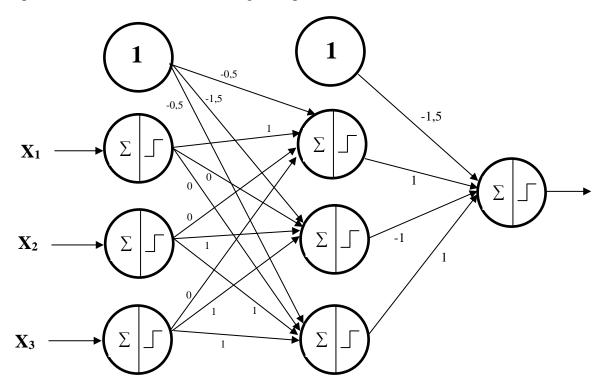
Cognome e Nome:

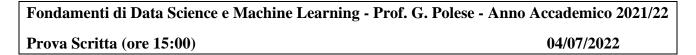
Matricola:

(Scrivere solo nello spazio bianco. Se necessario, usare il retro del foglio. Non sono ammessi elaborati su fogli diversi.)

Esercizio 1 (punti 6 su 30)

Ricavare l'espressione booleana calcolata dal seguente percettrone multilivello:





Esercizio 2 (punti 5 su 30)

Codificare il percettrone dell'esercizio 1 usando la libreria TensorFlow ed avviare una sessione con i seguenti valori: $X_1 = 1$, $X_2 = 0$ e $X_3 = 1$.

Esercizio 3 (punti 7 su 30)

Dati i seguenti punti in uno spazio bidimensionale: (2, 5)(3, 4)(6,8)(7,1)(1,2)(8,10)(1,5)(4,1)

Mostrare i passi di un algoritmo di clustering gerarchico bottom-up (mostrando ad ogni passo cluster e centroidi) per raggruppare i suddetti punti in 2 cluster, usando la distanza euclidea.

Fondamenti di Data Science e Machine Learning - Prof. G. Polese - Anno Accademico 2021/22 Prova Scritta (ore 15:00) 04/07/2022

Esercizio 4 (punti 6 su 30)

Dati i seguenti due schemi di database, di cui DB1 gestito dal nostro Dipartimento (Informatica) e DB2 gestito dalla segreteria studenti:

// DB1 Dipartimento

ESAMI2021(<u>Codice Corso</u>, <u>Studente</u>, Data, Voto, Lode) STUDENTE(Matricola, Cognome, Nome, Indirizzo, Cap, Città, Provincia)

// DB2 Segreteria

STUDENTE(<u>Matricola</u>, Codice Fiscale, Cognome, Nome, Telefono, CFU Superati) **ESAMI2022**(<u>Codice Corso</u>, <u>Studente</u>, Data, Voto, Lode)

Integrare i due schemi tramite una vista ed utilizzando l'approccio GAV.

Fondamenti di Data Science e Machine Learning - Prof. G. Polese - Anno Accademico 2021/22 Prova Scritta (ore 15:00) 04/07/2022

Esercizio 5 (punti 6 su 30)

Dato il seguente training set relativo ai comportamenti passati di clienti di un gestore di telefonia mobile. Quest'ultimo vorrebbe sfruttare tali dati per costruire un modello predittivo basato su un albero di decisione, al fine di prevedere i clienti che potrebbero cambiare gestore:

Soggetto	Età	Anni come Cliente	Ricariche Anno Precedente	Cambiato Gestore
1	30-50	1-3	300-400 €	NO
2	> 50	3-6	200-300 €	SI
3	30-50	> 6	300-400 €	NO
4	18-29	3-6	10-100 €	SI
5	30-50	1-3	300-400 €	NO
6	18-29	> 6	100-200 €	SI
7	18-29	1-3	200-300 €	NO
8	30-50	3-6	200-300 €	NO
9	> 50	> 6	100-200 €	SI
10	18-29	3-6	10-100 €	SI

Calcolare l'entropia a priori e, successivamente, determinare la scelta dell'attributo dipendente da collocare in ciascun livello durante la costruzione dell'albero di decisione, basandosi sul criterio di scegliere ad ogni livello l'attributo che più di tutti contribuisce a ridurre l'entropia.