

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 7 su 30)**

Dato il seguente dataset:

Item	X ₁	X ₂	X ₃	X ₄	Y
1	1	2	-2	-1	0
2	1	3	2	4	1
3	-3	1	-2	5	0
4	4	1	-3	2	1
5	-1	4	1	3	1
6	-9	8	-6	1	0

Disegnare una rete neurale (eventualmente anche con un solo neurone) che, presi in input i valori degli attributi predittivi X_1 , X_2 , X_3 , X_4 , di una riga, più un eventuale termine di bias, produca il valore Y della stessa riga. Si richiede quindi di specificare i pesi, l'eventuale termine di bias e le (o le) funzioni (o funzione) di attivazione.

Esercizio 2 (punti 6 su 30)

Usando il coefficiente di Pearson, determinare l'attributo predittivo maggiormente correlato all'attributo dipendente Y.

Esercizio 3 (punti 5 su 30)

Come l'esercizio 2, ma utilizzando FD, o eventualmente, RFD, invece del coefficiente di Pearson.

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(Codice Corso, Studente, Data, Voto, Lode)

STUDENTE(Matricola, Cognome, Nome, Indirizzo, Cap, Città, Provincia)

// DB2 Segreteria

STUDENTE(Matricola, Codice Fiscale, Cognome, Nome, Telefono, CFU Superati)

ESAMI2022(Codice Corso, Studente, Data, Voto, Lode)

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

Esercizio 5 (punti 6 su 30)

Dato il seguente training set relativo ai comportamenti passati di clienti di un fornitore di Gas:

Soggetto	Età	Anni come Cliente	Spesa Anno Precedente	Cambiato Fornitore
1	30-50	1-5	601-800 €	NO
2	> 50	5-10	801-1200 €	SI
3	30-50	> 10	301-600 €	NO
4	24-29	1-5	100-300 €	SI
5	30-50	5-10	301-600 €	NO
6	24-29	> 10	100-300 €	SI
7	24-29	1-5	100-300 €	NO
8	30-50	5-10	801-1200 €	NO
9	> 50	> 10	301-600 €	SI
10	24-29	1-5	100-300 €	SI

Nonostante i comportamenti dei suoi clienti potrebbero subire bruschi cambiamenti a causa dei bruschi aumenti recenti legati al conflitto Russa-Ucraina, il gestore 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.

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.