

Cognome e Nome:

Matricola:

Esercizio 1 (punti 5 su 30)

Disegnare un percettrone, eventualmente multilivello, con indicazione dei pesi e delle step function, che calcoli la funzione booleana $A \text{ OR } (B \text{ XOR } C)$.

Esercizio 2 (punti 5 su 30)

Data una tabella con 6 tuple. Si supponga che per una coppia di attributi X ed Y valgano le seguenti relazioni:

$$t_1[X] = t_2[X] = t_3[X] \neq t_4[X] = t_5[X] = t_6[X]$$

$$t_1[Y] \neq t_2[Y] = t_3[Y] = t_4[Y] = t_5[Y] \neq t_6[Y]$$

Calcolare il g3 error per una RFD $X \rightarrow Y$ che rilassi sull'extent.

Esercizio 3 (punti 6 su 30)

Data la seguente signature matrix:

Shingle	S ₁	S ₂	S ₃	S ₄
0	0	0	1	1
1	1	1	1	0
2	1	1	0	1
3	0	1	0	1
4	1	1	1	0
5	0	0	1	0

- Calcolare la similarità di Jaccard tra ogni coppia di colonne;
- Calcolare la signature di minhash per ogni colonna usando le seguenti 3 funzioni hash:

$$h_1(x) = (8x + 1) \bmod 6; h_2(x) = (9x + 3) \bmod 6; h_3(x) = (7x + 3) \bmod 6.$$

Mostrare l'evoluzione della matrice delle signature di minhash simulando l'esecuzione dell'algoritmo per il loro calcolo. Inoltre, calcolare le similarità di Jaccard tra tutte le coppie di signature di minhash.

Esercizio 4 (punti 7 su 30)

Dati i seguenti punti in uno spazio bidimensionale:

(3,1)(1, 1)(6,3)(3,2)(7,11)(9,1)(4,7)(2,3)(6,7)(5,1)(5,2)(4,4)

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

Esercizio 5 (punti 7 su 30)

Dato il seguente frammento di dataset:

Istanza	x_1	x_2	x_3	x_4
i1	-5	4	8	-1
i2	-3	5	3	4
i3	3	3	1	7
i4	1	2	3	-4
i5	-9	0	-2	1

Verificare se esistono valori dei 5 pesi w_i e del termine di bias che facciano in modo che un classificatore SVM lineare restituisca la classe positiva sulle prime 3 istanze e quella negativa sulle ultime 2.