

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 $A \text{ XOR } B \text{ AND } C$.

Esercizio 2 (punti 5 su 30)

Data una tabella con 6 tuple. Si supponga che su un attributo X valgano le seguenti relazioni:

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

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

dove il simbolo \approx indica che è possibile trovare una soglia ragionevole affinché due tuple siano simili, mentre ciò non vale per il simbolo \neq .

Dire se esiste una FD e/o una RFD $X \rightarrow Y$. Qualora valga una RFD che rilassi sull'extent, indicare il relativo g3 error.

Esercizio 3 (punti 6 su 30)

Data la seguente signature matrix:

Shingle	S_1	S_2	S_3	S_4
0	1	0	1	1
1	0	1	1	0
2	1	0	0	1
3	0	1	0	0
4	1	1	0	0
5	0	0	0	1

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

$$h_1(x) = (4x - 1) \bmod 6; h_2(x) = (2x + 3) \bmod 6; h_3(x) = (3x + 7) \bmod 6.$$

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:

(2,1)(1, 2)(5,2)(4,3)(6,10)(10,1)(3,6)(3,4)(5,6)(6,1)(4,1)(6,6)

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	1	-2	-9	1
i2	-6	2	9	-2
i3	3	7	2	6
i4	4	-2	2	-6
i5	-4	2	-2	2

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 2 istanze e quella negativa sulle ultime 3.