

Corso di Laurea Specialistica in Ingegneria Informatica
Data Mining e Scoperta di Conoscenza
Esame del 5 aprile 2006

Si consideri il seguente dataset:

Esercizio 1 (4 punti + 2).

Si consideri una 2-layer feedforward ANN con due nodi di input a e b , un nodo intermedio c ed un nodo di output d .

- (a) Descrivere la topologia della rete;
- (b) Si descriva analiticamente l'equazione dell'update dei pesi, giustificandola.
- (c) [OPZIONALE] considerando un'inizializzazione dei pesi al medesimo valore 0.1, si calcoli l'aggiornamento dei pesi dopo la prima iterazione dell'algoritmo BackPropagation, assumendo learning rate 0.3 e i seguenti dati di training:

a	b	d
1	0	1
0	1	0

Per i prossimi esercizi, si consideri il seguente dataset:

	A	B	C	D	E
1	1	0	1	1	0
2	1	1	0	1	1
3	1	0	1	1	0
4	0	1	0	1	0
5	1	0	1	0	1
6	0	1	1	1	0

Esercizio 2 (2 punti).

Si trovi il dendogramma dei clusters utilizzando l'algoritmo gerarchico Single-Linkage nell'ipotesi in cui si utilizzi la distanza di Jaccard.

Esercizio 4 (4 punti).

Assumendo un supporto del 50% e una confidenza dell'80%,

- (a) Trasformare il dataset in formato transazionale e applicare l'algoritmo Apriori per il calcolo degli itemset frequenti illustrando l'utilizzo di un Hash-Tree con bucket di dimensione 2
- (b) Calcolare le regole associative a partire dagli itemset frequenti calcolati nel punto (a).
- (c) Calcolare gli itemset frequenti multidimensionali sul dataset originario.