

COMPITO DI INTELLIGENZA ARTIFICIALE
25 Giugno 2003
Parte II – Durata 45 minuti (totale su punti 16)

Esercizio 1 (punti 9)

Si consideri il problema di stabilire se un'automobile è sportiva S o familiare F sulla base del colore della carrozzeria, della presenza o meno di un'antenna e del numero di ruote. Si vuole apprendere un albero di decisione per risolvere tale problema partendo dal seguente insieme di dati

Colore	Antenna	Ruote	Tipo
Rosso	Sì	quattro	S
Rosso	No	quattro	S
Rosso	No	quattro	S
Verde	No	quattro	S
Verde	No	sei	F
Blu	No	quattro	F
Blu	No	quattro	F
Blu	No	quattro	F
Blu	Sì	tre	F

- a) si indichi il guadagno di entropia del training set rispetto agli attributi Antenna e Colore e si scelga tra questi il primo attributo dell'albero. Si ricorda che.

$$info(S) = - \sum_{j=1}^k \frac{freq(C_j, S)}{|S|} \times \log_2 \left(\frac{freq(C_j, S)}{|S|} \right)$$

$$info_X(T) = \sum_{j=1}^n \frac{|T_j|}{|T|} \times info(T_j)$$

$$gain(X) = info(T) - info_X(T)$$

- b) si costruisca senza eseguire ulteriori calcoli l'albero decisionale che sarebbe costruito da c4.5 partendo dall'insieme di esempi precedente. Come condizione di terminazione non si usi quella adottata da c4.5 ma ci si fermi quando tutte le foglie contengono esempi omogenei .

Esercizio 2 (punti 7)

Descrivere come avviene la pianificazione non lineare.

In particolare, descrivere cosa si intende per causal link e come viene usato in un pianificatore non lineare.

SOLUZIONE

Esercizio 1

a) $\text{info}(T) = -4/9 * \log_2 4/9 - 5/9 * \log_2 5/9 = 0,991$

Calcoliamo il guadagno di entropia dell'attributo Antenna

$$\text{info}_{A}(T) = 2/9 \text{info}_{\text{si}}(T) + 7/9 \text{info}_{\text{no}}(T)$$

$$\text{info}_{\text{si}}(T) = 1$$

$$\text{info}_{\text{no}}(T) = -3/7 * \log_2 3/7 - 4/7 * \log_2 4/7 \approx 0.983$$

quindi

$$\text{info}_{A}(T) = 2/9 + 7/9 * 0.983 \approx 0.98$$

$$\text{Gain}(\text{Antenna}) = \text{info}(T) - \text{info}_{A}(T) = 0.01$$

Calcoliamo il guadagno di entropia dell'attributo Colore

$$\text{info}_{C}(T) = 3/9 \text{info}_{\text{Rosso}}(T) + 2/9 \text{info}_{\text{Verde}}(T) + 4/9 \text{info}_{\text{Blu}}(T)$$

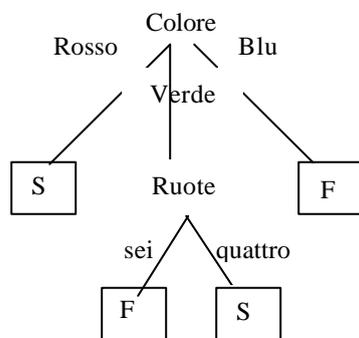
$$\text{info}_{\text{Rosso}}(T) = 0$$

$$\text{info}_{\text{Blu}}(T) = 0$$

$$\text{info}_{\text{Verde}}(T) = 1$$

$$\text{info}_{C}(T) = 0.2$$

$$\text{Gain}(\text{Colore}) = \text{info}(T) - \text{info}_{C}(T) = 0.791$$



Esercizio 2

Si calcolino i domini dopo la propagazione effettuata dal vincolo

x1 :: [6, 7]

x2 :: [6, 7]

x3 :: [8, 9]

x4 :: [8, 9]

x5 :: [10]