Soluzione Esercizi del Corso di Apprendimento Automatico

Prima Parte

Esercizio 2

Considerare il problema di apprendere il concetto *EnjoySport* visto a lezione, dove però si considerano i seguenti valori per gli attributi

$$
\begin{aligned}
\text{Sky} &\rightarrow \{Sunny, Cloudy\};\\
\text{AirTemp} &\rightarrow \{Hot, Cold\};\\
\text{Humidity} &\rightarrow \{Low, Normal, High\};\\
\text{Wind} &\rightarrow \{Weak, Strong\};\\
\text{Water} &\rightarrow \{Hot, Warm, Cold, VeryCold\};\\
\text{Forecast} &\rightarrow \{Same, Change\};
\end{aligned}
$$

a) definire un insieme minimo di esempi che, se presentato all'algoritmo *Candidate-Elimination*, produce il seguente Version Space:

$$
\begin{array}{llllllll}
S:\{< & Sunny & ? & High & Strong & Warm & ? & >\}\\
G:\{< & ? & ? & High & ? & ? & ? & >\}
\end{array}
$$

Risposta:

Una possibile sequenza di esempi è la seguente:

$$
\begin{array}{l}
1.\ (< S \quad H \quad H \quad S \quad W \quad S \quad >,\ Yes)\\
2.\ (< S \quad C \quad H \quad S \quad W \quad C \quad >,\ Yes)\\
3.\ (< S \quad C \quad L \quad S \quad W \quad C \quad >,\ No)
\end{array}
$$

b) dire quante ipotesi contiene il Version Space di sopra;

Risposta:

8. Si può dare la risposta sia enumerando esplicitamente le ipotesi che appartengono al Vesion Space, sia osservando che gli attributi che cambiano fra l'ipotesi in S e quella in G sono 3 ed avendo ogni attributo la possibilità di assumere solo 2 valori (un valore specifico, oppure ?), il numero totale di ipotesi nel Version Space è $2^3 = 8$.

c) dire qual'è il numero minimo di esempi addizionali richiesti per far convergere il Version Space al seguente concetto terget

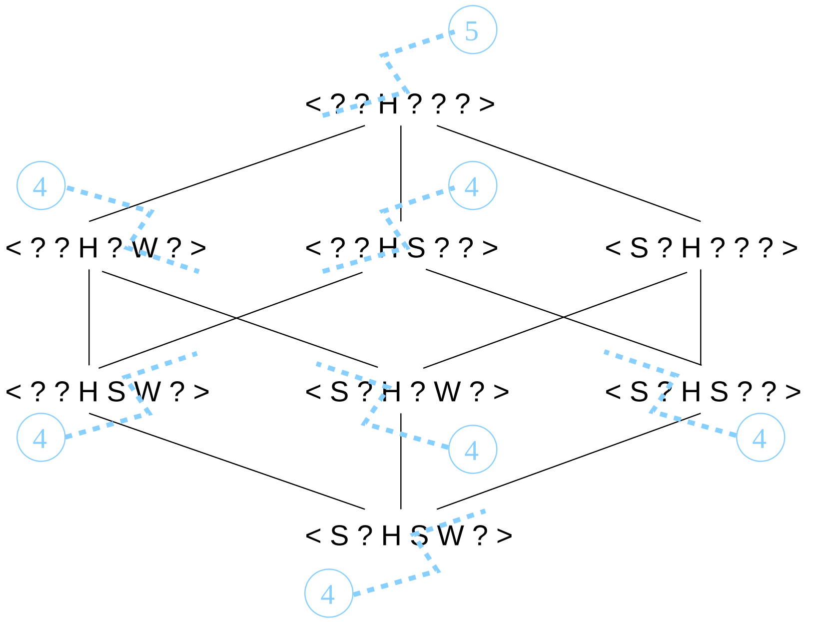
$$< \quad Sunny \quad ? \quad High \quad ? \quad ? \quad ? \quad >$$

Risposta:

una possibile sequenza è la seguente:

$$\begin{array}{l} 4.\ (< S \quad H \quad H \quad W \quad H \quad S \quad >, \quad Yes) \\ 5.\ (< C \quad H \quad H \quad W \quad H \quad S \quad >, \quad No) \end{array}$$

Di seguito viene mostrato come il Version Space cambia alla presentazione degli esempi 4 e 5.



Esercizio 2

Considerare il problema di apprendere il concetto *EnjoySport* visto a lezione, dove però si considerano i seguenti valori per gli attributi

$$\begin{aligned}
\text{Sky} &\to \{Sunny, Cloudy\}; \\
\text{AirTemp} &\to \{Hot, Cold\}; \\
\text{Humidity} &\to \{Low, Normal, High\}; \\
\text{Wind} &\to \{Weak, Strong\}; \\
\text{Water} &\to \{Hot, Warm, Cold, VeryCold\}; \\
\text{Forecast} &\to \{Same, Change\};
\end{aligned}$$

a) definire un insieme minimo di esempi che, se presentato all'algoritmo *Candidate-Elimination*, produce il seguente Version Space:

$$\begin{array}{lccccccl}
S: \{< & Sunny & ? & High & ? & Warm & ? & >\} \\
G: \{< & ? & ? & High & ? & ? & ? & >\}
\end{array}$$

Risposta:

Una possibile sequenza di esempi è la seguente:

$$\begin{array}{l}
1.\ (< S \quad H \quad H \quad W \quad W \quad S \quad >, \quad Yes) \\
2.\ (< S \quad C \quad H \quad S \quad W \quad C \quad >, \quad Yes) \\
3.\ (< S \quad C \quad L \quad S \quad W \quad C \quad >, \quad No)
\end{array}$$

b) dire quante ipotesi contiene il Version Space di sopra;

Risposta:

4. Si può dare la risposta sia enumerando esplicitamente le ipotesi che appartengono al Vesion Space, sia osservando che gli attributi che cambiano fra l'ipotesi in S e quella in G sono 2 ed avendo ogni attributo la possibilità di assumere solo 2 valori (un valore specifico, oppure ?), il numero totale di ipotesi nel Version Space è $2^2 = 4$.

c) dire qual'è il numero minimo di esempi addizionali richiesti per far convergere il Version Space al seguente concetto terget

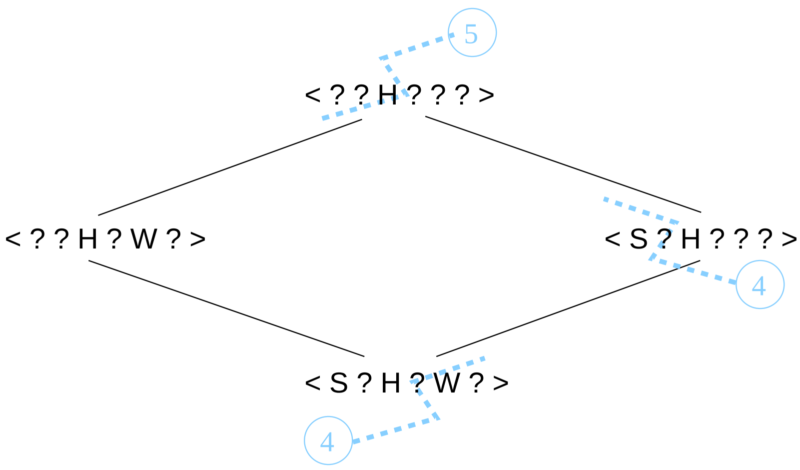
$$< \ ? \ ? \ High \ ? \ Warm \ ? \ >$$

Risposta:

una possibile sequenza è la seguente:

$$\begin{array}{l} 4.\ (<C\quad H\quad H\quad W\quad W\quad S\quad >,\quad Yes) \\ 5.\ (<C\quad H\quad H\quad W\quad H\quad S\quad >,\quad No) \end{array}$$

Di seguito viene mostrato come il Version Space cambia alla presentazione degli esempi 4 e 5.



Esercizio 2

Considerare il problema di apprendere il concetto *EnjoySport* visto a lezione:

a) definire un insieme minimo di esempi che, se presentato all'algoritmo *Candidate-Elimination*, produce il seguente Version Space:

$$\begin{array}{l} S: \{<\ ?\ \ ?\ \ Normal\ \ Strong\ \ Cool\ \ ?\ >\} \\ G: \{<\ ?\ \ ?\ \ ?\ \ Strong\ \ ?\ \ ?\ >\} \end{array}$$

Risposta:

Una possibile sequenza di esempi è la seguente:

$$\begin{array}{l} 1.\ (<S\quad W\quad N\quad S\quad C\quad S\quad >,\quad Yes) \\ 2.\ (<R\quad C\quad N\quad S\quad C\quad C\quad >,\quad Yes) \\ 3.\ (<R\quad C\quad N\quad W\quad C\quad C\quad >,\quad No) \end{array}$$

b) dire quante ipotesi contiene il Version Space di sopra;

Risposta:

4. Si può dare la risposta sia enumerando esplicitamente le ipotesi che appartengono al Vesion Space, sia osservando che gli attributi che cambiano fra l'ipotesi in S e quella in G sono 2 ed avendo ogni attributo la possibilità di assumere solo 2 valori (un valore specifico, oppure ?), il numero totale di ipotesi nel Version Space è $2^2 = 4$.

c) dire quanti esempi addizionali sono richiesti per far convergere il Version Space al seguente concetto target

$$< \ ? \ \ ? \ \ ? \ \ Strong \ \ Cool \ \ ? \ >$$

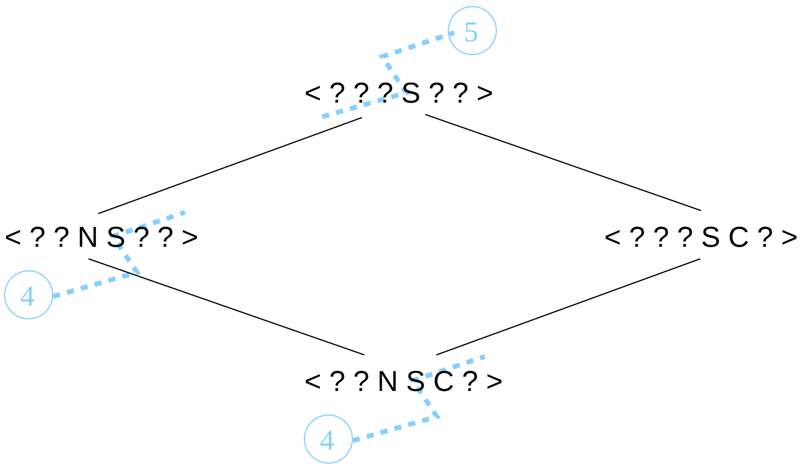
e fornire un insieme di tali esempi;

Risposta:

una possibile sequenza è la seguente:

$$\begin{array}{l} 4.\ (< R \quad C \quad H \quad S \quad C \quad C \quad >, \quad Yes) \\ 5.\ (< R \quad C \quad H \quad S \quad H \quad C \quad >, \quad No) \end{array}$$

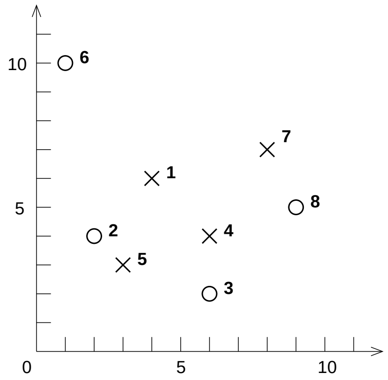
Di seguito viene mostrato come il Version Space cambia alla presentazione degli esempi 4 e 5.



Esercizio 2

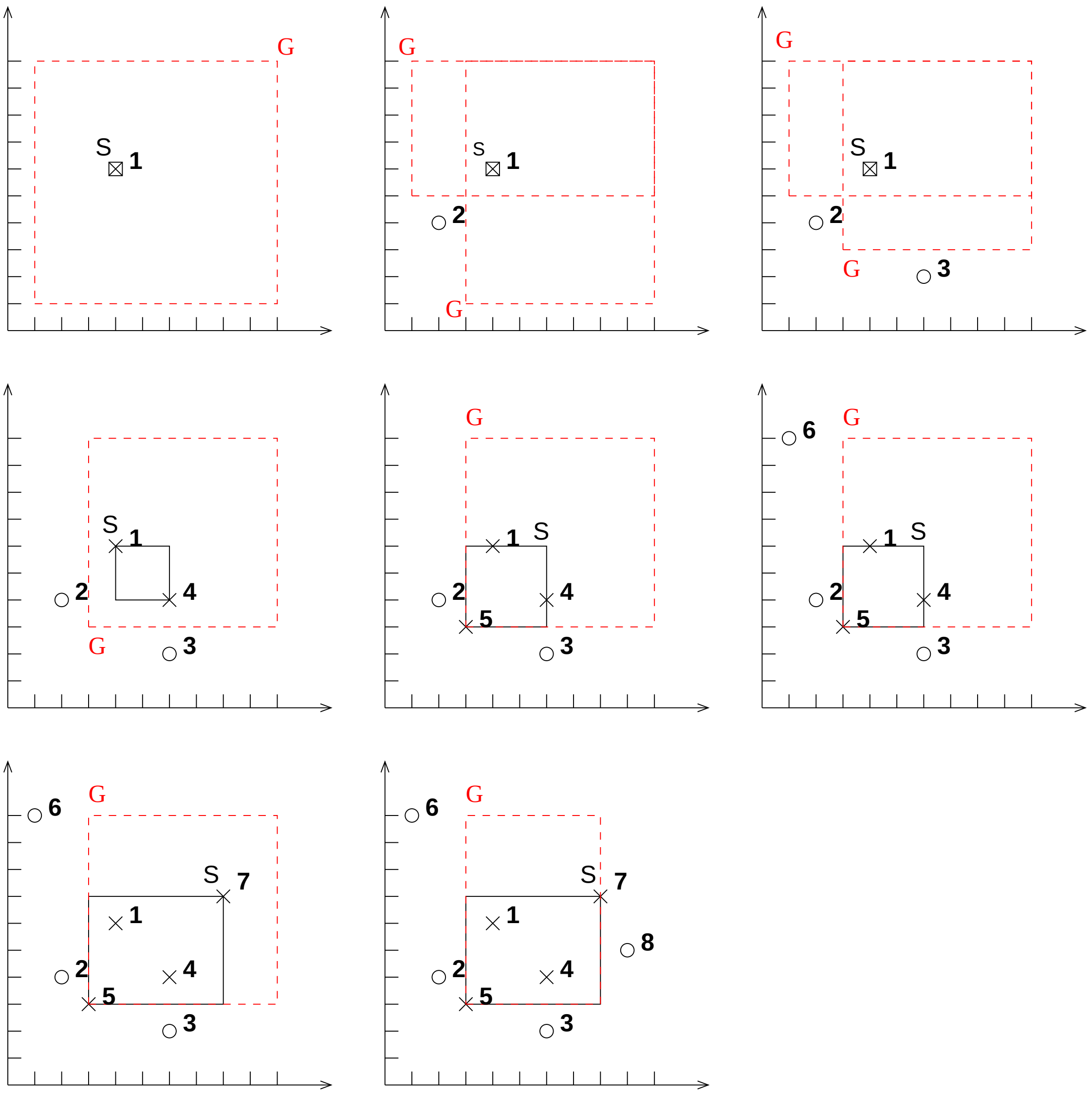
Sia dato uno spazio delle istanze costituito da punti sul piano a coordinate intere nell'intervallo [1,10], e uno spazio delle ipotesi dato da rettangoli con lati paralleli agli assi e vertici a coordinate intere (sempre nell'intervallo [1,10]), come visto a lezione.

Dato il seguente insieme di esempi



(dove le ics costituiscono istanze positive e i tondini istanze negative), e dove i numeri associati individuano la sequenza di presentazione degli esempi, descrivere come cambia il confine più generale e quello più specifico.

Risposta:



G
S
1
G
S
1
2
G
G
S
1
2
G
3
S
1
2
4
G
3
G
1
S
2
4
5
3
6
G
1
S
2
4
5
3
6
G
S
7
1
2
4
5
3
6
G
S
7
1
8
2
4
5
3

Esercizio 3

b) dato il seguente insieme di apprendimento

Esempio	*Target*	A_1 (5 *val.*)	A_2 (3 *val.*)	A_3 (4 *val.*)	A_4 (9 *val.*)
1	$+$	v_1	v_1	v_4	v_1
2	$+$	v_1	v_3	v_2	v_2
3	$-$	v_2	v_1	v_1	v_5
4	$-$	v_2	v_2	v_4	v_4
5	$-$	v_3	v_2	v_4	v_3
6	$+$	v_3	v_1	v_2	v_6
7	$+$	v_4	v_3	v_1	v_7
8	$-$	v_4	v_2	v_4	v_9
9	$+$	v_5	v_1	v_4	v_1
10	$-$	v_5	v_3	v_2	v_5

mostrare come ID3 (con GainRatio(S,A)) costruisce l'albero di decisione corrispondente. Per ogni attributo calcolare $GainRatio(S, A)$ e giustificare la scelta dell'attributo utilizzato ad ogni nodo interno.

Risposta:

Entropia: 1.000000

Attributo A_1

valore 1:[+,-] 2 0, valore 2:[+,-] 0 2, valore 3:[+,-] 1 1, valore 4:[+,-] 1 1, valore 5:[+,-] 1 1

Gain: 0.400000, Split: 2.321928, Ratio: 0.172271

Attributo A_2

valore 1:[+,-] 3 1, valore 2:[+,-] 0 3, valore 3:[+,-] 2 1

Gain: 0.400000, Split: 1.570951, Ratio: 0.254623

Attributo A_3

valore 1:[+,-] 1 1, valore 2:[+,-] 2 1, valore 3:[+,-] 0 0, valore 4:[+,-] 2 3

Gain: 0.039036, Split: 1.485475, Ratio: 0.026278

Attributo A_4

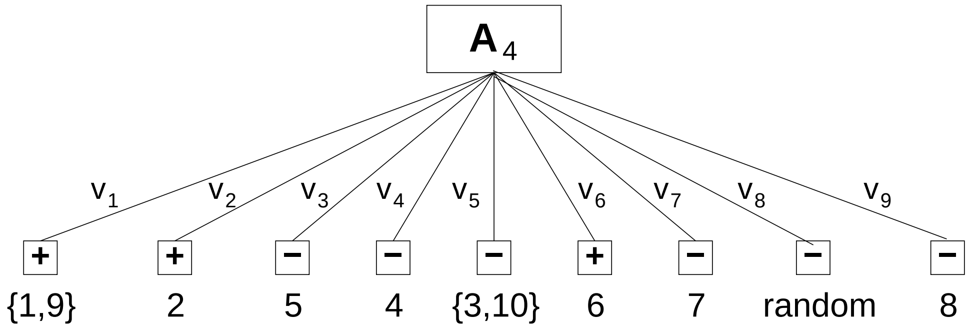
valore 1:[+,-] 2 0, valore 2:[+,-] 1 0, valore 3:[+,-] 0 1, valore 4:[+,-] 0 1, valore 5:[+,-] 0 2, valore 6:[+,-] 1 0, valore 7:[+,-] 1 0, valore 8:[+,-] 0 0, valore 9:[+,-] 0 1

Gain: 1.000000, Split: 2.921928, Ratio: 0.342240

Quindi si sceglie A_4

S_1 = {1[+],9[+]}, S_2 = {2[+]}, S_3 = {5[-]}, S_4 = {4[-]}, S_5 = {3[-],10[-]}, S_6 = {6[+]}, S_7 = {7[-]}, S_8 = vuoto, S_9 = {8[-]}

e l'albero finale è:



c) Mostrare cosa succede se si aggiunge l'esempio *(4,3,2,1)* con target **-**

Risposta:

Entropia: 0.994030

Attributo A_1

valore 1:[+,-] 2 0, valore 2:[+,-] 0 2, valore 3:[+,-] 1 1, valore 4:[+,-] 1 2, valore 5:[+,-] 1 1

Gain: 0.379950 Split: 2.299896 Ratio: 0.165203

Attributo A_2

valore 1:[+,-] 3 1, valore 2:[+,-] 0 3, valore 3:[+,-] 2 2

Gain: 0.335384 Split: 1.572624 Ratio: 0.213264

Attributo A_3

valore 1:[+,-] 1 1, valore 2:[+,-] 2 2, valore 3:[+,-] 0 0, valore 4:[+,-] 2 3

Gain: 0.007234 Split: 1.494919 Ratio: 0.004839

Attributo A_4

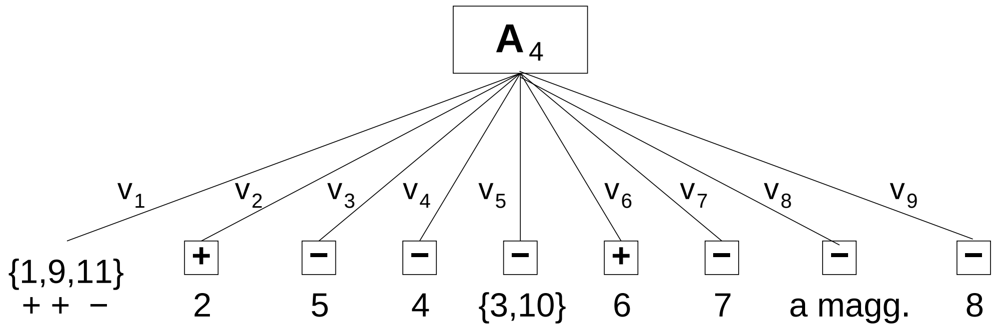
valore 1:[+,-] 2 1, valore 2:[+,-] 1 0, valore 3:[+,-] 0 1, valore 4:[+,-] 0 1,

valore 5:[+,-] 0 2, valore 6:[+,-] 1 0, valore 7:[+,-] 1 0, valore 8:[+,-] 0 0,

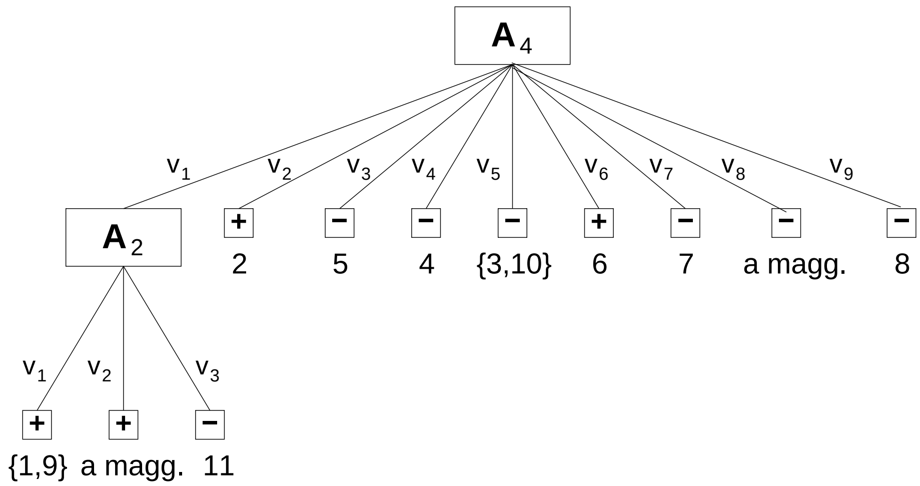
valore 9:[+,-] 0 1

Gain: 0.743586 Split: 2.845351 Ratio: 0.261334

Quindi si sceglie sempre A_4, però l'albero non è completo:



In particolare sid eve ancora elaborare l'insieme $S_1 = \{1[+],9[+],11[-]\}$, che risulta essere classificato correttamente da A_1, o A_2, o A_3. Supponendo di scegliere l'attributo che assume il numero minimo di valori distinti, cioé A_2, otteniamo l'albero finale:



Esercizio 3

b) dato il seguente insieme di apprendimento

Esempio	*Target*	A_1 (5 *val.*)	A_2 (3 *val.*)	A_3 (4 *val.*)	A_4 (9 *val.*)
1	$+$	v_1	v_1	v_4	v_1
2	$+$	v_1	v_3	v_2	v_2
3	$-$	v_2	v_1	v_1	v_5
4	$+$	v_2	v_2	v_4	v_4
5	$-$	v_3	v_2	v_4	v_3
6	$+$	v_3	v_1	v_2	v_6
7	$+$	v_4	v_3	v_1	v_9
8	$-$	v_4	v_2	v_4	v_9
9	$-$	v_5	v_1	v_4	v_1
10	$+$	v_5	v_3	v_2	v_5

mostrare come ID3 (con GainRatio(S,A)) costruisce l'albero di decisione corrispondente. Per ogni attributo calcolare $GainRatio(S, A)$ e giustificare la scelta dell'attributo utilizzato ad ogni nodo interno.

Risposta:

Entropia: 0.970951

Attributo A_1

valore 1:[+,-] 2 0, valore 2:[+,-] 1 1, valore 3:[+,-] 1 1, valore 4:[+,-] 1 1,

valore 5:[+,-] 1 1

Gain: 0.170951 Split: 2.321928 Ratio: 0.073624

Attributo A_2

valore 1:[+,-] 2 2, valore 2:[+,-] 1 2, valore 3:[+,-] 3 0

Gain: 0.295462 Split: 1.570951 Ratio: 0.188078

Attributo A_3

valore 1:[+,-] 1 1, valore 2:[+,-] 3 0, valore 3:[+,-] 0 0, valore 4:[+,-] 2 3

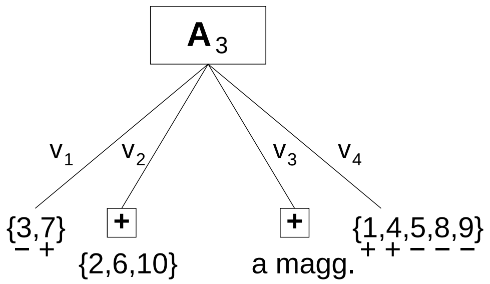
Gain: 0.285475 Split: 1.485475 Ratio: 0.192178

Attributo A_4

valore 1:[+,-] 1 1, valore 2:[+,-] 1 0, valore 3:[+,-] 0 1, valore 4:[+,-] 1 0, valore 5:[+,-] 1 1, valore 6:[+,-] 1 0, valore 7:[+,-] 0 0, valore 8:[+,-] 0 0, valore 9:[+,-] 1 1

Gain: 0.370951 Split: 2.721928 Ratio: 0.136282

Quindi si sceglie l'attributo A_3 per la radice:

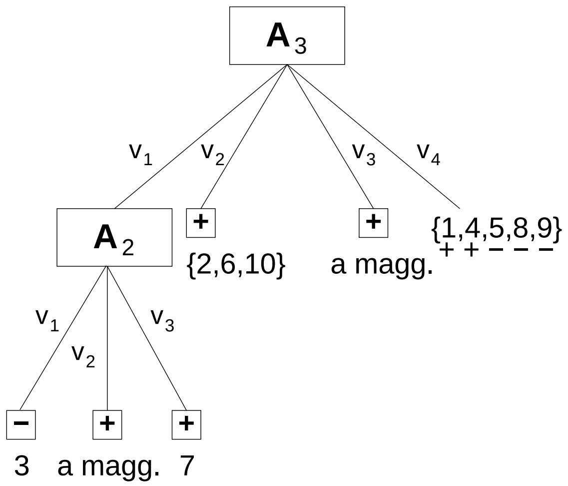


e rimangono i seguenti sottoinsiemi di esempi da elaborare:

S_1 = {3[-],7[+]}, S_4 = {1[+],4[+],5[-],8[-],9[-]}

poiché S_2 origina una foglia con etichetta +, mentre S_3, essendo vuoto, origina, a maggioranza, una foglia con etichetta +.

Si nota facilmente che S_1 è classificato correttamente da A_1 o A_2. Supponendo di scegliere l'attributo che assume il numero minimo di valori distinti, cioè A_2, si ottiene il seguente albero parziale:



dove l'etichetta della foglia nel mezzo è stata decisa risalendo alla radice, visto che gli esempi associati al nodo non mostrano il prevalere di una etichetta sull'altra.

Infine, S_4 è chiaramente classificato correttamente da A_1: