

# Machine Learning & Data Mining

Prova intermedia 8 maggio 2023

## 1) Concept learning

1.a) Dato uno spazio di ipotesi  $H$  e un insieme di esempi  $D$ , come si definisce il General Boundary  $G$  del Versione Space  $VS_{H,D}$ ?

1.b) Se gli esempi  $D$  sono tutti corretti (senza rumore o incompletezze) è possibile che la classificazione di una nuova istanza utilizzando l'ipotesi indotta da FIND-S per  $D$  e il version space  $VS_{H,D}$  dia risultati differenti? Se sì perché può accadere?

## 2) Alberi di decisione

2.1) Si consideri il seguente dataset  $D$  di 10 esempi con attributo classe "PlayTennis". Utilizzando come misura di impurità il Misclassification Error, determinare quale sia l'attributo migliore per il nodo radice dell'albero di decisione  $D$ .

Day	Outlook	Temperature	Humidity	Wind	PlayTennis
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes

2.2) Descrivere la procedura di Reduced Error Pruning per gestire il problema dell'overfitting.

## 3) Valutazione di algoritmi

3.1) Come viene stimata l'accuratezza con il metodo Bootstrap?

3.2) Definire le misure di Precision e Recall per un problema di classificazione binario.

## 4) Reti Neurali

4.1) Che cosa è e a cosa serve il "learning rate" nell'algoritmo della discesa del gradiente?

4.2) Come viene modificato il peso  $w_{ji}$  sull'arco da una unità di input  $i$  a una unità  $j$  nel livello nascosto di una rete neurale dall'algoritmo Backpropagation?

4.3) Scrivere la struttura (nodi e archi) di una rete neurale feed forward con 4 unità di input, due livelli nascosti con 3 unità e un livello di output con 4 unità.

## 5) Learning Bayesiano

Si consideri una nuova istanza  $x$  con valori binari di classificazione "yes" e "no", uno spazio di 4 ipotesi e un training set  $D$  tali che:  $P(h_1|D) = 0.40$ ,  $P(h_2|D) = 0.15$ ,  $P(h_3|D) = 0.1$ ,  $P(h_4|D) = 0.35$ .

$P(\text{"no"}|h_1) = 0$ ,  $P(\text{"yes"}|h_1) = 1$

$P(\text{"no"}|h_2) = 0$ ,  $P(\text{"yes"}|h_2) = 1$

$P(\text{"no"}|h_3) = 1$ ,  $P(\text{"yes"}|h_3) = 0$

$P(\text{"no"}|h_4) = 1$ ,  $P(\text{"yes"}|h_4) = 0$

Qual'è il valore  $v^*$  calcolato dal classificatore Bayesiano ottimo per  $x$ ?

Motivare la risposta facendo i conti con la formula che determina come ricavare  $v^*$ .