

Machine Learning & Data Mining

Prova intermedia 10 maggio 2024

1) Concept learning

1.a) E' possibile che gli algoritmi Find-S e Candidate Elimination diano esattamente lo stesso risultato? In quali situazioni? Motivare la risposta in modo chiaro e conciso.

1.b) Quali sono i principali limiti dell'algoritmo FIND-S?

2) Alberi di decisione

2.1) Si consideri il seguente dataset di 10 esempi D1-D10 con attributo classe "PlayTennis", in cui i valori No e Yes **vanno invertiti** (ad esempio per D1, PlayTennis vale Yes anziché No). Utilizzando come misura di impurità il Misclassification Error, determinare quale sia l'attributo migliore per il nodo radice dell'albero di decisione.

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 brevemente un metodo probabilistico per gestire il valore mancante per un attributo di un record nel dataset di training.

3) Valutazione di algoritmi

3.1) Come è definito l'intervallo di confidenza per $Error_S(h)$?

3.2) Come è fatta Confusion Matrix per un problema di Classificazione in cui l'attributo classe ha 3 possibili valori?

4) Reti Neurali

4.1) Cosa sono gli iperparametri di una rete neurale? (rispondere in modo chiaro e conciso, indicando tre iperparametri differenti.

4.2) Scrivere la struttura di una rete neurale con un livello di unità nascoste che consente di apprendere qualsiasi formula booleana (logica proposizionale) di 2 variabili A e B.

4.3) Che differenza c'è tra discesa del gradiente batch e stocastica (incrementale)? Rispondere in modo conciso.

5) Learning Bayesiano

5.1) Come è definita l'ipotesi h_{ML} con massima verosimiglianza (maximum likelihood) per un dataset di esempi D

5.2) Sotto quale assunzione questa ipotesi è quella migliore probabilmente?