

Esercizi proposti DTs

Si prenda un altro dataset tra quelli messi a disposizione da Weka e applicare l'algoritmo J48 dei DTs e l'algoritmo di Naive Bayes (non è argomento d'esame, ma per chi vuole si trova al capito 12.6 e 20.2.2 del libro AIMA). Applicare un 10 fold cross-validation per entrambi gli algoritmi.

Si richiede di consegnare una presentazione power point o un pdf in cui si descrive:

1. Il dataset scelto: qual è, cosa contiene, qual è l'obiettivo (cosa si vuole classificare/predire), eventuali modifiche al dataset (pre-processing) e qualche statistica come la divisione in classi o altri attributi rilevanti.
2. Perché applicare il Decision Tree a questo dataset ed eventualmente, se lo studente lo ha studiato, perché Naive Bayes.
3. Visualizzazione dell'albero risultante per il DT J48.
4. Quali sono i risultati:
 - a. Tabella contenente i valori di Precision, Recall ed F-measure
 - b. Breve discussione dei risultati (quale approccio è migliore e perché)
5. Una piccola sperimentazione in cui viene eliminato un attributo, riapplicato J48, mostrato i nuovi risultati e confrontati con quelli precedenti, e visualizzazione del nuovo albero con una piccola discussione.

INFO:

Per l'esercizio si può utilizzare direttamente Weka *OPPURE* è possibile farlo direttamente in Python come mostrato a lezione.

Inviare un unico file compresso contenente codice (python book) e soluzione agli esercizi (pdf; pptx; word),

- chiamato <cognome>.zip
- all'indirizzo hromei@ing.uniroma2.it
- oggetto "[IA] esercitazione DTs e NB".

Per esempio il mio file si chiamerebbe: `hromei.zip`

N.B.: Si ricorda agli studenti che le specifiche devono essere seguite *tassativamente*, pena la non correzione dell'esercitazione.