

# B003725 Intelligenza Artificiale (2016/17)

Studente: Tommaso Scarlatti (5784154) — <2017-01-10 Tue>

## Elaborato assegnato per l'esame finale

### Istruzioni generali

Il lavoro svolto dovrà essere inviato per email due giorni prima della data dell'esame orale, includendo:

1. Sorgenti e/o files ausiliari sviluppati, evidenziando se necessario le parti riprese da altre fonti (che dovranno essere **opportunamente citate**) e le parti sviluppate personalmente.
2. Se necessario, un file README che spieghi come compilare o far eseguire i programmi sottomessi (eventualmente indicando dipendenze da pacchetti software non standard).
3. Una breve relazione (massimo 4 pagine in formato pdf) che descriva il lavoro ed i risultati sperimentali. Non è necessario ripetere in dettaglio i contenuti del libro di testo o di eventuali articoli, è invece necessario che vengano fornite informazioni sufficienti a *riprodurre* i risultati riportati.

L'elaborato sarà oggetto di discussione durante l'esame orale.

**Importante:** va evitato di allegare files eseguibili (inclusi files .jar o .class generati da Java) al messaggio email, al fine di evitare il filtraggio automatico da parte del software antispam di ateneo!

---

### Alberi di decisione

Nella prima parte di questo elaborato si sviluppa del codice (in un linguaggio di programmazione a scelta) per l'apprendimento di alberi di decisione come esposto in classe e descritto in R&N 2009 §18.3, utilizzando l'entropia come misura di impurità. Nel passo base, l'algoritmo accetta un parametro  $m$  per controllare la complessità dell'albero: se il numero di errori è inferiore ad  $m$ , si crea una foglia invece di continuare la ricorsione.

Nella seconda parte, si applica il codice ad almeno tre data sets scelti a piacere dal repository [MLData](#), e si producono delle curve (al variare di  $m$ ) che mostrano l'errore sul training set e sul test set in funzione della complessità dell'albero (misurata dal numero di nodi interni).