

B003725 Intelligenza Artificiale (2017/18)

Studente: Iacopo Erpichini (5766356) — <2018-02-27 Tue>

Elaborato assegnato per l'esame finale

Istruzioni generali

Il lavoro svolto sarà oggetto di discussione durante l'esame orale e dovrà essere sottomesso per email due giorni prima dell'esame, includendo:

1. Links ad un repository **pubblico** su [github](#), [gitlab](#), o [bitbucket](#), dove reperire sorgenti e materiale sviluppato. In alternativa è accettabile allegare un file zip.
2. Un file README che spieghi:
 - come usare il codice per riprodurre i risultati sottomessi
 - se vi sono parti del lavoro riprese da altre fonti (che dovranno essere **opportunamente citate**)
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.

Importante: negli eventuali allegati va evitato di inserire files eseguibili (inclusi files `.jar` o `.class` generati da Java), al fine di evitare il filtraggio automatico da parte del software antispam di ateneo!

Potatura di regole negli 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. Si implementa quindi una semplice strategia di pruning sulle regole corrispondenti all'albero (DNF dei cammini dalla radice alle foglie) basata sull'errore sul validation set (si veda [Mitchell 1997, cap. 3](#), pw: ai15).

Nella seconda parte, si applica il codice per l'apprendimento di alberi di decisione ad almeno tre data sets scelti a piacere dal repository [MLData](#), confrontando i risultati ottenuti prima e dopo il pruning.