

B003725 Intelligenza Artificiale (2017/18)

Studente: Manuel Pepe (5776848) — <2018-01-09 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!

Alberi di decisione con dati mancanti

Nella prima parte di questo esercizio si implementa, in un linguaggio di programmazione a scelta, l'algoritmo di apprendimento per alberi di decisione descritto in classe. Si implementa quindi la strategia descritta nella sezione 3.7.4 in [Mitchell \(1997\)](#) (pw: ai15) per gestire valori mancanti. Nella seconda parte, si applica l'algoritmo a problemi di classificazione scegliendo a piacere almeno tre data sets (reperibili dall'[UCI Machine Learning Repository](#)), dopo aver rimosso (casualmente e uniformemente, con probabilità p) alcuni valori negli attributi. Si riportano infine le accuratezze, misurate con 10-fold cross-validation, ottenute al variare di p , per $p = 0, 0.1, 0.2, 0.5$.