

Data Mining e Scoperta di Conoscenza

Progetto 12

Implementazione dell'algoritmo CAR (Class Association Rules)

1. Si estenda Weka con una nuova classe `weka.classifiers.CAR.java`, che implementi il classificatore **CAR** i cui dettagli sono riportati nel seguente articolo *"Integrating Classification and Association Rule Mining"* fornito insieme al progetto.
2. Si applichi l'algoritmo implementato sui dati UCI confrontando l'accuratezza dell'algoritmo
 - a. Rispetto a un insieme di algoritmi già presenti in Weka
 - b. Al variare dei parametri dell'algoritmo
3. Si studi il dataset UCI "Heart Disease Databases" (caratterizzato da un alto tasso di valori mancanti), e si implementi una strategia per l'imputazione dei valori mancanti (implementando il filtro (`weka.filters.unsupervised.CARFilter.java`) che faccia uso di regole associative. Si valuti come varia l'accuratezza di tecniche di classificazione tradizionali (alberi di decisione, classificatori IBk, Naive Bayes) se si utilizza la tecnica implementata

NOTE PER L'ESECUZIONE DEL PROGETTO

1. Scrivi un rapporto di circa 10 pagine in cui
 - a. Descrivi analiticamente l'algoritmo che hai implementato.
 - b. Commenti le parti essenziali del codice Java che hai scritto, e metti in un'appendice l'intero codice
 - c. commenti e illustri graficamente e quantitativamente gli esperimenti effettuati.
2. Prepare delle slides Powerpoint (non più di 10 slides) in cui riassumi gli esiti del progetto