

Data Mining e Scoperta di Conoscenza

Progetto 3

Implementazione dell'algoritmo di clustering BIRCH

1. Si estenda Weka con una nuova classe `weka.clusterers.birch.java`, che implementi l'algoritmo di clustering **BIRCH** i cui dettagli sono riportati nel seguente articolo "*BIRCH: An efficient Data Clustering Method for Very Large Databases*" fornito insieme al progetto.
2. Si applichi l'algoritmo implementato:
 - a. sul dataset UCI
 - b. sui dati forniti contestualmente al progetto valutando come varia l'accuratezza predittiva
 - a) rispetto agli altri algoritmi di clustering forniti in Weka.
 - b) al variare dei parametri dell'algoritmo.

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