

ALGORITMA KAMILA UNTUK *CLUSTERING* DATA TIPE CAMPURAN

Dedi Rosadi

Laboratorium Komputasi Statistika, Departemen Matematika, FMIPA UGM
Email: dedirosadi@ugm.ac.id

Abstrak

Pada data berukuran besar, seringkali ditemukan data yang merupakan campuran data bertipe kontinu dan bertipe kategorik. Pada paper ini dibahas salah satu algoritma yang dapat digunakan untuk mengkluster data tipe campuran, yang disebut algoritma KAMILA. Algoritma KAMILA merupakan gabungan dari metode algoritma k-means dan Gaussian-multinomial mixture models, dimana nantinya algoritma KAMILA dapat menangani kelemahan dari masing-masing metode tersebut. Pada paper ini juga diberikan contoh kasus dari algoritma ini untuk keperluan mengelompokkan siswa berdasarkan prestasinya. Komputasi akan dilakukan dengan software R.

Kata Kunci: *clustering, KAMILA, data tipe campuran, k-means, Gaussian-multinomial mixture models*

