



Lembar Praktikum 4
Mata Kuliah Data Mining
Semester Genap Tahun Akademik 2018/2019

Topik: Teknik-teknik dalam praproses data dalam R

Deskripsi Tugas

1. Bukalah file data forestfire.csv. Atribut apa saja yang memiliki *missing value*? Isilah nilai yang kosong tersebut dengan R. Tampilkan kode program beserta penjelasan hasilnya!
2. Lakukan diskretisasi data forestfire.csv pada atribut DC (4 bin) dan RH (3 bin)! Gabungkan hasil diskretisasi dengan data hasil praproses nomor 1. Tampilkan dan jelaskan hasilnya!
3. Lakukan normalisasi pada atribut FPMC dan DMC serta lakukan transformasi dengan membuat atribut baru (new_atribut) yang merupakan penjumlahan keduanya!
4. Database anggota suatu komunitas dibuat oleh dua orang yang berbeda (file database1.csv dan database2.csv). Buatlah database baru yang merupakan integrasi dari kedua database tersebut dan terapkan praproses jika masih memungkinkan untuk meningkatkan kualitas data tersebut!

Ketentuan penulisan laporan:

1. Jawaban dituliskan dalam file .doc atau .docx dilengkapi dengan nama kelompok dan anggota kelompok.
2. Laporan diunggah dalam LMS sesuai jadwal yang ditentukan.
3. Ketentuan nama fail laporan: LKP<no LKP>_Paralel<kelas paralal>_Kelompok<no kelompok>.docx, contoh: LKP4_Paralel1_Kel1.docx