

PERTEMUAN 5

TEKNIK PRAPROSES DATA PADA WEKA

LAPORAN PENDAHULUAN

1. Apakah semua data dengan tipe.csv atau .arff dapat diproses untuk data mining dengan menggunakan Weka?
2. Mengapa perlu dilakukan praproses data?
3. Apakah setiap akan melakukan tahapan data mining harus dilakukan praproses data?
4. Apakah semua teknik praproses data harus dilakukan ketika melakukan praproses data?

TUGAS PRAKTIKUM

1. Jawablah pertanyaan-pertanyaan berikut:
 - a. Buka dataset german-credit.arff lalu identifikasi atribut apa saja yang mengalami missing value!.
 - b. Lakukan pembersihan missing value dalam atribut tersebut!
 - c. Jelaskan apa saja yang dapat dilakukan untuk menangani atribut missing value!
2. Jelaskan secara singkat perbedaan pembersihan data ReplaceMissingValues pada atribut yang bersifat kualitatif dengan atribut yang bersifat kuantitatif?
3. Jelaskan perbedaan pengaturan ignoreClass pada teknik ReplaceMissingValues! Gunakan data german-credit.arff dan hilangkan 2 instance pertama pada atribut class.
4. Jika nantinya dataset german-credit.arff akan dilakukan deteksi outlier, apakah perlu melakukan normalisasi data pada salah satu atau beberapa atribut pada dataset tersebut? Jelaskan secara singkat alasan anda!
5. Pada dataset german-credit.arff, lakukan diskretisasi data pada atribut credit_amount dan jelaskan bagaimana hasil diskretisasi data tersebut!
6. Pada kasus apakah diperlukan teknik reduksi data? Jelaskan secara singkat!