

Tugas Sains Data

Klasifikasi menggunakan ID3

Isa Albanna, S.Si., M.Si

Klasifikasi

Keterangan Dataset

- 1. Age:** usia pasien (dalam tahun)
- 2. Anemia:** Penurunan sel darah merah atau hemoglobin
- 3. High blood pressure:** Jika pasien menderita hipertensi
- 4. Creatinine phosphokinase:** Tingkat enzim CPK dalam darah (mcg/L)
- 5. Diabetes:** Jika pasien menderita diabetes
- 6. Ejection fraction:** Persentase darah yang meninggalkan jantung pada setiap kontraksi
- 7. Sex:** Wanita atau pria
- 8. Platelets:** Trombosit dalam darah (kiloplatelet/mL)
- 9. Serum creatinine:** Tingkat kreatinin dalam darah (mg/dL)
- 10. Serum sodium:** Tingkat natrium dalam darah (mEq/L)
- 11. Smoking:** Jika pasien merokok
- 12. Time:** Periode tindak lanjut (dalam hari)
- 13. (target) death event:** Jika pasien meninggal selama masa tindak lanjut

Target machine Learning adalah Membuat model untuk memprediksi kemungkinan pasien meninggal karena gagal jantung dari data informasi yang telah diberikan.

Dari informasi tersebut, susunlah pohon keputusan menggunakan python dan lakukan pengujian dari data testing yang anda ambil. Kerjakan secara kelompok dengan maksimum anggota adalah 3 anak. Kirimkan jawaban anda dalam PDF dan sertakan link file atau kode python yang nantinya bisa saya akses. Link dapat berupa kode python secara offline atau kode yang terintegrasi dalam google-colabs

Data Set dapat diunduh pada file berikut: (dikirim melalui Telegram Group)