Sebuah gambar berisi gambar garis, grafik vektor

Deskripsi dibuat secara otomatis

**LABORATORIUM KOMPUTASI STATISTIKA DAN MATEMATIKA**

**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM**

**UNIVERSITAS GADJAH MADA**

**Sekip Utara BLS 21, Yogyakarta 55281**

**KETENTUAN DAN PERATURAN**

1. Jangan lupa berdoa sebelum mengerjakan.
2. Jawaban diketik dengan font Times New Roman ukuran 12pt, line spacing 1.15, margin normal, dan ukuran kertas A4.
3. Lembar jawaban terdiri dari soal, jawab, sintaks, output, interpretasi output, tabel (apabila diminta)
4. **Dilarang meng-*copypaste* atau menjiplak hasil jawaban kelompok lain (plagiarisme)**. Apabila ketahuan meng-*copypaste* atau menjiplak, maka nilai akan di bagi sebanyak jumlah praktikan yang hasil jawabannya sama.
5. **Dilarang bertanya pada Asisten Praktikum mengenai jawaban**.
6. Jangan lupa mencantumkan output program Python dan interpretasi.
7. Yang perlu dicantumkan di dalam bagian Pembahasan adalah ketikan ulang syntax, screenshot output (Pastikan Notebook Nya ada Output), dan interpretasi output.
8. **Jawaban dikumpulkan dalam bentuk .ipnyb (Sintaks )** **dengan format nama file: Assigment4\_NIF (contoh : Assignment4\_19777).**
9. Submit file di Google Form paling lambat hari **Senin, 8 Mei 2023 pukul 23.59** WIB. Keterlambatan pengumpulan akan ada pengurangan nilai.
10. **Keterlambatan pengumpulan akan ada pengurangan nilai akhir (-1 per menit)** .

**Format pengerjaan laporan :**

1. Soal
2. Jawaban

**Asisten Praktikum :**

Iqbal Hanif Anggita Adi 19542

      Ahmad Habib Hasan  Zein (20277)

     Muhammad Zaki Nurkholis 19733

Ahmad Habib Hasan Zein 20277

**Dosen Pengampu :**

    Dr. Adhitya Ronnie Effendie, S.Si., M.Si., M.Sc.

Mohamad Fahruli Wahyujati, S.Si., M.Si.

**[SOAL: 100]**

Dengan menggunakan dataset pada file “bab6\_adult\_data.csv” lakukan analisis klasifikasi untuk memprediksi pendapatan individu dari suatu kota. Analisis terdiri dari:

1. *Exploratory Data Analysis* (EDA),
2. Data preprocessing,
3. Feature Engineering (Jika diperlukan),
4. Modelling (Menggunakan Random Forest),
5. serta lakukan improvement (0.01 / 1% minimal) terhadap model yang telah diajarkan boleh melakukan beberapa improvement (*Feature selection*, *hyperparameter tunning*, *etc*.)
6. Carilah variable mana saja yang berpengaruh pada model (*feature importance*) boleh dalam bentuk visualisasi atau table.

metric evaluasi yang digunakan adalah F1-Score. Pengerjaan disertai dengan penjelasan dari setiap langkah-langkah analisis (comment dan/atau markdown)