Nama : Fawzy Imam Hamdani Gunamukti

NIM : 231011401030

KLS : 05TPLE017

Penjelasan Machine Learning Pertemuan 5

1. Membaca data hasil pra-pemrosesan dari file CSV.

2. Memisahkan data menjadi fitur (X) dan target (y) untuk klasifikasi

kelulusan.

3. Membagi dataset menjadi train, validation, dan test (70%-15%-15%)

dengan stratifikasi agar seimbang.

4. Menyiapkan pipeline pra-pemrosesan numerik:

- Mengisi nilai kosong (imputasi median).

- Melakukan standarisasi fitur numerik (StandardScaler).

5. Melatih model Logistic Regression (baseline):

- Menggunakan class\_weight=‘balanced’ agar data tidak bias.

- Mengevaluasi model dengan F1-score dan classification report

pada data validasi.

6. Melatih model Random Forest sebagai model lanjutan (improvement):

- Menggunakan 300 pohon (n\_estimators=300) dan

max\_features=‘sqrt’.

- Mengevaluasi performa menggunakan F1-score pada data validasi.

7. Melakukan optimasi hyperparameter dengan GridSearchCV:

- Parameter yang diuji: max\_depth dan min\_samples\_split.

- Menggunakan StratifiedKFold (5 lipatan) dan scoring=‘f1\_macro’.

- Menampilkan parameter terbaik dan hasil cross-validation

terbaik.

8. Mengevaluasi model terbaik (best\_rf) pada data validasi dan test:

- Menghitung F1-score, classification report, dan confusion

matrix.

- Menghitung ROC-AUC dan menggambar kurva ROC.

9. Menyimpan model terbaik ke file ‘model.pkl’ menggunakan joblib.

10. Membangun REST API sederhana dengan Flask:

- Endpoint ‘/’ → menampilkan pesan bahwa API aktif.

- Endpoint ‘/predict’ (POST) → menerima input JSON, memprediksi

hasil, dan mengembalikan label prediksi:

- 1 → “LULUS TEPAT WAKTU”

- 0 → “TIDAK LULUS TEPAT WAKTU / TERLAMBAT”

- Menyertakan juga probabilitas prediksi dalam persen.

11. Menjalankan Flask server di port 5000 untuk menerima request

prediksi model.

Kesimpulan: Kode ini mencakup seluruh alur machine learning dari

pembacaan data, pelatihan model, optimasi parameter, evaluasi,

penyimpanan model, hingga pembuatan REST API untuk prediksi kelulusan

mahasiswa.

Notes:

Untuk melakukan testing Hit End point bisa menggunakan Postman,

atau buat file python sendiri. Yang saya lakukan di pertemuan 5 ini Adalah membuat file python request dan kirim body Json ke endpoint FLASK.

