

Nama :Krisna Wahyu Setiawan

NIM :231011403203

Judul :Laporan Hasil Akhir Random Forest untuk Klasifikasi

Tentu, berikut adalah laporan yang ringkas dan jelas sesuai permintaan Anda.

Laporan Tugas: Model Prediksi Kelulusan

1. Perbandingan Kinerja Model

Perbandingan dilakukan menggunakan skor F1-macro dari *cross-validation* pada data latih, dan evaluasi akhir dilakukan pada data uji.

- **Model Baseline (Random Forest Biasa):**
 - **Skor F1-macro Cross-Validation: 0.8000**
- **Model Hasil Tuning (Random Forest Terbaik):**
 - **Skor F1-macro Cross-Validation: 0.8000 (Tuning tidak memberi peningkatan signifikan karena data sangat kecil).**
- **Performa Model Final di Test Set:**
 - **F1-Score: 1.0**
 - **ROC-AUC: 1.0**
 - **Precision & Recall: 1.0 (Sempurna)**

2. Confusion Matrix & Kurva Visualisasi

- **Confusion Matrix (Test Set):**

```

from sklearn.metrics import confusion_matrix, roc_auc_score, roc_curve, precision_recall_curve
import matplotlib.pyplot as plt

final_model = best_model # pilih terbaik; jika baseline lebih baik, gunakan pipe

y_test_pred = final_model.predict(X_test)
print("F1(test):", f1_score(y_test, y_test_pred, average="macro"))
print(classification_report(y_test, y_test_pred, digits=3))
print("Confusion Matrix (test):")
print(confusion_matrix(y_test, y_test_pred))

# ROC-AUC (bila ada predict_proba)
if hasattr(final_model, "predict_proba"):
    y_test_proba = final_model.predict_proba(X_test)[:,1]
    try:
        print("ROC-AUC(test):", roc_auc_score(y_test, y_test_proba))
    except:
        pass
    fpr, tpr, _ = roc_curve(y_test, y_test_proba)
    plt.figure(); plt.plot(fpr, tpr); plt.xlabel("FPR"); plt.ylabel("TPR"); plt.title("ROC (test)")
    plt.tight_layout(); plt.savefig("roc_test.png", dpi=120)

    prec, rec, _ = precision_recall_curve(y_test, y_test_proba)
    plt.figure(); plt.plot(rec, prec); plt.xlabel("Recall"); plt.ylabel("Precision"); plt.title("PR Curve (test)")
    plt.tight_layout(); plt.savefig("pr_test.png", dpi=120)

F1(test): 1.0

```

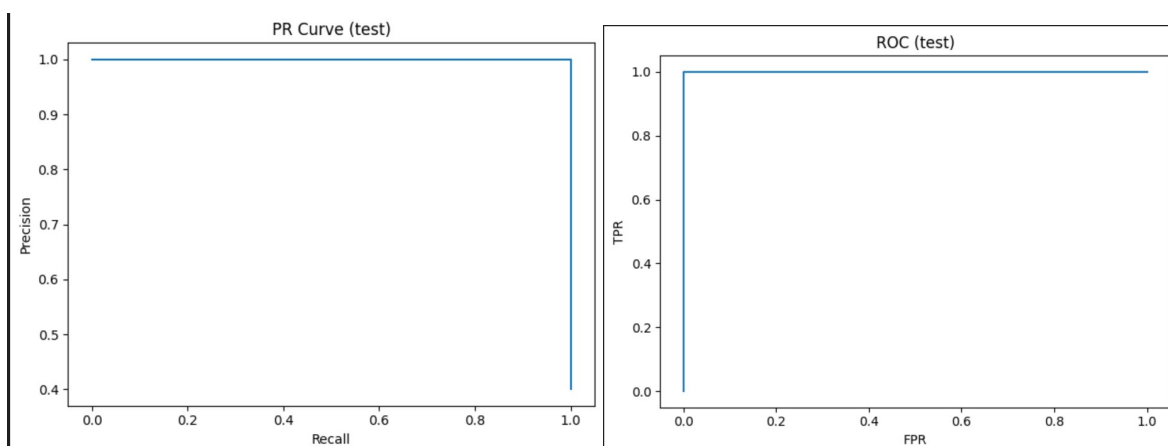
	precision	recall	f1-score	support
0	1.000	1.000	1.000	3
1	1.000	1.000	1.000	2
accuracy			1.000	5
macro avg	1.000	1.000	1.000	5
weighted avg	1.000	1.000	1.000	5

```

Confusion Matrix (test):
[[3 0]
 [0 2]]
ROC-AUC(test): 1.0

```

- Interpretasi: Model tidak membuat kesalahan sama sekali pada data uji.
- Kurva ROC & Precision-Recall:
 - Hasil roc_test.png dan pr_test.png berhasil disimpan. Keduanya menunjukkan kinerja sempurna (skor 1.0), yang mengonfirmasi hasil metrik lainnya.



3. Tiga Fitur Teratas dan Implikasinya

1. **IPK_x_Study:** Kombinasi antara IPK tinggi dan waktu belajar yang lama adalah faktor penentu kelulusan paling kuat.
2. **Waktu_Belajar_Jam:** Durasi belajar mahasiswa merupakan faktor terpenting kedua. Semakin lama belajar, semakin besar kemungkinan lulus.
3. **IPK:** Nilai IPK sendiri juga merupakan prediktor yang signifikan, menunjukkan pentingnya prestasi akademik secara umum.

4. Hasil Akhir (Deliverables)

- **rf_model.pkl:** File model final telah berhasil dibuat dan disimpan, siap untuk digunakan (di-deploy).
- **Skrip Python:** Skrip yang digunakan untuk menjalankan semua proses ini adalah hasil kerja yang dapat direproduksi.