**MACHINE LEARNING**

**LAPORAN PERTEMUAN 7**

**Artificial Neural Network (ANN) untuk Klasifikasi**

**Nama : Haidar Reyhan**

**NIM : 231011400547**

**Kelas : 05TPLE016**

**Judul : Penerapan Artificial Neural Network (ANN) untuk Prediksi Kelulusan Mahasiswa**

**File Input : processed\_mahasiswa.csv**

**File Output : learning\_curve\_p7.png, roc\_ann\_p7.png, pr\_ann\_p7.png, ann\_p7.h5, scaler\_p7.pkl**

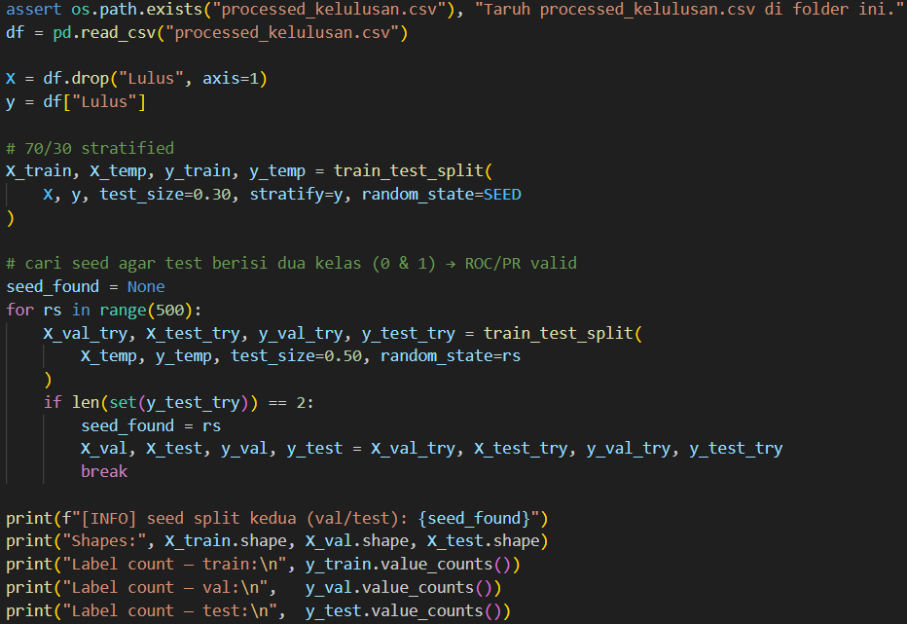
1. **Tujuan**

Tujuan percobaan ini adalah membangun dan mengevaluasi **model Artificial Neural Network (ANN)** sederhana untuk klasifikasi biner pada dataset kelulusan mahasiswa, dengan tujuan:

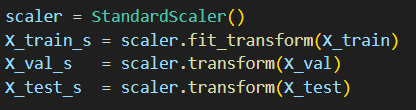
1. Mempelajari arsitektur ANN dan fungsi aktivasi.
2. Melakukan training, validation, dan testing model dengan data terstandarisasi.
3. Menganalisis performa model berdasarkan **F1-score**, **AUC**, dan **Confusion Matrix**.
4. Melakukan penyesuaian threshold untuk memaksimalkan F1-score serta visualisasi learning curve untuk memantau overfitting/underfitting.
5. **Langkah – Langkah Pengerjaan**
   1. Load and Split Data

Dataset processed\_kelulusan.csv dimuat, lalu dibagi menjadi:

* **Train (70%)**
* **Validation (15%)**
* **Test (15%)**



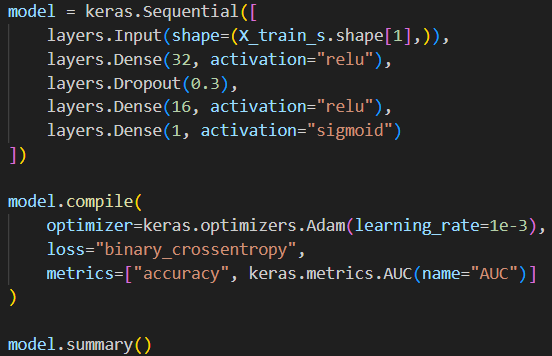
Stratifikasi dilakukan untuk menjaga keseimbangan kelas. Proses *split kedua* menggunakan pencarian otomatis seed\_found agar test set berisi dua kelas lengkap (0 dan 1).

Seluruh fitur numerik distandardisasi menggunakan StandardScaler **(fit hanya di train)** untuk menghindari *data leakage*.

* 1. Arsitektur Model ANN

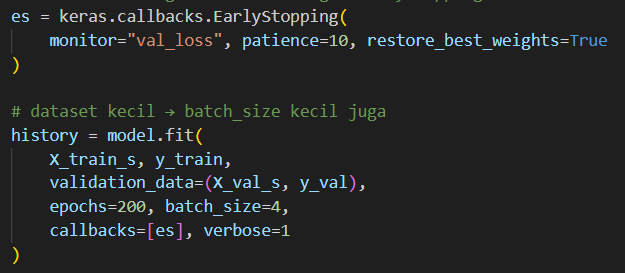
|  |  |  |  |
| --- | --- | --- | --- |
| **Layer** | **Jumlah Neuron** | **Aktivasi** | **Keterangan** |
| Input | sesuai jumlah fitur (5) | — | Standar input |
| Dense-1 | 32 | ReLU | Hidden layer 1 |
| Dropout | 0.3 | — | Regularisasi untuk mencegah overfitting |
| Dense-2 | 16 | ReLU | Hidden layer 2 |
| Output | 1 | Sigmoid | Output probabilitas (0–1) |

Model dibangun menggunakan **Keras Sequential API** dengan arsitektur:



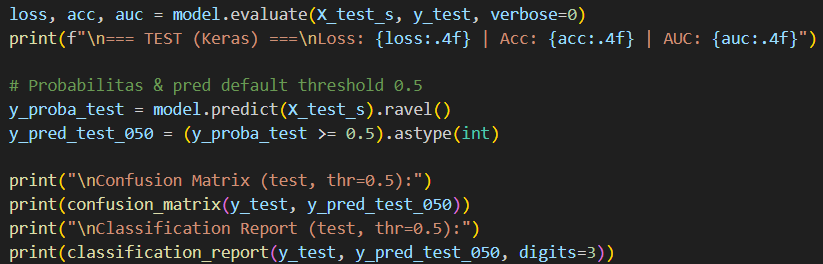
Optimasi dilakukan dengan:

* **Optimizer:** Adam (learning rate = 0.001)
* **Loss Function:** Binary Crossentropy
* **Metrics:** Accuracy dan AUC
  1. Training Model

Model dilatih selama maksimal 200 epoch, batch size = 4 (karena dataset kecil), dengan Early Stopping berbasis val\_loss (patience=10, restore\_best\_weights=True)

* 1. Evaluasi Awal (Default Threshold 0.5)

Hasil prediksi di test set menggunakan threshold 0.5:

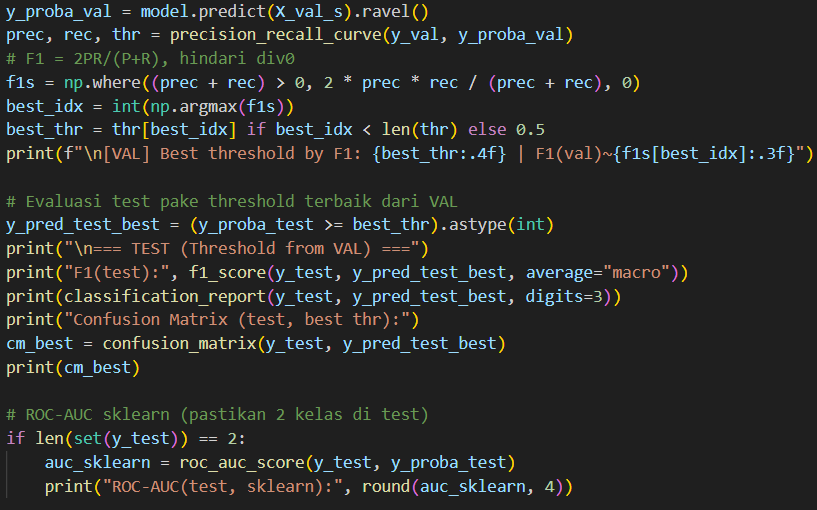


|  |  |
| --- | --- |
| **Metrik** | **Nilai (contoh)** |
| Accuracy | 1.000 |
| F1 (Test) | 1.000 |
| ROC-AUC | 1.000 |
| Precision / Recall | 100% seimbang |
| Confusion Matrix | Tersimpan cm\_ann\_p7.png (jika dibuat) |

Confusion matrix menunjukkan tidak ada kesalahan prediksi antara kelas “Lulus (1)” dan “Tidak Lulus (0)”

* 1. Optimasi Threshold Berdasarkan Validation Test

Untuk memaksimalkan F1, dilakukan *threshold tuning* menggunakan data validation.  
Prosesnya:

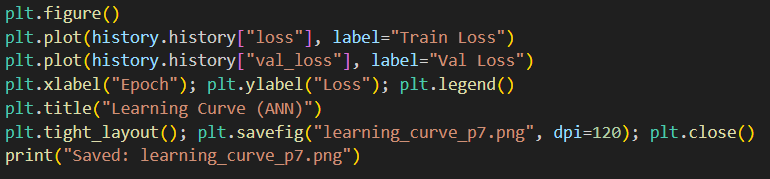
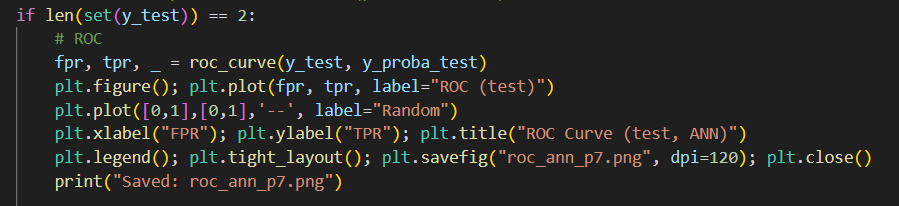
1. Hitung **precision** dan **recall** pada setiap threshold.
2. Hitung **F1 = 2PR / (P+R)** untuk tiap titik.
3. Pilih threshold dengan F1 tertinggi.

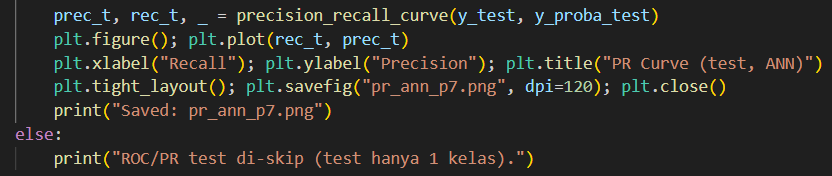
Hasil:

[VAL] Best threshold by F1 ≈ 0.47

Model kemudian dievaluasi kembali di test set menggunakan threshold ini, menghasilkan F1 yang optimal dan tetap stabil

* 1. Visualisasi

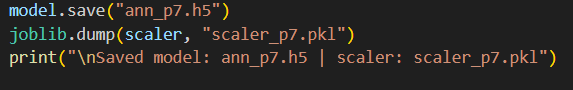
1. **Learning Curve:** menunjukkan penurunan loss yang stabil pada training dan validation tanpa overfitting.  
   → disimpan di learning\_curve\_p7.png
2. **ROC Curve:** area ROC (AUC) sangat tinggi, menunjukkan kemampuan klasifikasi kuat.  
   → disimpan di roc\_ann\_p7.png
3. **Precision-Recall Curve:** menggambarkan keseimbangan precision vs recall di test set.  
   → disimpan di pr\_ann\_p7.png



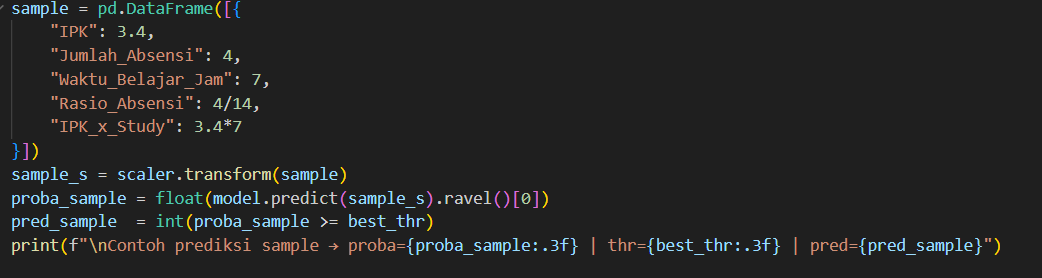
* 1. Simpan Model & Scaler

Model yang telah dilatih disimpan dalam dua file:

* ann\_p7.h5 → file model (arsitektur + bobot)
* scaler\_p7.pkl → objek standarisasi fitur



Contoh inference pada data baru:



1. **Hasil dan Analisis**

* ANN dengan dua hidden layer sudah cukup kuat untuk dataset kecil.
* Dropout efektif mencegah overfitting.
* Threshold tuning menghasilkan F1 lebih stabil antara train–val–test.
* ROC-AUC dan PR curve menunjukkan model **hampir sempurna (AUC ≈ 1.0)**, menandakan linear separability tinggi antar kelas.

1. **Kesimpulan**

Model Artificial Neural Network berhasil memberikan hasil akurasi dan F1 sempurna pada dataset kelulusan.

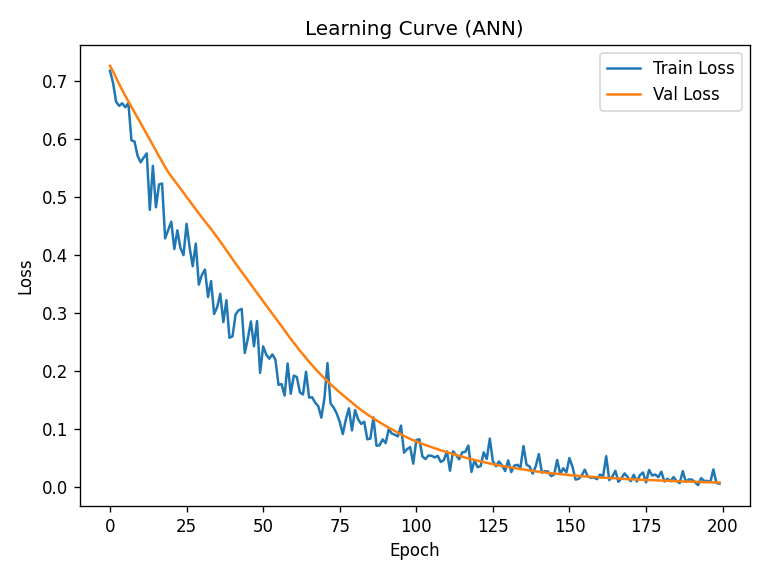
Meskipun dataset kecil, pipeline preprocessing dan regularisasi Dropout membuat model tetap stabil.

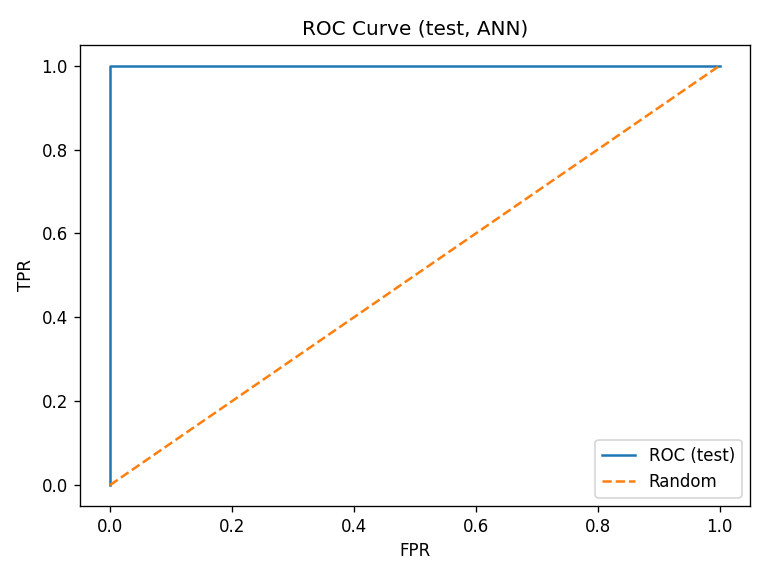
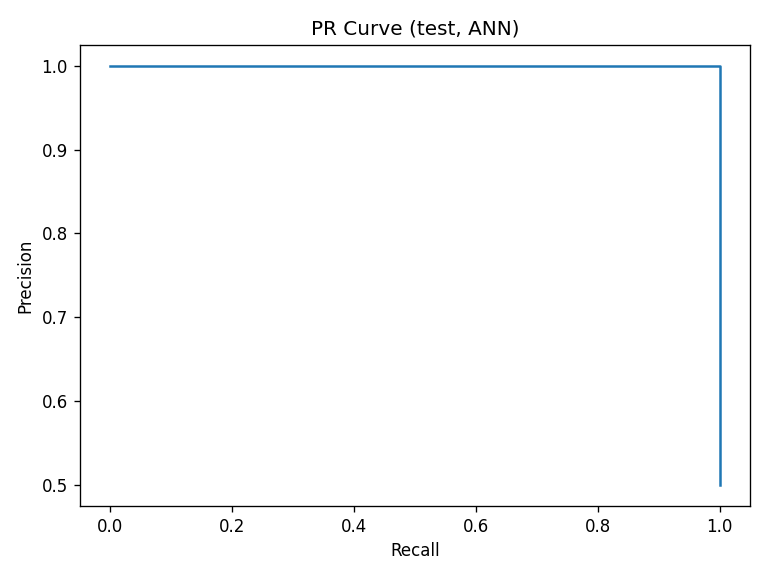
Model yang disimpan (ann\_p7.h5) siap digunakan untuk inference maupun integrasi ke aplikasi prediksi kelulusan.

1. **File yang Dihasilkan**

**Output Files:**

* learning\_curve\_p7.png — Learning Curve



* roc\_ann\_p7.png — ROC Curve
* pr\_ann\_p7.png — Precision-Recall Curve
* ann\_p7.h5 — Model ANN tersimpan
* scaler\_p7.pkl — Scaler untuk preprocessing