**Tugas 2: Machine Learning – Decision Tree dan Random Forest**

**Muhammad Zaidan Ramdhan - 0110222040**

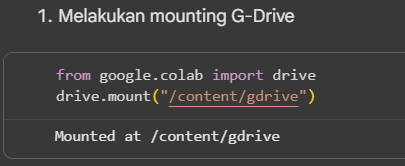
Teknik Informatika, STT Terpadu Nurul Fikri, Depok

\*E-mail: muha22040ti@student.nurulfikri.ac.id

Laporan ini bertujuan melakukan penerapan algoritma *Decision Tree* untuk melakukan klasifikasi pada dataset Iris. Dataset ini terdiri dari empat fitur numerik yaitu *Sepal Length (Cm)*, *Sepal Width (Cm)*, *Petal Length (Cm)*, dan *Petal Width (Cm)*, serta satu kolom target berupa nama spesies bunga (*Species*). Karena kolom target masih berbentuk teks, maka dilakukan proses *label encoding* agar data target dapat diubah menjadi bentuk numerik sehingga dapat digunakan oleh model *machine learning*. Selanjutnya, dataset dibagi menjadi dua bagian yaitu 80% data training dan 20% data testing menggunakan fungsi train\_test\_split dari *scikit-learn*.

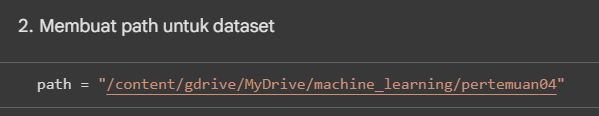
Model *Decision Tree Classifier* kemudian dilatih menggunakan data training untuk mempelajari pola hubungan antara fitur-fitur numerik dengan label spesies bunga. Setelah proses pelatihan selesai, model diuji menggunakan data testing untuk mengevaluasi performanya. Evaluasi dilakukan dengan menghitung nilai akurasi, menampilkan classification report, serta membuat confusion matrix untuk melihat hasil prediksi setiap kelas. Hasil evaluasi menunjukkan bahwa model mampu melakukan klasifikasi dengan baik terhadap data uji yang diberikan.

**1. Tugas mandiri – membuat model Decision Tree**



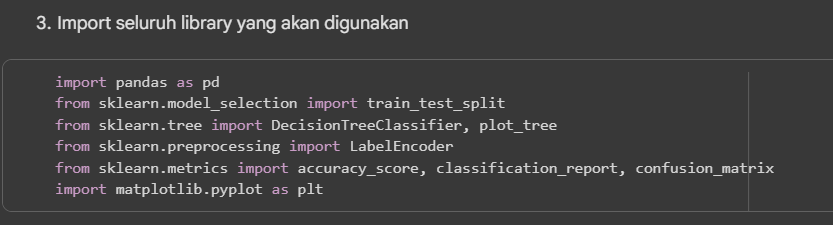
*Gambar 1. 1*

Pada *Gambar 1.1* di atas, merupakan sebuah code untuk mounted atau menghubungkan google colab dengan google drive.



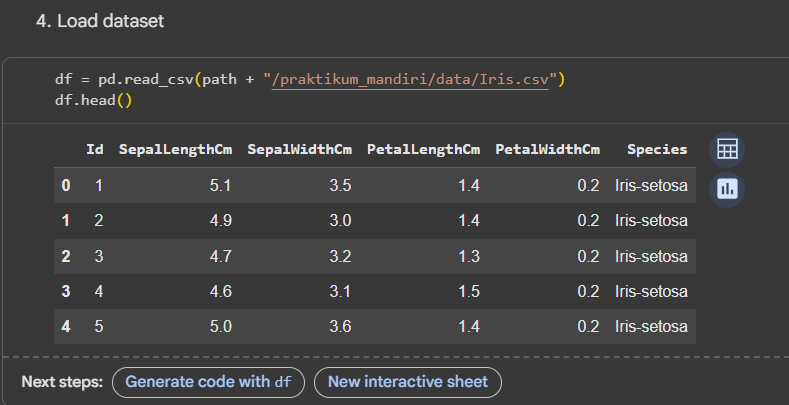
*Gambar 1. 2*

Pada *Gambar 1.2* di atas, kita membuat sebuah path untuk mengatur folder yang akan digunakan, dalam case ini ita menggunakan folder pertemuan04.



*Gambar 1. 3*

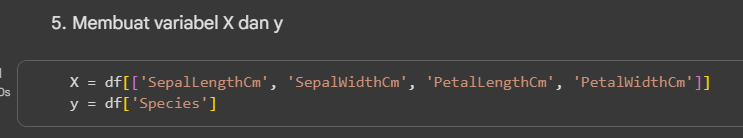
Pada *Gambar 1.3* di atas, merupakan sebuah code untuk import beberapa library yang akan kita gunakan.



*Gambar 1. 4*

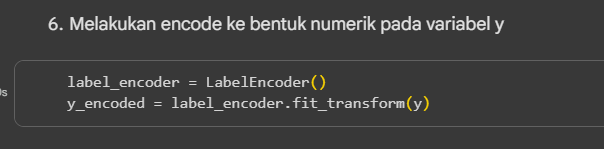
Pada *Gambar 1.4* di atas, merupakan code untuk menampilkan dataset pada sebuah tabel menggunakan dataset calonpembelimobil.csv dan hanya menampilkan 5 data.

• **df = pd.read\_csv(path + ‘/praktikum\_mandiri/data/Iris.csv)** merupakan sebuah code untuk membaca file csv yang terdapat pada folder data dengan diikuti file Iris.csv.



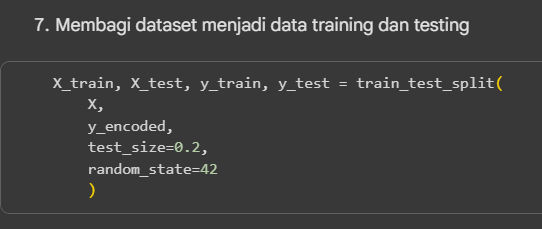
*Gambar 1. 5*

Pada *Gambar 1.5* di atas, merupakan sebuah code untuk memisahkan variabel X dan Y.



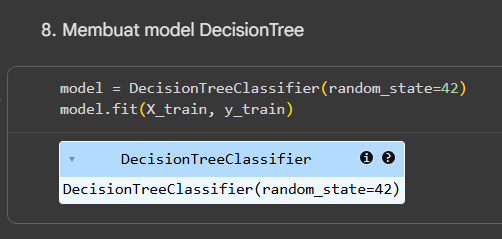
*Gambar 1. 6*

Pada *Gambar 1.6* di atas, merupakan sebuah code untuk melakukan encode column dataset menjadi numerik.



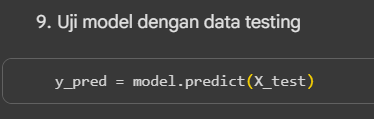
*Gambar 1. 7*

Pada *Gambar 1.7* di atas, merupakan sebuah code untuk melakukan pembagian data training dan data testing.



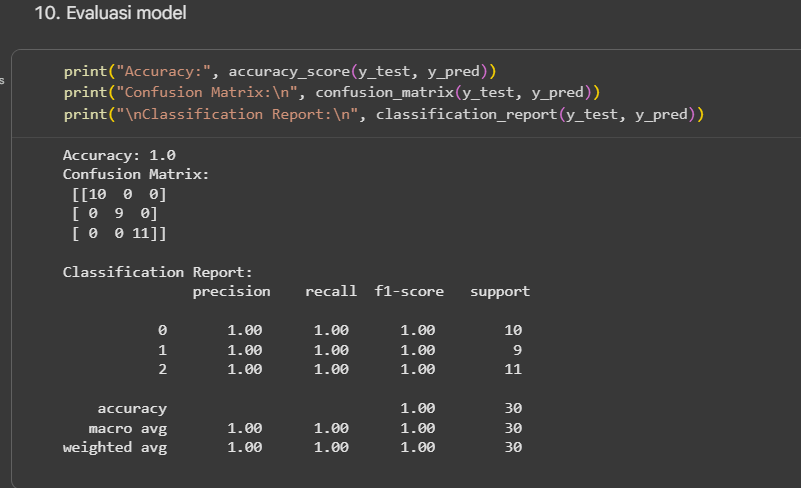
*Gambar 1. 8*

Pada *Gambar 1.8* di atas, membuat model DecisionTree.



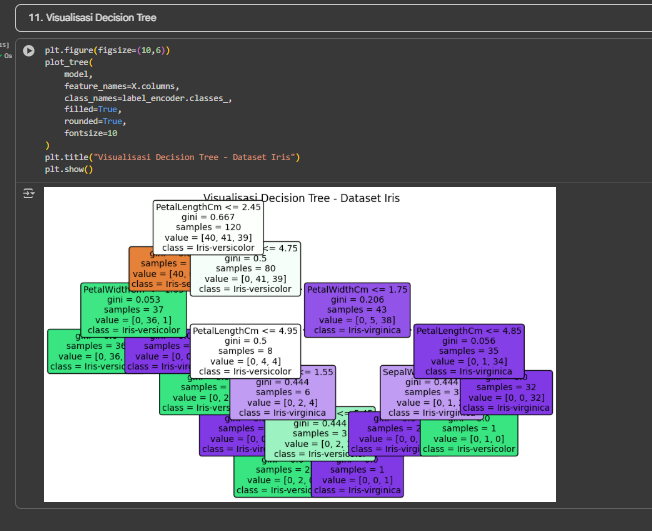
*Gambar 1. 9*

Pada *Gambar 1.9* di atas, merupakan sebuah code untuk menguji model dengan data testing.



*Gambar 1. 9*

Pada *Gambar 1.9* di atas, merupakan sebuah code untuk evaluasi model dengan data uji atau testing.



*Gambar 1. 10*

Pada *Gambar 1.10* di atas, merupakan sebuah kode membuat visualisasi dari decision tree.

Link github praktikum:

<https://github.com/MuhZaidanRamdhan/machine-learning/blob/main/pertemuan05/praktikum_05/notebooks/praktikum_05.ipynb>

Link github tugas:

<https://github.com/MuhZaidanRamdhan/machine-learning/blob/main/pertemuan05/praktikum_mandiri/notebooks/latihan_05.ipynb>