Tugas 5 : Implementasi Decision Tree untuk klasifikasi Iris

**Syayid Muhammad Akbar - 0110222086 1\***

1 Teknik Informatika, STT Terpadu Nurul Fikri, Depok

\*E-mail: [syay22086@student.nurulfikri.ac.id](mailto:syay22086@student.nurulfikri.ac.id)

1. Read CSV in python using pandas (Praktikum)

Langkah pertama untuk melakukan analisis data atau pemrosesan data menggnakan machine learning adalah membaca data itu sendiri, menggunakan read\_csv (karena data berbentuk CSV)

A screenshot of a computer

AI-generated content may be incorrect.

*Gambar 1. Load data*

Hasil setelah kita running akan menghasilkan data tersebut:

A screenshot of a graph

AI-generated content may be incorrect.

2. Gain information from dataset

Langkah selanjutnya yaitu melihat informasi dasar yang dimiliki oleh dataset. Dalam pandas dapat menggunkan info(), untuk mengekstraksi informasiyang terdapat dalam dataset.

A screenshot of a computer

AI-generated content may be incorrect.

Setelah mengetahui tipe data dari setiap kolom maka, Langkah selanjutnya dan tidak kalah penting yaitu melakukan pengecheckkan terhadap missing value dari data tersebut.

A screenshot of a computer

AI-generated content may be incorrect.

*Gambar 5. Load data mandiri dataset*

**3. Modeling**

Bagian terpenting yaitu modelling menggunkana decision tree, algoritma ini termasuk kedalan jenis machine learning supervised yang dapat mengklasifikasikan secara otomatis berdasarkan gini yang sudah ditentukan.

A screenshot of a computer

AI-generated content may be incorrect.

Sebelum melakukan klasifikasi kita akan merubah kolom species dengan integer, menggunkan metode encoder, atau merubah string/object kedalam bentuk numerik tanpa menghilangkan essence mereka.

A screenshot of a computer program

AI-generated content may be incorrect.

Spliting data, setelah melakukan encoder kita dapat membagi atau splitting data kedalan 2 bagian yaitu fitur dan target, Dimana fitur berisi kolom kolom X dan target berisi kolom Y.

Kita juga men-setting bagaimana ksistem pembagian yang akan dilakukan computer, 20% data test dan 80% data train.

A screen shot of a computer

AI-generated content may be incorrect.

Hasil setelah kita menerapkan decision tree untuk melakukan klasifikasi, dapat dilihat bahwa model tersebut dapat menghasilkan akurasi sebesar 0.93

Dibawah merupakan hasil dari nilai perbandingan actual value dan predicted value

A screenshot of a computer

AI-generated content may be incorrect.