

**Nama** : Muhammad Raihan Butar-Butar

**NIM** : 1103213077

## **Decision Tree**

Decision Tree (Pohon Keputusan) adalah model prediktif dalam bidang analisis data yang menggambarkan serangkaian keputusan dan konsekuensinya dalam bentuk pohon. Model ini digunakan untuk melakukan klasifikasi dan regresi pada data dengan mengambil keputusan berdasarkan fitur-fitur tertentu. Berikut adalah bagian-bagian yang menyusun Decision Tree:

- Simpul Akar: Mewakili pertanyaan awal untuk memulai proses pengambilan keputusan.
- Simpul Keputusan: Mewakili jawaban/keputusan yang memungkinkan untuk pertanyaan tersebut.
- Simpul Daun : Simpul yang tidak memiliki cabang lagi. Simpul inilah yang merupakan hasil akhir dari decision tree.
- Cabang: Merupakan garis yang menghubungkan satu simpul dengan simpul lainnya.

Proses pembuatan decision tree dibagi menjadi beberapa tahap. Pertama pengumpulan dan persiapan data. Pada tahap ini, data yang diperlukan untuk membuat model decision tree dikumpulkan dan dipersiapkan. Data yang diperlukan terdiri dari variabel target atau dependen yang akan diprediksi, dan variabel prediktor atau independen yang digunakan sebagai acuan dalam membuat keputusan. Data yang diberikan dapat berupa data numerik atau kategorikal.

Selanjutnya, Pembentukan decision tree. pada tahap ini, decision tree dibentuk dari data yang telah dikumpulkan. Proses ini dilakukan dengan menggunakan algoritma decision tree seperti ID3, C4.5, CART, atau CHAID. Algoritma-algoritma tersebut akan memilih variabel prediktor yang paling informatif untuk dijadikan pemisah pada setiap cabang dari decision tree.

Setelah decision tree terbentuk, dilakukan pruning atau pemangkasan pada cabang-cabang yang tidak signifikan atau tidak memberikan kontribusi pada prediksi. Proses pruning bertujuan untuk menghindari overfitting atau kelebihan fitting pada model decision tree.

Setelah decision tree terbentuk dan dipangkas, model decision tree dievaluasi dengan menggunakan data yang belum dipakai selama pembuatan decision tree. Evaluasi dilakukan dengan menghitung akurasi, presisi, recall, F1 score, dan metrik lainnya untuk menilai dan performa model decision tree.

Setelah decision tree terbukti memiliki performa yang baik, model decision tree dapat digunakan untuk memprediksi nilai target atau dependen pada data baru. Proses ini dilakukan dengan menggunakan model decision tree yang telah dibentuk dan diuji sebelumnya.