

Nama : Muhammad Alfi Syahrin  
NIM : 1103213144  
Kelas : TK-45-05

## **Catatan Decision Dan Classification Tree**

Decision tree secara general membuat pernyataan dan membuat keputusan berdasarkan pernyataan tersebut benar (True) atau salah (False). Ketika decision tree mengklasifikasikan sesuatu menjadi kategori, maka disebut sebagai pohon klasifikasi. Dan ketika decision tree melakukan prediksi nilai numerik, maka disebut sebagai pohon regresi. Pada decision tree, kita dapat menambahkan data numerik dengan data “Ya” atau “Tidak”. Pada decision tree, bagian teratas disebut sebagai root atau akar. Root akan memberikan panah ke branches atau cabang dan diteruskan oleh branches sampai leaf atau daun yang merupakan titik terakhir dan tidak memiliki panah ke bawahnya yang akan berperan sebagai output. Penentuan akar dan cabang berdasarkan tabel yang disediakan dapat dilakukan dengan cara melihat hubungan antar kolom. Dalam video ini, hubungan antar kolom dapat kita lihat menggunakan metode Gini Impurity. Gini Impurity memiliki rumus  $\text{Impurity} = 1 - (\text{probability of "Yes"})^2 - (\text{probability of "No"})^2$ . Root berdasarkan hubungan antara salah satu kolom dengan kolom target dengan Gini Impurity terkecil. Apabila cabang masih memiliki 2 kemungkinan yang terjadi, maka bisa dilanjutkan lagi menggunakan kolom lainnya yang memiliki gini impurity terkecil. Namun, jika data kolom target terlalu sedikit untuk dilakukan pemecahan lagi (Contoh data “Yes” ada 3 dan data “No” ada 1), bisa terjadi adanya “Overfit”. Jika hal tersebut terjadi, kita dapat membiarkan data tersebut mengikuti mayoritas dari semua datanya sebagai output.