

Nama : Ilham Khodar Trijaya  
NIM : 1103213084  
Kelas : TK-45-04

Video ini berisi mengenai cara membangun dan mengevaluasi *Random Forest* atau “hutan acak”. Decision Tree mudah dibuat, digunakan, dan diinterpretasikan. Namun dalam prakteknya, Decision Tree kurang baik untuk pembelajaran statistik. Oleh karena itu, diperlukannya “hutan acak” karena menggabungkan kesederhanaan dan fleksibilitas Decision Tree sehingga lebih akurat.

Langkah Membuat “Hutan Acak” “

1. Membuat dataset Bootstrap

Bootstrap bekerja dengan melakukan resampling pada dataset yang sudah ada, yaitu dengan mengambil sampel berulang kali secara acak dengan pengembalian dari data tersebut. Proses ini dilakukan untuk mensimulasikan distribusi populasi berdasarkan sampel yang ada, sehingga dapat digunakan untuk berbagai keperluan, dataset tersebut bisa berisi data numerik atau kategorikal, dan bisa berasal dari berbagai bidang seperti keuangan, kesehatan, atau penelitian sosial.

2. Membuat Decision Tree dengan dataset Bootstrap

Menggunakan metode “Hutan Acak” kita hanya perlu menggunakan subset variabel. Alih-alih mempertimbangkan keempat variabel, kita secara acak memilih 2 variabel.

Random forest memang menggunakan subset acak (random subset) untuk membangun setiap pohon individual di dalam forest. Ini merupakan salah satu aspek kunci yang membedakannya dari pohon keputusan (decision tree) biasa dan berkontribusi pada kemampuannya untuk mengurangi overfitting dan meningkatkan generalizability.

Dalam evaluasi performa Random Forest, Out-of-Bag (OOB) error memainkan peran penting. OOB error mengacu pada kesalahan rata-rata yang dibuat oleh Random Forest pada data yang tidak digunakan untuk melatih pohon individual di dalam forest.