

Nama : Alfian Mohamad Firdaus

NIM : 1103204002

## **Rangkuman Random Forest StatQuest**

Stat Quest membahas Random Forest dalam serangkaian video, di mana dalam bagian pertama ini, pembuat video, Josh Dharma, menjelaskan langkah-langkah membangun dan mengevaluasi Random Forest. Random Forest merupakan kombinasi dari pohon keputusan (Decision Trees) yang memiliki tingkat akurasi lebih tinggi daripada model Decision Trees tunggal. Langkah pertama dalam membangun Random Forest adalah membuat dataset bootstrap dengan mengambil sampel acak dari dataset asli dengan pengembalian (duplicates allowed). Kemudian, pohon keputusan dibangun menggunakan dataset bootstrap ini, namun hanya menggunakan subset acak dari variabel atau kolom pada setiap langkahnya.

Langkah selanjutnya adalah menggunakan Random Forest yang telah dibuat untuk melakukan prediksi pada data baru. Data tersebut dijalankan melalui setiap pohon keputusan dalam Random Forest, dan hasilnya dihitung berdasarkan mayoritas suara dari semua pohon tersebut. Dalam mengevaluasi kinerja Random Forest, Josh menjelaskan konsep Out-of-Bag (data yang tidak termasuk dalam dataset bootstrap). Out-of-Bag samples digunakan untuk mengestimasi tingkat akurasi model, dan Out-of-Bag error dihitung sebagai proporsi sampel yang salah diklasifikasikan.

Selanjutnya, Josh membahas pengaturan optimal pada Random Forest, khususnya jumlah variabel yang digunakan pada setiap langkahnya. Pengaturan ini diuji dengan mencoba beberapa nilai, dan dipilih yang memberikan akurasi paling tinggi. Video ini memberikan gambaran lengkap tentang cara membangun, menggunakan, dan mengevaluasi Random Forest untuk meningkatkan performa model prediktif.