

Nama : M Rayyan Aqiilah Manna

NIM : 1103210210

## **Random Forests**

Random Forests adalah suatu metode dalam machine learning yang termasuk dalam kategori ensemble learning. Ensemble learning merupakan pendekatan di mana beberapa model pembelajaran digabungkan untuk meningkatkan kinerja dan ketepatan prediksi dibandingkan dengan menggunakan satu model tunggal.

Random forests terbentuk dari decision tree. Tree memiliki satu aspek yang menghalanginya untuk menjadi alat predictive learning, yaitu ketidak akuratannya. Tree akan bekerja dengan baik saat menggunakan data yang membentuknya tetapi kurang fleksibel untuk mengklasifikasikan sampel baru. Pada random forests, kita menggabungkan simplisitas decision tree dengan fleksibilitas lebih sehingga akan sangat membantu keakuratannya.

Berikut adalah langkah-langkah membuat random forests :

### **1. Membuat bootstrapped dataset**

Untuk membuat bootstrapped dataset, kita akan mengambil sampel acak dari dataset yang kita analisis. Pemilihan dapat dilakukan lebih dari satu kali pada data yang sama.

### **2. Membuat decision tree**

Langkah selanjutnya, kita akan membuat decision tree dari bootstrapped dataset yang baru saja kita buat. Tetapi, kita hanya akan menggunakan subset variabel atau kolom pada setiap step.

### **3. Ulangi step 1 dan 2 berkali kali sampai kita memiliki sekitar 100 tree.**

4. Ketika kita sudah memiliki banyak tree, kita dapat mengklasifikasi sampel baru menggunakan random forest dengan mengikuti tree yang sudah dibentuk tadi.

Chest Pain	Good Blood Circ.	Blocked Arteries	Weight	Heart Disease
Yes	No	No	168	

...and now we want to know if they have heart disease or not.

Pada gambar tersebut kita ingin mengetahui apakah sampel tersebut termasuk heart disease atau tidak. Kita dapat mengikuti alur per tree sehingga menemukan jawaban ya atau tidak, yang akan dicatat jumlahnya. Jika "Ya" memiliki hasil lebih banyak daripada "Tidak" setelah melalui semua tree yang dibuat, maka kolom tersebut dapat diisi dengan "Ya".