

**Nama: Irawan Mardiansyah**

**Keals: TK-45-04**

**NIM: 1103210174**

Keunggulan utama decision tree yaitu kemudahan dalam pembuatan, penggunaan serta interpretasinya. Selain punya keunggulan decision tree juga punya kelemahan, yaitu terletak pada akurasi. Model ini bekerja sangat baik pada data yang sedang digunakan untuk pembuatannya, namun kurang efektif saat mengklasifikasikan sampel yang baru.

Random Forest menggabungkan decision tree dengan fleksibilitas, menghasilkan peningkatan akurasi yang signifikan. Keuntungan random forest yaitu adanya penggunaan sampel bootstrap dan subkumpulan variabel yang bisa menghasilkan beragam pohon, yang dapat meningkatkan efektivitas random forest dibandingkan decision tree individual. Ada beberapa langkah-langkah ketika saat membuat Random Forest, berikut langkah – langkah nya:

1. Langkah pertama, yaitu membuat kumpulan data bootstrap. Kumpulan data bootstrap dibuat dengan memilih data secara acak memakai data asli dengan ukuran yang sama.
2. Langkah kedua, yaitu membuat decision tree. Cara membuat decision tree yaitu menggunakan dataset bootstrap dengan sekumpulan kategori variabel acak pada setiap langkahnya. Misalnya adalah, hanya dua variabel yang dipertimbangkan pada setiap langkah tersebut.
3. Langkah ketiga, yaitu mengulang langkah 1 dan 2. Langkah 1 dan 2 akan diulang beberapa kali agar bisa menghasilkan beragam pohon.

Random Forest adalah metode yang lebih efektif daripada decision tree individual karena memiliki akurasi yang lebih tinggi dan fleksibilitas yang lebih baik ketika mengklasifikasikan sampel baru. Ada beberapa pengoptimalan random forest yang bisa dilakukan, yaitu:

- a. Menentukan jumlah dari setiap variabel: Di langkah pertama pembuatan pohon hanya dua variabel yang digunakan untuk membuat keputusan pada setiap langkah.
- b. Membandingkan error out-of-bag: Error out-of-bag dapat dibandingkan untuk random forest dengan dua variabel per langkah dan random forest dengan tiga variabel per langkah.
- c. Pengujian pengaturan yang berbeda: Berbagai pengaturan harus diuji dan menggunakan metode random forest dengan akurasi tertinggi yang dipilih.
- d. Proses optimasi:
  - Membuat random forest.
  - Memperkirakan keakuratan dari metode random forest.
  - Mengubah jumlah variabel yang digunakan pada setiap langkah.
  - Mengulangi beberapa langkah berkali-kali.
  - Memilih random forest mana yang akurasi paling tinggi.