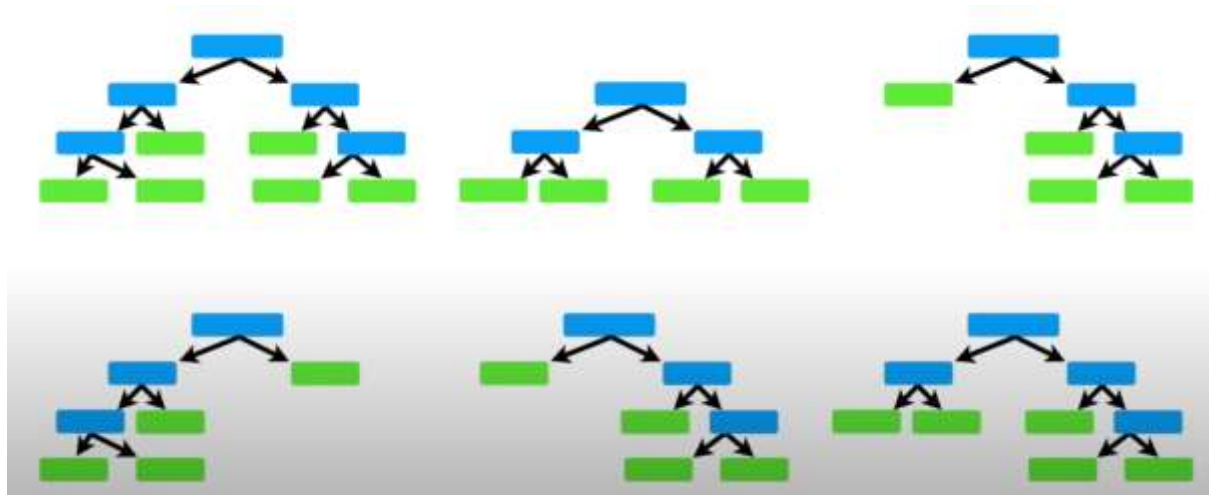


Nama: Arif Muhammad Laduni
NIM: 1103213003

Random Forest

Random Forest sendiri dibangun berdasarkan dari *decision trees*. *Decision trees* sendiri mudah dibuat, mudah digunakan, dan mudah diinterpretasikan. Namun dalam praktiknya, mereka tidak terlalu bagus karena mereka tidak fleksibel. Ketika harus mengklasifikasi sebuah sampel baru. Maka, pada *random forest* menggabungkan kesederhanaan pohon *decision trees* dengan fleksibilitas menghasilkan peningkatan besar dalam akurasi.



Pembuatan *random forest* terdapat beberapa tahap yang mana pada Langkah pertama yaitu menciptakan *bootstrap* dataset. *Bootstrap* dataset dibuat dengan memilih sampel secara acak dari data asli, dengan memperbolehkan pemilihan sampel yang sama lebih dari satu kali. Setelah itu, dibangun *Decision Tree* menggunakan *bootstrap* dataset tersebut, tetapi hanya menggunakan subset acak dari variabel atau kolom pada setiap langkahnya. Penggunaan *Random Forest* dilakukan dengan menjalankan data baru melalui setiap *Decision Tree* yang telah dibangun. Setiap pohon memberikan vote untuk kategori yang sesuai, dan hasil akhir diambil dari mayoritas suara.

Untuk mengevaluasi keberhasilan *Random Forest*, digunakan *out-of-bag* dataset. Ini adalah data yang tidak termasuk dalam *bootstrap* dataset karena diizinkan adanya duplikasi. *Out-of-bag* dataset digunakan untuk mengukur akurasi model, dan *out-of-bag* error dihitung sebagai proporsi sampel yang salah diklasifikasikan oleh *Random Forest*.