

Random Forest

Random Forest adalah teknik pembelajaran mesin yang menggunakan banyak pohon keputusan untuk membuat prediksi yang akurat. Random Forest menggunakan subset data pelatihan dan subset fitur untuk membangun setiap pohon secara acak. Random Forest mengumpulkan hasil dari setiap pohon saat membuat prediksi dan menggunakan mayoritas suara untuk masalah klasifikasi atau rata-rata untuk masalah regresi untuk membuat prediksi akhir. Metode ini mengurangi overfitting dan meningkatkan kinerja dan kestabilan model, membuatnya populer dalam banyak aplikasi prediksi dan klasifikasi.

Langkah Langkah membuat dataset :

- Membuat bootstrap dataset dengan menggunakan dataset asli secara acak, duplikasi diperbolehkan dalam pengambilan data ini
- Membuat decision tree menggunakan bootstrap dataset namun dengan memilih kolom secara acak untuk simpul di dalam pohonnya, decision tree ini dibuat berkali kali dengan semua kemungkinan yang ada dari bootstrapped dataset.
- lakukan evaluasi dengan menggunakan data out of bag dataset, dan hitung seberapa akurat kesimpulan yang dihasilkan dari semua decision tree yang ada.