

Rangkuman Video: Random Forest

Kinanti Rahayu Az-Zahra 1103210052 TK 45 04

Hutan acak merupakan kumpulan pohon keputusan yang memiliki kemampuan mengatasi keterbatasan fleksibilitas prediksi yang dimiliki oleh pohon keputusan tunggal. Pohon keputusan dikenal sebagai alat analisis data yang sederhana dalam pembuatan, penggunaan, dan interpretasi. Meskipun demikian, pohon keputusan memiliki kelemahan karena terbatas dalam kemampuan mengklasifikasikan sampel baru dengan baik. Hutan acak memadukan kelebihan sederhana pohon keputusan dengan fleksibilitas yang lebih tinggi, yang menghasilkan peningkatan akurasi prediksi yang signifikan.

Berikut merupakan Langkah langkah cara pembuatan *Random Forest*;

- Langkah pertama adalah membuat kumpulan data bootstrap dengan memilih sampel secara acak dari kumpulan data asli dengan penggantian.
- Langkah kedua adalah membuat pohon keputusan menggunakan dataset bootstrap, tetapi hanya mempertimbangkan subset variabel acak pada setiap langkah.

Ketika terdapat pasien baru dengan data pengukuran, proses prediksi dimulai dengan menjalankan data melalui seluruh pohon yang ada dalam hutan acak. Setiap pohon memberikan suara untuk kategori tertentu berdasarkan data tersebut, dan kategori yang mendapat suara terbanyak dianggap sebagai prediksi akhir untuk pasien tersebut.

Untuk mengukur keakuratan hutan acak, digunakan proporsi sampel di luar kantong yang diklasifikasikan dengan benar oleh model. Metode out-of-bag (OOB) diterapkan dengan menguji sampel yang tidak dimasukkan dalam dataset bootstrap, memberikan gambaran objektif tentang seberapa baik hutan acak dapat mengklasifikasikan data baru.

Upaya untuk meningkatkan akurasi hutan acak dapat dilakukan dengan mengoptimalkan jumlah variabel yang digunakan per langkah pembangunan pohon. Eksperimen dilakukan dengan mencoba berbagai pengaturan dan memilih kombinasi variabel yang memberikan akurasi terbaik, membantu meningkatkan performa model secara keseluruhan.

Dengan menggabungkan keunggulan pohon keputusan dengan fleksibilitas prediksi, hutan acak mampu meningkatkan akurasi dalam memprediksi kategori sampel baru. Proses pengoptimalan model melibatkan pembuatan beberapa pohon dengan dataset bootstrap yang berbeda dan subset variabel yang beragam, menciptakan pendekatan prediktif yang lebih efektif dan handal. Dengan demikian, penerapan hutan acak menjadi strategi yang kuat untuk meningkatkan keakuratan prediksi dalam konteks pemodelan dan analisis data.