

Muhammad Fauzan Aldi

1103210049

TK45-05

Rangkuman

Random Forest Part-1 Building, Using and Evaluating

Dalam video ini, Josh Starmer dari channel StatQuest menjelaskan tentang Random Forests, pertama Josh menjelaskan tentang decision tree atau pohon keputusan, decision tree meskipun mudah dibuat dan diinterpretasikan juga memiliki keterbatasan fleksibilitas dalam mengklasifikasikan sampel baru, yang dapat memengaruhi akurasi. Dengan mengutip "The Elements of Statistical Learning," josh menjelaskan bahwa meskipun decision tree berfungsi dengan baik tetapi tetap tidak fleksibel ketika harus mengklasifikasikan sampel baru, tetapi baik nya adalah Random Forests menggabungkan decision tree yang simple dengan fleksibilitas, menghasilkan peningkatan akurasi yang signifikan. Kemudian yang kedua proses pembuatan Random Forest yaitu pertama membuat dataset bootstrap dengan mengambil data random dari dataset yang asli, kemudian yang kedua membuat decision tree dari subset variabel secara random, yang ketiga ulangi step pertama dan kedua tadi dengan membuat dataset bootstrap dan membuat decision tree nya. Yang ketiga menjelaskan tentang penilaian akurasi dilakukan menggunakan sampel out of bag, yang tidak termasuk dalam bootstrap mengukur sejauh mana model dapat mengklasifikasikan sampel secara benar. Dengan contoh seorang pasien baru, Josh menunjukkan bagaimana data pasien dapat dijalankan melalui setiap decision tree yang tadi dibuat, dan hasilnya diakumulasikan untuk membuat keputusan akhir. Kemudian yang keempat atau yang terakhir, Josh menjelaskan tentang evaluasi akurasi model Random Forest menggunakan konsep sampel out of bag, dalam memproses sampel yang tidak termasuk dalam dataset bootstrap, dan bagaimana hasilnya diukur untuk mengestimasi tingkat akurasi secara keseluruhan.