

Nama : Alfian Mohamad Firdaus

NIM : 1103204002

Rangkuman Decision-Tree StatQuest

StatQuest menjelaskan dengan jelas konsep dasar dari pohon keputusan (decision trees) dan bagaimana cara membangunnya dari data mentah. Pohon keputusan digunakan untuk klasifikasi dan prediksi, dengan menggabungkan pernyataan dan keputusan berdasarkan kebenaran pernyataan tersebut. Pohon keputusan terdiri dari simpul akar (root node), simpul internal (internal nodes), dan simpul daun (leaf nodes). Setiap simpul mewakili pertanyaan atau pernyataan, dan hasil dari pohon keputusan diambil dari simpul daun.

Penjelasan dilanjutkan dengan contoh pembangunan pohon keputusan menggunakan data mentah, dimana kita mengukur ketidakmurnian (impurity) dari setiap simpul untuk menentukan pertanyaan terbaik untuk ditempatkan di akar pohon. Ketidakmurnian diukur menggunakan berbagai metode seperti *gini impurity*, dan pemilihan pertanyaan di akar pohon didasarkan pada metode yang menghasilkan ketidakmurnian terendah. Setelah pohon dibangun, kita dapat menggunakan pohon tersebut untuk memprediksi kategori output dari data baru dengan menjalankan data tersebut melalui pohon keputusan sesuai dengan aturan yang telah ditentukan.

Terakhir, StatQuest juga menyoroti pentingnya mencegah overfitting dengan teknik seperti pruning atau menetapkan batasan pertumbuhan pohon. Penanganan overfitting penting karena dapat mengakibatkan model yang kurang akurat dalam memprediksi data baru. Dengan memahami konsep dasar dan teknik praktis dalam membangun dan menggunakan pohon keputusan, kita dapat membuat model yang lebih efektif dan akurat dalam memecahkan masalah klasifikasi dan prediksi.