

Nama : Afif Ibadurrahman Jalaluddin

Nim : 1103210221

Kelas : TK4503

Decision Tree-StatQuest

Video tersebut membahas tentang *decision trees* dan *classification trees*. *decision trees* adalah model yang membuat pernyataan dan mengambil keputusan berdasarkan kebenaran pernyataan tersebut. Jika pohon tersebut digunakan untuk mengelompokkan data ke dalam kategori, disebut sebagai *classification trees*; sedangkan jika untuk memprediksi nilai numerik, disebut sebagai *regression tree*.

decision trees dapat menggunakan kombinasi data numerik dan ya/tidak (binary) dalam satu pohon. Saat membangun pohon, kita mulai dari simpul atas yang disebut *root*, kemudian bercabang ke *branches* dan berakhir di simpul *leaves*. Setiap simpul memiliki arah cabang yang menunjukkan keputusan (*true/false*) berdasarkan pernyataan yang diberikan.

Dalam konteks *classification trees*, *impurity* dari simpul *leaves* dapat diukur dengan berbagai metode, salah satunya adalah *genie impurity*. Pemilihan pernyataan dan *impurity* diukur untuk membangun pohon agar dapat memprediksi kategori dengan akurasi tinggi.

Pada akhirnya, video tersebut menjelaskan bahwa terdapat dua metode umum untuk mengatasi *overfitting* pada *decision trees*, yaitu *pruning* dan pembatasan pertumbuhan pohon dengan menetapkan batas jumlah data pada setiap simpul *leaves*. Proses pembangunan pohon juga melibatkan uji coba nilai-nilai tertentu menggunakan metode *cross-validation* untuk memilih parameter yang paling optimal.

Secara keseluruhan, *decision trees* adalah alat yang berguna dalam klasifikasi dan prediksi berdasarkan data yang diberikan.