

Nama : Jhose Immanuel Sembiring
NIM : 1103202047
Kelas : TK44G4

Decision and Classification Trees

Decision Trees merupakan representasi struktur hierarkis yang mengambil keputusan berdasarkan kondisi true atau false.

Classification Trees digunakan untuk mengategorikan data atau memprediksi nilai numerik.

Struktur Decision Trees

Decision Trees terdiri dari serangkaian node yang mewakili keputusan dan cabang yang mengarah pada hasil akhir. Node akhir disebut daun yang menunjukkan klasifikasi akhir setelah proses pengambilan keputusan. Pembentukan pohon klasifikasi dimulai dari node akar dan kemudian memisahkan data berdasarkan kondisi yang diberikan hingga mencapai daun. Proses ini memungkinkan pengkategorian data berdasarkan serangkaian keputusan yang diambil.

Pembangunan Decision Trees

Dalam pembangunan pohon klasifikasi, evaluasi impuritas daun dilakukan menggunakan metode seperti Gini Impurity. Data numerik seperti usia memerlukan ambang batas (thresholds) untuk memisahkan data dengan efektif. Proses ini memungkinkan identifikasi fitur-fitur yang paling informatif dalam pembentukan pohon.

Pemangkasan dan Pembatasan Decision Trees

Dalam mengatasi masalah overfitting, yaitu ketika pohon terlalu kompleks dan terlalu cocok dengan data latih, bisa dilakukan melalui pemangkasan atau pembatasan pertumbuhan pohon. Pemangkasan melibatkan penghapusan bagian dari pohon untuk mengurangi kompleksitas, sementara pembatasan pertumbuhan mengatur ambang batas untuk impuritas daun atau ukuran sampel. Teknik validasi seperti cross-validation membantu menentukan efektivitas model pohon. Selain itu, tersedia pula sumber belajar tambahan seperti panduan belajar StatQuest dan opsi dukungan untuk yang tertarik dalam eksplorasi statistik dan pembelajaran mesin lebih lanjut.