

Muhammad Harits

1103223159

Bab 6: Decision Trees (Pohon Keputusan)

Ringkasan Bab

Bab ini memperkenalkan Decision Trees, salah satu algoritma ML yang paling intuitif dan mudah dipahami. Model ini menjadi dasar dari algoritma ensemble yang sangat kuat, terutama Random Forest.

Konsep Utama

- Cara kerja Decision Tree
Model belajar dengan membangun serangkaian aturan keputusan sederhana (if-else) yang membagi data menjadi kelompok-kelompok semakin homogen.
- Model "White Box"
Tidak seperti SVM atau Neural Network yang sulit diinterpretasikan (black box), Decision Tree mudah dipahami. Kita bisa melihat langsung bagaimana model memutuskan suatu prediksi.
- Algoritma CART (Classification and Regression Trees)
Digunakan oleh Scikit-Learn. CART membagi data secara rekursif menjadi dua subset dengan mencari fitur dan threshold terbaik untuk menghasilkan node paling "murni".
- Impurity Measures (Ukuran Kemurnian)
 - Gini Impurity (default)
 - Entropy (berbasis teori informasi)
- Regularisasi untuk mencegah overfitting
Hyperparameter penting:
 - max_depth
 - min_samples_split
 - min_samples_leaf
- Kelemahan
Sangat mudah overfitting dan sensitif terhadap sedikit perubahan data.
- Decision Tree untuk Regresi
Bukan memprediksi kelas, tetapi memprediksi nilai rata-rata dari data pada node daun.