

Chapter 6 : Summary

Diedrick Darrell Darmadi - 1103223031

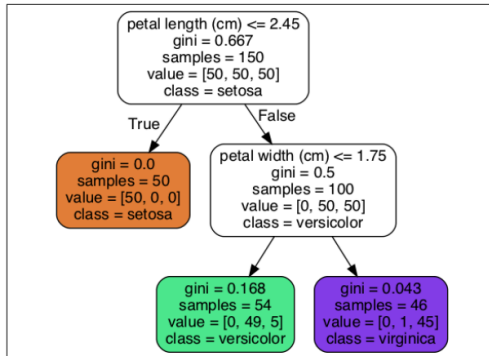


Figure 6-1. Iris Decision Tree

Bab 6 membahas konsep dasar Decision Tree sebagai model yang bekerja dengan membelah ruang fitur menjadi area-area yang lebih kecil berdasarkan aturan sederhana. Contoh awal ditunjukkan dengan model yang memisahkan data iris menggunakan kondisi berbasis fitur seperti petal length untuk membuat keputusan klasifikasi.

Selanjutnya dijelaskan bagaimana algoritma memilih fitur dan threshold terbaik menggunakan kriteria impurity seperti Gini impurity atau entropy. Proses pembelahan berulang digambarkan sebagai struktur pohon dengan node keputusan dan leaf node yang memuat prediksi

Bab ini kemudian membahas proses pembuatan model yang terlalu kompleks jika semua pembelahan dibiarkan hingga kedalaman maksimal, sehingga menyebabkan overfitting. Contoh visual keputusan yang sangat mengikuti data ditunjukkan untuk menggambarkan masalah ini

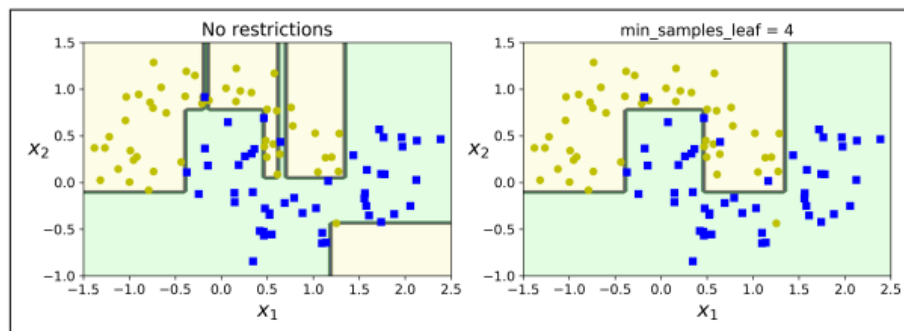


Figure 6-3. Regularization using *min_samples_leaf*

Untuk mengatasi overfitting, digunakan teknik pruning seperti membatasi kedalaman pohon, membatasi jumlah sampel minimum dalam sebuah node, atau melakukan *cost complexity pruning*. Grafik perbandingan pohon yang dipangkas dan tidak dipangkas ditampilkan untuk menunjukkan dampaknya terhadap generalisasi.

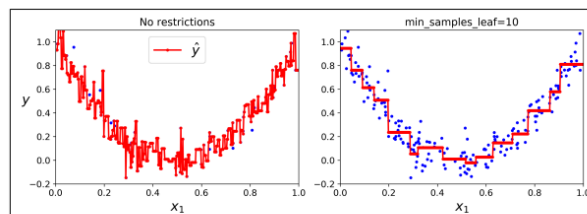


Figure 6-6. Regularizing a Decision Tree regressor

Bab ini juga membahas bagaimana decision tree bekerja pada regresi, di mana pohon membagi ruang data dan memberikan nilai rata-rata pada setiap leaf. Contoh fungsi regresi non-linear ditampilkan untuk memperlihatkan bagaimana model menangkap pola tetapi rentan terhadap overfitting. Terakhir,

dijelaskan limitasi decision tree seperti sensitivitas terhadap rotasi data dan instabilitas terhadap perubahan kecil pada dataset, serta bagaimana penggunaan ensemble akan dibahas pada bab selanjutnya untuk mengatasi kekurangan ini.