Machine Learning (WiSe 2025/2026)

Author: Suvansh Shukla
Matriculation No. 256245

Assignment 4 Task 4.4

Overfitting in decision trees is an issue where a decision tree model starts to "over-correct" itself to match outliers and noisy data.
A good way to figure out if your decision tree is overfitted is by seeing how many levels it has. According to "Occam's Razor" we can attribute too many levels in a decision tree to overfitting.

Decision trees are "greedy" models, meaning they keep on bending/breaking to accomodate all possible data points even though the target
function may not necessarily exhibit those characteristics.

Overfitted decision trees may have the following characteristics:

- too complex
- too accomodating of noise
- doesn't stop early or properly, keeps on going

Overfitting may be prevented in decision trees by:

- limiting the number of levels/branches they can have
- by using an "mix of experts" approach, where specific subtrees are created/chosen to handle features for specific data
- remove irrelevant features
- prune weaker branches
- set a minimum samples per split or per leaf