To simplify the visibility of the decision tree in the iris dataset classification task, I decided to encode Decision Tree Classifier for the Iris Dataset, code the values of each of the four input features (petal width, petal length, sepal length and sepal width) by Lucas Herranz Gancedo (M11351802) to one of four labels (a, b, c or d). The strategy for encoding follows below rules for all features: class 0: Setosa $a: MIN \leq val < \frac{MIN + MEAN}{2}, \; b: \frac{MIN + MEAN}{2} \leq val < MEAN,$ class 1: Versicolor class 2: Virginica Assuming all samples in final tree Level 4 are classified as: class1, the decision tree would misclassify 7 samples only. Obtaining a classification accuracy of: $c: MEAN \leq val < rac{MAX + MEAN}{2}, \ d: rac{MAX + MEAN}{2} \leq val < MAX$ $100*(1-rac{7}{150})=95.33\%$ Level #0 Root Node {**0**: 50, **1**: 50, **2**: 50} Splitting on feature petal width $a:0.10 \leq val < 0.65$ $b:0.65 \leq val < 1.20$ $petal\ width \in d$ $c: 1.20 \leq val < 1.85$ $petal\ width \in b$ $d: 1.85 \leq val \leq 2.50$ Level #1 Leaf Node Level #1 Node Level #1 Leaf Node Level #1 Leaf Node **(0**: 50, **1**: 0, **2**: 0) {**0**: 0, **1**: 10, **2**: 0} {0: 0, 1: 40, 2: 16} **(0**: 0, **1**: 0, **2**: 34**)** Splitting on feature petal length $petal\ length \in c$ $a:1.00 \leq val < 2.38$ $petal\ length \in b$ $b: 2.38 \leq val < 3.76$ $petal\ length \in d$ $c: 3.76 \leq val < 5.33$ $d:5.33 \leq val \leq 6.90$ Level #2 Leaf Node Level #2 Node Level #2 Leaf Node {**0**: 0, **1**: 1, **2**: 0} {**0**: 0, **1**: 39, **2**: 8} **(0**: 0, **1**: 0, **2**: 8**)** Splitting on feature sepal length $sepal\ length \in b$ $a: 4.30 \le val < 5.07$ $b:5.07 \leq val < 5.84$ $sepal\ length \in c \hspace{5mm} sepal\ length \in d$ $sepal\ length \in a$ $c: 5.84 \le val < 6.87$ $d:6.87 \leq val \leq 7.90$ Level #3 Leaf Node Level #3 Leaf Node Level #3 Node Level #3 Leaf Node {0: 0, 1: 0, 2: 1} {0: 0, 1: 14, 2: 0} {0: 0, 1: 23, 2: 7} {**0**: 0, **1**: 2, **2**: 0} Splitting on feature sepal width $sepal\ width \in b$ $a: 2.00 \le val < 2.53$ $b: 2.53 \leq val < 3.06$ $sepal\ width \in a$ $sepal\ width \in c$ $c: 3.06 \le val < 3.73$ $d: 3.73 \le val \le 4.40$

Level #4 Leaf Node

(0: 0, **1**: 6, **2**: 0)

Level #4 Leaf Node

{**0**: 0, **1**: 3, **2**: 1}

Level #4 Leaf Node

{**0**: 0, **1**: 14, **2**: 6}