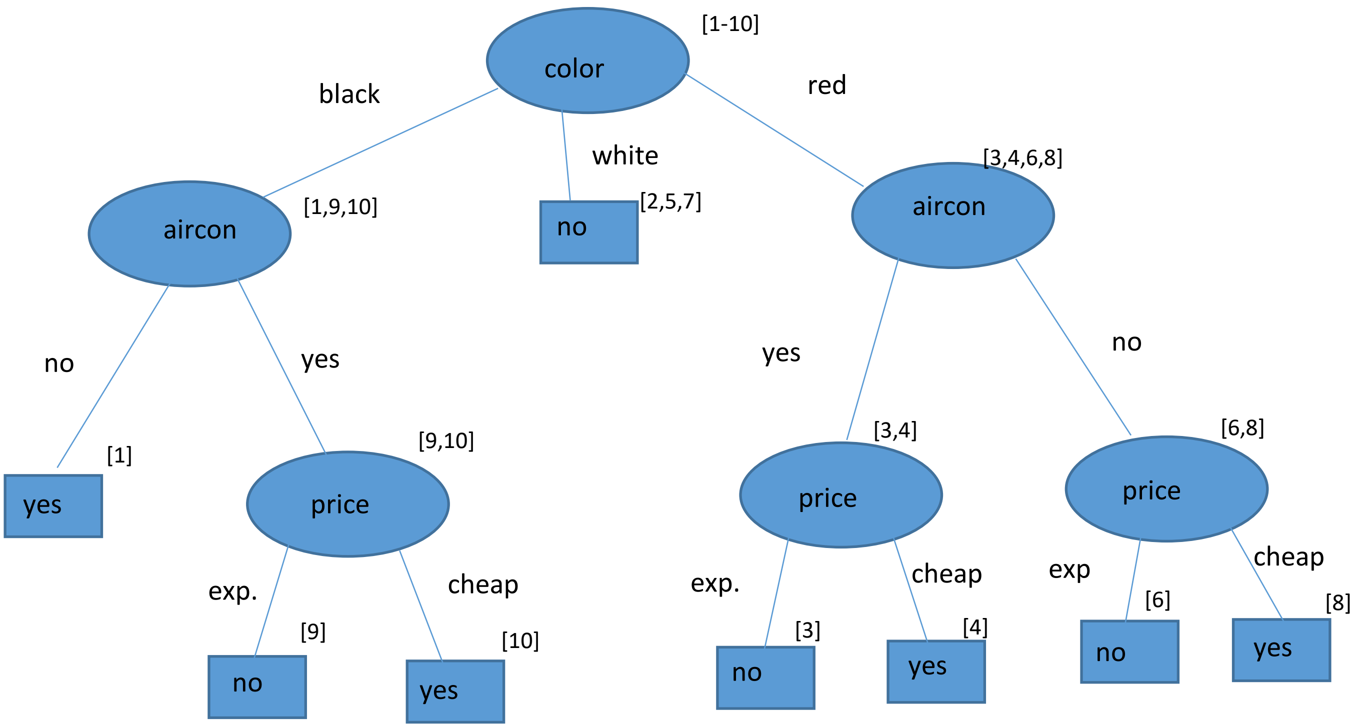
Assignment 5

Gruppe: Ole Elija Dziewas, Marius Ludwig Bachmeier, Meike Valentina Bauer

Aufgabe 1.1:



Aufgabe 1.2:

$$H(S) = -\frac{3}{10} \cdot \log_2\left(\frac{3}{10}\right) - \frac{3}{10} \cdot \log_2\left(\frac{3}{10}\right) - \frac{4}{10} \cdot \log_2\left(\frac{4}{10}\right) = 1{,}5710$$

$$H(S_{black}) = -\frac{2}{3} \cdot \log_2\left(\frac{2}{3}\right) - \frac{1}{3} \cdot \log_2\left(\frac{1}{3}\right) = 0{,}9183$$

$$H(S_{white}) = -1 \cdot \log_2(1) = 0$$

$$H(S_{red}) = -\frac{1}{2} \cdot \log_2\left(\frac{1}{2}\right) - \frac{1}{2} \cdot \log_2\left(\frac{1}{2}\right) = 1$$

$$Gain(S, colour) = H(S) - [\frac{3}{10} \cdot H(S_{black}) + \frac{3}{10} \cdot H(S_{white}) + \frac{4}{10} \cdot H(S_{red})]$$

$$Gain(S, colour) = 0{,}8955$$

Aufgabe 1.3:

Assumption: The entropy of each attribute is employed to choose the attribute in the step "Choose A from attributes" using the ID3 algorithm as shown in the Exercise Slides on slide 3.

The statement "The trees resulting from ID3 for the original set and the copy set are identical for each ID3 run" holds true.

Given that the data rows of the copy set only differ from the original set in that they were permutated in a random fashion, the relative frequencies and thus the entropy of the attributes remain the same. Permutation, by definition, does not add or subtract, but merely permutate, meaning shift, rows (in this case) around. Moreover, permutation also doesn't change the values of the attributes in this case, as only entire rows get permutated. This way no absolute and hence no relative frequency is ever manipulated, leading to the entropy not being changed and as a result leading to the decision trees of the original and the copy set being the very same.

Aufgabe 2.1:

1. Step:

$$\vec{w} = (0\ 0\ 0)^T$$

$$\vec{x_1} = (1\ 2\ 4)^T$$

$$o(\vec{x_1}) = sgn(0) = -1$$

$$\vec{w} = (1\ 2\ 4)^T$$

2. Step:

$$\vec{w} = (1\ 2\ 4)^T$$

$$\vec{x_2} = (1\ \ 1\ -2)^T$$

$$o(\vec{x_1}) = sgn(1 + 2 - 8) = -1$$

$$\vec{w} = (2\ 3\ 2)^T$$

3. Step:

$$\vec{w} = (2\ 3\ 2)^T$$

$$\vec{x_3} = (1\ 6\ 2)^T$$

$$o(\vec{x_1}) = sgn(2 + 18 + 4) = +1$$

$$\vec{w} = (1\ \ -3\ 0)^T$$

4. Step:

$$\vec{w} = (1 \quad -3\ 0)^T$$

$$\vec{x_4} = (1 \quad -3 \ \ 3)^T$$

$$o(\vec{x_1}) = sgn(1 + 9 + 0) = \ +1$$

$$\vec{w} = (0 \ \ 0 \quad -3)^T$$

Aufgabe 2.2:

Given the numeric dataset there is no linear function that is able to split up the dataset in a way the points could be classified correctly. Therefore no matter how often the weights are trained, the perceptron model is not able to solve this problem for all instances correctly.