HOCHSCHULE LANDSHUT

Artificial Intelligence I

Lab 6 - Winter Semester 2023 / 2024
https://moodle.haw-landshut.de/course/view.php?id=10282

1. Given the datasets

- S_1 = (yes, yes, yes, no, no, yes, yes, yes, yes, no, no)
- S_2 = (yes, yes, yes, no, no, no) => 50/50 => am höchsten

compute the corresponding entropy values. Which one has the greater entropy and thus more uncertainty?

2. Given the following data set for the skiing classification problem

Day	Snow_Dist	Weekend	Sun	Skiing
1	≤ 100	yes	yes	yes
2	≤ 100	yes	yes	yes
3	≤ 100	yes	no	yes
4	≤ 100	no	yes	yes
5	> 100	yes	yes	yes
6	> 100	yes	yes	yes
7	> 100	yes	yes	no
8	> 100	yes	no	no
9	> 100	no	yes	no
10	> 100	no	yes	no
11	> 100	no	no	no

generate the corresponding decision tree using the ID3-algorithm.

$IG(S, Snow\text{-}Dist) = I(S) - \frac{4}{11} I(S_{\leq 100}) - \frac{7}{11} I(S_{>100}) = 0{,}994 - \frac{4}{11} \cdot 0 - \frac{7}{11} \cdot 0{,}863 = 0{,}445$

$I(S) = -\frac{6}{11} \log_2\left(\frac{6}{11}\right) - \frac{5}{11} \log_2\left(\frac{5}{11}\right) = 0{,}994$

$I(S_{\leq 100}) = 0 = -0 \cdot \log_2(0) - \frac{4}{4} \log_2\left(\frac{4}{4}\right)$ (weil immer yes bei Skiing)

$I(S_{>100}) = -\frac{5}{7} \log_2\left(\frac{5}{7}\right) - \frac{2}{7} \log_2\left(\frac{2}{7}\right) = 0{,}863$

$IG(S, Weekend) = 0{,}994 - \frac{7}{11} I(Weekend_{yes}) - \frac{4}{11} I(Weekend_{no}) = 0{,}994 - \frac{7}{11} \cdot 0{,}863 - \frac{4}{11} \cdot 0{,}811 = 0{,}15$

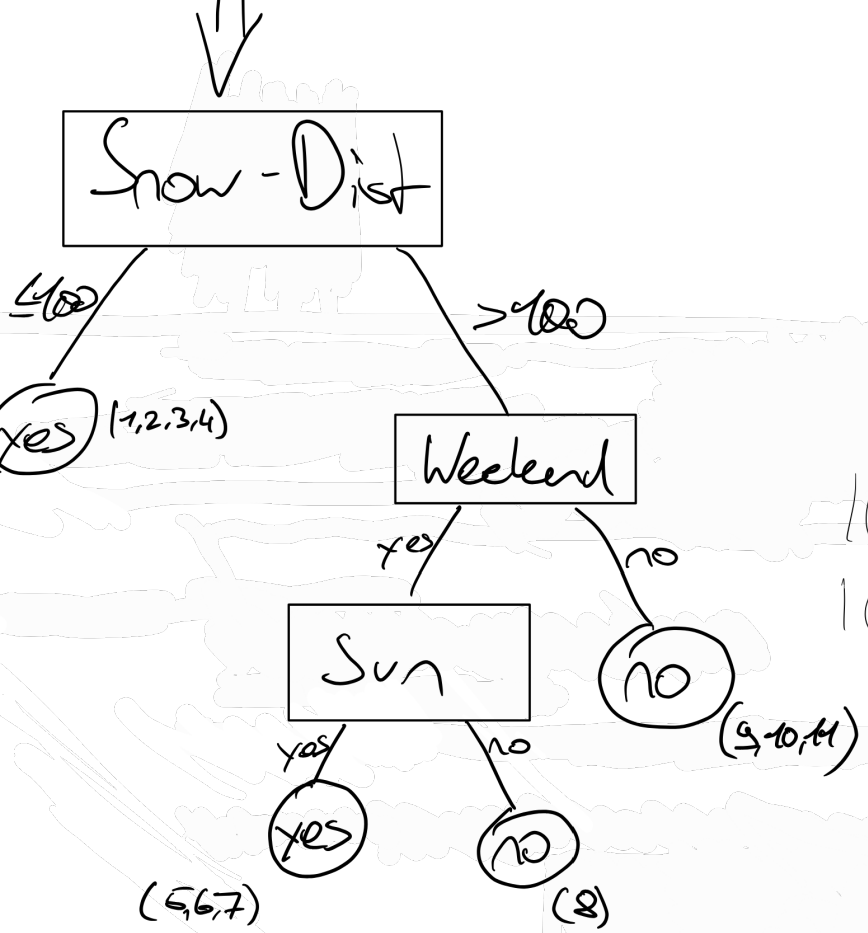
$-\frac{5}{7} \log_2\left(\frac{5}{7}\right) - \frac{2}{7} \log_2\left(\frac{2}{7}\right) = 0{,}863$

$-\frac{3}{4} \log_2\left(\frac{3}{4}\right) - \frac{1}{4} \log_2\left(\frac{1}{4}\right) = 0{,}811$

$IG(S, Sun) = 0{,}994 - \frac{8}{11} I(Sun_{yes}) - \frac{3}{11} I(Sun_{no}) = 0{,}994 - \frac{8}{11} \cdot 0{,}954 - \frac{3}{11} \cdot 0{,}918 = 0{,}049$

$-\frac{5}{8} \log_2\left(\frac{5}{8}\right) - \frac{3}{8} \log_2\left(\frac{3}{8}\right) = 0{,}954$

$-\frac{1}{3} \log_2\left(\frac{1}{3}\right) - \frac{2}{3} \log_2\left(\frac{2}{3}\right) = 0{,}918$



Weekend	Sun	Skiing
yes	yes	yes
yes	yes	yes
yes	yes	no
yes	no	no
no	yes	no
no	yes	no
no	no	no

$I(S) - \frac{4}{7}\left[I(W_{yes}) - \frac{3}{7} I(W_{no}) = 0,292\right.$

$I(S_{now-Dist}, Sun) = 0,17$