

## Task 1

a) yes the person goes sailing.

b) Gini Impurity

$$1 - (\text{probability of "yes"})^2 - (\text{probability of "no"})^2$$

left Node

$$1 - \left(\frac{4}{7}\right)^2 - \left(\frac{3}{7}\right)^2 = 0.375$$

right Node

$$1 - \left(\frac{8}{8}\right)^2 - \left(\frac{0}{8}\right)^2 = 0$$

because it's pure

$$\frac{4}{12} \cdot 0.375 + \frac{8}{12} \cdot 0 = 0.125 \quad \text{Impurity of Decision}$$

c)

Outlook	Wind	Temp	Storage	Sailing
sunny	5	30	yes	+
sunny	3	25	yes	+
rain	12	15	yes	+
overcast	15	2	no	-
rain	16	25	yes	+
sunny	14	18	yes	+
rain	8	5	no	-
sunny	9	30	yes	+
overcast	14	5	no	-
sunny	1	2	no	-
rain	4	25	yes	+
rain	14	24	yes	+
sunny	11	20	yes	+
sunny	2	18	yes	+
overcast	8	22	yes	+
overcast	13	24	yes	+

given: 5+temp = yes



$$1 - \left(\frac{2}{5}\right)^2 - \left(\frac{3}{5}\right)^2 = 0.48$$

$$1 - \left(\frac{7}{9}\right)^2 - \left(\frac{2}{9}\right)^2 = 0.54567$$

$$\frac{3}{12} \cdot 0.48 + \frac{9}{12} \cdot 0.54567 = 0.37036 \quad \text{Impurity}$$

Comparison: Both a higher impurity than the one in Task b!  
Task should get build node  
now lowest impurity and  
then work on with newly calculated lowest impurity.

Task b has a lower impurity so it should get taken as level 2 as criterion 1