1.

Bias measures the deviation between the expected prediction and the true label, which shows the fitting ability of learning algorithm

Variance shows How model measure the spread of data

Bias-Variance trade off show the relationship of bias and variance, increasing the bias, will decrease the Variance

Decrease the Variance, will increase the bias, so we aim to find the trade off of a model

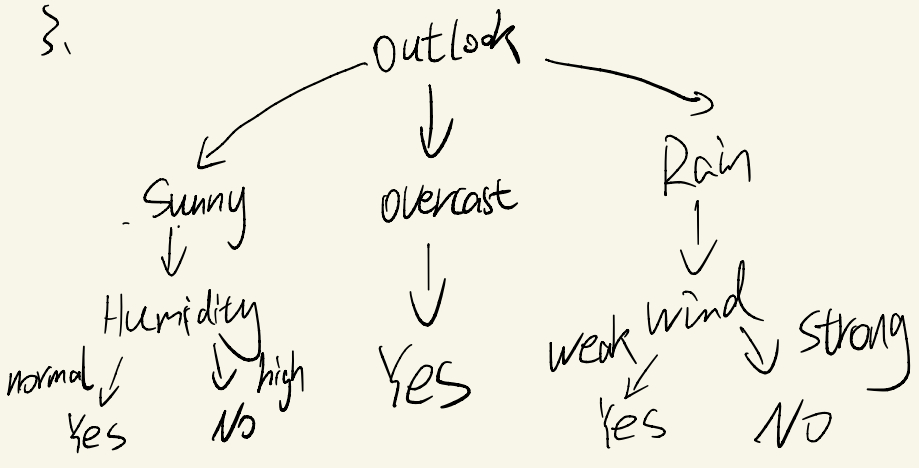
2.

$$\text{precision} = \frac{50}{50+30} = \frac{5}{8}$$

$$\text{recall} = \frac{50}{50+40} = \frac{5}{9}$$

$$F_1 = \frac{2 * \text{precision} * \text{Recall}}{\text{precision} + \text{Recall}} = \frac{10}{17}$$

3.



$$Entropy(D) = -(\frac{6}{10}\log_2\frac{6}{10} + \frac{4}{10}\log_2\frac{4}{10}) \approx 0.971$$

$$Entropy(outlook, Sunny) = -(\frac{1}{4}\log_2\frac{1}{4} + \frac{3}{4}\log_2\frac{3}{4}) \approx 0.811$$

$$Entropy(outlook, overcase) = -\log_2 1 = 0$$

$$Entropy(outlook, Rain) = -(\frac{3}{4}\log_2\frac{3}{4} + \frac{1}{4}\log_2\frac{1}{4}) \approx 0.811$$

$$Gain(D, outlook) = Entropy(D) - E(outlook, Sunny) - E(outlook, overcase) - E(outlook, Rain)$$
$$= 0.322$$

Same as ↑

$Gain(D, Temp) \approx 0.096$

$Gain(D, Humidity) \approx 0.125$

$Gain(D, wind) \approx 0.092$ we choose outlook as node

so, $Entropy(D_{Rain}) = 0.811$

$Gain(D_{Rain}, Temp) = 0.811 - 0 - 0.5 = 0.311$

$Gain(D_{Rain}, Humidity) = 0.811 - 0 - 0.689 = 0.1325$

$Gain(D_{Rain}, Wind) = 0.811$, choose wind →

Rain	weak	Yes
Rain	weak	Yes
Rain	weak	Yes

Rain	Strong	No

Same as ↑ $Entropy(D_{sunny}) = 0.811$

$Gain(D_{sunny}, Temp) = 0.971$

$Gain(D_{sunny}, Humidity) > 0.971$

$Gain(D_{sunny}, wind) < 0.971$ choose Humidity

↓

sunny	high	No
sunny	high	No
sunny	high	No

sunny	normal	Yes

4.

$P(w_1 | d_{1,1}(x)=1) = \frac{4}{7}$

$P(w_1 | d_{2,1}(x)=1) = \frac{1}{2}$ $\quad class1 = \frac{4}{7} * \frac{1}{2} * 0$

$P(w_1 | d_{3,2}(x)=1) = 0$

$P(w_2 | d_{1,1}(x)=1) = \frac{30}{70} = \frac{3}{7}$

$P(w_2 | d_{2,1}(x)=1) = \frac{1}{2}$ $\quad class2 = 0.24$

$P(w_2 | d_{3,2}(x)=1) = 1$

class 2 is better than class 1