(a). Run :

Yes	No	Total
8	7	15

$$E(s) = \sum p(x) \log_2 \frac{1}{p(x)}$$

$$E(s) = -\frac{7}{15}\log_2 \frac{7}{15} - \frac{8}{15}\log_2 \frac{8}{15} = 0.99679$$

$$IG(s, \text{outlook}) = H(s) - \sum p(x) * H(x) \qquad \boxed{H(s) = 0.99679}.$$

Out = Rainy	Out = Overcast	Out = Sunny	Total
6	2	7	15

$$P(S_{rainy}) = \frac{6}{15} \quad P(S_{overcast}) = \frac{\cancel{7}2}{15} \quad P(S_{sunny}) = \frac{7}{15}$$

$$\begin{tikzcd} & \text{1yes} & & \text{0yes} & & \text{7yes} \\ \text{6 Rainy} \arrow[ur] \arrow[dr] & & \text{2overcast} \arrow[ur] \arrow[dr] & & \text{7 sunny} \arrow[ur] \arrow[dr] & \\ & \text{5No} & & \text{2No} & & \text{0No} \end{tikzcd}$$

$$H(S_{rainy}) = -\frac{5}{6}\log_2\left(\frac{5}{6}\right) - \frac{1}{6}\log_2\left(\frac{1}{6}\right) = 0.6500$$

$$H(S_{overcast}) = -\log(1) - 0 = 0$$

$$H(S_{sunny}) = -\log(1) - 0 = 0$$

$$IG(s, \text{outlook}) = 0.99679 - \frac{6}{15} \times 0.65 \quad \therefore \boxed{IG(s, \text{outlook}) = 0.73679}.$$

$$IG(s, \text{temp}) = H(s) - \sum p(x) * H(x) \qquad H(s) = 0.99679$$

temp = mild	temp = cold	temp = hot	total
6	6	3	15

$$P(S_{mild}) = \frac{6}{15} \quad P(S_{cold}) = \frac{6}{15} \quad P(S_{temp}) = \frac{3}{15}$$

$$\begin{tikzcd} & \text{4yes} & & \text{2yes} & & \text{2yes} \\ \text{6 mild} \arrow[ur] \arrow[dr] & & \text{6 cold} \arrow[ur] \arrow[dr] & & \text{3 hot} \arrow[ur] \arrow[dr] & \\ & \text{2No} & & \text{4No} & & \text{1No} \end{tikzcd}$$

$$H(S_{mild}) = -\frac{4}{6}\log_2\left(\frac{4}{6}\right) - \frac{2}{6}\log_2\left(\frac{2}{6}\right) = 0.918296$$

$$H(S_{cold}) = -\frac{4}{6}\log_2\left(\frac{4}{6}\right) - \frac{2}{6}\log_2\left(\frac{2}{6}\right) = 0.918296$$

$$H(S_{hot}) = -\frac{1}{3}\log_2\left(\frac{1}{3}\right) - \frac{2}{3}\log_2\left(\frac{2}{3}\right) = 0.918296$$

$$IG(s, \text{temp}) = 0.99679 - \frac{12}{15}(0.918296) - \frac{3}{15}(0.98296) = \boxed{IG(s, \text{temp}) = 0.07849}.$$

$IG(S, hum) = H(S) - \Sigma p(x) * H(x)$

$H(S) = 0'99679.$

hum = high	hum = norm	total
6	9	15

$p(S_{high}) = \frac{6}{15}$ $p(S_{norm}) = \frac{9}{15}$

$$\begin{tikzcd} & 2yes \\ 6\,high \arrow[ur] \arrow[dr] & \\ & 4\,No \end{tikzcd} \qquad \begin{tikzcd} & 6yes \\ 9\,Norm \arrow[ur] \arrow[dr] & \\ & 3\,No \end{tikzcd}$$

$H(S_{high}) = -\frac{4}{6}\log_2\left(\frac{4}{6}\right) - \frac{2}{6}\log_2\left(\frac{2}{6}\right) = 0'918296$

$H(S_{nom}) = -\frac{6}{9}\log_2\left(\frac{6}{9}\right) - \frac{3}{9}\log_2\left(\frac{3}{9}\right) = 0'918296.$

$\therefore \boxed{IG(S, Norm) = 0'07849}.$

∴ the first node will be outlook. (highest info gain).

	Yes	No	Total
Rainy :	1	5	6

$H(S_{rainy}) = -\frac{5}{6}\log_2\left(\frac{5}{6}\right) - \frac{1}{6}\log_2\left(\frac{1}{6}\right) = 0'65$ $\boxed{H(S_{rainy}) = 0'65}.$

$IG(S_{rainy}, temp) = H(S_{rainy}) - \Sigma p(x) * H(x).$

temp = mild	temp = cold	temp = hot	total
2	3	1	6

$p(S_{mild}) = \frac{2}{6}$ $p(S_{cold}) = \frac{3}{6}$ $P(S_{hot}) = \frac{1}{6}$

$$\begin{tikzcd} & 1yes \\ 2\,mild \arrow[ur] \arrow[dr] & \\ & 1\,No \end{tikzcd} \qquad \begin{tikzcd} & 0yes \\ 3cold \arrow[ur] \arrow[dr] & \\ & 3\,No \end{tikzcd} \qquad \begin{tikzcd} & 0yes \\ 1hot \arrow[ur] \arrow[dr] & \\ & 1\,No \end{tikzcd}$$

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

$H(S_{cold}) = -\log 1 - 0 = 0$

$H(S_{hot}) = -\log 1 - 0 = 0$

$IG(S_{rainy}, temp) = 0'65 - \left(1 \times \frac{2}{6}\right) = 0'3166$ $\boxed{IG(S_{rainy}, temp) = 0'3166}.$

$IG(S_{rainy}, hum) = H(S_{rainy}) - \sum p(x) * H(x)$

$H(S_{rainy}) = 0'65$

hum = high	hum = normal	total
4	2	6

$P(S_{high}) = \frac{4}{6}$ $\quad P(S_{norm}) = \frac{2}{6}$

$$\begin{tikzcd}[row sep=tiny] & 0\text{yes} & & & 1\text{yes} \\ 4\ \text{high} \arrow[ur] \arrow[dr] & & & 2\ \text{Normal} \arrow[ur] \arrow[dr] & \\ & 4\ \text{No} & & & 1\ \text{No}\end{tikzcd}$$

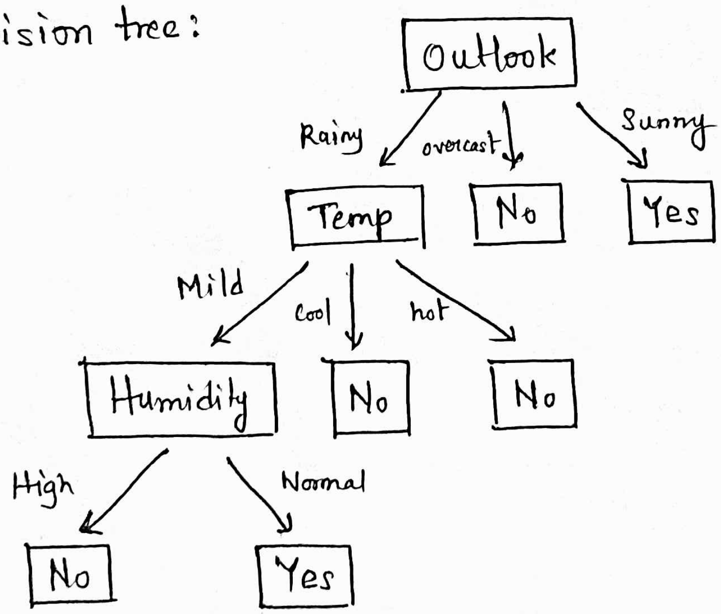
$H(S_{high}) = \log 1 - 0 = 0$

$H(S_{nomal}) = -\frac{1}{2}\log\frac{1}{2} - \frac{1}{2}\log\frac{1}{2} = 1$

$IG(S_{rainy}, nomal) = 0'65 - (1 \times \frac{2}{6}) = 0{\cdot}3166$ $\quad \boxed{IG(S_{rainy}, hum) = 0'3166}$

Therefore, the next decision node can be any of humidity or temperature.

∴ The decision tree:



(b) If tomorrows, outlook is overcast, the temperature is Cool and the Humidity level is high, the runner will mot go for a run.

2.(a)

Accuracy -

```
Training/validation accuracy for minimum node entropy 0.010000 is 1.000 / 0.863
Training/validation accuracy for minimum node entropy 0.050000 is 0.999 / 0.863
Training/validation accuracy for minimum node entropy 0.100000 is 0.997 / 0.865
Training/validation accuracy for minimum node entropy 0.200000 is 0.990 / 0.867
Training/validation accuracy for minimum node entropy 0.400000 is 0.979 / 0.861
Training/validation accuracy for minimum node entropy 0.800000 is 0.919 / 0.856
Training/validation accuracy for minimum node entropy 1.000000 is 0.871 / 0.840
Training/validation accuracy for minimum node entropy 2.000000 is 0.596 / 0.600
Test accuracy with minimum node entropy 0.200000 is 0.872
```

Error Rate –

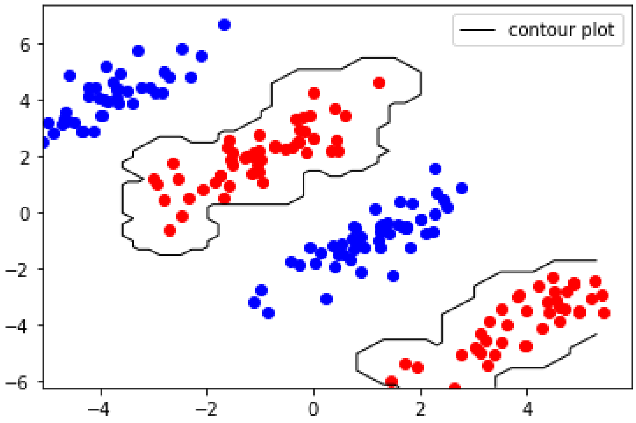
```
Training/validation error rate for minimum node entropy 0.010000 is 0.000 / 0.137
Training/validation error rate for minimum node entropy 0.050000 is 0.001 / 0.137
Training/validation error rate for minimum node entropy 0.100000 is 0.003 / 0.135
Training/validation error rate for minimum node entropy 0.200000 is 0.010 / 0.133
Training/validation error rate for minimum node entropy 0.400000 is 0.021 / 0.139
Training/validation error rate for minimum node entropy 0.800000 is 0.081 / 0.144
Training/validation error rate for minimum node entropy 1.000000 is 0.129 / 0.160
Training/validation error rate for minimum node entropy 2.000000 is 0.404 / 0.400
Test error rate with minimum node entropy 0.200000 is 0.128
```

2.(b) With the increase in the minimum node entropy, the validation accuracy decreases for both validation and training dataset. If there was a high error-rate on both training and validation data that would indicate that the model may be too simple. But that does not happen in the above case. If the error-rate is slightly increased on training data compared to validation data, thus the model is not too complex as it seems.

3.(a) sigma = 0.001

Accuracy on the simulated data is 1.00

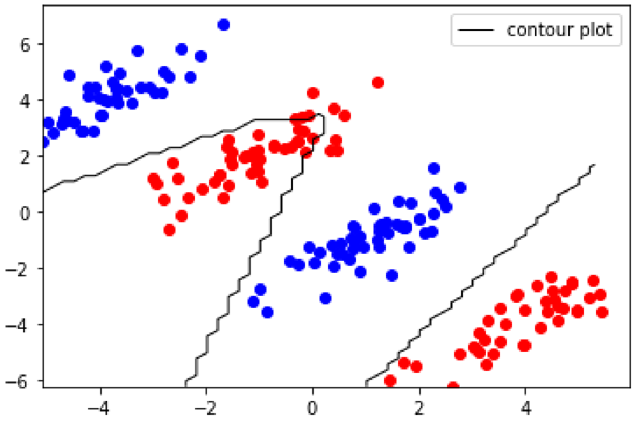
Decision Boundary –



sigma = 0.1

Accuracy on the simulated data is 0.95

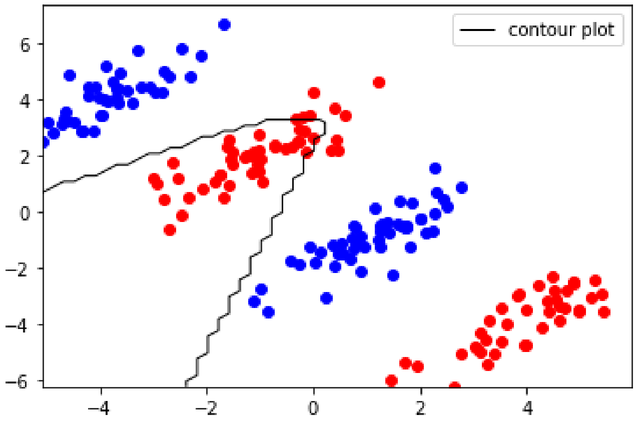
Decision Boundary –



sigma = 1

Accuracy on the simulated data is 0.72

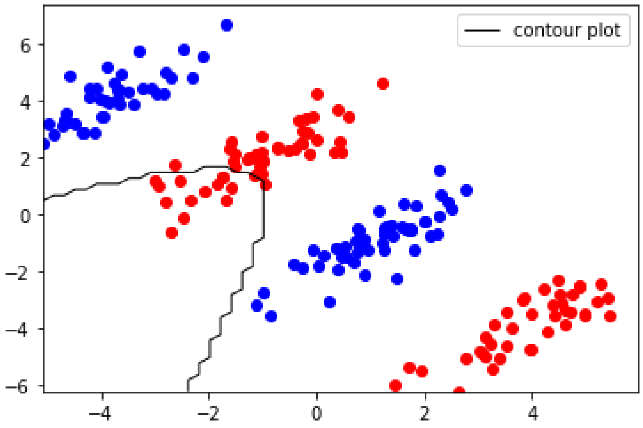
Decision Boundary –



sigma = 10

Accuracy on the simulated data is 0.57

Decision Boundary –



3.(b) Accuracy for sigma pool= (0.001,0.001,0.005) – 0.51 (on second dataset containing 7 and 9)

Accuracy for sigma pool = (0.1,0.1,0.5) – 0.83 (on second dataset containing 7 and 9)

Accuracy for sigma pool= (1,1,5) – 0.96 (on second dataset containing 7 and 9)

Accuracy for sigma pool= (10,1,5) – 0.96 (on second dataset containing 7 and 9)

With increasing sigma, the decision boundaries are not accurate as can be seen from the above contour plots. The accuracy of the model on the simulated data decreases as the sigma is increased. Thus, we have to optimize sigma so that the validation accuracy increases again it does not overfit to the training data set.

Whereas on decreasing sigma the accuracy on the training set does not increase but rather decreases (for the second dataset containing 7 and 9) whereas it does not seem to change on the first dataset (containing 4 and 9) which remains at a low value throughout.

3.(c) sigma pool= (0.001,0.001,0.005):

Accuracy on the digit classification-49 is 0.50

Accuracy on the digit classification-79 is 0.51

sigma pool = (0.1,0.1,0.5):

Accuracy on the digit classification-49 is 0.50

Accuracy on the digit classification-79 is 0.83

sigma pool= (1,1,5):

Accuracy on the digit classification-49 is 0.50

Accuracy on the digit classification-79 is 0.96

sigma pool= (10,1,5):

Accuracy on the digit classification-49 is 0.5

Accuracy on the digit classification-79 is 0.96