Pattern Recognition

HW3: Ensemble Methods

Part 1, Coding

Question 1

Gini of data is 0.4628099173553719
Entropy of data is 0.9456603046006401

Question 2

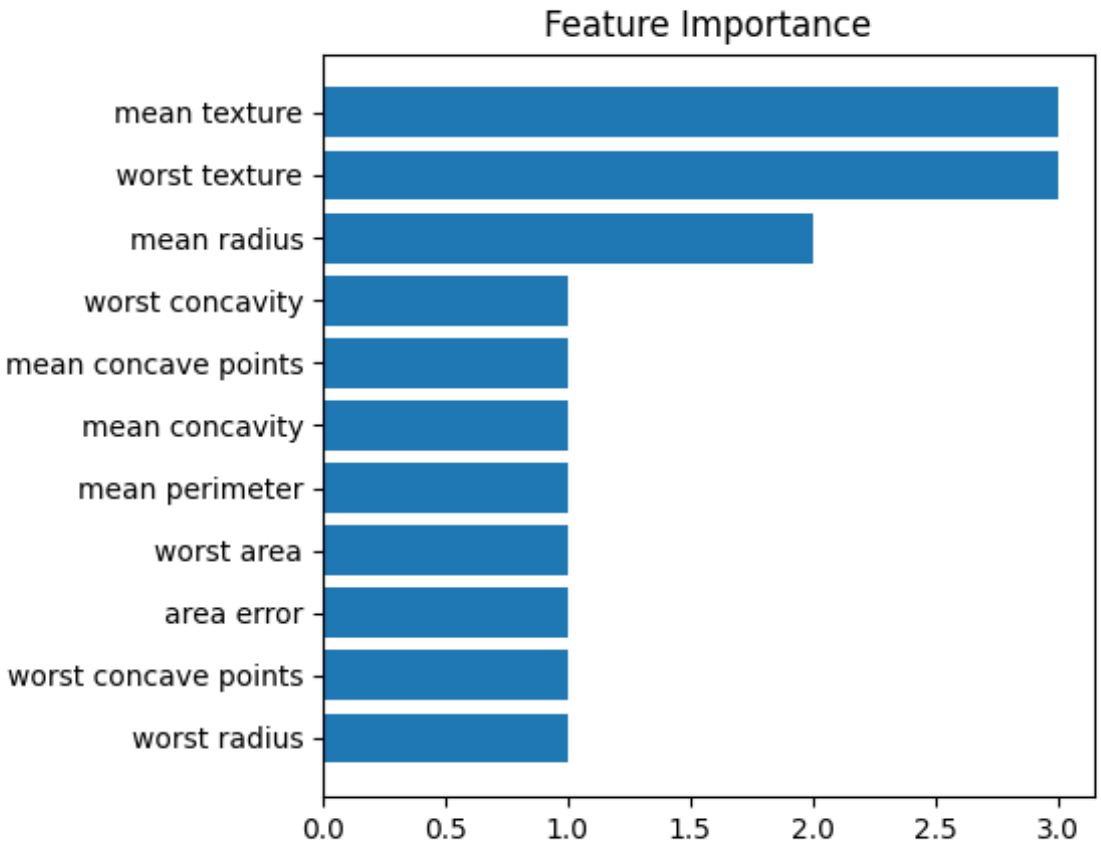
Question 2.1

Accuracy of clf_depth3: 0.9370629370629371
Accuracy of clf_depth10: 0.9300699300699301

Question 2.2

Accuracy of clf_gini: 0.9370629370629371
Accuracy of clf_entropy: 0.951048951048951

Question 3



Question 4

Question 4.1

Accuracy of clf_10tree: 0.972027972027972
Accuracy of clf_100tree: 0.9440559440559441

Question 4.2

Accuracy of clf_random_features: 0.951048951048951
Accuracy of clf_all_features: 0.965034965034965

Part 2, Questions

Question 1

Let $w = w_n^{(m)}$ and $I = I(y_m(x_n) \neq t_n)$

Let $\frac{\partial E}{\partial \alpha_m} = 0$

$$\left(\frac{1}{2}\ln e \cdot e^{\frac{\alpha_m}{2}} - \left(-\frac{1}{2}\right)\ln e \cdot e^{-\frac{\alpha_m}{2}}\right) \Sigma wI +$$

$$\left(-\frac{1}{2}\right)\ln e \cdot e^{-\frac{\alpha_m}{2}} \Sigma w = 0$$

$$\left(e^{\frac{\alpha_m}{2}} + e^{-\frac{\alpha_m}{2}}\right) \Sigma wI - e^{-\frac{\alpha_m}{2}} \Sigma w = 0$$

$$e^{\frac{\alpha_m}{2}} \Sigma wI = e^{-\frac{\alpha_m}{2}} (\Sigma w - \Sigma wI)$$

$$e^{\alpha_m} = \frac{\Sigma w - \Sigma wI}{\Sigma wI}$$

$$\alpha_m = \ln\left(\frac{\Sigma w - \Sigma wI}{\Sigma wI}\right) = \ln\left(\frac{1 - \frac{\Sigma wI}{\Sigma w}}{\frac{\Sigma wI}{\Sigma w}}\right)$$

Therefore, $\alpha_m = \ln\left(\frac{1 - \varepsilon_m}{\varepsilon_m}\right)$

Question 2

		C_1	C_2
model A	predict C_1	300	100
	predict C_2	100	300
modeB	predict C_1	200	0
	predict C_2	200	400

Misclassifications:

model A: $100 + 100 = 200$

model B: $0 + 200 = 200$

So both misclassification rates are the same

model A

$$Q_1(T) = -\frac{3}{4}\ln\left(\frac{3}{4}\right) - \frac{1}{4}\ln\left(\frac{1}{4}\right)$$

$$Q_2(T) = -\frac{1}{4}\ln\left(\frac{1}{4}\right) - \frac{3}{4}\ln\left(\frac{3}{4}\right)$$

$$C_A(T) = Q_1(T) + \lambda|T| + Q_2(T) + \lambda|T|$$

$$= 1.12467 + 2\lambda|T|$$

model B

$$Q_1(T) = -1\ln 1 - 0$$

$$Q_2(T) = -\frac{1}{3}\ln\left(\frac{1}{3}\right) - \frac{2}{3}\ln\left(\frac{2}{3}\right)$$

$$C_B(T) = 0.63651 + 2\lambda|T|$$

Clearly, $C_B(T) < C_A(T)$

Question 3

Sum of square error $E = \sum_{n=1}^{N} (t - t_n)^2$

E is minimized when $\frac{dE}{dt} = 0$

$$\frac{dE}{dt} = d \sum_{n=1}^{N} (t - t_n)^2 \frac{1}{dt} = 0$$

$$d \sum_{n=1}^{N} (t^2 - 2tt_n + t_n^2) \frac{1}{dt} = \sum_{n=1}^{N} 2t - \sum_{n=1}^{N} 2t_n = 0$$

$$Nt - \sum_{n=1}^{N} t_n = 0$$

$$t = \frac{\sum_{n=1}^{N} t_n}{N}$$

Where $\frac{\sum_{n=1}^{N} t_n}{N}$ is the mean of $\{t_n\}_{n=1\sim N}$