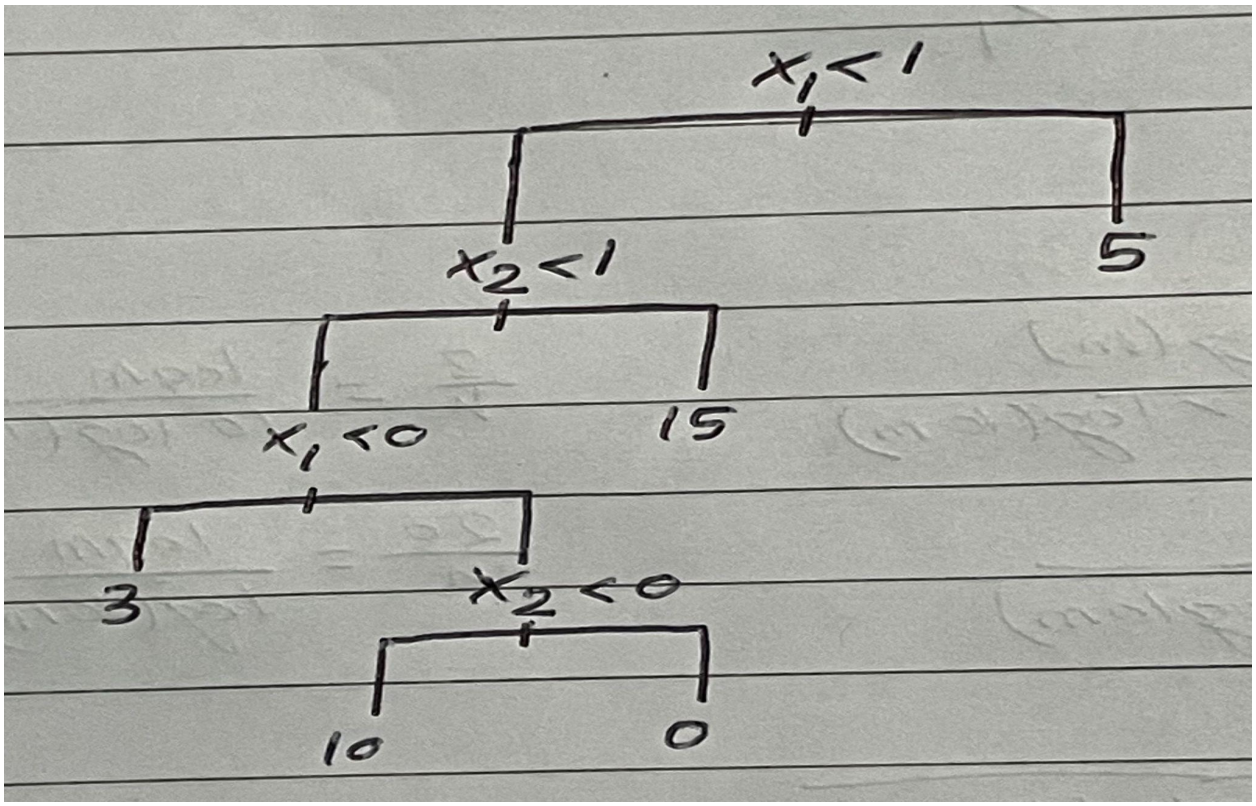
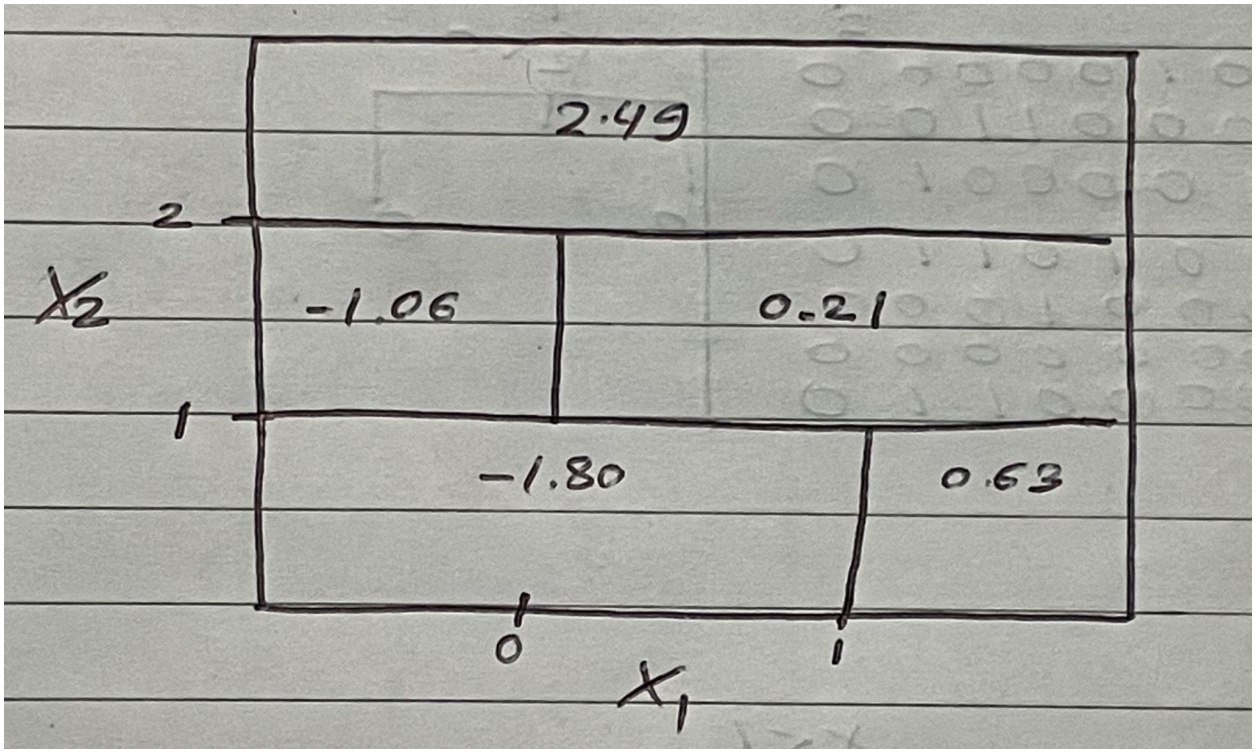
Q.no. 1



Q.no. 2



Q.no. 3

If a Decision Tree is overfitting the training set, it may be a good idea to decrease max_depth, since this will limit the model from fitting very closely to the train data and hence regularize the model. In other words, decreasing the max_depth will restrict/constrain the freedom of the decision trees by allowing it to have only a predetermined number of parameters which in turn reduces the overfitting.

Q.no. 4

The computational complexity of training a Decision Tree is O(n × m log(m)).
Let k be the time taken for training 20 million instances.

So,

$$\frac{2}{k} = \frac{n * m\, log(m)}{n * 10m * log(10m)}$$

$$\frac{2}{k} = \frac{log(m)}{10 * log(10m)}$$

Since m = 2 * 10^6,

$$\frac{2}{k} = \frac{log(2 * 10^6)}{10 * log(10 * 2 * 10^6)}$$

$$\frac{2}{k} = \frac{6.301}{10 * 7.301}$$

Therefore,
k = 23.17 hours
The training time is approximately 23.2 hours.