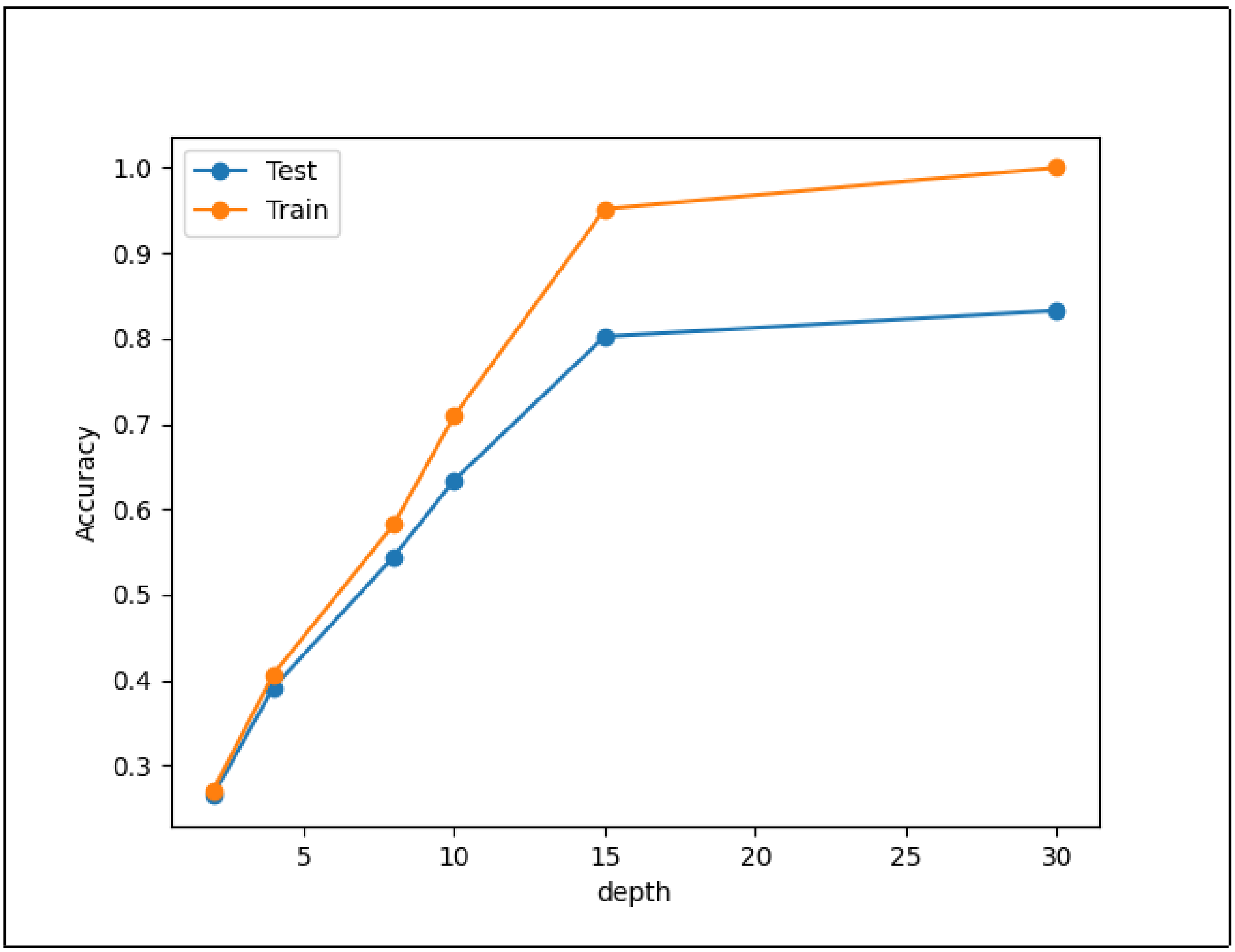
Question 1: 1
Question 2: 3
Theory 6

Question 1:

(b)

```
Testing Accuracy using entropy 0.8242811501597445
Training Accuracy using entropy 1.0
Best Testing Accuracy is 0.8326487144 with the depth 30
```



(c) We get less accuracy as compare to part a and b because we have max-depth of decision tree as 3

```
Accuracy 0.3591967138293017
```

(d)

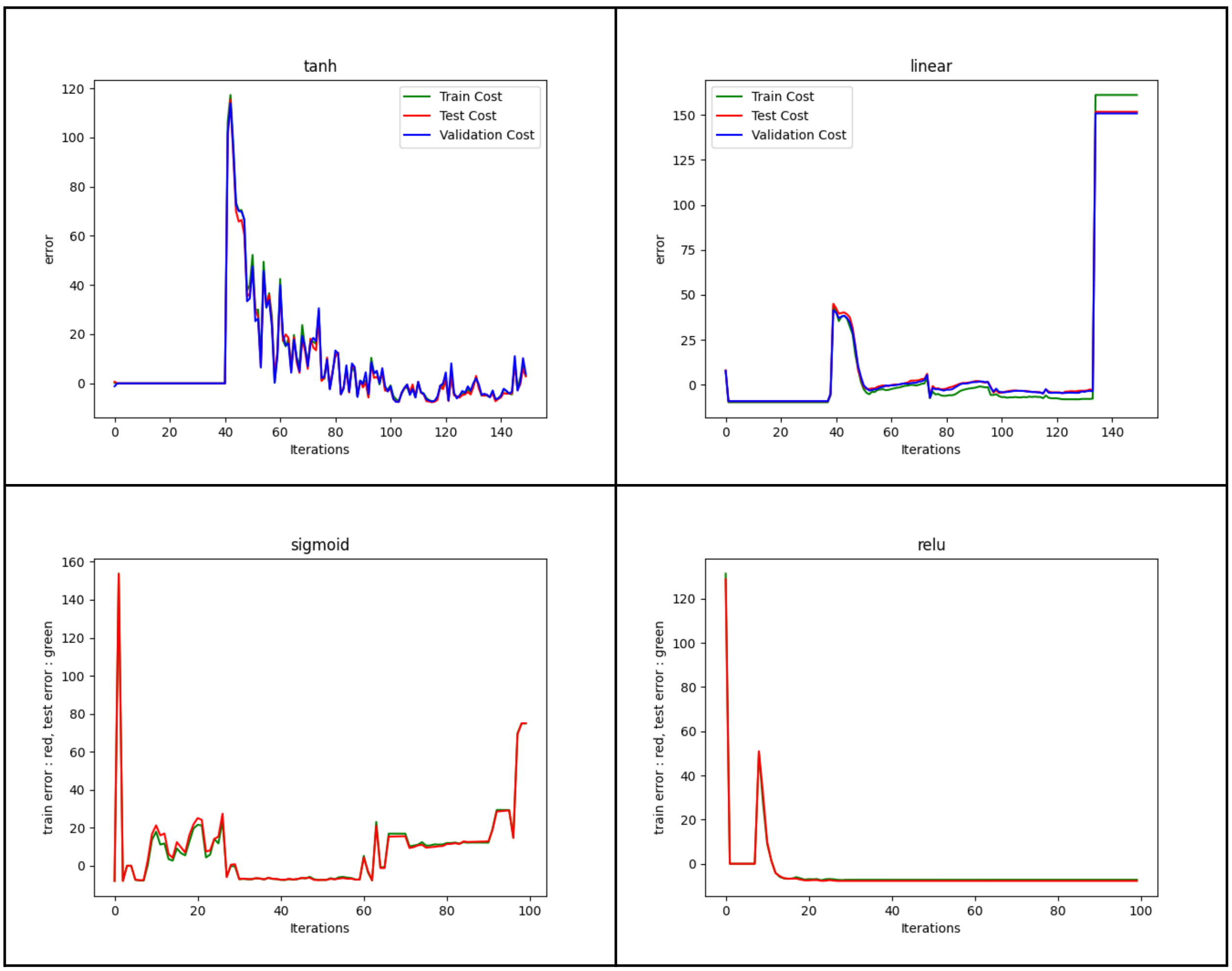
```
Accuracy at testing: 0.397 at depth: 4 Stumps: 20
Accuracy at training: 0.414 at depth: 4 Stumps: 20
Accuracy at validation: 0.412 at depth: 4 Stumps: 20
Accuracy at testing: 0.397 at depth: 4 Stumps: 50
Accuracy at training: 0.413 at depth: 4 Stumps: 50
Accuracy at validation: 0.407 at depth: 4 Stumps: 50
Accuracy at testing: 0.398 at depth: 4 Stumps: 100
Accuracy at training: 0.414 at depth: 4 Stumps: 100
Accuracy at validation: 0.409 at depth: 4 Stumps: 100
Accuracy at testing: 0.603 at depth: 8 Stumps: 20
Accuracy at training: 0.648 at depth: 8 Stumps: 20
Accuracy at validation: 0.622 at depth: 8 Stumps: 20
Accuracy at testing: 0.610 at depth: 8 Stumps: 50
Accuracy at training: 0.659 at depth: 8 Stumps: 50
Accuracy at validation: 0.635 at depth: 8 Stumps: 50
Accuracy at testing: 0.618 at depth: 8 Stumps: 100
Accuracy at training: 0.664 at depth: 8 Stumps: 100
Accuracy at validation: 0.636 at depth: 8 Stumps: 100
Accuracy at testing: 0.748 at depth: 10 Stumps: 20
Accuracy at training: 0.817 at depth: 10 Stumps: 20
Accuracy at validation: 0.760 at depth: 10 Stumps: 20
Accuracy at testing: 0.748 at depth: 10 Stumps: 50
Accuracy at training: 0.820 at depth: 10 Stumps: 50
Accuracy at validation: 0.760 at depth: 10 Stumps: 50
Accuracy at testing: 0.757 at depth: 10 Stumps: 100
Accuracy at training: 0.832 at depth: 10 Stumps: 100
Accuracy at validation: 0.771 at depth: 10 Stumps: 100
Accuracy at testing: 0.886 at depth: 15 Stumps: 20
Accuracy at training: 0.988 at depth: 15 Stumps: 20
Accuracy at validation: 0.894 at depth: 15 Stumps: 20
Accuracy at testing: 0.900 at depth: 15 Stumps: 50
Accuracy at training: 0.995 at depth: 15 Stumps: 50
Accuracy at validation: 0.909 at depth: 15 Stumps: 50
Accuracy at testing: 0.905 at depth: 15 Stumps: 100
Accuracy at training: 0.996 at depth: 15 Stumps: 100
Accuracy at validation: 0.916 at depth: 15 Stumps: 100
Accuracy at testing: 0.898 at depth: 20 Stumps: 20
Accuracy at training: 0.997 at depth: 20 Stumps: 20
Accuracy at validation: 0.901 at depth: 20 Stumps: 20
Accuracy at testing: 0.907 at depth: 20 Stumps: 50
Accuracy at training: 0.999 at depth: 20 Stumps: 50
Accuracy at validation: 0.915 at depth: 20 Stumps: 50
Accuracy at testing: 0.919 at depth: 20 Stumps: 100
Accuracy at training: 1.000 at depth: 20 Stumps: 100
Accuracy at validation: 0.923 at depth: 20 Stumps: 100
Accuracy at testing: 0.897 at depth: 30 Stumps: 20
Accuracy at training: 0.997 at depth: 30 Stumps: 20
```

```
Accuracy at validation: 0.904 at depth: 30 Stumps: 20
Accuracy at testing: 0.909 at depth: 30 Stumps: 50
Accuracy at training: 0.999 at depth: 30 Stumps: 50
Accuracy at validation: 0.920 at depth: 30 Stumps: 50
Accuracy at testing: 0.913 at depth: 30 Stumps: 100
Accuracy at training: 1.000 at depth: 30 Stumps: 100
Accuracy at validation: 0.925 at depth: 30 Stumps: 100
```

(e)

```
Accuracy: 0.815 at Estimator: 4
Accuracy: 0.813 at Estimator: 8
Accuracy: 0.816 at Estimator: 10
Accuracy: 0.812 at Estimator: 15
Accuracy: 0.810 at Estimator: 20
```

Question 2:



3.
Activation function for the output layer is softmax.The reason behind it is that we require probabilities to do classification. While, other activation functions like tanh gives range (-1,1). Further, softmax is a generalized version of sigmoid. It makes logits sum upto 1 which is
required for classification

Relu	Sigmoid	Linear	tanh	Leaky relu	softmax
0.7323	0.09	0.68	0.80	0.93	0.7

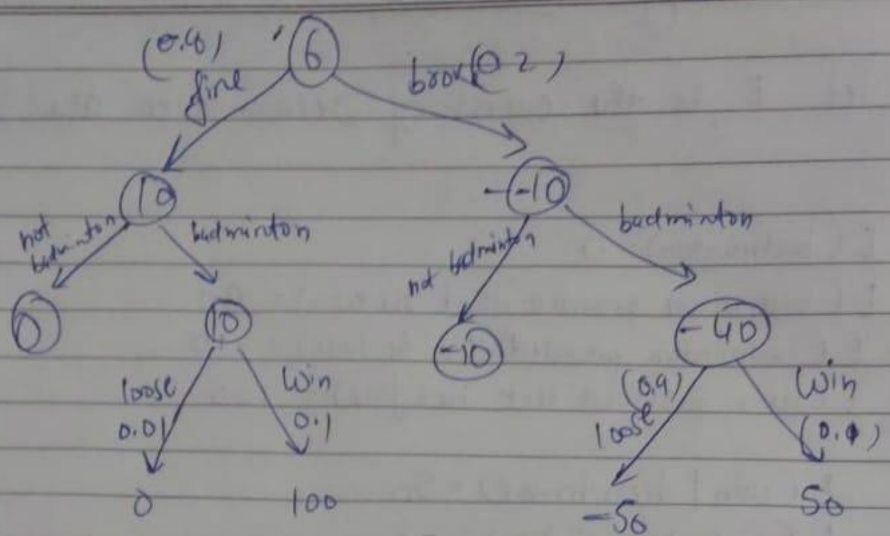
4.

```
Using Sklearn
 ReLU
 Testing accuracy: 0.9257142857142857
 Linear
 Testing accuracy: 0.09428571428571429
 Sigmoid
 Testing accuracy: 0.10285714285714286
 tanh
 Testing accuracy: 0.9114285714285715
```

5.Adjusting learning rate puts crucial effect on convergence of cross-entropy. Here, with a high learning rate such as 0.1 , the model can minimize the loss but unable to converge.Decreasing the learning rate to 0.01, we still can reach minima , again unable to converge smoothly. Training loss is somewhat providing required results, but validation loss still not converging effectively.Further, decreasing the learning rate by 10 fold to 0.001,we can see our training and validation loss, 0both are at minima and converged smoothly and effectively.

Theory

Date........................



D.T with perfect along the legs.

(2) Using our analysis ~~and~~ on decision tree, we can say that best ~~possible~~ scenario will be to play badminton and win given that he is fine and expected utility is 100.

(3) let E_{inf} be the Event state that we have info
Similarly E_{noinfo} — — we have no info.

~~(4) let E be the~~

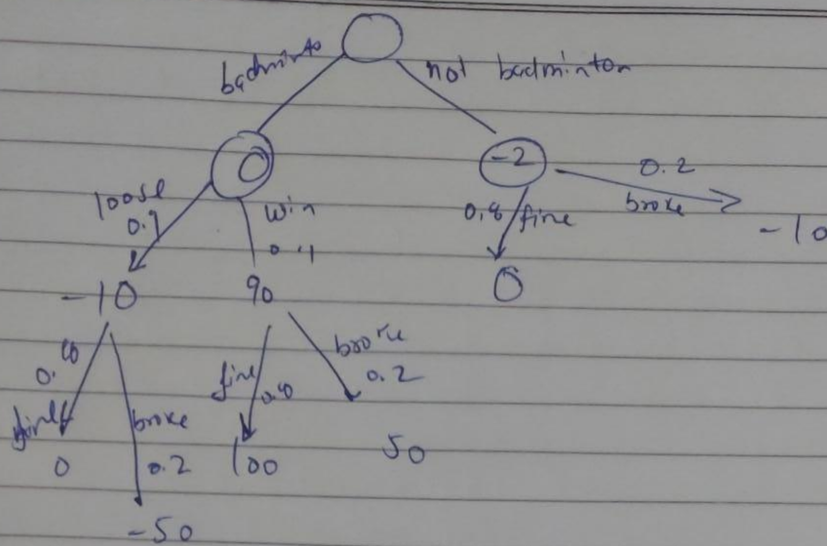
$$E\left(\begin{matrix}\text{perfect info}\\ \text{about leg}\end{matrix}\right) = 6 - 0 = 6$$

(4) ~~let~~ $E\left(\begin{matrix}\text{perfect info}\\ \text{about winning}\end{matrix}\right) = 7.2 - 0 = 7.2$

(5) Yes, we can use DT to find the output.

Spiral Teacher's Sign..................

Date



not badminton
-2
0.2
broke
-10
0.8
fine
0
win
0.1
90
broke
0.2
fine
100
50
broke
0.2
-50