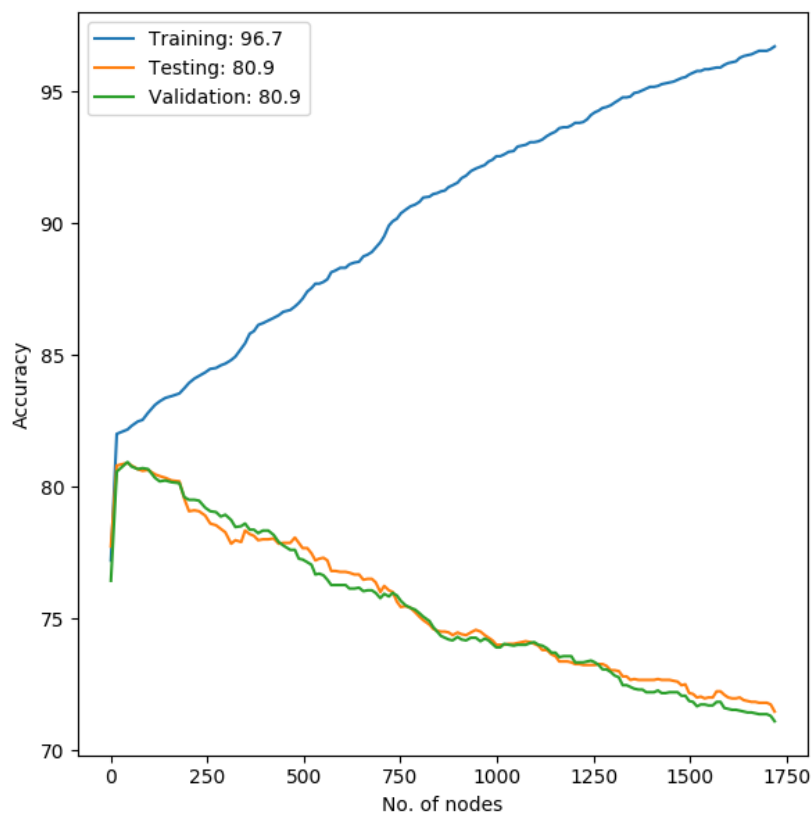


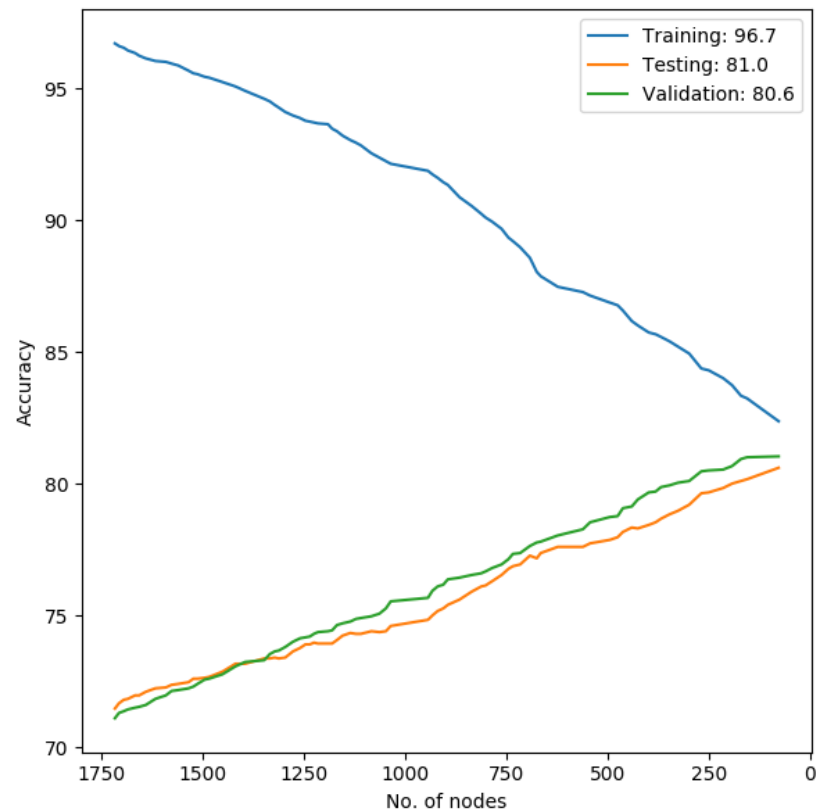
Problem 1: Decision Trees

Answer to the problem goes here.

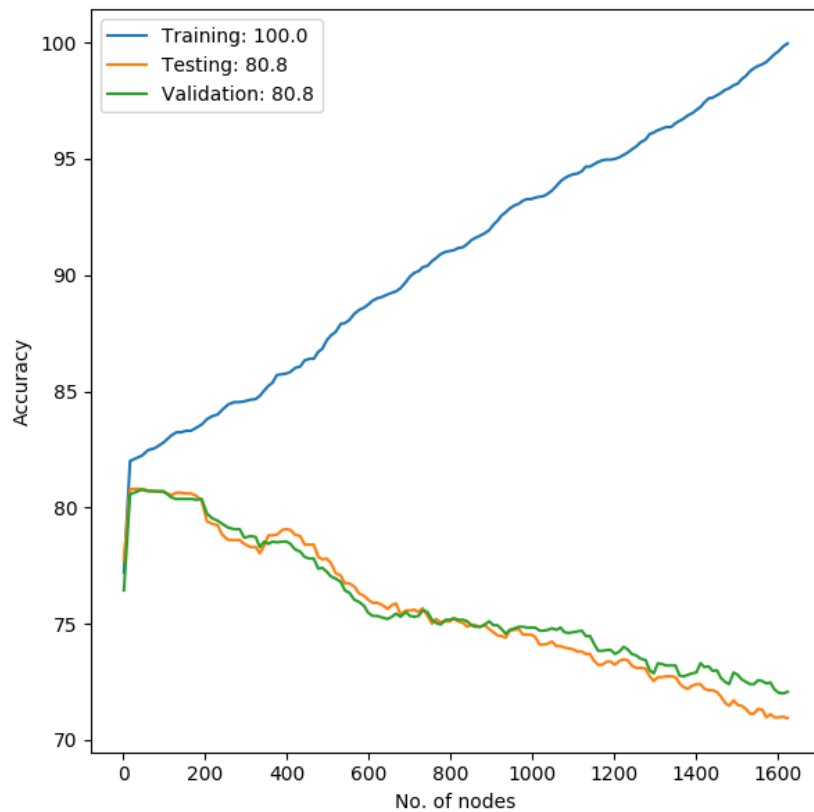
1. With the increase in the number of nodes, the tree starts overfitting. This is because as the number of nodes increases (that are considered for prediction), the training accuracy increases while the validation and testing accuracy decreases.



2. When we prune the tree, the number of nodes decreases. Because of this the accuracy on the training set reduces, while there is increase in accuracies of testing and validation set.



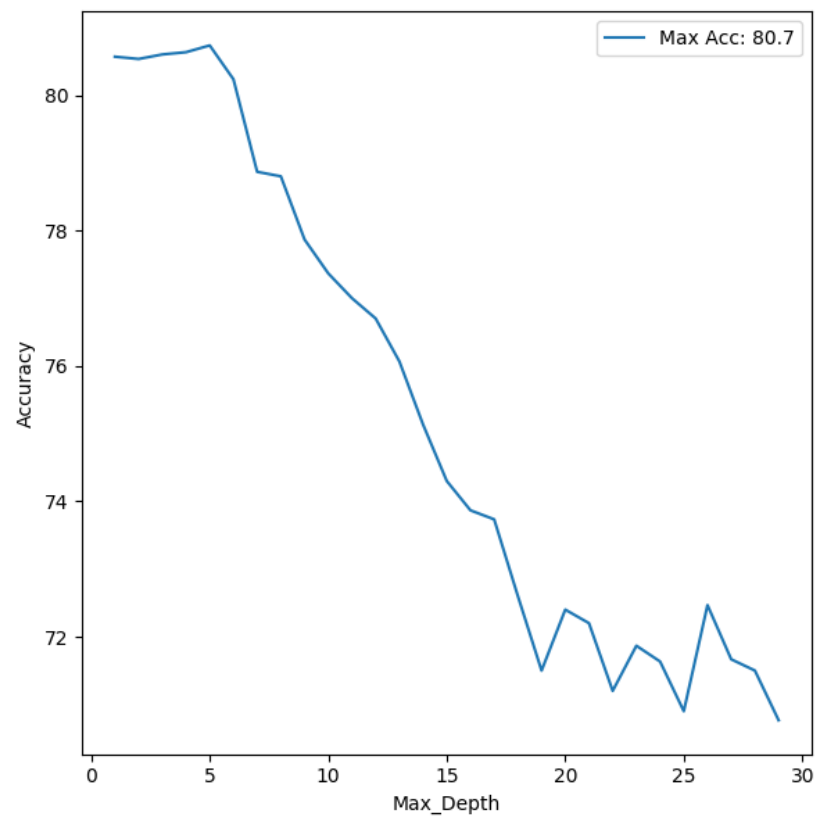
3.



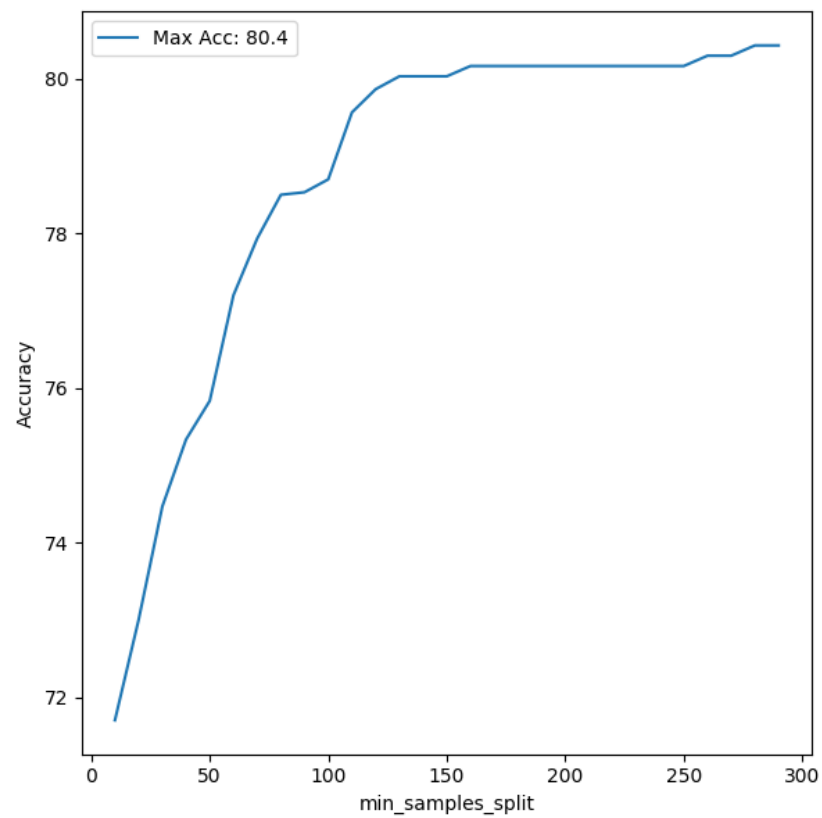
```
height of tree = 18
no of nodes in the tree = 1623
Train data Accuracy: 99.944
Test data Accuracy: 70.98
Validation data Accuracy: 72.45
att= 0 splits at= [120000.0, 255000.0, 280000.0, 340000.0, 425000.0]
att= 4 splits at= [27.5, 30.0, 36.0]
att= 11 splits at= [14835.0, 17441.0, 18537.0]
att= 12 splits at= [0.0, 1e-05, 1695.0]
att= 13 splits at= [0.0, 1e-05]
att= 14 splits at= [158.0, 784.0, 1663.0, 5277.0]
att= 15 splits at= [17787.5, 18762.0, 20361.0]
att= 16 splits at= [14799.0, 17328.0]
att= 17 splits at= [1200.0, 2015.0]
att= 18 splits at= [1808.0, 3000.0]
att= 19 splits at= [5000.0, 7000.0]
```

```
att= 20 splits at= [466.0]  
att= 21 splits at= [3500.0, 23738.0]  
att= 22 splits at= [3498.0, 5000.0]
```

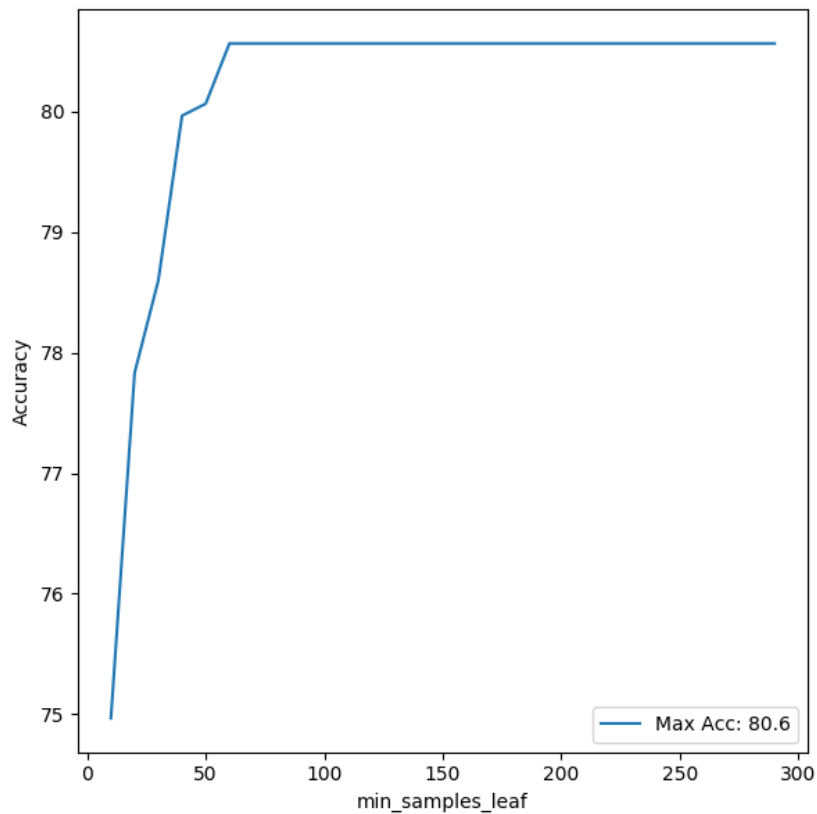
4.



As the max_depth parameter is increased, the accuracy on the validation set decreases. This suggests overfitting.



with the Increase in the `min_sample_split` parameter prevents the tree from growing too much, therefore the validation accuracy increases.



The `min_sample_leaf` has a similar effect on the tree as the `min_sample_split` parameter as it prevents the tree from growing too much hence preventing it from overfitting. It can be concluded that by decreasing the number of nodes in the tree the validation accuracy increases.

Problem 2: Neural Networks

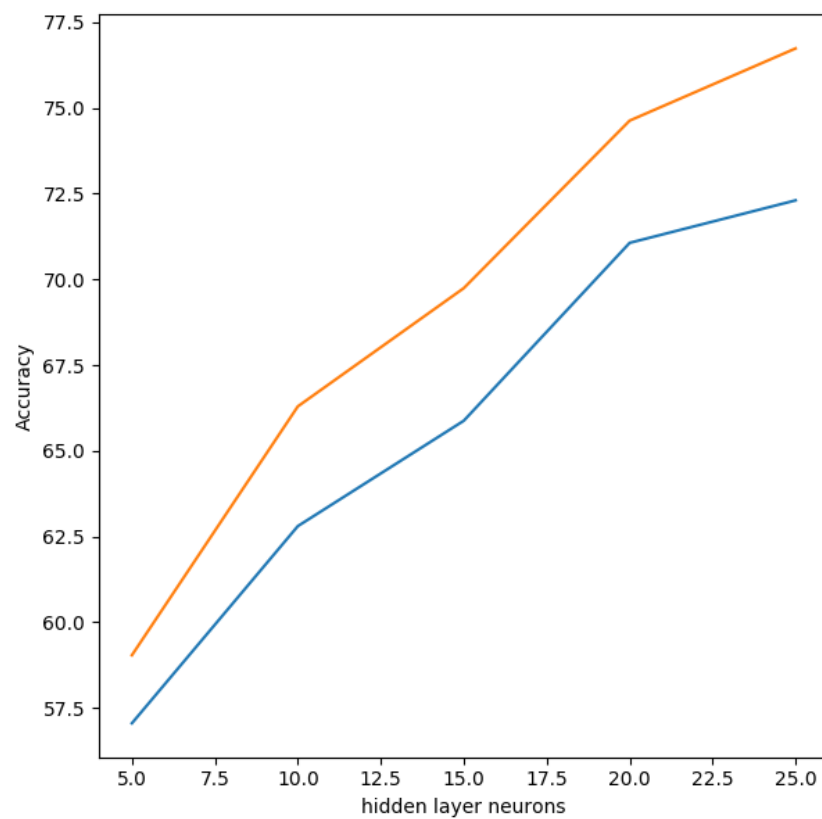
1. [csv files link](#).
2. Neural Network implemented.
3. Single hidden layer
Stopping Criteria: Number of epochs = 200
Accuracy for each hidden unit: 5,10,15,20,25
For training data-

- 62.98

- 66.25
- 69.11
- 73.33
- 81.18

For testing data-

- 60.8163
- 63.1418
- 64.6877
- 68.994
- 78.1292



confusion matrix for each hidden layer:
For hidden layers=5

Assignment 3

```
[[9723 2770    0    0    0    0    0    0    0    0]
 [5152 5447    0    0    0    0    0    0    0    0]
 [ 460  746    0    0    0    0    0    0    0    0]
 [ 147  366    0    0    0    0    0    0    0    0]
 [  46   47    0    0    0    0    0    0    0    0]
 [  43   11    0    0    0    0    0    0    0    0]
 [   8   28    0    0    0    0    0    0    0    0]
 [   0    6    0    0    0    0    0    0    0    0]
 [   5    0    0    0    0    0    0    0    0    0]
 [   3    2    0    0    0    0    0    0    0    0]]
```

For hidden layers=10

```
[[10039 2454    0    0    0    0    0    0    0    0]
 [ 3818 6781    0    0    0    0    0    0    0    0]
 [  225  981    0    0    0    0    0    0    0    0]
 [   45  468    0    0    0    0    0    0    0    0]
 [   39   54    0    0    0    0    0    0    0    0]
 [   43   11    0    0    0    0    0    0    0    0]
 [    2   34    0    0    0    0    0    0    0    0]
 [    0    6    0    0    0    0    0    0    0    0]
 [    4    1    0    0    0    0    0    0    0    0]
 [    4    1    0    0    0    0    0    0    0    0]]
```

For hidden layers=15

```
[[10304 2189    0    0    0    0    0    0    0    0]
 [ 3548 7051    0    0    0    0    0    0    0    0]
 [  215  991    0    0    0    0    0    0    0    0]
 [   30  483    0    0    0    0    0    0    0    0]
 [   73   20    0    0    0    0    0    0    0    0]
 [   49    5    0    0    0    0    0    0    0    0]
 [    1   35    0    0    0    0    0    0    0    0]
 [    0    6    0    0    0    0    0    0    0    0]
 [    4    1    0    0    0    0    0    0    0    0]
 [    5    0    0    0    0    0    0    0    0    0]]
```


For hidden layers=20

```
[[10722 1771 0 0 0 0 0 0 0 0]
 [ 2874 7725 0 0 0 0 0 0 0 0]
 [  121 1085 0 0 0 0 0 0 0 0]
 [   31  482 0 0 0 0 0 0 0 0]
 [   72   21 0 0 0 0 0 0 0 0]
 [   48    6 0 0 0 0 0 0 0 0]
 [    0   36 0 0 0 0 0 0 0 0]
 [    0    6 0 0 0 0 0 0 0 0]
 [    5    0 0 0 0 0 0 0 0 0]
 [    5    0 0 0 0 0 0 0 0 0]
 0      0      0]]
```

For hidden layers=25

```
[[11642  851 0 0 0 0 0 0 0 0]
 [ 2162 8437 0 0 0 0 0 0 0 0]
 [   42 1164 0 0 0 0 0 0 0 0]
 [   28  485 0 0 0 0 0 0 0 0]
 [   81   12 0 0 0 0 0 0 0 0]
 [   50    4 0 0 0 0 0 0 0 0]
 [    0   36 0 0 0 0 0 0 0 0]
 [    0    6 0 0 0 0 0 0 0 0]
 [    4    1 0 0 0 0 0 0 0 0]
 [    5    0 0 0 0 0 0 0 0 0]]
```