Name : Neeladri Das (18CS10036)
Avijit Mandal (18CS30010)

As our data is very big, contains about 40,000 instances so it takes about 12 minutes to build the Decision Tree, taking 10 Random splits would require approximately 120 minutes and finding the average accuracy over different heights would take 12~14 hours. So all our calculations and reports are based on the first **8000 instances of data.**

1.Best Tree Obtained:

Here is tree is allowed to go Upto As max possible over 10 Random shuffles
Height : 35
Accuracy: 80.375

2. Best Depth:

Over 20 inputs taken from user of max_depth , such that if max_depth is given -1 then max_depth possible build else upto max_depth tree is build (if possible else stopped at a particular point information gain is found to be 0)

Best Depth : 27
Accuracy : 80.625

3. Pruning:

For pruning we used reduced error pruning where we made a tree of best depth obtained from second part and trained it with **adult.data [4000:8000]** with training data 80% and validation data 20%

ALL THE GRAPH PLOTS ARE PLOTTED IN Untitled.ipynb