

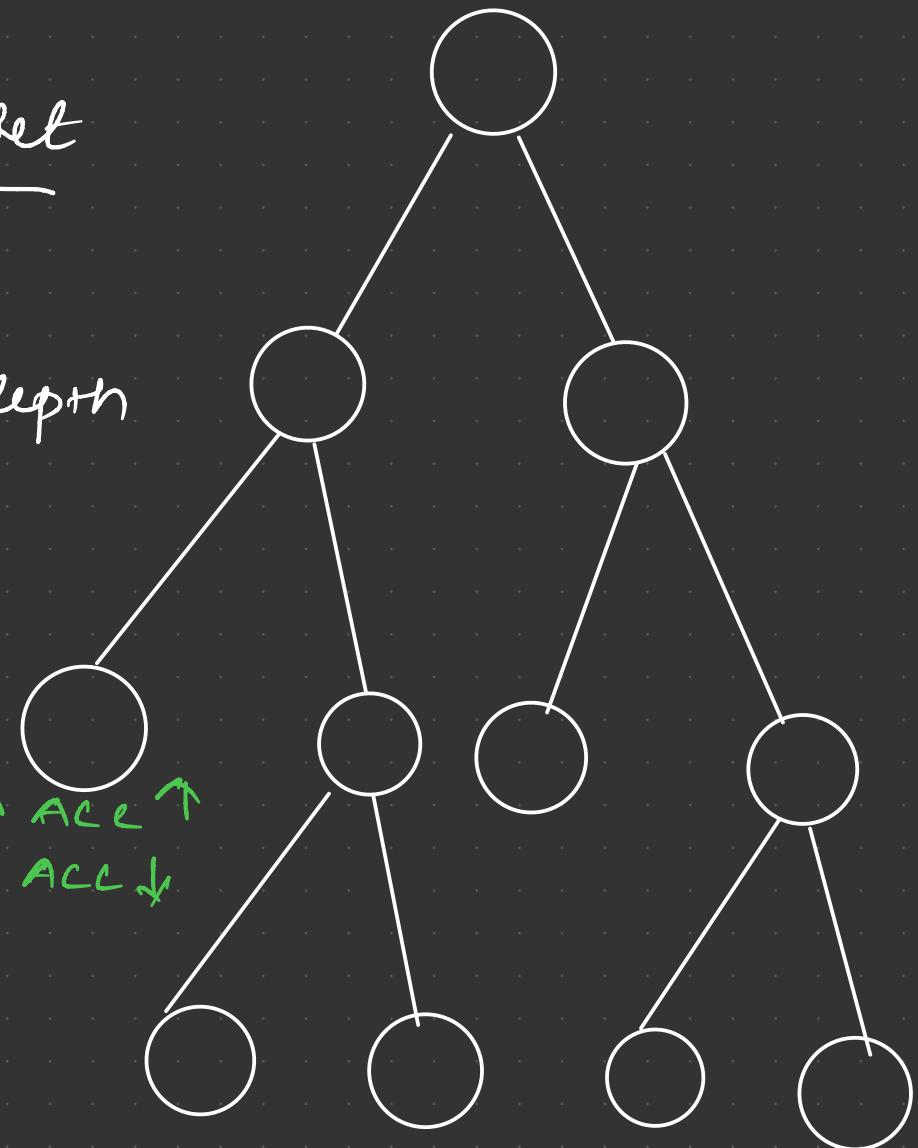
Post Pruning and Pre Pruning

Training Dataset

whenever we split to its depth it may result in Overfitting



low Bias Train ACC ↑
high Variance Test ACC ↓



we might have have got higher training accuracy, however the test accuracy might be low

how can we prevent the overfitting scenario?

Reduce overfitting

Post pruning

Pre Pruning

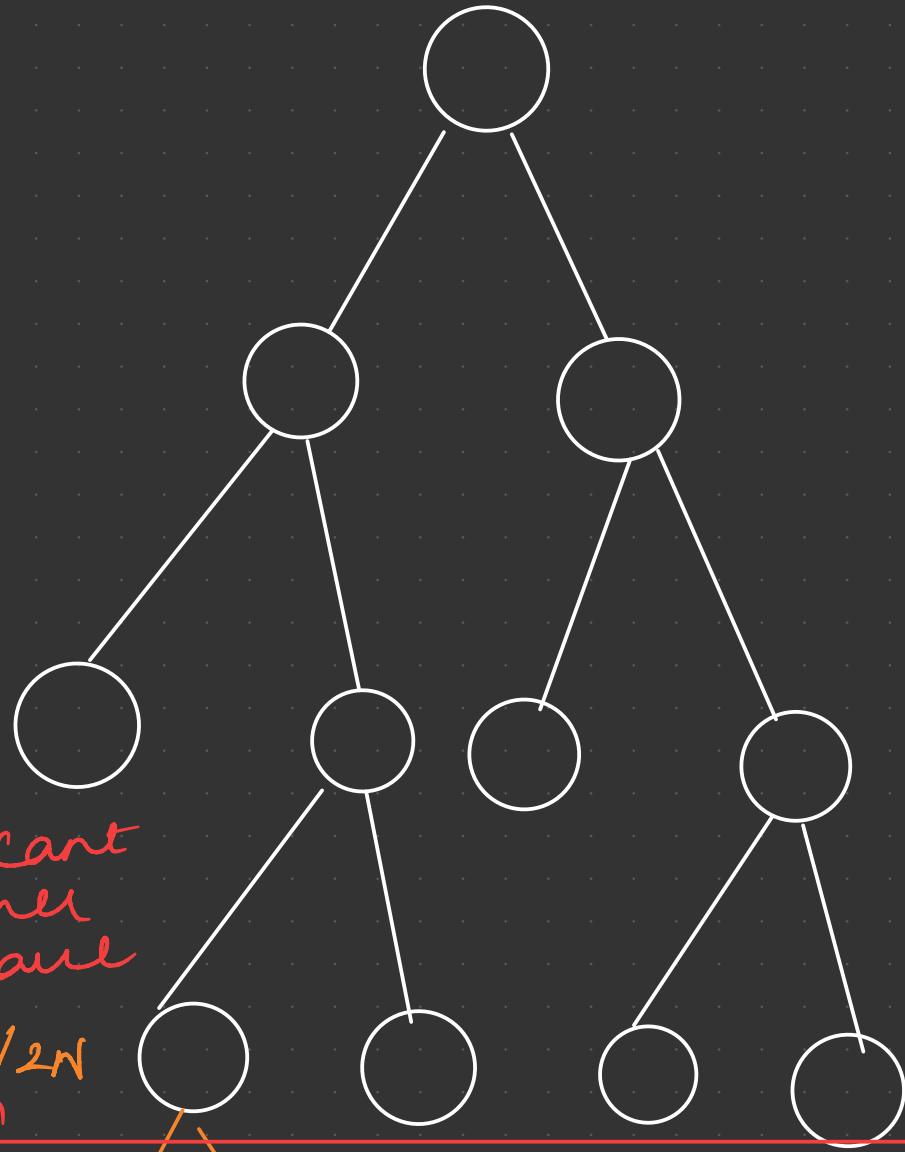
These are the two techniques to prevent overfitting

Post Pruning

I will construct the complete decision tree, after that I will start pruning it up.

what is pruning?

Pruning is the process of selective removal of nodes from the constructed tree.



$\frac{9y}{10n}$ $\frac{9y}{2n}$ $\frac{9y}{2n}$

Post Pruning



construct the decision tree



Prune it with respect to
its depth

Note:

Post pruning should be used for the smaller data sets.

Pre Pruning

In Pre Pruning technique we will not be constructing the tree entirely -

While constructing the decision tree we can play around with some of the parameters, so that we can prun the tree while constructing them.

Some of the parameters are,

- Max features
 - Max depth
 - Split.
- } these are hyperparameters

Generally, with pre pruning we tune with hyper parameter.

→ Hyper Parameter tuning while constructing decision tree.