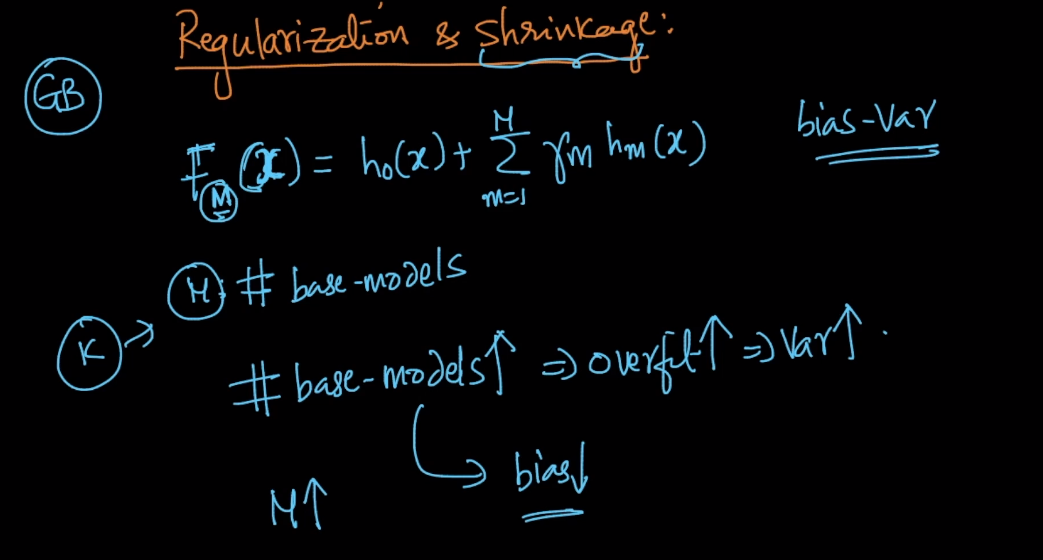
**Regularization by Shrinkage**

In boosting we calculate Fm(x) for mth model as shown below

But if we have more number of base models i.e no. of models increases then it starts overfitting i.e variance increases although bias decreases but variance increases

So to avoid this we use technique shrinkage

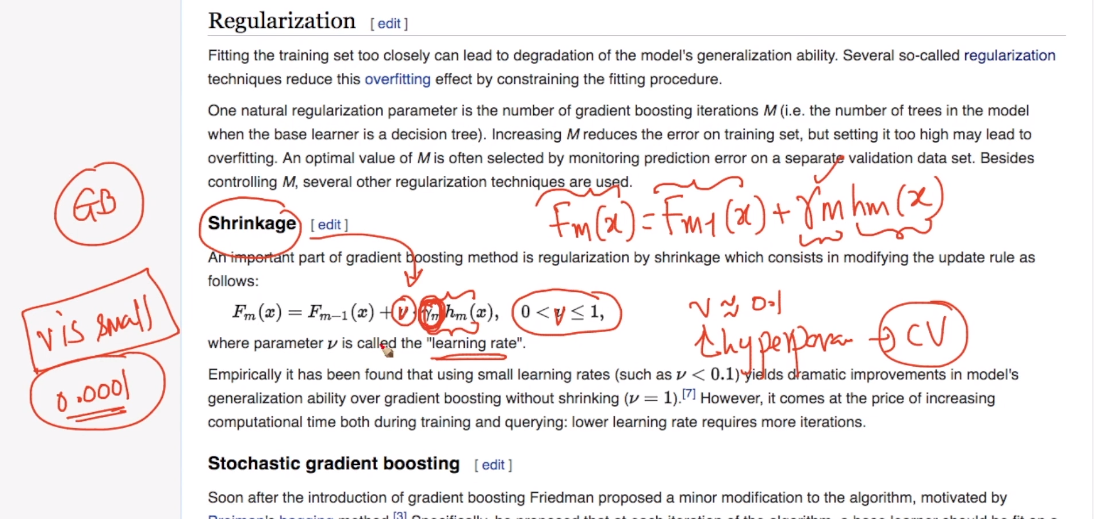


In shrinkage we use one more hyperparameter v which is learning rate.

Also v is small value.

So by applying v we are reducing weight which we found using gradient boosting gamma\_m \* hm(x)

i.e we are reducing gamma\_m by v so that we don’t overfit

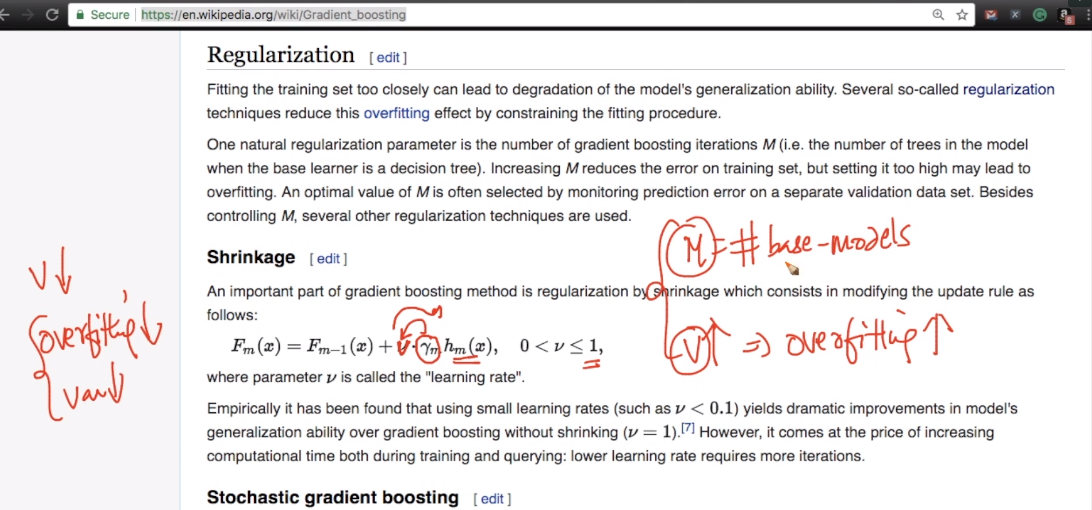


So if v is small or decreases we give less weightage to this models(gamma\_m \* hm(x)) therefore we are reducing variance and also problem of overfitting.

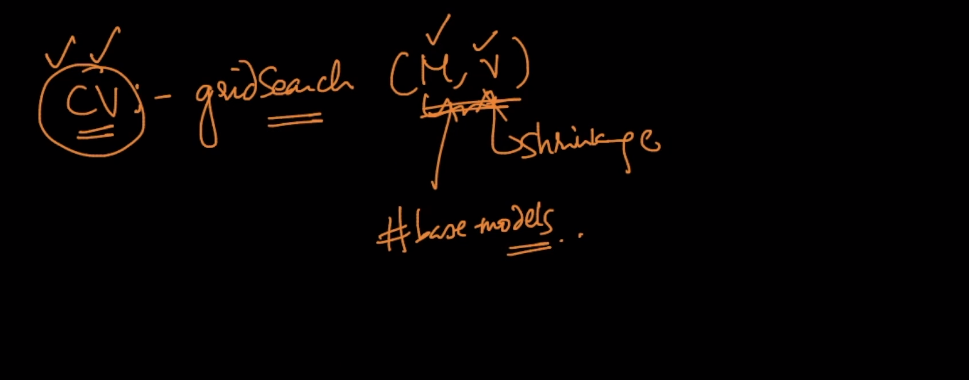
And as V increases it increase variance and thus it increases overfitting.

So we have to choose v carefully using cross validation

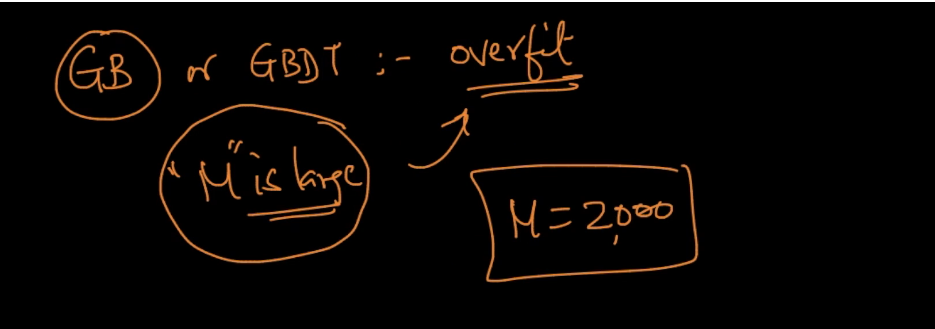
Therefore in boosting we have two hyperparameters M(no. of base models) , v(shrinkage)



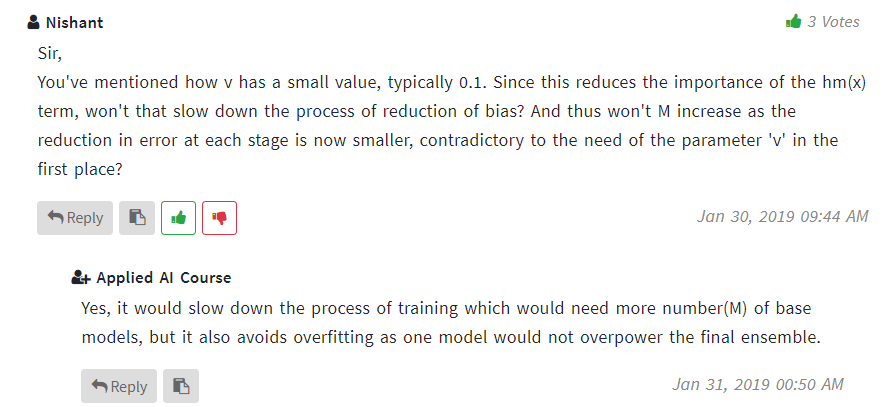
Ideally we should use gridsearch cross validation to find correct M and v

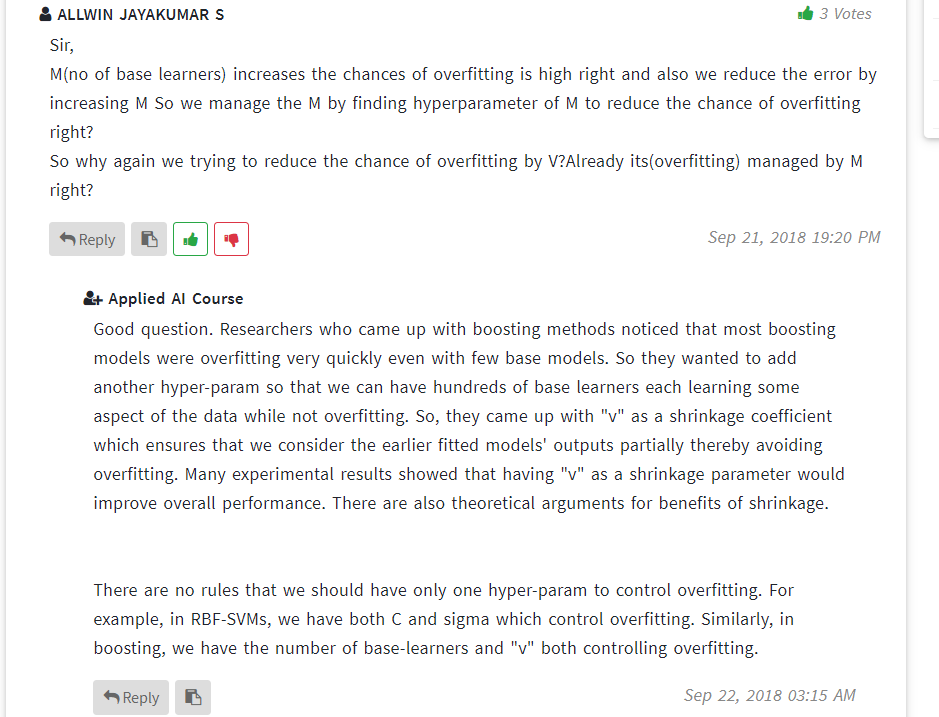


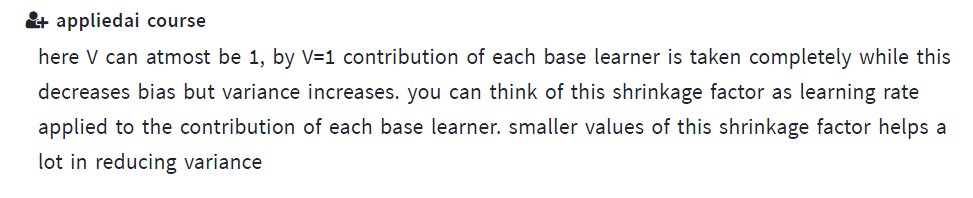
If M is large then GB or GBDT might overfit therefore we use another hyperparameter shrinkage as well



Comments:







Comments :

<http://blog.kaggle.com/2017/01/23/a-kaggle-master-explains-gradient-boosting/>