INDIAN INSTITUTE OF TECHNOLOGY MADRAS

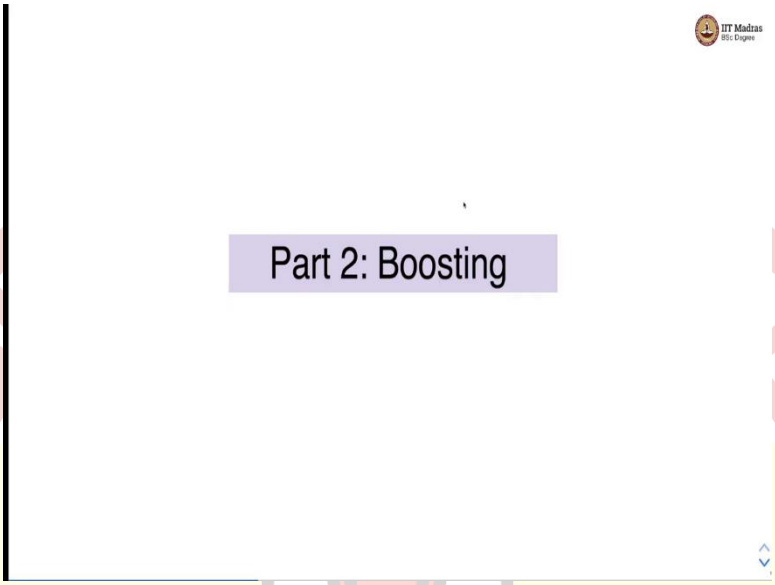
सिद्धिर्भवति कर्मजा

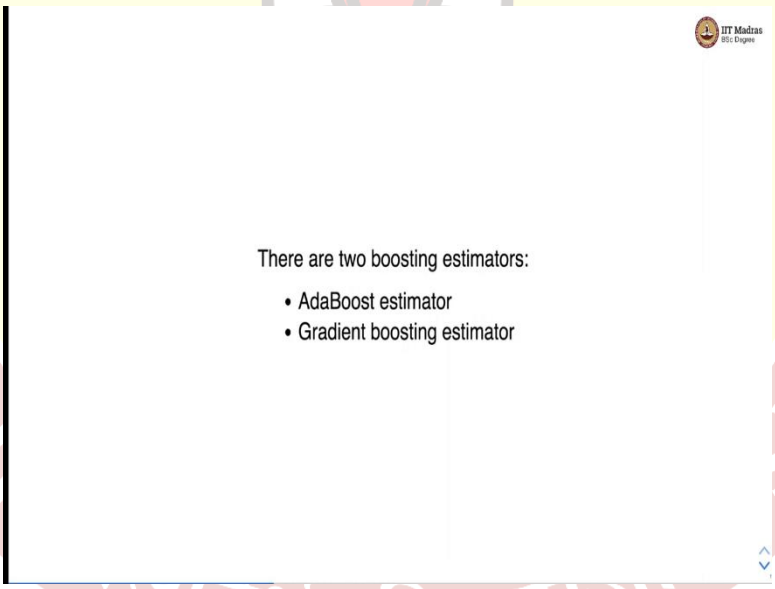
IIT Madras

ONLINE DEGREE

Machine Learning Practice
Professor. Ashish Tendulkar
Indian Institute of Technology, Madras
Boosting: AdaBoost, GradientBoost

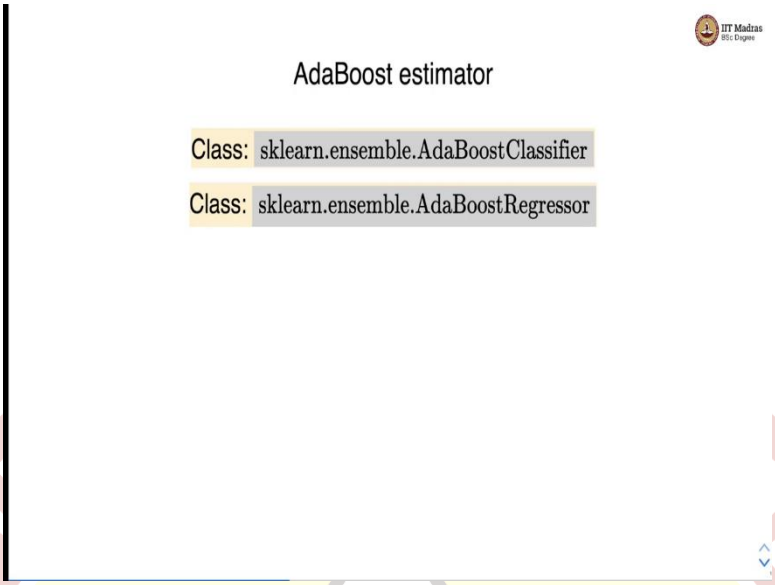
(Refer Slide Time: 0:10)

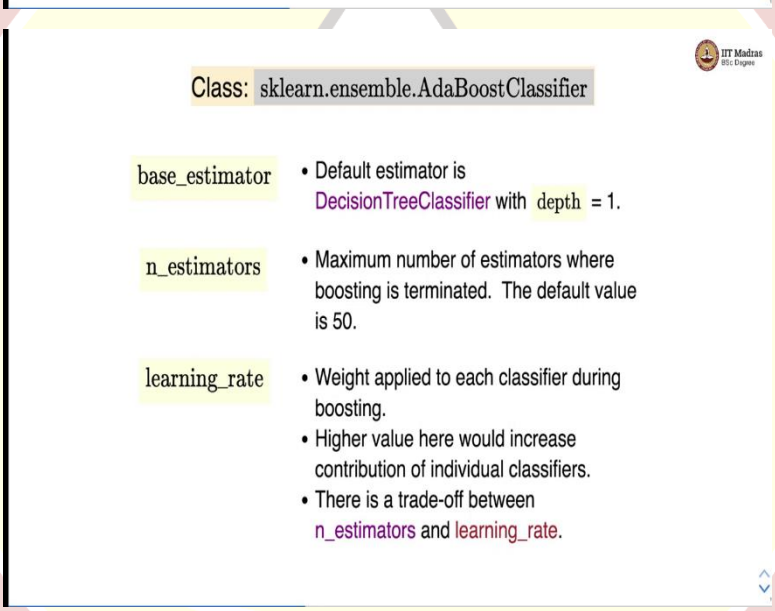




Namaste! Welcome to the next video of Machine Learning Practice Course. In this video, we will discuss about boosting techniques as implemented in sklearn. There are 2 boosting estimators AdaBoost estimator and Gradient boosting estimator that are of our interest as far as this course is concerned.

(Refer Slide Time: 0:33)

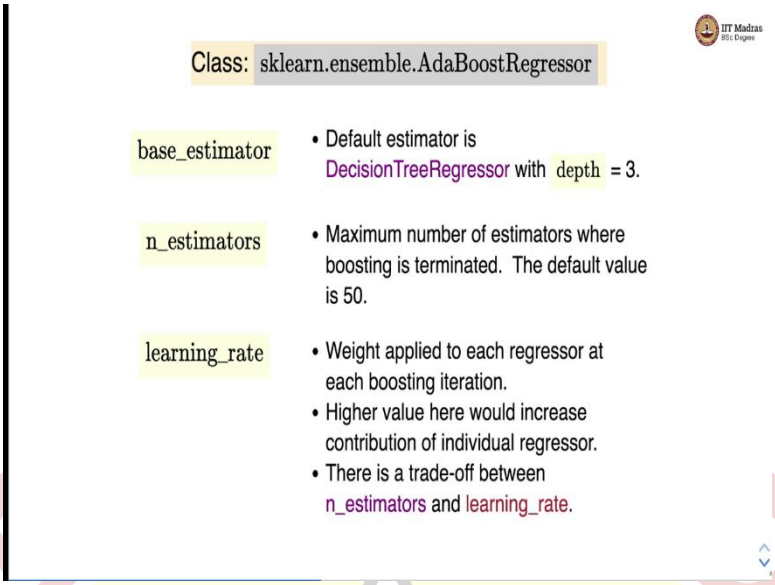




AdaBoost estimators there are 2 of them, one is AdaBoostClassifier, and AdaBoostRegressor and they are implemented as part of sklearn.ensemble model. AdaBoostClassifier takes base _estimator parameter that tells us what kind of estimator we will use to start the boosting procedure. The default estimator is DecisionTreeClassifier with depth = 1, then there is a parameter called n _estimators that provides the maximum number of estimators where boosting is terminated, the default value is 50.

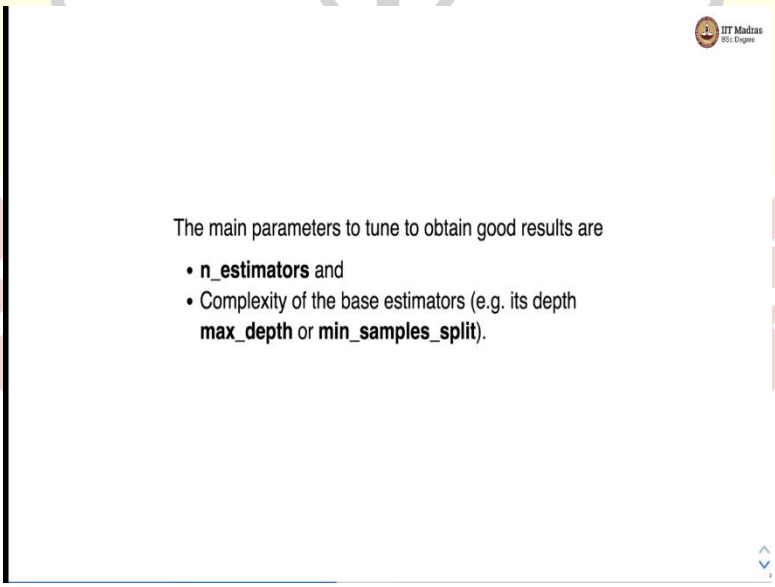
Then there is a learning _rate, which specifies the weight to be applied to each classifier during the boosting. Higher value for learning _rate would increase the contribution of individual classifiers. There is a trade-off between n _estimators parameter and learning _rate parameter.

(Refer Slide Time: 1:38)



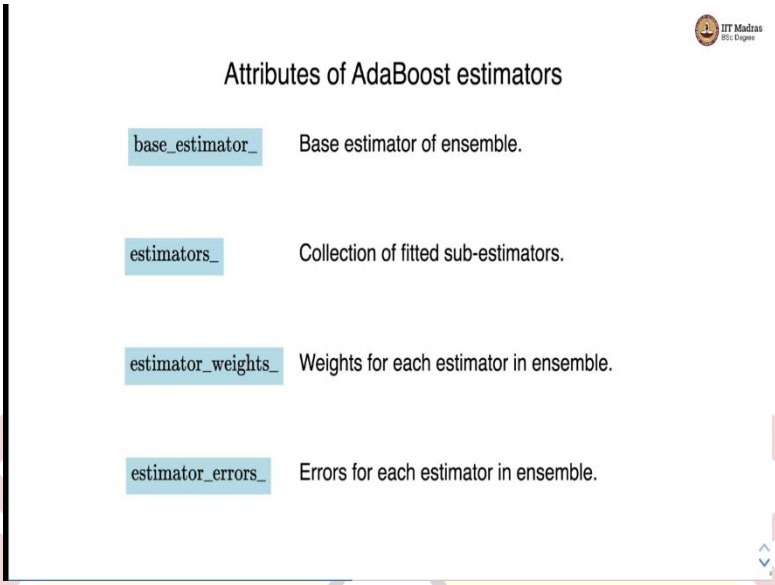
Let us look at AdaBoostRegressor. It also has a base _estimator. And the default base _estimator is DecisionTreeRegressor with depth = 3. Number of estimators are set to default value of 50 and they specify the maximum number of estimators where the boosting will be terminated. And learning _rate, which is weight applied to each regressor, higher value here would increase contribution of individual regressor and there is a trade-off between n _estimators and learning _rate.

(Refer Slide Time: 2:13)



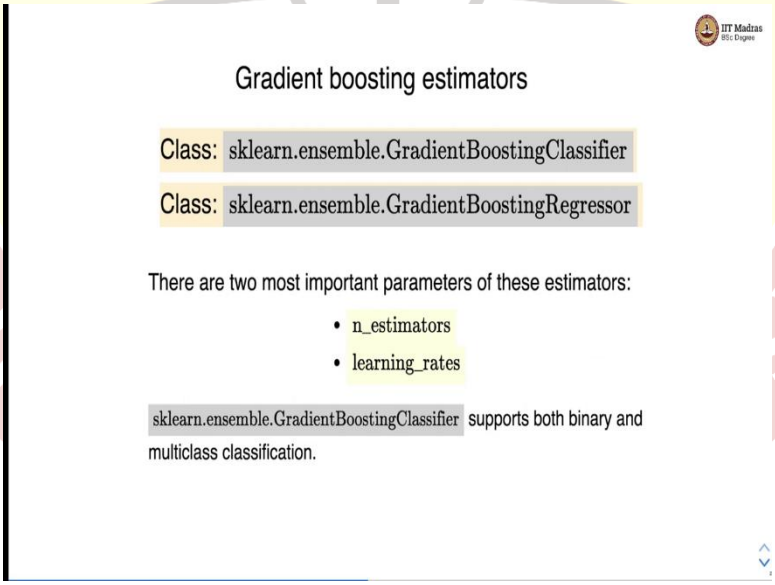
There are a couple of parameters that we can tune to obtain good results in case of AdaBoost estimators. One is the number of estimators and second is complexity of base _estimator, where the complexity is defined in terms of the depth of the tree, and minimum samples required to split a node.

(Refer Slide Time: 2:34)



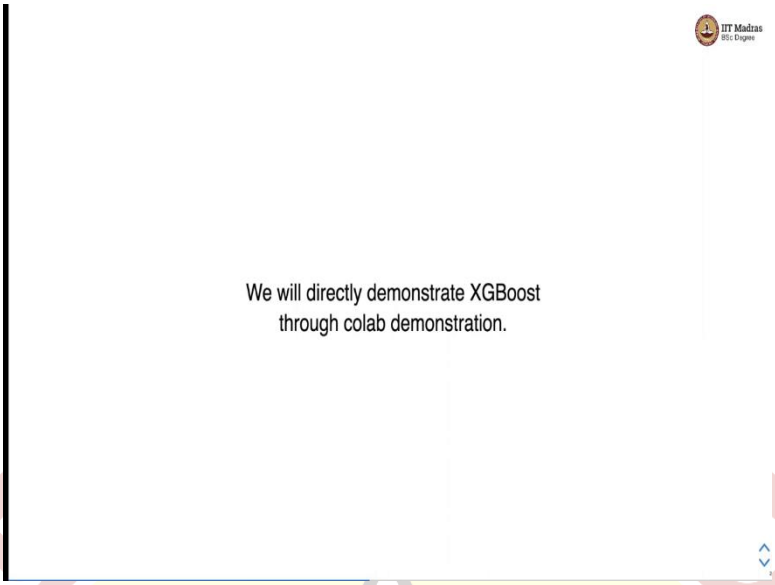
Let us look at attributes of AdaBoost estimators. So, there is an attribute which is base _estimator_ that provides basic estimator of the ensemble. Estimator_ is a collection of fitted sub-estimators. An estimator_weights_ provides the weights of each estimator in ensemble. An estimator _errors_ provides errors of each estimator in ensemble.

(Refer Slide Time: 3:05)



Let us look at gradient boosting estimators, there are there is a GradientBoostingClassifier and GradientBoostingRegressor. And there are 2 important parameters in these estimators again, which is number of estimators where the boosting will be terminated under learning _rate. GradientBoostingClassifier supports both binary and multiclass classification.

(Refer Slide Time: 3:31)



And finally, there is one more boosting method which is XGboost which we will demonstrate directly through collab. So, in this video, we look at various boosting estimators in sklearn. So, AdaBoost and gradient boosting is implemented in sklearn, whereas XGBoost is not directly implemented in sklearn. And we will be implementing an XGboost with some other library. And we will see that through the collab demonstration.