Adaboost: you can have many base estimators like random forest, with this difference that you select some random columns and rows for the first estimator. For the second, you give random rows and columns and data that first estimator did wrong. For the third estimator, you give it the data that second estimator failed plus random columns and rows. You go forward in this way. The when you get a data to predict its label. You vote number of estimators that predicted x with rate of y%. then you calculate what class is the answer. In Adaboost you can select any estimator to be :random\_forest, decision tree or svm.

Random forest: In random forest you choose different columns for every estimator, and you do every estimate by a different decision tree. At the end you vote for the label. Random forest is for the times that you have many features.

Svm: is by plotting different lines and dividing the surface in different sections in 2 dimension problem.

Decision tree: divide the problem to different parts and make it smaller and smaller to solve the classification problem.

Measures for classification problem:

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | P | |
|  |  | y | n |
| E | y | TP | FN |
| n | FP | TN |

1- Accuracy= TP+TN/ Total

2- Recall= TP/TP+FN ---- if really it is x, what is the probability that it guesses x.

3- Precision= TP/TP+FP----- if it predicts x, how much probable that it guessed right.

4- Confusion Matrix--- that has all these information above inside.

Some scientists believe that if you make your target data normal or uniform distribution like, you will have a more accurate results in regression.