**DIFFERENCE BETWEEN BAGGING BOOSTING AND STACKING**

Bagging aims to decrease variance, boosting aims to decrease bias, and stacking aims to improve prediction accuracy. Bagging and boosting combine homogenous weak learners. Stacking combines heterogeneous solid learners. Bagging trains models in parallel and boosting trains the models sequentially.

All three are techniques of ensemble learning, here we will start to know how these technique works-

**Bagging**

Let there be a class and a teacher have 5 problems or we can say that we have a sample set with some data points.

There are 5 problems and the teacher creates 3 sample spaces or models (a,b,c) to solve problems-

* a has been assigned to solve problems-(1,4,5,3)
* b has been assigned to solve problems-(1,2,3,4)
* c has been assigned to solve problems-(2,1,3,4,5)

Here replacement and repetition are allowed in assigning problems. So we create these bootstrap samples to solve problems. Now, all 3 sample spaces solve problems in parallel.

* a answers A.
* b answers A.
* c answers B.

So, here our answer will be with maximum vote or the answer that gets maximum vote= A

**Boosting**

* A teacher has 5 problems (1,2,3,4,5). He calls a student to solve these 5 problems. The student solves the problems and his output is (1,2) which is the same or matched as the prediction. But 3,4,5 problems give different outputs or the prediction is wrong.
* Now, he calls another student 2 to solve (3,4,5) problems. Now, the output and prediction for 3 are matched but for (4,5) mismatched. Again teacher calls student 3 to solve problems (4,5).
* It means in comparison to Bagging, the approach of Boosting is to work sequentially. Here the problems are not independent and have dependencies on others. Here the models will not work parallelely to solve problems.
* So, it means first a weak learner comes to solve the problems and gives some output and based on his output the next weak learner comes to solve the problem and gives some output. Now, again the same process starts till we get the actual result and our prediction matches.

In this way, our learning becomes strong as there is a dependency on each other.

**Stacking**

* It is a mixture of both Bagging and Boosting. Here, we take 3 sample spaces and give the question-

(1,2,3) to a

(2,3,4,5) to b

(1,2,3) to c

* So, here we make 3 samples and assign them to a, b, and c. All solve the problem and the answer that comes from that we don’t do voting.
* Here we create a dataset of the answer given by (a,b,c). Now, we call a topper or the most intelligent among them to solve that answer dataset.
* On, this basis we give our output and in stacking it is called a Meta Model.

**Summary:**

* The three sample spaces we create to solve problems are our models. These models will work on the problems and train them in this way like logistic regression, decision boost, or Knn boost or etc.
* But at last, our output will come in terms of Meta Model which works best among all results will be our output.
* So, stacking came up with an idea to solve the complexities of **Bagging and boosting.**