

Do you want
to know about

Bootstrap Sampling?

-letthedataconfess →

@letthedataconfess

What is Bootstrap Sampling?

Bootstrap Sampling is a statistical technique that involves drawing of sample data repeatedly with replacement from a data source to estimate population parameters using samples.

Sampling with replacement means that a data point is added back in the sample, after it is drawn from it once.



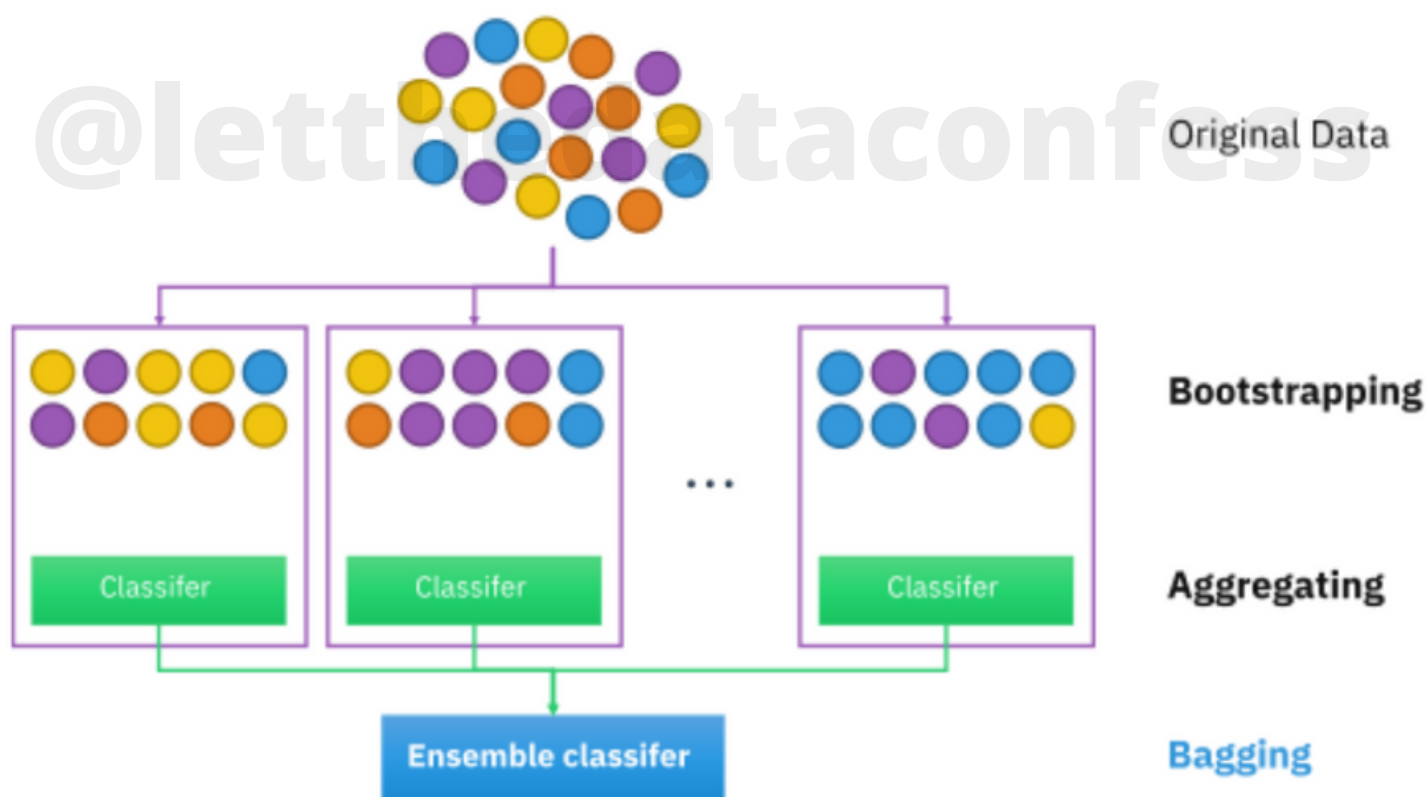
What is the need of Bootstrap Sampling?

- 1.** When we have to estimate a parameter of a large population, we can take the help of Bootstrap Sampling. We take a small sample from the population and calculate the statistics, then add the samples drawn back into the population. Then, we repeat this procedure 'n' number of times and take the average statistic as the estimate of the population parameter.
- 2.** It helps in avoiding overfitting and improves the stability of machine learning algorithms.



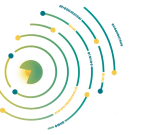
Where is Bootstrap Sampling used in Machine Learning?

Bootstrap sampling is used in machine learning ensemble algorithm called bagging (also called bootstrap aggregating).



It is used in many ensemble machine learning algorithms like random forests, AdaBoost, gradient boost, and XGBoost.





How is it done?



Select the sample size

Select an observation from the training data randomly

Now add this observation to the previously selected sample

The samples not selected are usually referred to as the “out-of-bag” samples. For a given iteration of Bootstrap resampling, a model is built on the selected samples and is used to predict the out-of-bag samples.

For getting better results, it is always better to increase the number of repetitions.



Ace Your Next DATA SCIENCE INTERVIEW

with our
Mock Interviews

CLICK HERE 

BOOK NOW

OR

Visit: www.letthedataconfess.com



S c h e d u l e Y o u r M o c k I n t e r v i e w N o w !



www.letthedataconfess.com



@letthedataconfess

LET THE DATA CONFESS

Understand | Learn | Code | Implement

