**NAME- ARNAV GUPTA**

**USN- ENG19CS0041**

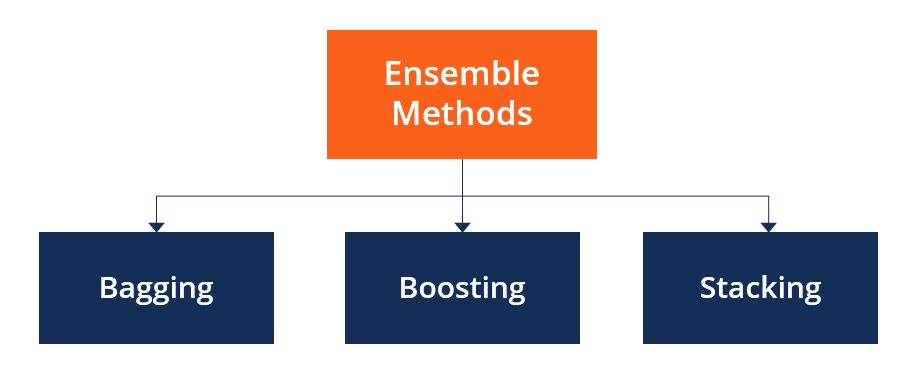
**CLASS- 6 SEM, A SECTION**

**SUBJECT- DATA SCIENCE**

**ENSEMBLE METHODS**

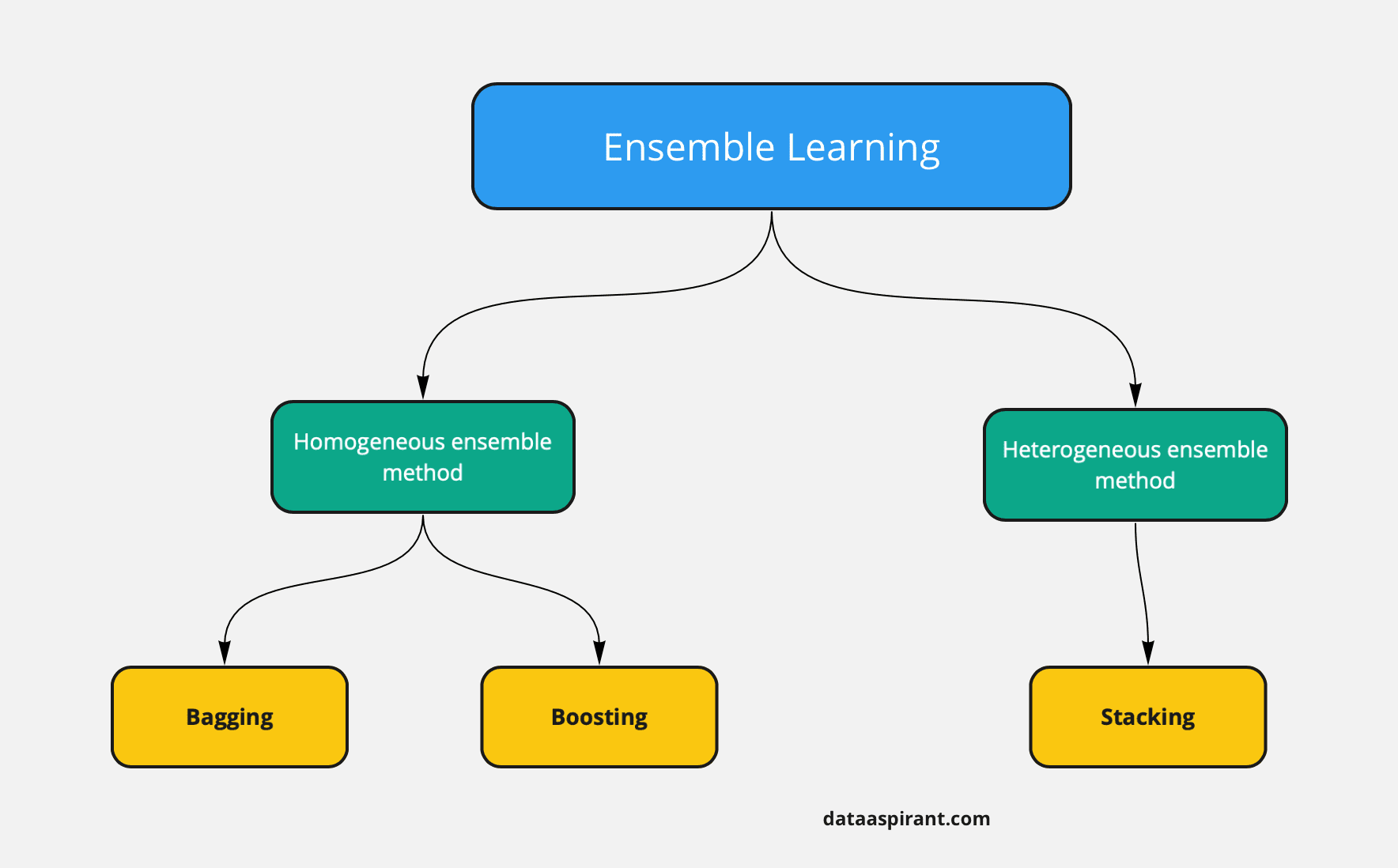
Ensemble Methods is a machine learning method that combines a few basic models to produce one accurate predictive model.

Ensemble methods are techniques which aims at improving the accuracy of results in models by integrating multiple models instead of a single model. The integrated models increase the accuracy of the results significantly which has boosted the popularity of ensemble methods in [machine learning](https://courses.corporatefinanceinstitute.com/courses/machine-learning-python-fundamentals).



### Categories of Ensemble Methods:

1. **Sequential ensemble techniques** generate base learners in a sequence, e.g., Adaptive Boosting. The sequential generation of base learners builds up the dependence between the base learners. The performance of the model is then improved by allocating higher weights to previously mislead learners.
2. In **parallel ensemble techniques**, base learners are yield in a parallel format, e.g., [random forest](https://corporatefinanceinstitute.com/resources/knowledge/other/random-forest/). Parallel methods harness the parallel generation of base learners to comforts independence between the base learners. The independence of base learners significantly decreases the error due to the application of averages.



### Main Types of Ensemble Methods

1. **BAGGING:** Bagging, a short form of bootstrap integration, is widely used in classification and regression. Increase the accuracy of the models by using decision trees, which greatly reduces variability. Variation reduction increases accuracy, eliminates overfitting, which is a challenge for many speculative models.

Bagging is advantageous as weak learners are integrated to form one strong learner who is more stable than one individual learner. It also eliminates any variations, thus reducing the overfitting of models. Another limitation is that it is computationally expensive. Therefore, it can lead to more likelihood of models where the proper bagging procedure is ignored.

1. **BOOSTING:** Boosting is an essemble method that learns from past prediction errors to make better predictions for the future. This approach incorporates a few weak learners who make one strong learner, thus greatly improving modeling.

Boosting works by arranging weak learners in sequence, so that weak learners learn from next learner in sequence in order to build better guessing models.

1. **Stacking:** Stacking is also another ensemble method, is often referred to as stacked generalization. It works by allowing a training algorithm to ensemble several other similar learning algorithm predictions.

Stacking has been successfully implemented in regression, density estimations, distance learning, and classifications. It can also be used to measure the error rate involved during bagging.

### Variance Reduction

* Ensemble methods are ideal for reducing the variance in models, hence increasing the accuracy of predictions.
* The variance is eliminated when many models are integrated to form a single prediction that is chosen from all other possible predictions from the combined models.
* An ensemble of models combines various models to ensure that the resulting prediction is the best possible, based on the consideration of all predictions.

### ****Summary****

* **Ensemble methods aims at improving predictability in models by integrating several models to make one very dependable model.**
* **The most admired ensemble methods are boosting, bagging, and stacking.**
* **Ensemble methods are good-fit for regression and classification, where they reduce bias and variance to improve the accuracy of models.**