|  |  |
| --- | --- |
| **Bagging** | **Boosting** |
| * **Bagging helps to decrease the model’s variance.**  1. **Bagging** is the simplest way of combining predictions that belong to the same type  * In **Bagging** each model receives equal weight * In **Bagging** each model is built independently * In **Bagging** different training data subsets are randomly drawn with replacement from the entire training dataset. * **Bagging** tries to solve over-fitting problem * If the classifier is unstable (high variance), then we should apply **Bagging**. * **Bagging** is extended to Random forest model | * **Boosting helps to decrease the model’s bias** * **Boosting** is a way of combining predictions that belong to the different types. * **Boosting** models are weighted according to their performance. * **Boosting** new models are influenced by performance of previously built models. * In **Boosting** every new subsets contains the elements that were misclassified by previous models. * **Boosting** tries to reduce bias. * If the classifier is stable and simple (high bias) then we should apply **Boosting**. * **Boosting** is extended to **Gradient boosting**. |

* **Bagging** and **boosting** are similar in that they are both ensemble techniques, where a set of weak learners are combined to create a strong learner that obtains better performance than a single one.
* These methods are designed to improve the stability and the accuracy of Machine Learning algorithms. Combinations of multiple classifiers decrease variance, especially in the case of unstable classifiers, and may produce a more reliable classification than a single classifier.

<https://www.kaggle.com/prashant111/bagging-vs-boosting>

**Boosting**

|  |  |
| --- | --- |
| **Partitioning of data** | **Higher vote to misclassified samples** |
| **Goal to achieve** | **Increase accuracy** |
| **Methods used** | **Gradient descent** |
| **Functions to combine single model** | **Weighted majority vote** |
| **Example** | **Ada Boost** |

**Bagging**

|  |  |
| --- | --- |
| **Partitioning of data** | **Random** |
| **Goal to achieve** | **Minimum variance** |
| **Methods used** | **Random subspace** |
| **Functions to combine single model** | **Weighted average** |
| **Example** | **Random Forest** |

