Ensemble Learning

Ensemble Learning

This method helps in improving machine learning results by combining several models. This contributes positively to predictive accuracy. The primary concept behind ensemble models is to combine weak learners to form active learners. Bagging and Boosting are two types of ensemble learning.

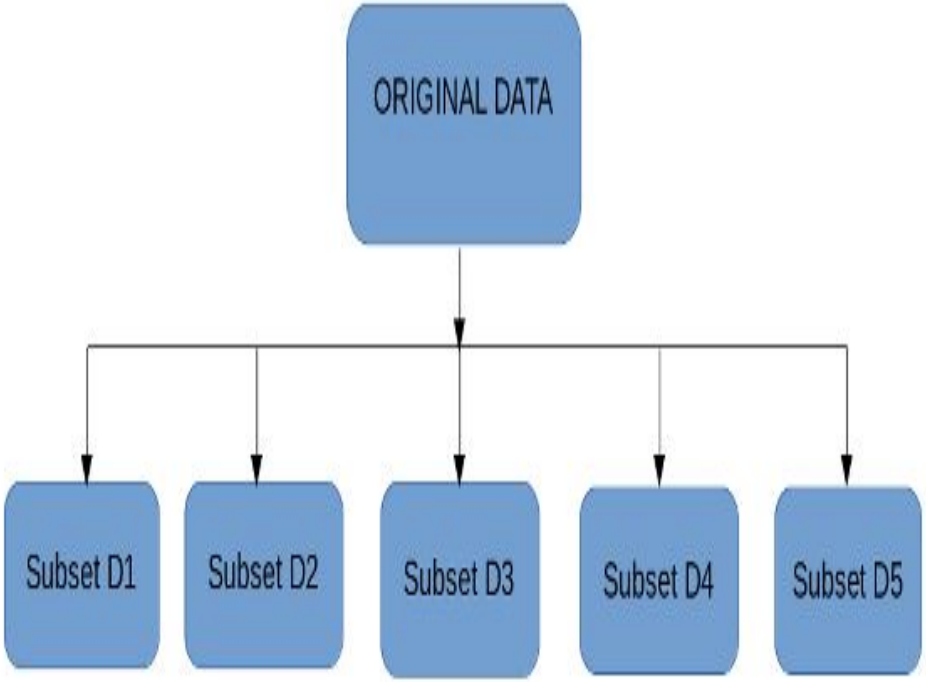
Bagging

- The idea behind bagging is combining the results of multiple models (for instance, all decision trees) to get a generalized result. Here’s a question: If you create all the models on the same set of data and combine it, will it be useful? There is a high chance that these models will give the same result since they are getting the same input. So how can we solve this problem? One of the techniques is bootstrapping.

Bagging

- Bootstrapping is a sampling technique in which we create subsets of observations from the original dataset, **with replacement**.
- Bagging (or Bootstrap Aggregating) technique uses these subsets (bags) to get a fair idea of the distribution (complete set). The size of subsets created for bagging may be less than the original set.

Bagging



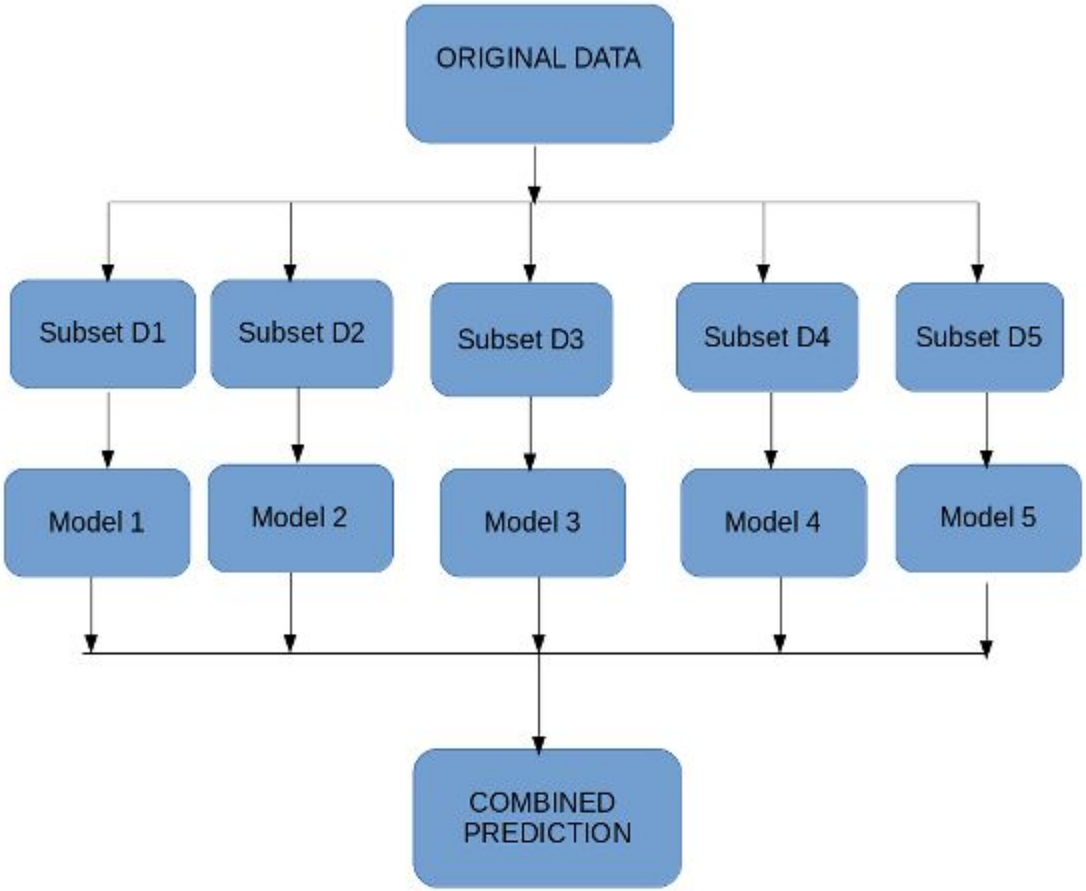
Multiple subsets are created from the original dataset, selecting observations with replacement.

A base model (weak model) is created on each of these subsets.

The models run in parallel and are independent of each other.

The final predictions are determined by combining the predictions from all the models

Bagging

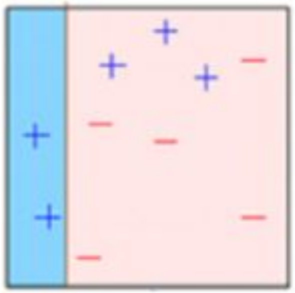

ORIGINAL DATA
Subset D1
Subset D2
Subset D3
Subset D4
Subset D5
Model 1
Model 2
Model 3
Model 4
Model 5
COMBINED
PREDICTION

Boosting

- Before we go further, here’s another question for you: If a data point is incorrectly predicted by the first model, and then the next (probably all models), will combining the predictions provide better results? Such situations are taken care of by boosting.
- Boosting is a sequential process, where each subsequent model attempts to correct the errors of the previous model. The succeeding models are dependent on the previous model.
- Let’s understand the way boosting works in the below steps.

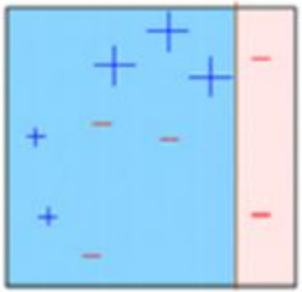
Boosting

- A subset is created from the original dataset.
- Initially, all data points are given equal weights.
- A base model is created on this subset.
- This model is used to make predictions on the whole dataset



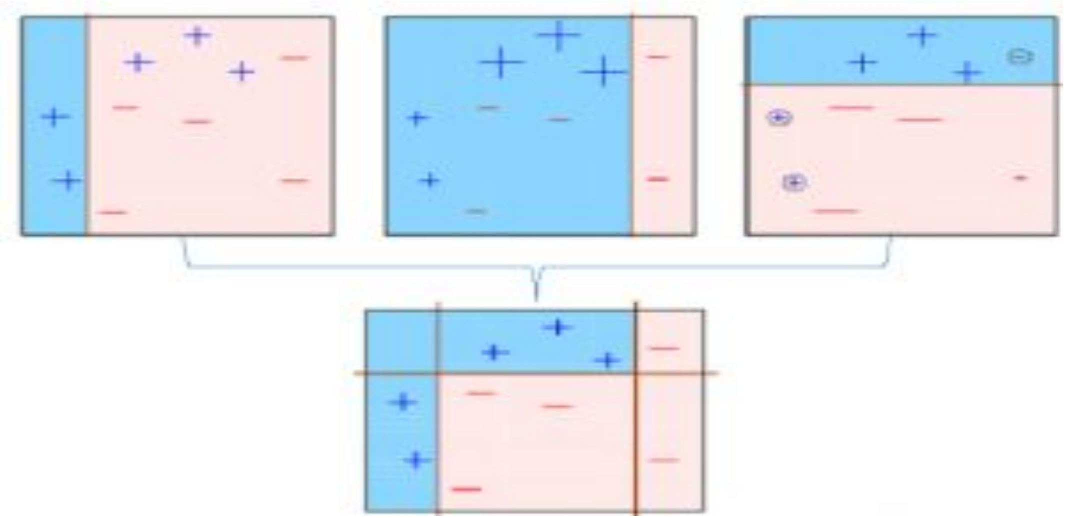
Boosting

- Errors are calculated using the actual values and predicted values.
- The observations which are incorrectly predicted, are given higher weights. (Here, the three misclassified blue-plus points will be given higher weights)
- Another model is created and predictions are made on the dataset. (This model tries to correct the errors from the previous model)

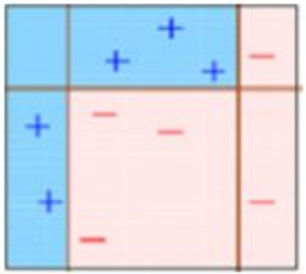


Boosting

- Similarly, multiple models are created, each correcting the errors of the previous model.
- The final model (strong learner) is the weighted mean of all the models (weak learners)



- Thus, the boosting algorithm combines a number of weak learners to form a strong learner. The individual models would not perform well on the entire dataset, but they work well for some part of the dataset. Thus, each model actually boosts the performance of the ensemble.



Algorithms based on Bagging and Boosting

- Bagging and Boosting are two of the most commonly used techniques in machine learning.
- Bagging algorithms:
  - Bagging meta-estimator
  - Random forest
- Boosting algorithms:
  - AdaBoost
  - GBM
  - XGBM
  - Light GBM
  - CatBoost

Bagging

Bagging stands for Bootstrap aggregating, which combines several models for better predictive results. In statistical classification and regression, bagging improves the stability and accuracy of machine learning algorithms by decreasing the variance and reducing the chances of overfitting.

Steps involved in bagging

- The original dataset is divided into multiple subsets, selecting observations with replacements. This process of random sampling is called bootstrapping.
- A base model is created on each of the subsets.
- The subsets are independent of each other; hence the training of each model is done in parallel.
- We derive the final prediction by combining the predictions from all the models.

The most common implementation of bagging ensemble learning is the Random Forest .Thus, the technique makes the random selection of features rather than using all features to develop trees.

Boosting

Boosting involves building a strong classifier from several weak classifiers using the weak models in series. The first step is to build a model from the training set. Then we create the second model, which tries to correct the error incurred while training the first. This process is continued while adding new models until the maximum number of models is reached, or the training data is finished.

Gradient Boosting or Adaboost is an implementation of boosting. AdaBoost stands for Adaptive Boosting and is a technique that combines multiple weak classifiers into a single strong classifier. Gradient boosting uses the gradient descent algorithm that minimizes any differentiable loss function.

Difference Between Bagging and Boosting

Bagging	Boosting
The original dataset is divided into multiple subsets, selecting observations with replacement.	The new subset contains the components mistrained by the previous model.
This method combines predictions that belong to the same type.	This method combines predictions that belong to the different types.
Bagging decreases variance.	Boosting decreases bias.
Base classifiers are trained parallelly.	Base classifiers are trained sequentially.
The models are created independently.	The model creation is dependent on the previous ones.

Similarities Between Bagging and Boosting

- Both are ensemble methods that improve the stability of the machine learning model.
- Both generate one learner from multiple learners.
- The final decision is by combining the predictions of the N learners.
- Both algorithms help in dealing with the bias-variance trade-off.
- Both can be used to solve classification as well as regression problems.

Conclusion

- Ensemble learning helps in improving machine learning results by combining several models.
- Bagging involves fitting many decision trees on different samples of the dataset and averaging the predictions.
- Boosting involves adding ensemble members sequentially to correct the predictions made by prior models and outputs a weighted average of the predictions.