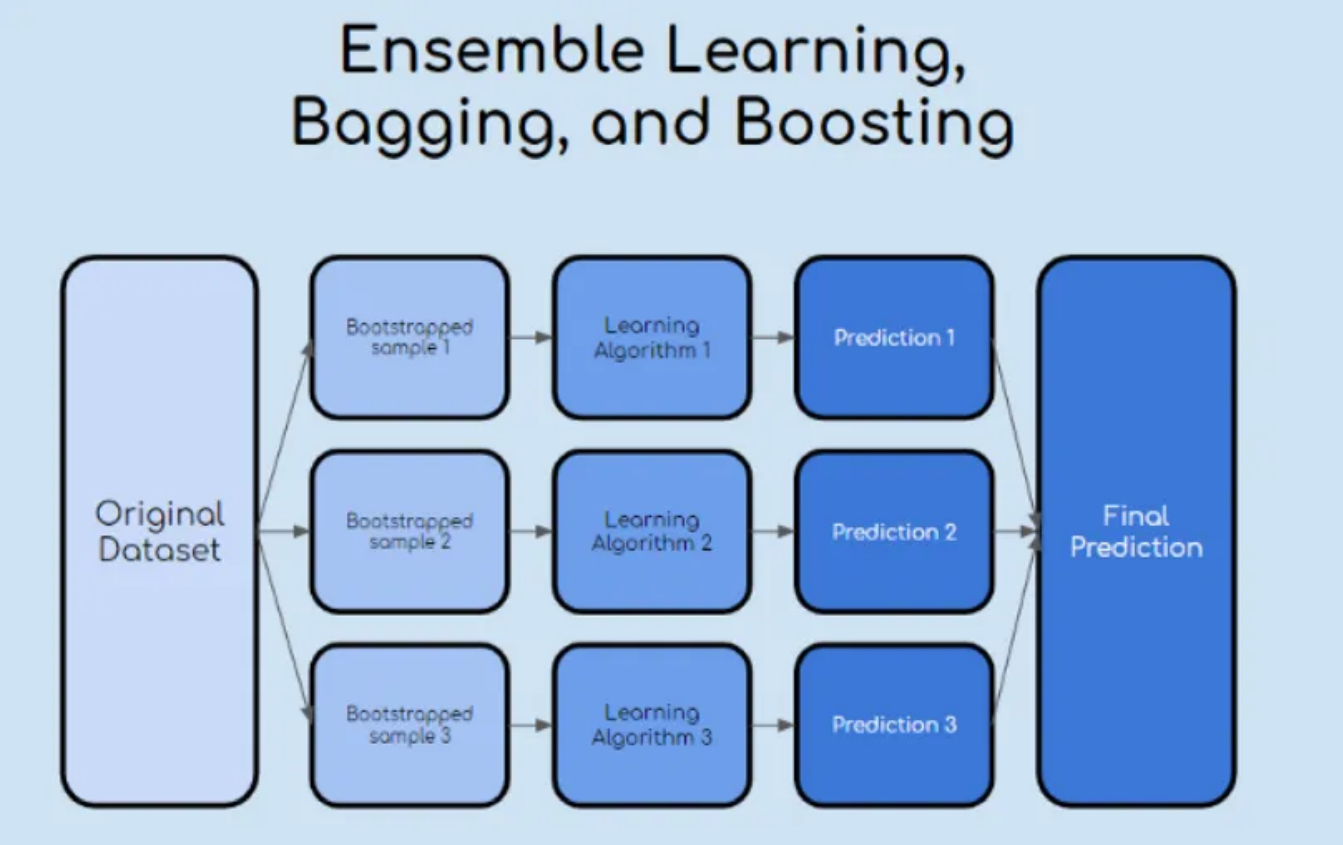
Everything you need to know about Ensemble Learning

Great Learning · Follow
4 min read · Apr 29, 2021

Contributed by: Venkat

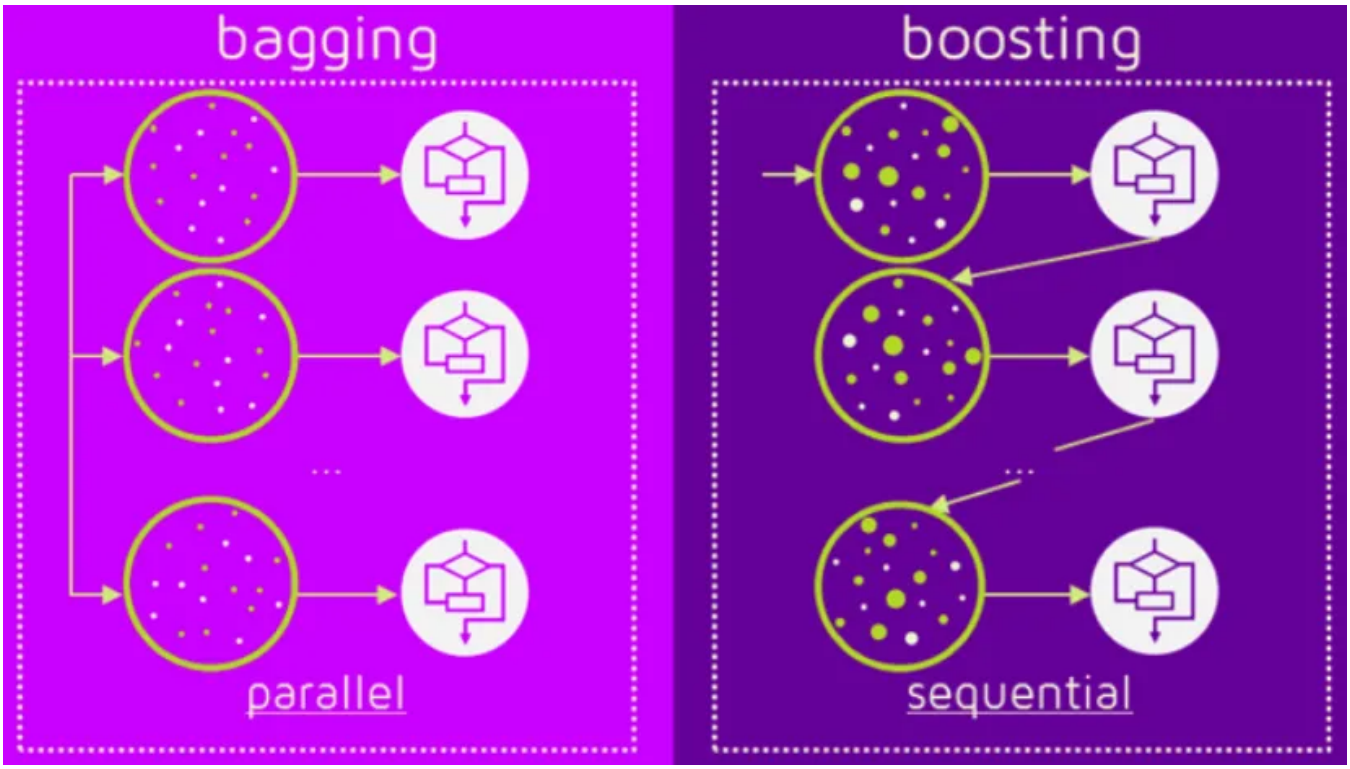
What is Ensemble Learning?

It is a powerful method to improve the performance of the model. It combines the predictions from multiple models. The results are improved by using this ML algorithm.



Types of Ensemble Learning

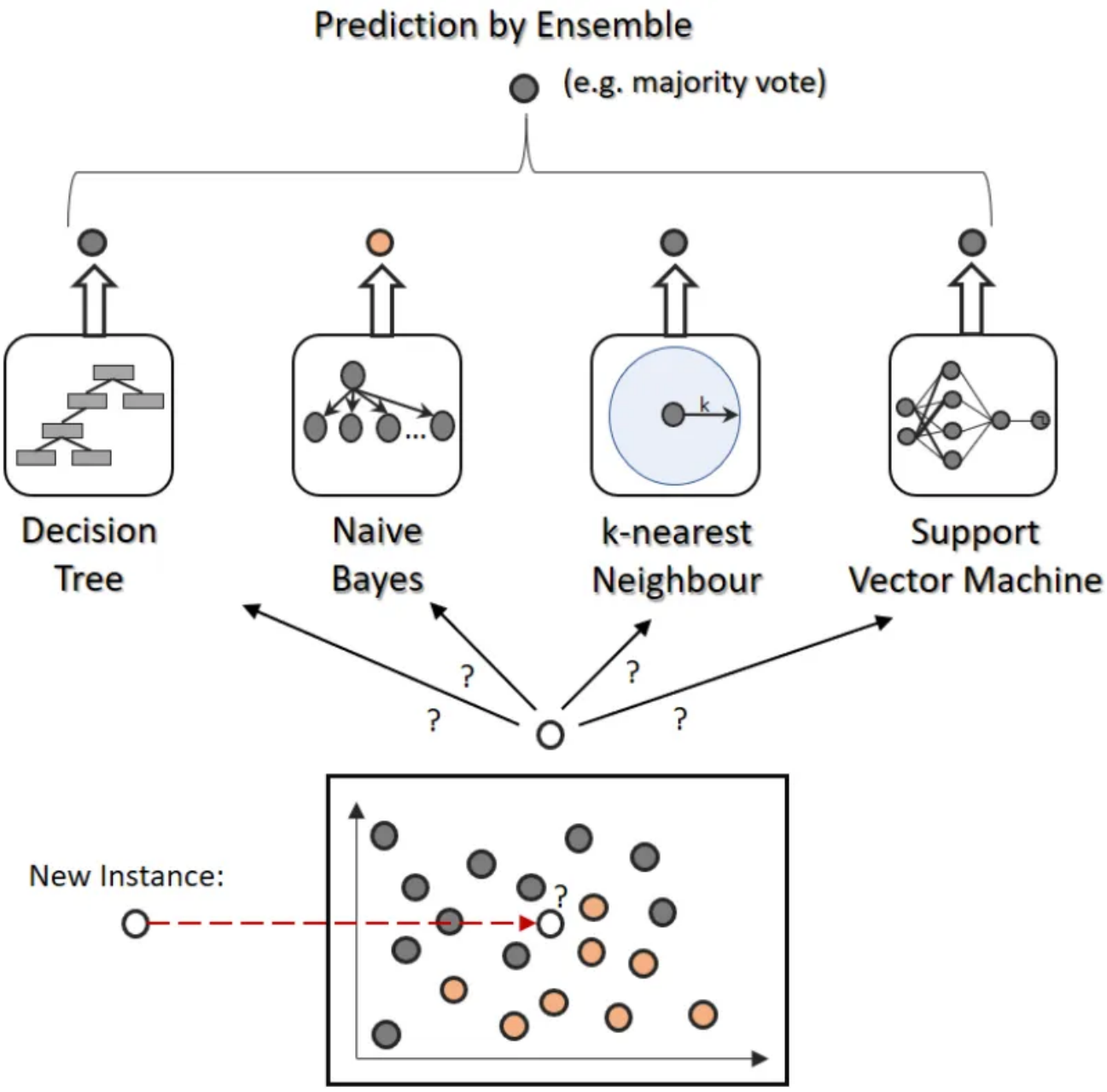
- Bagging or Bootstrap Aggregation — Random Forest
- Boosting — AdaBoost, XG Boost and Gradient Boost
- Stacking



Why do we use Ensemble Techniques?

These techniques help in reducing the variance (bagging), bias (boosting) and improve predictions.

- **Bagging**: Random Forest handles overfitting, training time is less, sometimes increases the bias, reduces variance, independent Parallel classifiers.
- **Boosting**: Gradient Boosting can overfit (avoided by tuning the parameters), reduce bias. They are called sequential classifiers.



How does boosting algorithms work?

Step 1: Reading the data

Step 2: Assigns equal weight to each sample observation

Step 3: False predictions are assigned to the next base learner with a higher weightage on these incorrect predictions.

Step 4: Repeat Step 2 until the algorithm can correctly classify the output.

AdaBoost and XG Boost are generic. XG Boost is better because it works for generic loss functions while Adaboost is for classification with exponential loss.

How Gradient Boosting Works? (Iterative Corrections)

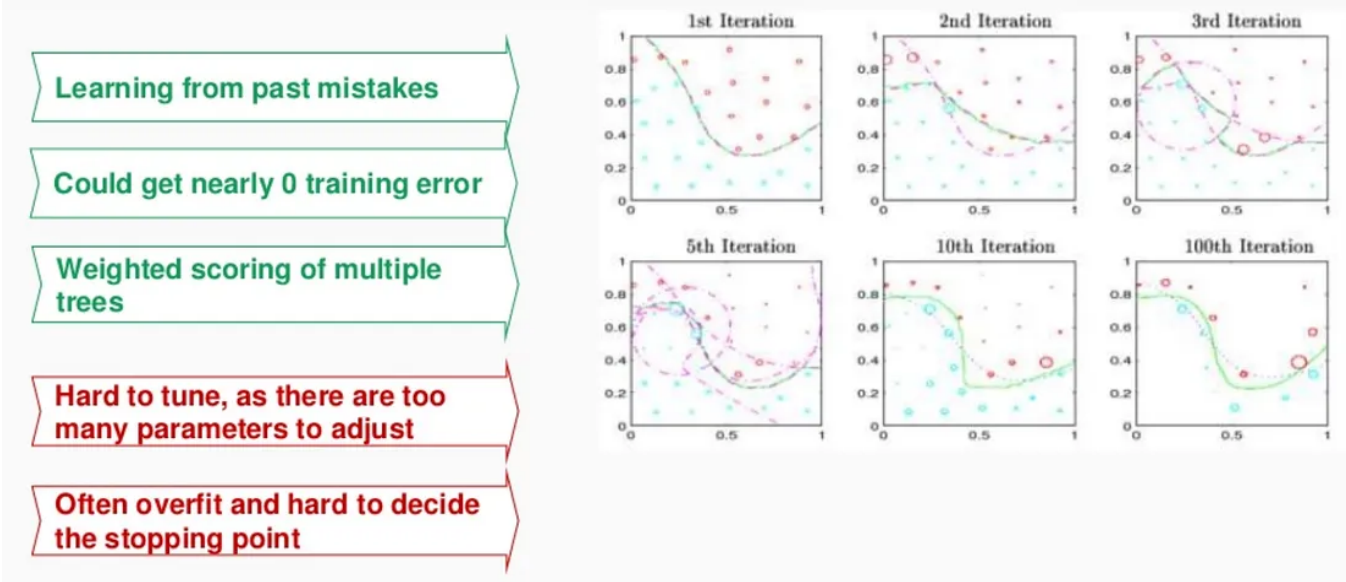
Step 1: Compute the average of the target column

Step 2: Calculate the residual based on actual and predicted observations

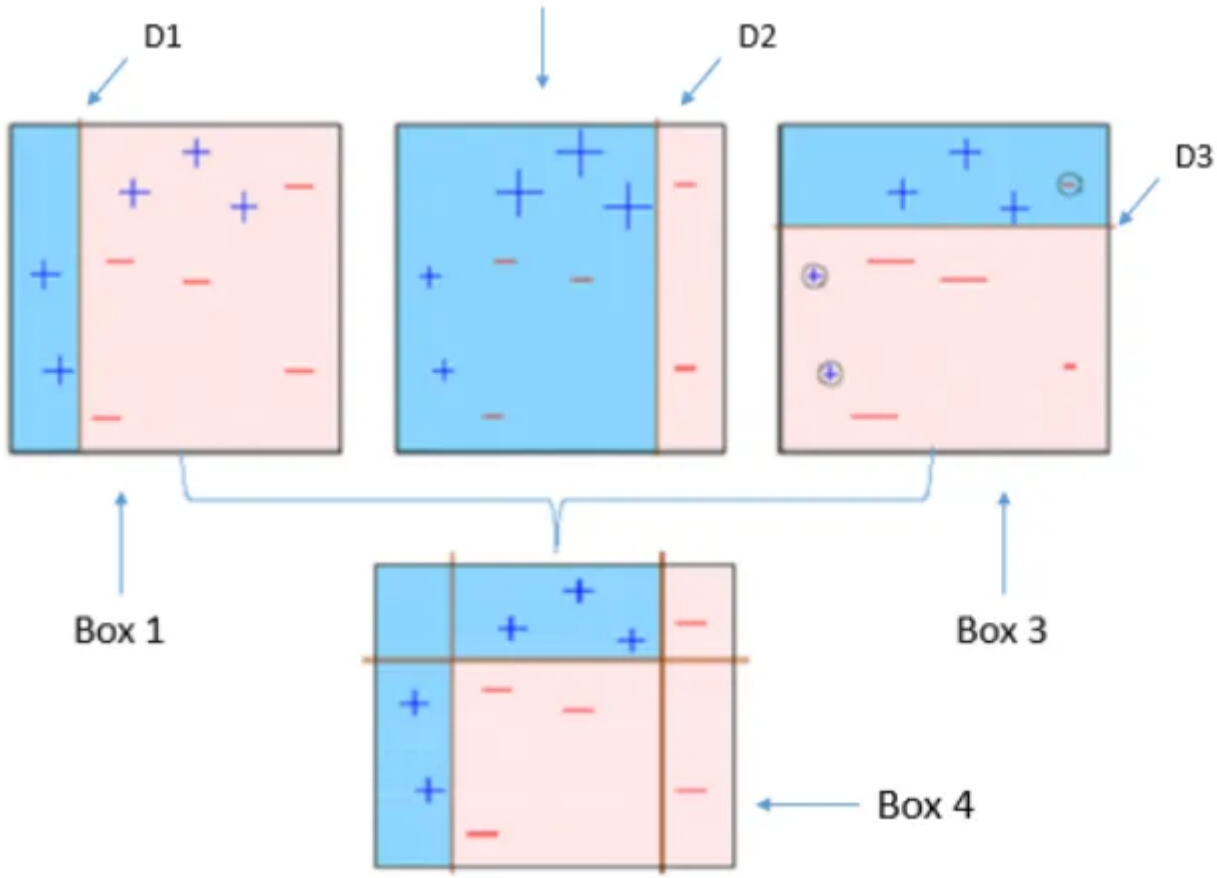
Step 3: Fit the model on residual as target and get the predicted residual

Step 4: Update the default prediction by the new residual

Step 5: Optimizing the Loss functions of previous learner



How does AdaBoost Work?



Step 1: Each observation is weighted equally for the first decision stump.

Step 2: Mis-class observation are assigned to higher weights

Step 3: A new decision stump is drawn by considering the observation with higher weights as more significant.

Step 4: Again if any observations are misclassified, they are given higher weights.

Step 5: The process continues until the observation falls in the right classes.

XG Boosting

Advanced version of gradient boosting that is designed to focus on computational speed and model efficiency.

Random Forest (Bootstrap Aggregation)

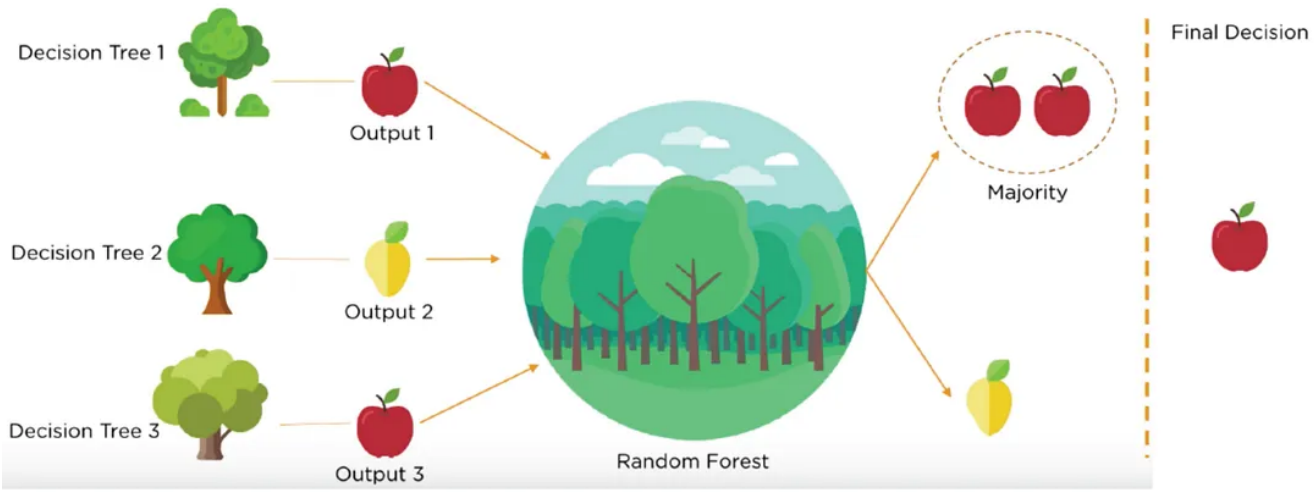
It is a bagging method with trees as weak learners. Each tree is fitted on a bootstrap sample considering only a subset of variables randomly chosen.

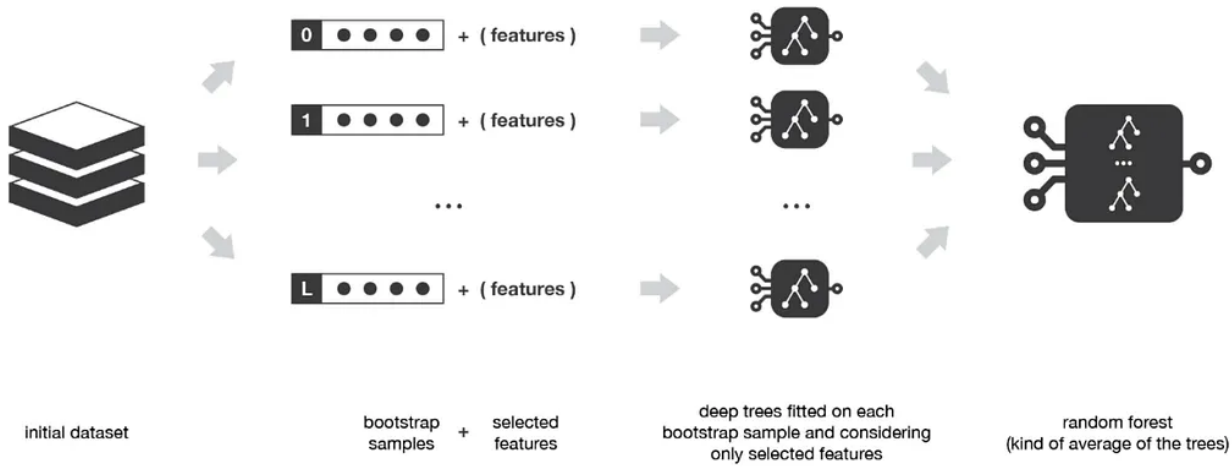
Note: Samples are drawn with replacement

It decreases the variance in the prediction by generating additional data for training from the dataset using combinations with repetitions to produce multi-sets of the original data.

How Random Forest Work (Majority Wins)

It is a method that operates by constructing multiple decision trees during the training phase. The decision of the majority of the trees is chosen by the random forest as the final decision.

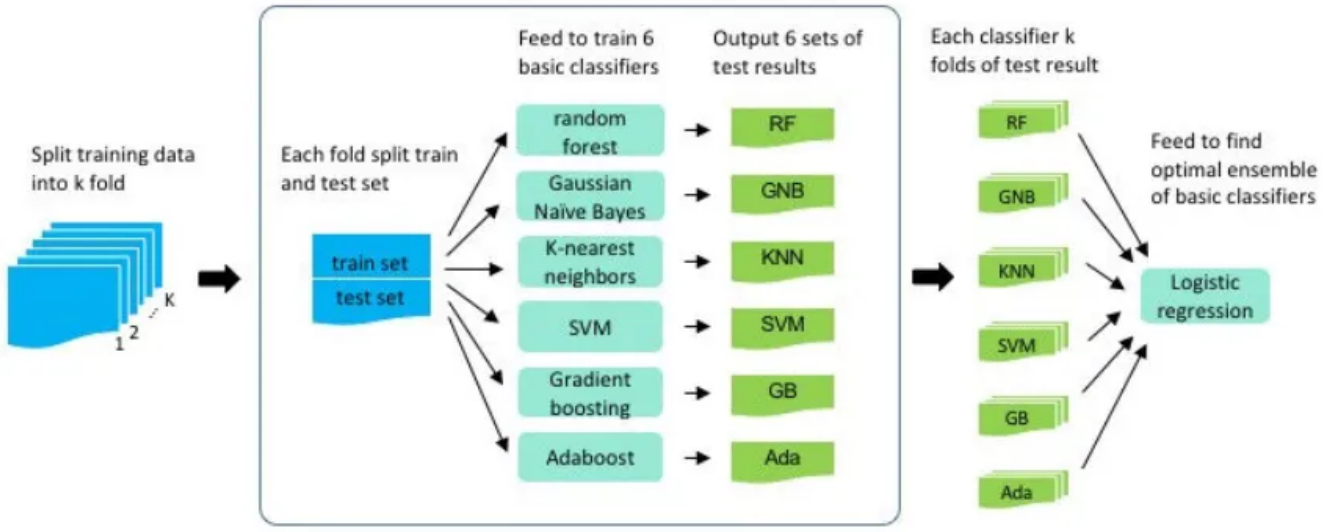




Conclusion:

- Random forests are an effective tool in prediction.
- Forests give results competitive with boosting.
- Random inputs and random features produce good results in classification and less in regression. For larger data sets, we can gain accuracy by combining random features with boosting.

Stacking: Combines the predictions from two or more ML algorithms.



Advantages of Stacking

- Unlike bagging, the models are typically different (e.g. not all decision trees) and fit on the same dataset.
- Unlike boosting, a single model is used to learn how to best combine the predictions from the contributing models.