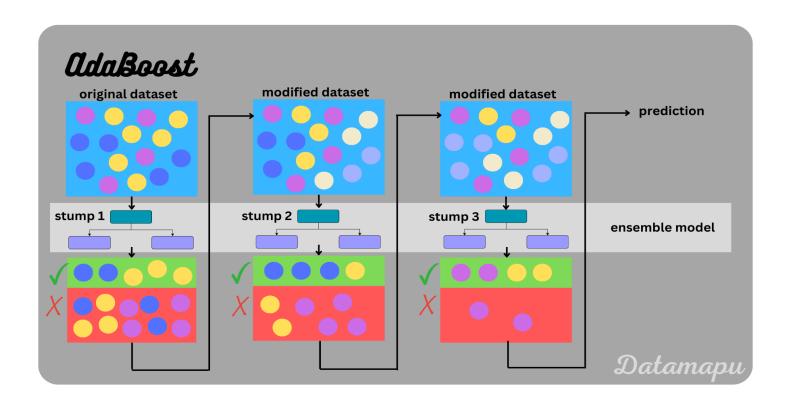
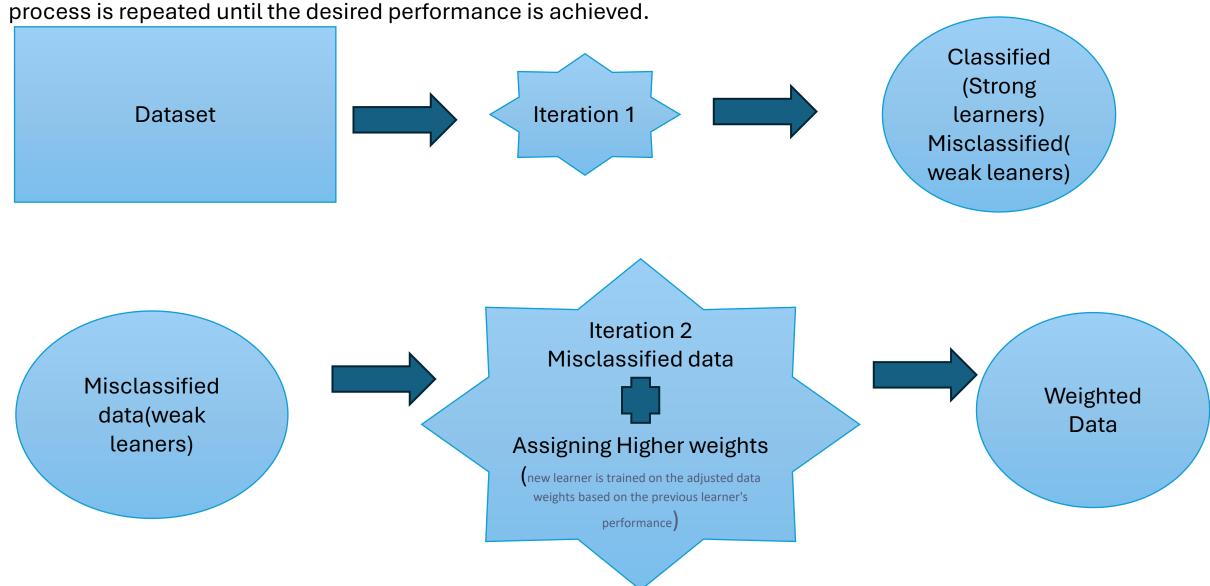
Ada Boosting Regression

Ada Boosting is a boosting algorithm. It will combines the prediction from the multiple weak learner, by using the stamp decision to form a strong learner. It is a meta-estimator that begins by fitting a regressor on the original dataset and then fits additional copies of the regressor on the same dataset but where the weights of instances are adjusted according to the error of the current prediction.



The weights of the samples are adjusted at each iteration. the data points that were previously misclassified, assigning higher weights to them in subsequent iterations, thus progressively improving the overall prediction accuracy. This process is repeated until the desired performance is achieved.



Ada Boost Working principle: Step-2 Step-3 Step-1 **Training** Train a weak Update the Calculate set(assign weights(Adjust the learner(using error rate on equal weights weights of the data points current data the training based on their classification to all the data results. Misclassified data weights) points receive higher weights, data. points) while correctly classified data points receive lower weights). Step-5 Continue steps 2-Step-6 4, training new weak learners and **Final Prediction** updating weights until a predefined number of

iterations is

reached.

Step-4

combine the predictions from all the weak learners, where each prediction is weighted based on the accuracy of the corresponding weak learner

Final Prediction:

$$\hat{y}(x) = \sum_{m=1}^M \alpha_m h_m(x)$$

where:

- $\hat{y}(x)$ is the final prediction for the input x,
- α_m is the weight for the m-th weak learner (which is based on the performance of $h_m(x)$),
- $h_m(x)$ is the prediction of the m-th weak learner.

Pros:

- 1. Improves Weak learners
- 2. Handles complex data.
- 3. Works well in small dataset
- 4. Ada boost regressor is used with decision tree, linear models etc.

Cons:

- 1. This will lead to overfitting.
- 2. Requires Tuning and proper choice of weak learners to perform well.
- 3. Performance depends on the weak learners.
- 4. Limited Performance on Highly Correlated Data and Instability in Some Cases.