Ada Boost Algorithm

1. What is adaptive boosting

Ans: Ada boosting / adaptive boosting is a machine learning algorithm. It considers a series of classifiers & combines the votes of each individual classifier.

2. Explain the Process of Ada Boost.

Ans: Ada boost algorithm follows 4 steps.

Step - 1: Weights are assigned to each training instance.

Step - 2: A series of k classifiers are iteratively learned.

Step - 3: "M_i" is called a classifier. After a "M_i" is learned, the weights are updated to allow the subsequent classifier. "M_{i+1}" classifier pays more attention to the instances that were miss-classified by "M_i".

Step - 4: The final boosted classifier, M^* combines the votes of each individual classifiers.

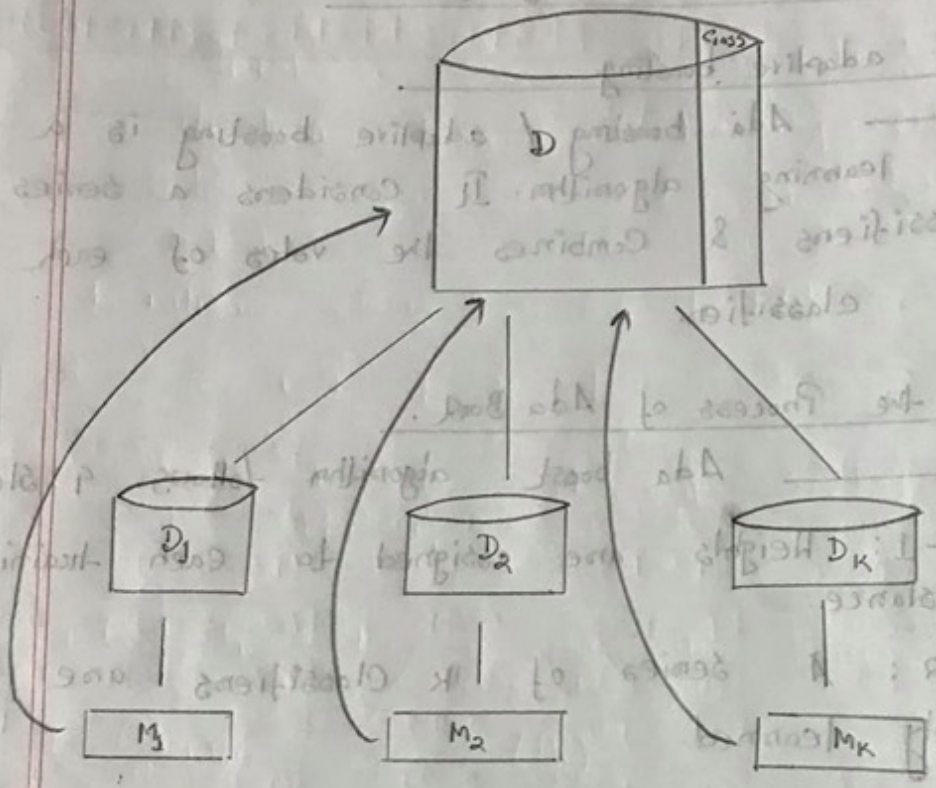

Fig: Ada Boost Algorithm

3. Define Learning Scheme

Ans:

Learning scheme means using a classifier to train a machine learning model.

4. Define Selection & Replacement technique

Ans

Selection & Replacement are Part of ada boost. Selection means selecting the instances with large weight. Replacement means updating the weights.

5. Define Ensemble learning

Ans:

Ensemble learning means using multiple learning algorithms to gain better Predictive Performance.

6. Write the difference between bagging & boosting

Ans

Bagging	Boosting
a. Individual models are built separately.	a. Each new model is influenced by the performance of those built previously.
b Equal weight is given to all models.	b. Weight is assigned considering the contribution of a model's Performance.

7. Write down the Ada Boost Algorithm.

Ans:

Input: Training Data D, number of iterations K & a learning scheme.

Output: Ensemble Model, M^*.

Method:

1. initialise weight $x_i \in D$ to $1/d$
2. for $i = 1$ to K do
3. Sample D with replacement according to instance weight to obtain D_i.
4. Use D_i, & learning scheme to derive a model, M_i
5. Compute error (M_i)
6. if error $(M_i) \geq 0.5$ then
7. go back to step 3 & try again
8. end if
9. for each Correctly classified $x_i \in D$ do
10. multiply weight of x_i by $\left(\frac{\text{error}(M_i)}{1 - \text{error}(M_i)}\right)$
11. end for
12. normalize weight of instances,
13. end for

To use M^* to classify a new instance, X_{new}.

1. initialize weight of each class to zero
2. for $i = 1$ to n do
3. $w_i = \log \frac{1 - \text{error}(M_i)}{\text{error}(M_i)}$ // This is the weight of the classifiers vote.
4. ~~Voted~~ $c = M_i(X_{new})$ // This is the class Prediction by M_i (model)
5. add w_i to weight for class c
6. end for
7. return class with largest weight.

---

8. What is Normalising Weight.

Ans: Updating the weight of the sample by the current weak classifier in a training at each stage

9. Why Normalising Weight is important.

Ans By normalising weights, the weights of misclassified instances are increased & the weights of correctly classified instances are decreased.

Error Rate → 50%. Mode
Acc

- Error rate যদি 50%. এর বেশি হয়, মানে M_i (M_i মানে classifier/Model) যদি 50%. এর বেশি instance কে miss-classify করে তাহলে ঐ M_i বাদ দিয়ে, আবার ০ থেকে নতুন করে M_i বানাতে হবে।
- Accuracy বাড়ালে learning scheme change করতে হবে।
- Correctly যখন একটা instance এর model classified হয়, তারপর
  - => যারা সঠিকভাবে classify হচ্ছে তাদের weight কমায় দিব।
  - => যারা সঠিকভাবে classify হয় নাই, তাদের weight বাড়ায় দিব,
- Ada boost প্রতিটা row / instance কে weight দেয়,
- আমার কাছে যদি N সংখ্যক instance থাকে, তাহলে প্রতিটাকে weight দিব 1/N or 1/D.
- Standard way হল একদম শুরুতে সবগুলো instance এর weight 1 দেওয়া,
- প্রথম বারে সবই weight manually দিয়ে দেওয়া হয়, পরবর্তীতে accuracy এর উপর base করে classifier/model কে weight দেওয়া হয়।

Normalisation of Weight = $\frac{\text{sum of old weights}}{\text{sum of new weights}}$

Correctly Classified Instance এর weight update.

= $\frac{\text{error}}{\text{accuracy}}$

Classifier vote, $w_i = \frac{\text{accuracy}}{\text{error}}$.

10. Write the advantages of Ada Boost

Ans: The advantages of Ada boost are

a. Very simple to implement

b. Feature selection on very large sets of features.

c. Ada boost adjusts ~~adapti~~ adaptively the errors of the weak ~~hypothese~~ hypotheses by weak learner.

d. Reduces variance.

e. Fairly good generalization.

Ada boost, noise & outliers এর প্রতি sensitive.

11. What is Outliers.

Ans: Extreme values that deviate from other observations on data.

ছোট Simulation =>

ধরি,

আমার কাছে 5 টা instances আছে

x_1, x_2, x_3, x_4, x_5

Ada boost এতে,

প্রথমে সবগুলো weight দিয়ে দেই।

Step: 1

x_1 , x_2 , x_3 , x_4 , x_5

$\downarrow$ $\downarrow$ $\downarrow$ $\downarrow$ $\downarrow$

$\omega_1 = 1$ $\omega_2 = 1$ $\omega_3 = 1$ $\omega_4 = 1$ $\omega_5 = 1$

ধরি,

Ada boost একটা learning scheme use করে, M_1 নামক classifier/model তৈরী করেছে,

যেখানে,

Classified হয়েছে => x_1, x_3, x_5

Misclassified হয়েছে => x_2, x_4

তাহলে, M_1 এর Accuracy = $3/5 = 0.6$

তাহলে M_1 টিকে যাবে, এখন weight Updat করবো, Update এর সময় আগের weight এর সাথে সাথে যুক্ত করবো।

Error / accuracy ? যেগুলো সঠিকভাবে classified

উদাহরণ

x_1 , x_2 , x_3 , x_4 , x_5

$\omega_1 = 1$ $\omega_2 = 1$ $\omega_3 = 1$ $\omega_4 = 1$ $\omega_5 = 1$

New weight

$\omega_1 = 1 \times \frac{215}{315} = 0.67$ $\omega_2 = 1 = 1$ $\omega_3 = 1 \times \frac{215}{315} = 0.67$ $\omega_4 = 1 = 1$ $\omega_5 = 1 \times \frac{215}{315} = 0.67$

Weight Normalization $= \frac{\text{Sum of old weights}}{\text{Sum of new weights}}$

$= \frac{1+1+1+1+1}{0.67+1+0.67+1+0.67}$

$= \frac{5}{4.01}$

$= 1.24$

Step: 2 তে
এসব weight হবে

$\omega_1 = 0.67 \times 1.24 = 0.83$

$\omega_2 = 1 \times 1.24 = 1.24$

$\omega_3 = 0.67 \times 1.24 = 0.83$

$\omega_4 = 1 \times 1.24 = 1.24$

$\omega_5 = 0.67 \times 1.24 = 0.83$

যেগুলোর Correctly Classified সেগুলোর weight একই হবে,

কতবার iteration চালাবো, সেটা আমরা fix করে দিব,

এক্ষেত্রে,

M_1, M_2, M_n সংশ্লিষ্ক classifier থেকে decision নেবার সময় weighted majority voting নেওয়া হবে,

ধরি,

M_1 এর → Accuracy = 70%.
→ Error = 30%.

M_2 এর → Accuracy = 50%.
→ Error = 50%.

M_3 এর → Accuracy = 60%.
→ Error = 40%.

তাহলে,

M_1 এর $w_1 = \log \frac{0.7}{0.3} = 0.36$ // $\log \frac{\text{Accuracy}}{\text{Error}}$

M_2 " $w_2 = \log \frac{0.5}{0.5} = 0$

M_3 " $w_3 = \log \frac{60}{40} = 0.17$

ধরি, M_1 এর label হল yes
M_2 " " হল no
M_3 " " " no

তাহলে একই label no হল M_2 এবং M_3

এই দুইটা যোগ করলে

$$M_2\omega_2 + M_3\omega_3 = 0 + 0.17 = 0.17$$

এখানে,

$$M_1\omega_1 > (M_2\omega_2 + M_3\omega_3)$$

অর্থাৎ, Decision আমরা yes. এবং Ada boost larget weight টা return করবে।