Nihal Patil Follow

Aug 16 · 2 min read · Listen

Adaboost Cheat Sheet

The adaptive boosting technique became formulated by way of yoav freund and robert schapire, who gained the gödel prize for their work. Adaboost works on improving the areas where the bottom learner fails. The bottom learner is a gadget getting to know algorithm that is a vulnerable learner and upon which the boosting approach is applied to show it right into a sturdy learner. Any gadget getting to know set of rules that be given weights on schooling facts may be used as a base learner. In the instance taken under, choice stumps are used as the base learner.

- When nothing works Boosting does.
- Supervised ML algo.
- Ensemble sequential approach.
- Adaptive Boosting.
- One of most amazing idea in ML.
- Why we call it booting => we boost the weak learner i.e. decision stump.
- What is Weak Learner? => It performs better than random model and it comprises of 2 leaf and 1 decision node.
- Every decision stump takes an attribute. Decision stumps called as shallow trees.
- Decision stump accuracy is near by 0.6
- Due to sequential approach nth model will be dependent of (n-1)th model which is made by previous stump mistake taking in the account.
- All weak learners contribute in order to make a strong learner.
- Sample wt. = 1/n

- Total Error(TE) = (Num Misclassified Data pts)/(Total Num Data pts)
- Model Pf(Performance) (MP)= 1/2 (log (1-TE)/TE)
- NSW[Misclassified] = old wt. * e^(MP)
- NSW[Classified] = old wt. * e^(-MP)
- We change distribution of data on each iteration.
- Due to bucket system our model chooses more misclassified values rather than correctly classified values.

Advantages and Disadvantages :

Advantages:

1. Time complexity is less.
2. Provides better accuracy than many other ML models.
3. Less likely to be overfitted.
4. Classification and regression.
5. No scaling is required.
6. The accuracy of weak classifiers can be improved by using Adaboost.
7. Non Parametric.
8. Easy to understand and implement.

Disadvantages :

1. The main disadvantage of Adaboost is that it needs a quality dataset.
2. Slower than XG-Boost.
3. Outliers and noisy data should be avoided.

Photo by Bill Jelen on Unsplash

Towards Data Science | Artificial Intelligence | Data Science | Boosting | Machine Intelligence