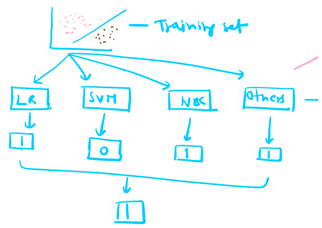
Ensemble Learning

Friday, September 24, 2021 7:01 AM

Decision Tree - Different subset of training set

Bagging, Boosting,



— Voting classifiers

— Diverse predictions

Bagging & Pasting ⇒ ① Use multiple / different algorithms
② Use same algorithm on random subsets of the training data — bagging / Pasting

Random patches & Subspaces ⇒ max features & bootstrap_features

Random Forest ⇒ max-features = "auto" -

Extra Trees ⇒ not best val → random / 0 threshold

Gradient Boosting ⇒

Ensemble ⇒ DT1 + DT2 + DT3 + DT4 + __ DT100

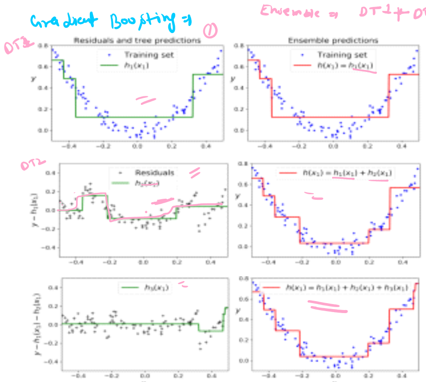


Figure 7-9. In this depiction of Gradient Boosting, the first predictor (top left) is trained normally, then each consecutive predictor (middle left and lower left) is trained on the previous predictor's residuals; the right column shows the resulting ensemble's predictions

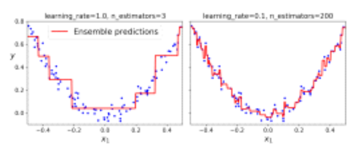


Figure 7-10. GBRT ensembles with not enough predictors (left) and too many (right)

X-G Boost | Light GBM | Cat Boost Algorithm

$\frac{1}{m}$ ⇒ C 64% - 36%

m - size of the data — Training set

64% - Within the bag
36% - OOB

$n \le m$ = LR on n sample
$c \le C$ = No. of columns

Bagging / Boosting

$\left(1 - \frac{1}{m}\right)^n$

≈ 36%

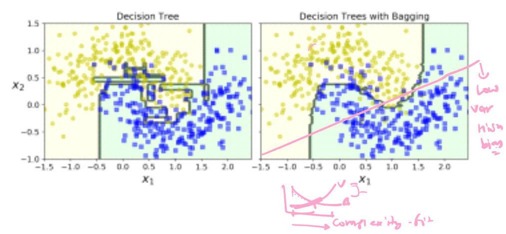
m = 100
⇒ n_1 = 64%
⇒ n_2 = 36%
100 = m_1

10 training set (64 / 36)
DT - 100 - 36

Bagging - 10 DT
① DT - 100 - 64 - Within / 36 - OOB
② DT - 100 - 64 - With / 36 OOB
③
⑩ DT - 100 - 64% / 36%

50 / 100
50 / 100
Max voting

Random Samples ⇒
DT
100 - Dimension
n-estimator = 100
DT
parallel
x_1
gini
>5 <5
⑤



`bootstrap=False` and `max_samples=1.0`

`bootstrap_features` to `True` and/or `max_features`