

Ensemble Learning

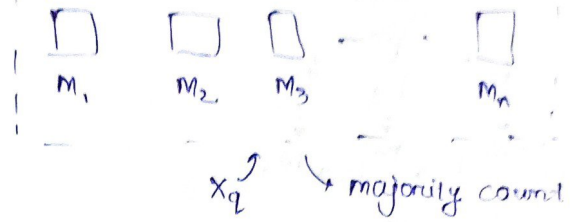
7:34 PM
mon, may 5

Wisdom of the crowd

70% → Y
30% → N

1. Base model must different
either data different w/ same model
Same data diff model

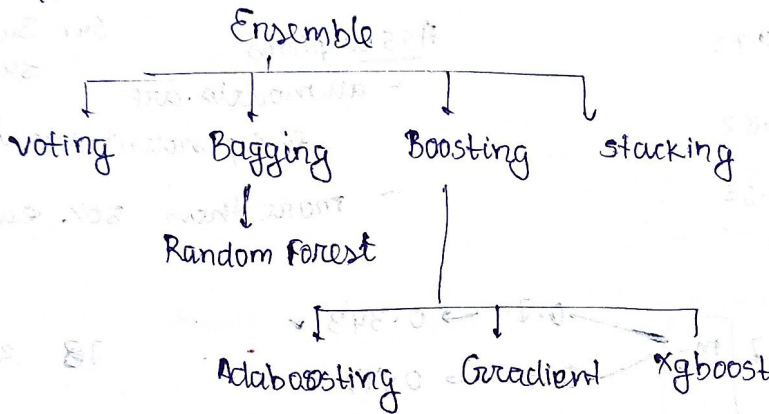
Classification



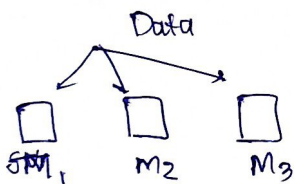
Regression

$$\frac{81 + 32 + 41 + 75}{4} = \text{find the mean}$$

Types (Imp)



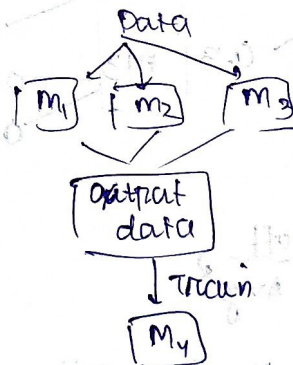
voting



Classification
- majority count

Regression
- mean

stacking

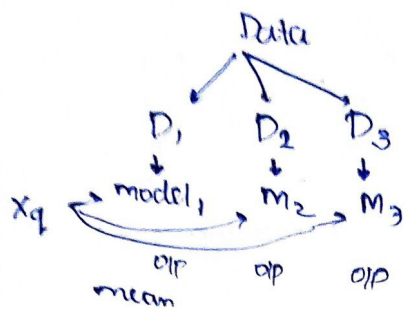


m_1	m_2	m_3	Result

- Now model M_4 understands which Result is most of time correct
- Acc to that M_4 provides weightage.

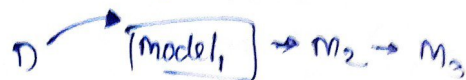
Bagging

- Stand for Bootstrapped Aggregation
- same model different data



Boosting

- It learns from previous wrong prediction

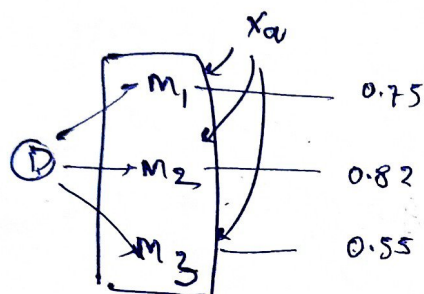


Benefits

- 1- improve performance
- 2- Bias variance
- 3- Robust

when to use
Always ←

Voting Ensemble

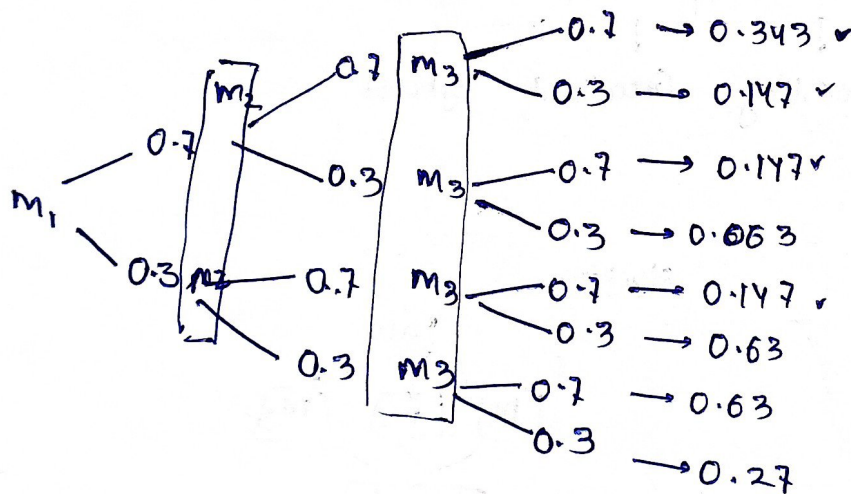


P-1

Sun Jun 15
8:44 am

Assumptions

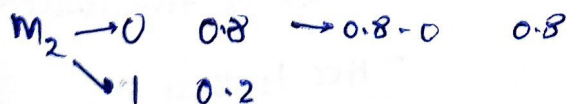
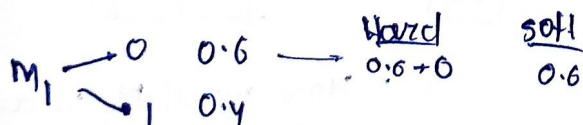
- all models are ^{system} independent are in nature
- more than 50% each model



78 2 in collect
7 more

Hard
Direct ams

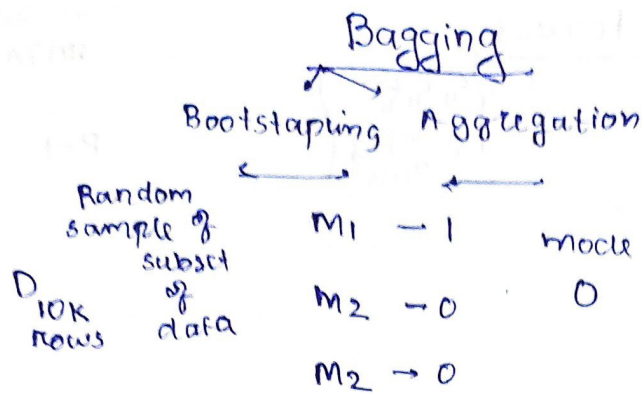
Soft Avg



$$\frac{0.6+0.8}{2} = 0.7 \rightarrow 0$$

$P-2/P-3$
 classifier
 Regumina

diff model
diff para.



P-1
Sun Jan 15
9:29 am

- overfitting
- low bias low variance

Types

Pasting - Row sampling w/o replace

Random Subspaces - columns sampling (w/, w/o) replace

Random Patches - Both Row, col Sampling

Bagging - Row sampling w/ replace

Bagging Classifiers

Out of Bag Sample

- * During sample time 63% point model seen 37% are unseen.

Tip

Bagging > Pasting

- row sample 0.5 - 5.
- high dimensional data need only

P-2

Bagging Regressor

P-3