

OOB-Score

out-of-bag Score

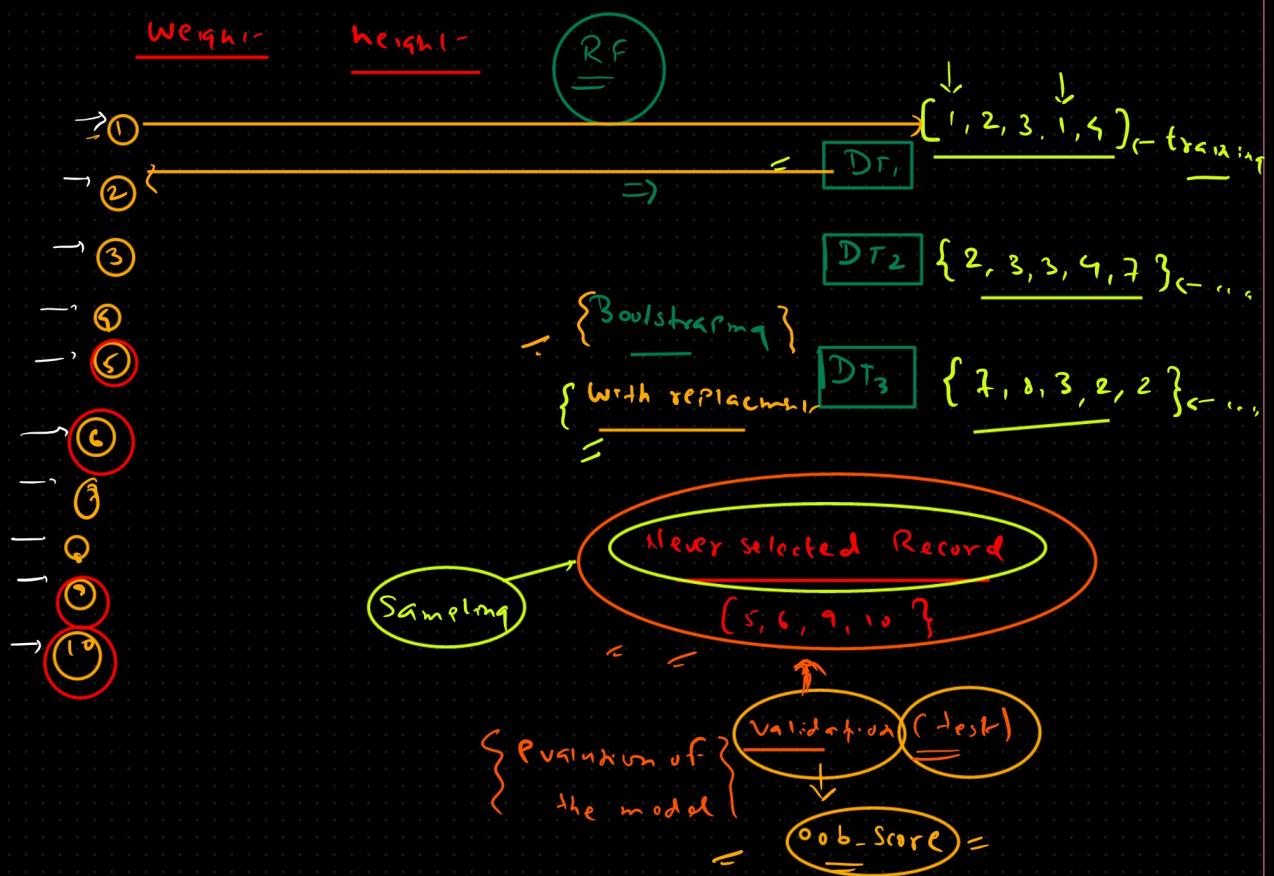
RF \Rightarrow ① Bootstrap ② Aggregation



Sampling

① With replacement -

② Without Rel.



Bagging and Random forest

Bagging

① Bootstrap
Aggregation

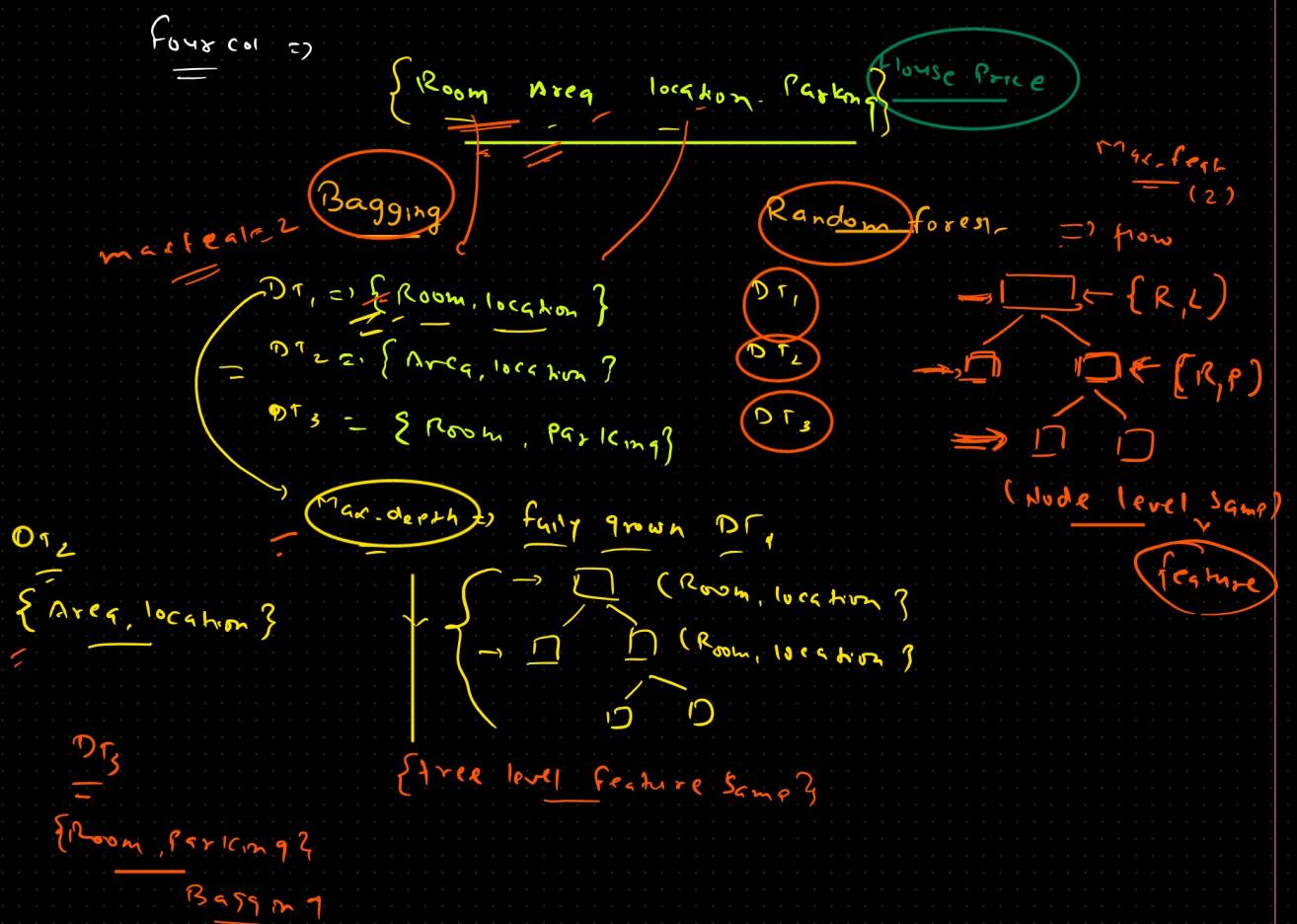
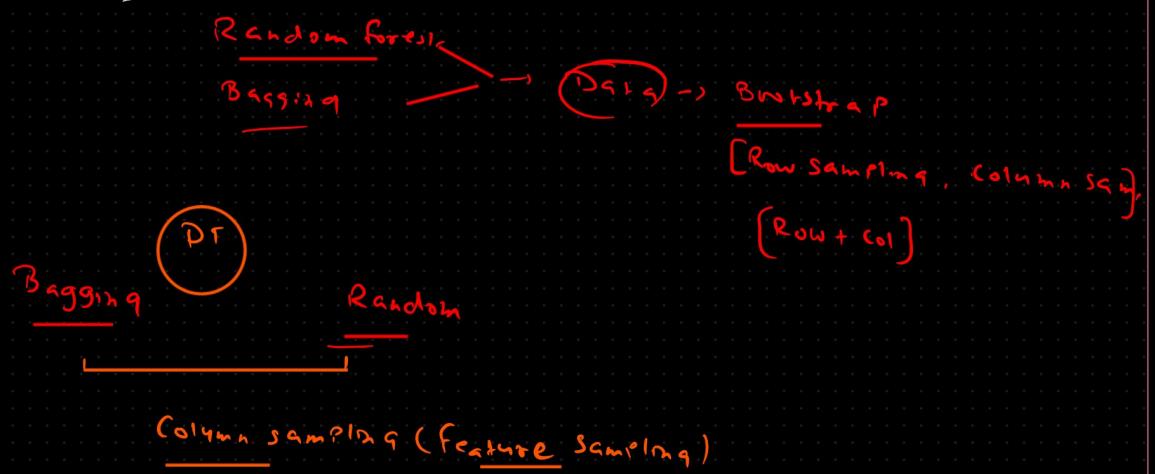
{ ② any sort of a
Estimator [lenet, LR, log, svm, ...]}

Random forest

① Bootstrap Agg.

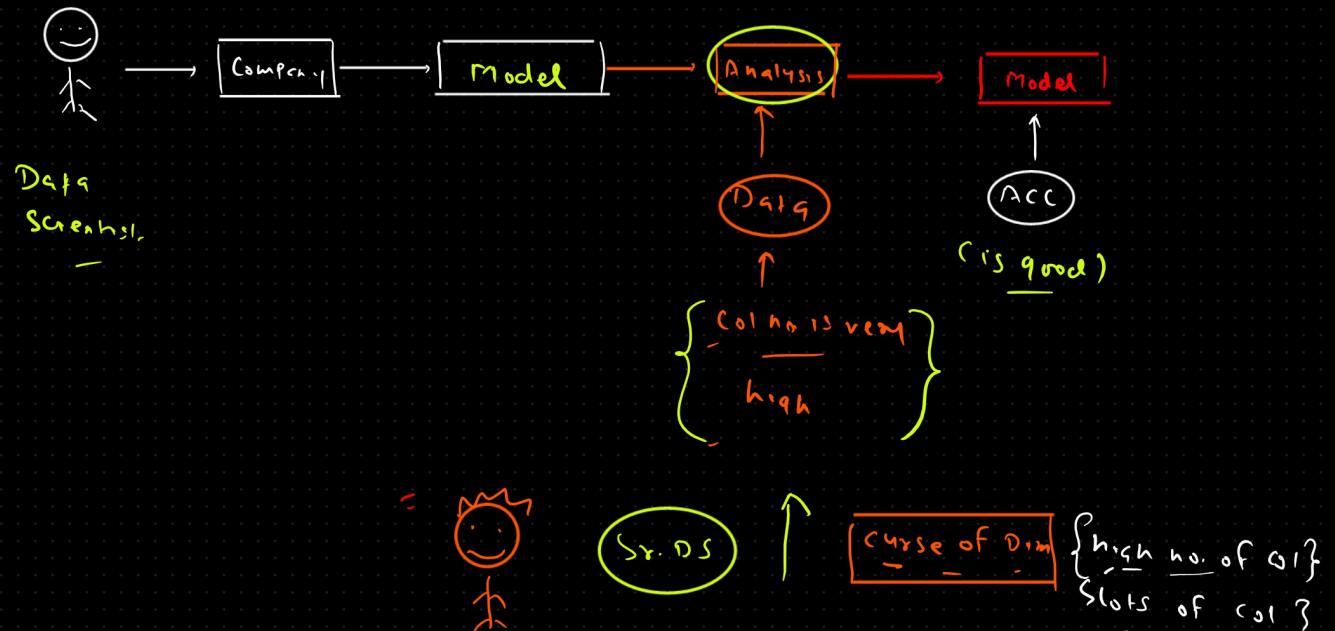
{ ② Special case of the
bagging \rightarrow OR }

Question \Rightarrow Bagging technique \Rightarrow $DT \leftarrow$ estimator \Rightarrow Random forest

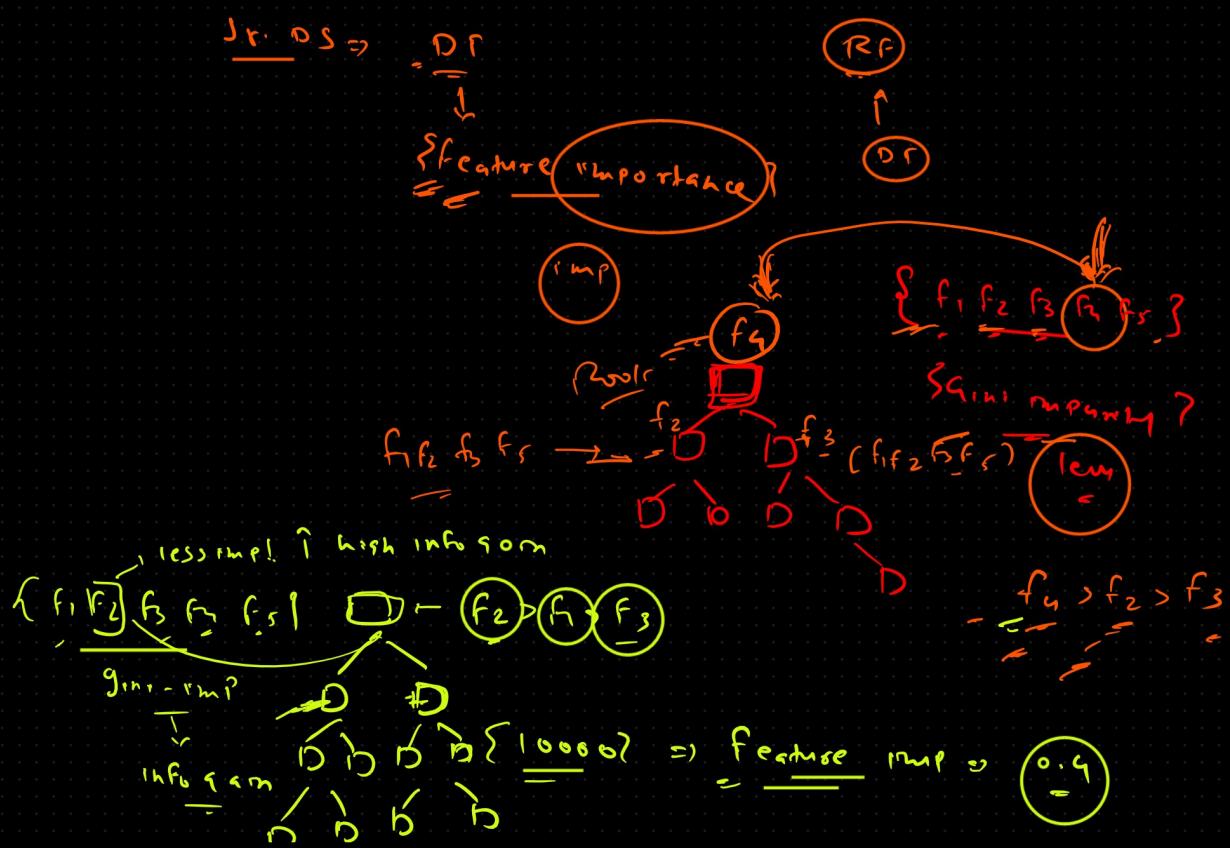
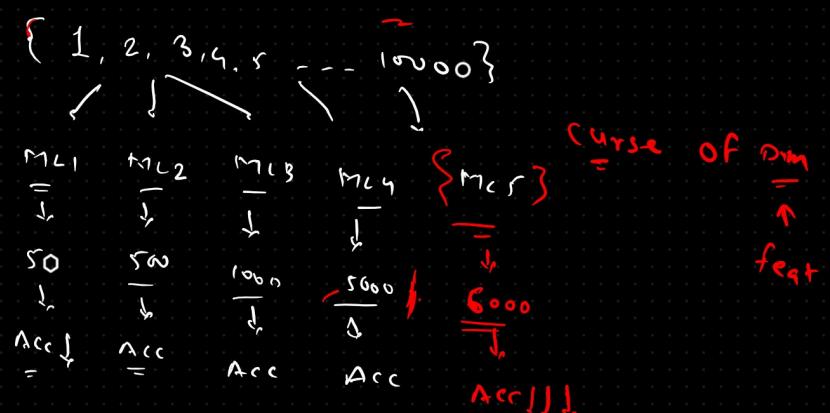


Feature importance \Rightarrow feature selection

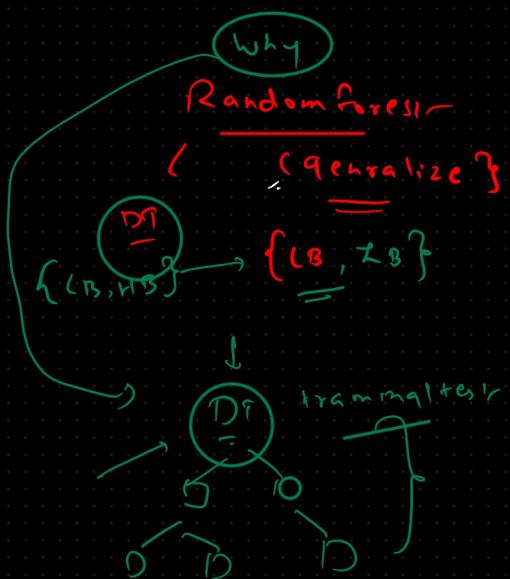
\downarrow
columns



Lots of data (lots of feature) \Rightarrow it is not true ML Learning case



Selected important feature

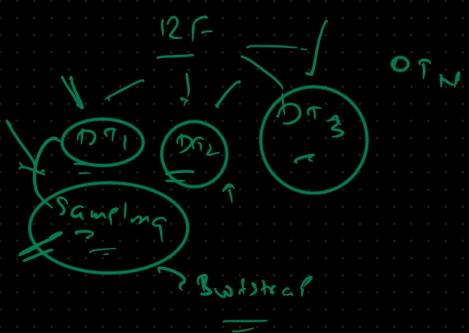


Mcprob

overfitting underfitting

$\{LB, HB\}$ $\{LB, HV\}$

\rightarrow fully grown DT \rightarrow log reg
 \rightarrow Svm \rightarrow Lin reg
 \rightarrow kNN



Stacking

① Bagging

Homogenous

Parallel

Bootstrap + Agg

② Boosting

Homogenous

sequential

③ Voting

Heterogeneous

Parallel

④ Stacking

Heterogeneous tech
entention of voting

2 - stages

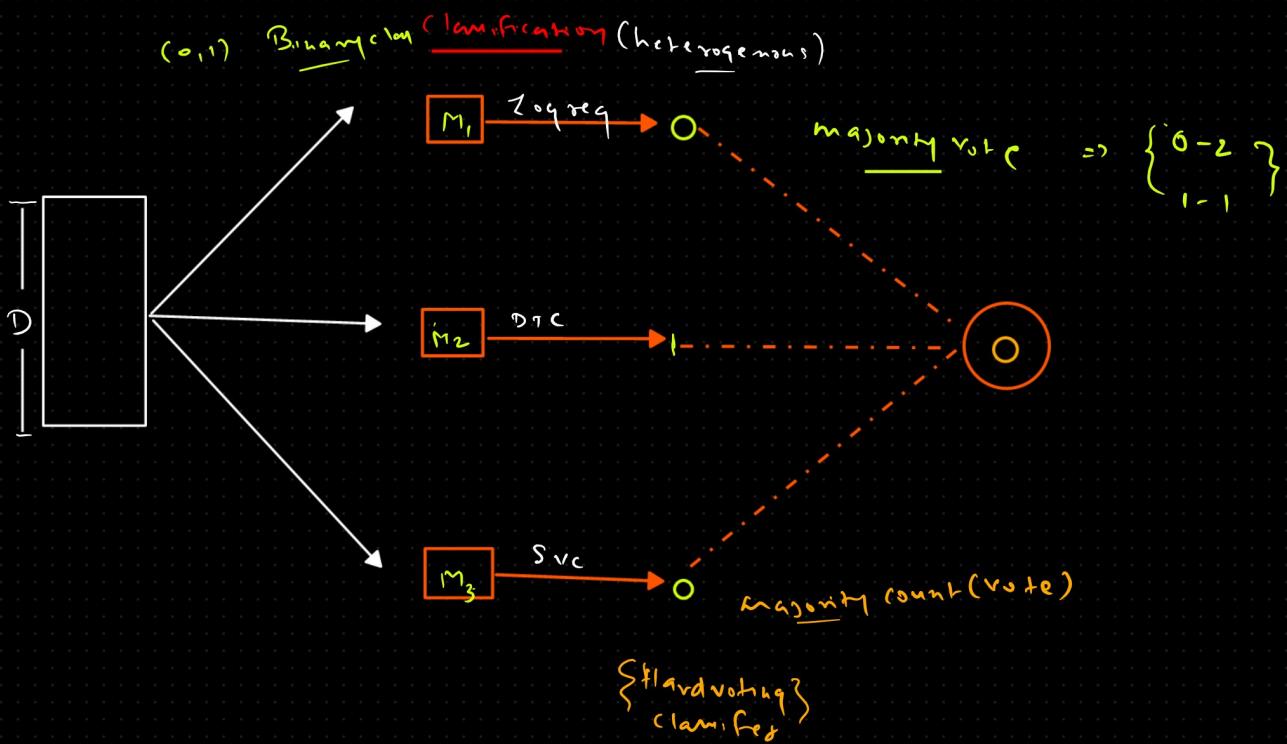
① Base model

② Meta model

Ensemble technique = { technique of arranging the Models }

Complex - Arch.

=
Composition



Soft classif. (Probability)

$$M_1 \begin{cases} 0 \rightarrow 0.6 \\ 1 \rightarrow 0.4 \end{cases}$$

$$M_2 \begin{cases} 0 \rightarrow 0.7 \\ 1 \rightarrow 0.3 \end{cases}$$

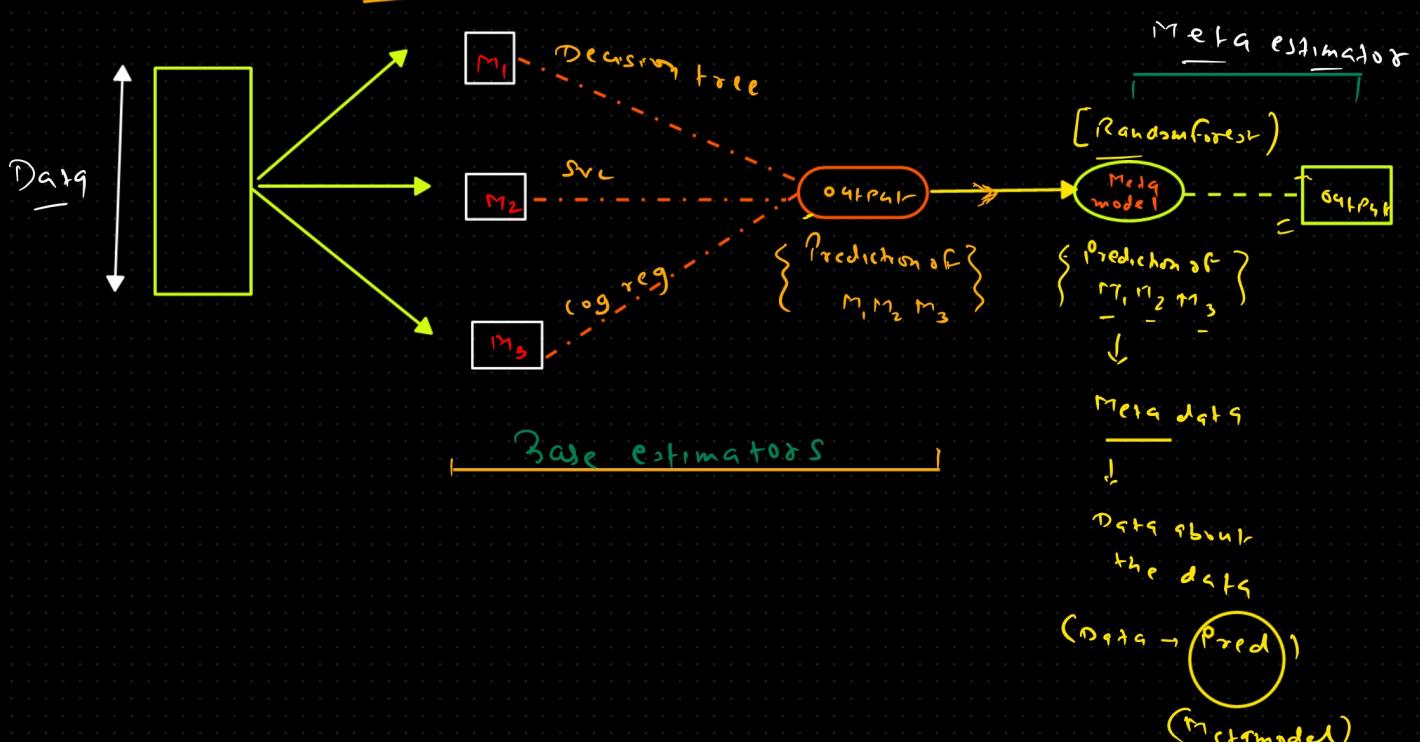
$$M_3 \begin{cases} 0 \rightarrow 0.3 \\ 1 \rightarrow 0.7 \end{cases}$$

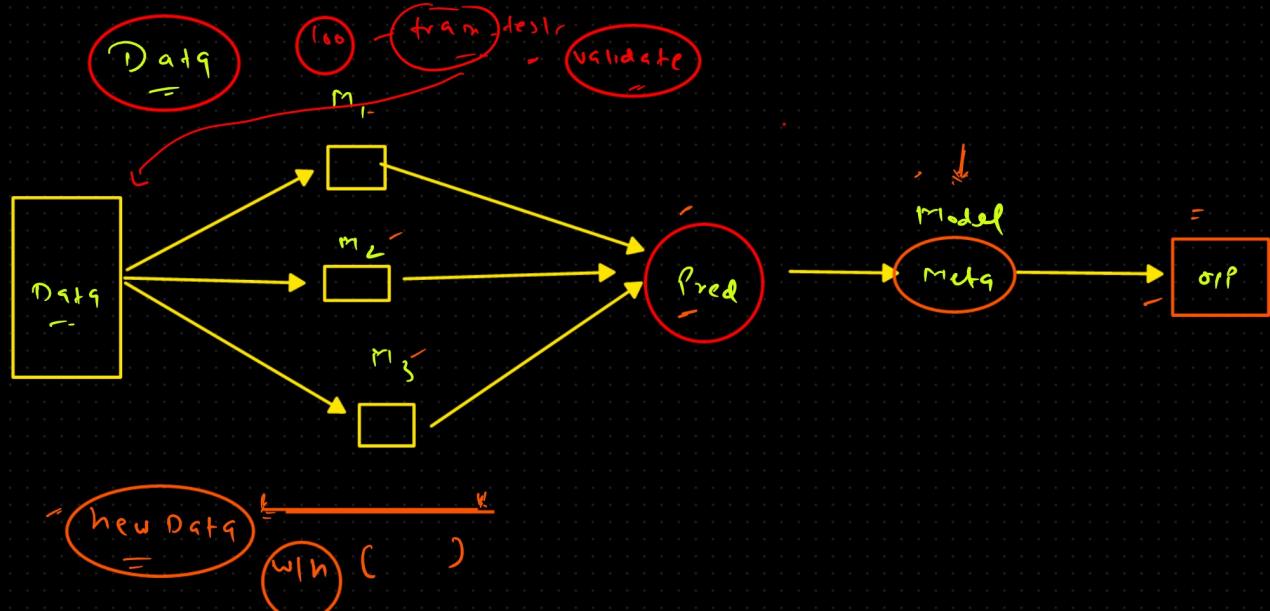
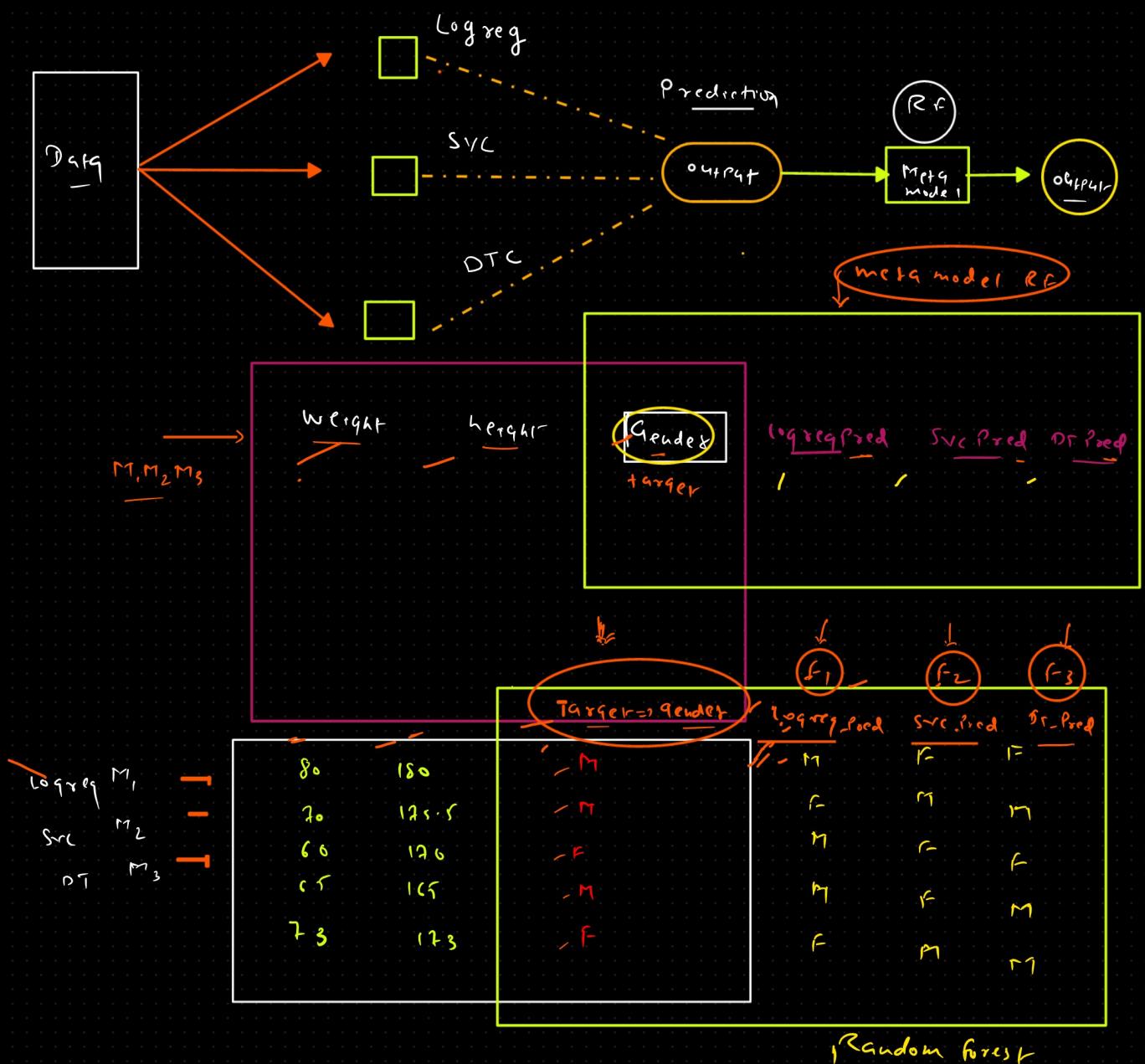
Stacking

① Base level

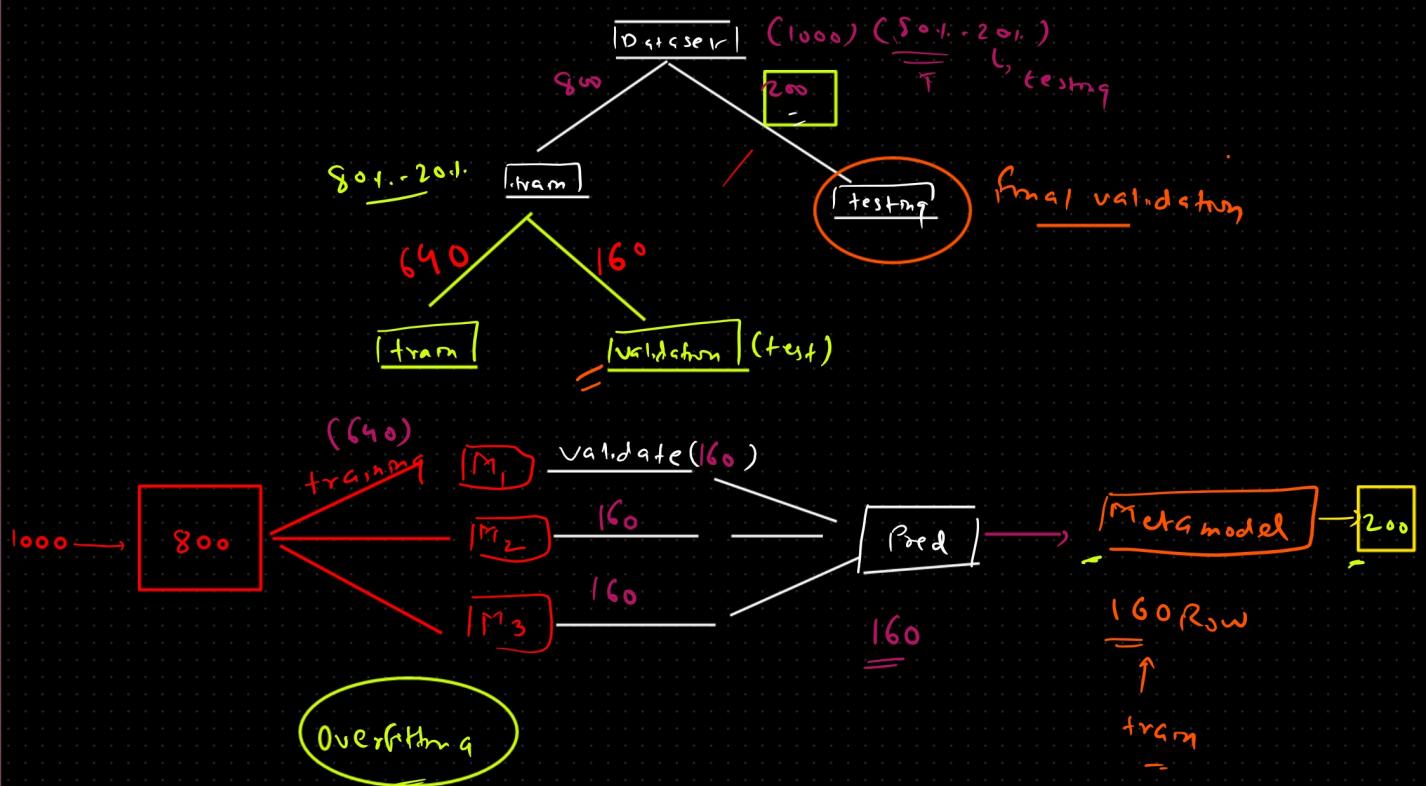
② Meta level \leftarrow on meta data \leftarrow data about the data

heterogeneous (M_1, M_2, M_3 , Diff.)





Blending \Rightarrow Data Division



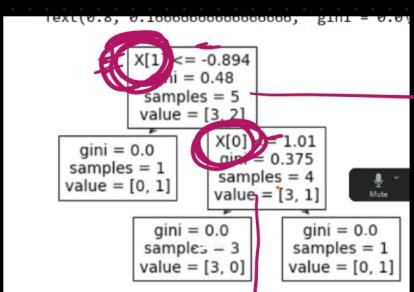
Infogain \Rightarrow Gini X

$$\text{Infogain} \Rightarrow \text{Impurity of parent node} = \left[\frac{\text{sample} \times \text{right imp}}{TNS} + \frac{\text{sample} \times \text{left imp}}{TNS} \right]$$

$$0.98 - \left[\frac{9}{5} \times 0.375 + 0 \right]$$

$$0.98 - \left[\frac{9 \times 0.375}{5} \right] \Rightarrow 0.98 - 0.3 \Rightarrow 0.18$$

Infogain \Rightarrow



0.18

$$\frac{9}{5} \left(0.375 - [0] \right)$$

$$1.8 \times 0.375 = 0.8 \times 0.75 = 0.30$$

0.30

$$\text{Normalized} = \frac{0.3}{0.18 + 0.30} \approx \frac{0.3}{0.48}$$

(2nd Feature)
= 0.625

$$= \frac{0.18}{0.48} = 0.375$$

- x(0)

1st