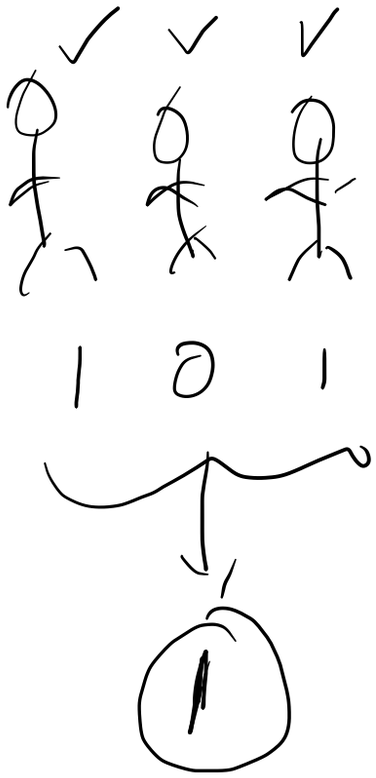
Strong

leader

↑

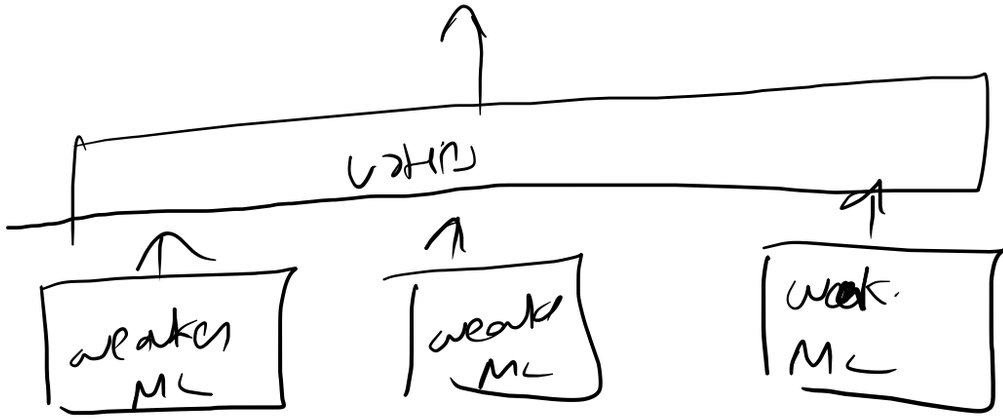
differ



→ 100%

Higher accuracy

→

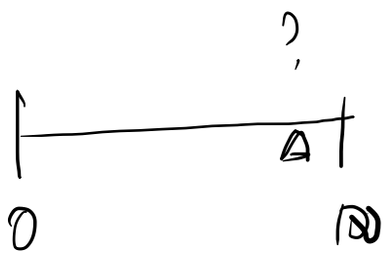


ensembling
- Bagging
- Boosting
- Stacking

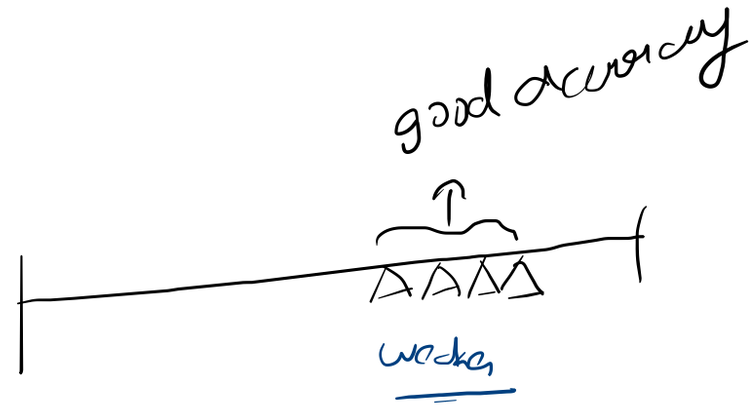
① confusion matrix

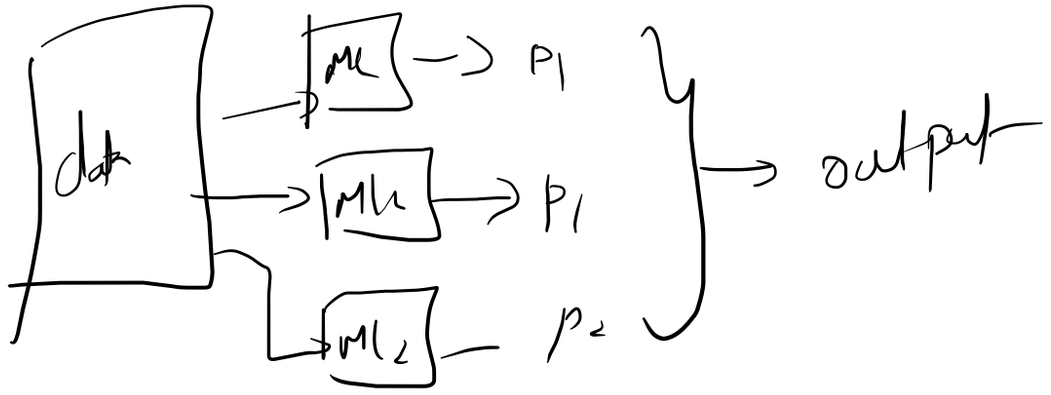
ⓐ Rmse ← [residue]

↳ construct multiple weak learners



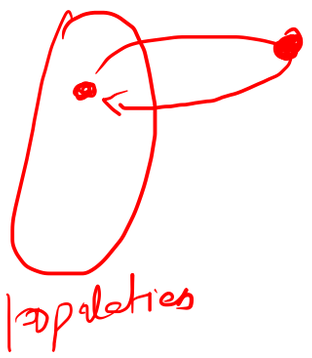

very different

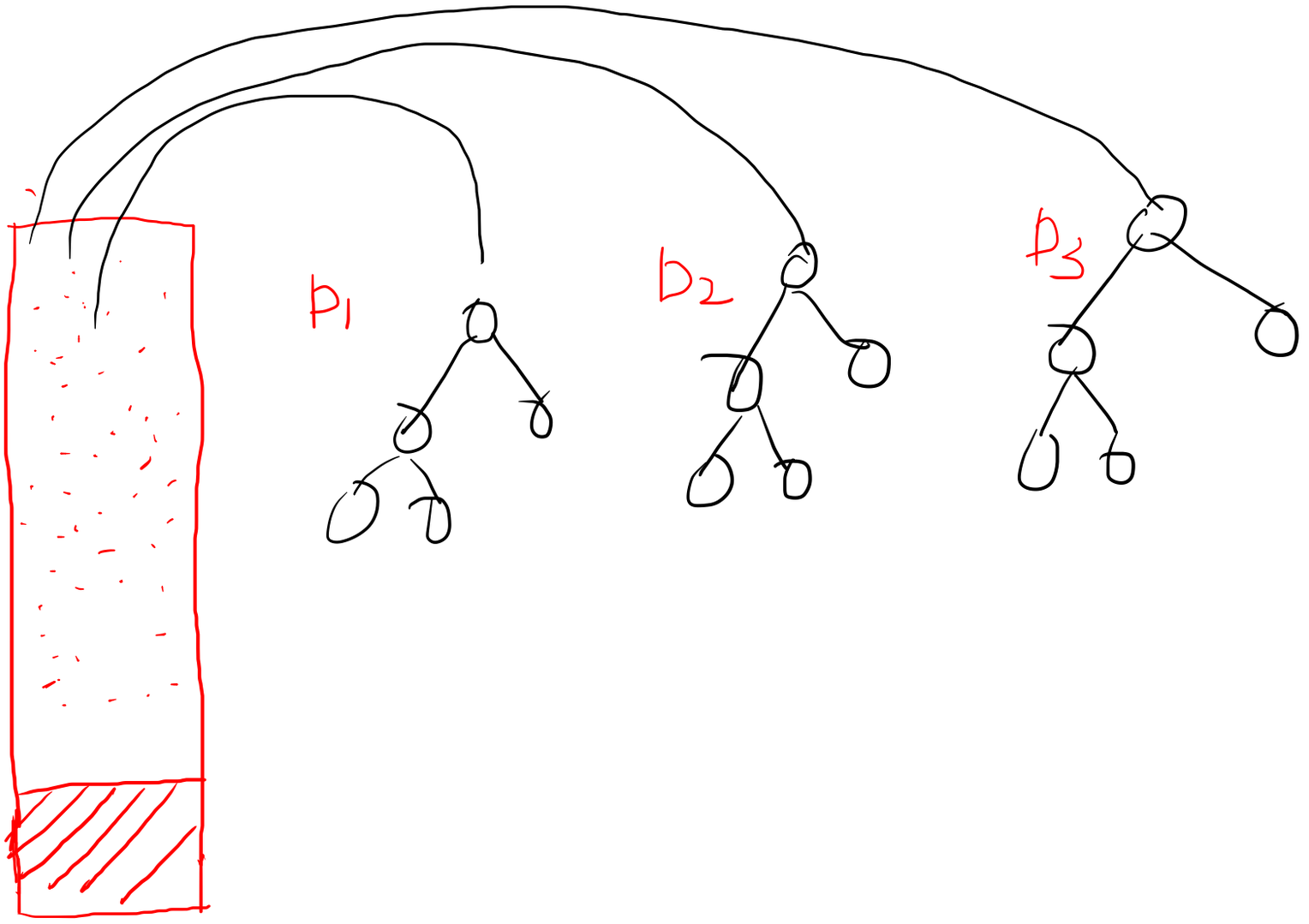


ensemble → Bagging [Random forest]

Bootstrapping




Train data →

100R SOC

Test data

$P_1 = P_2 = P_3$

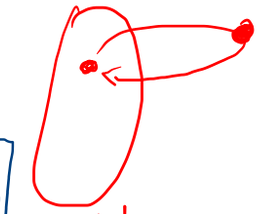
ensembled → Bagging [Random forest]

n. estimator

Bootstrapping

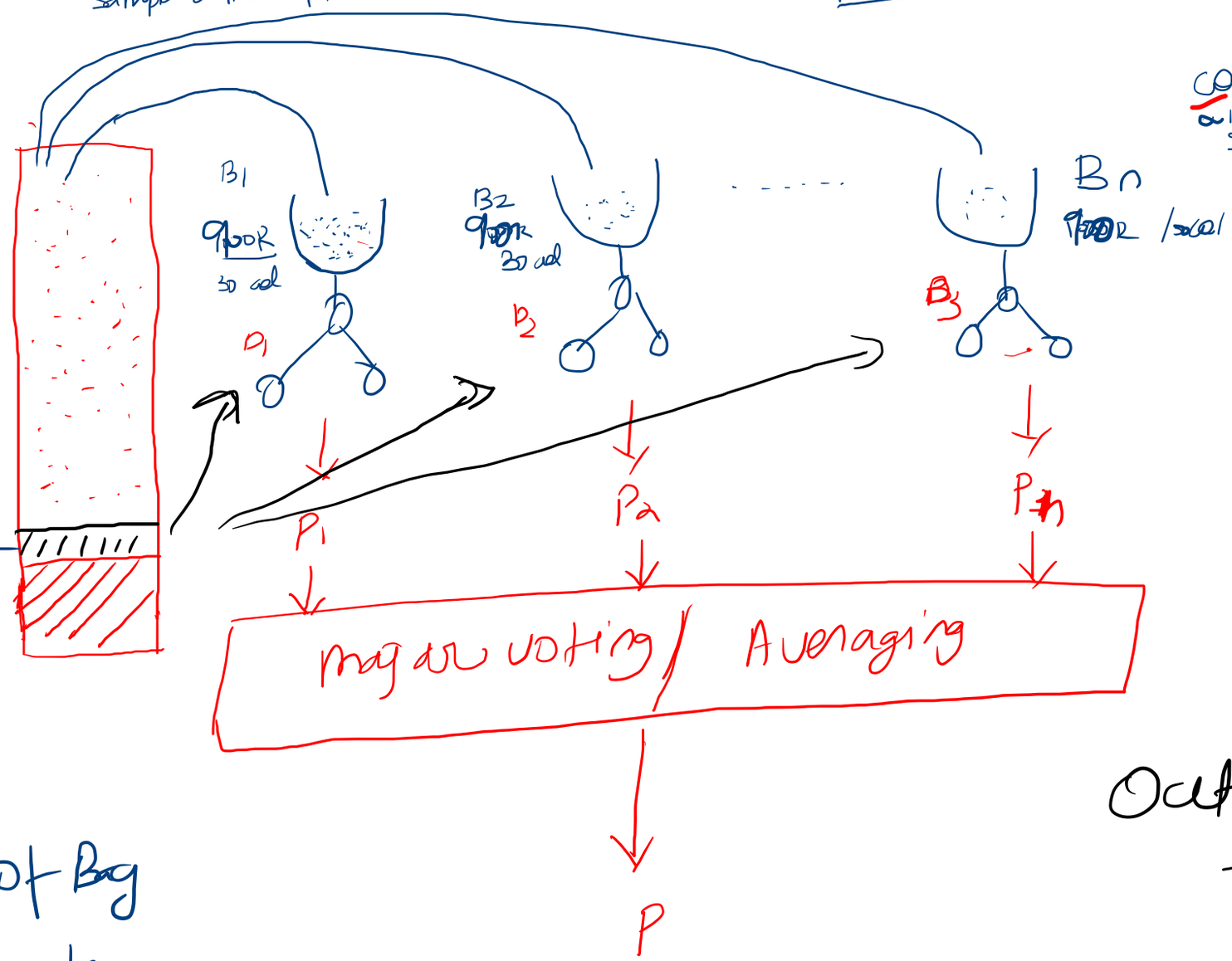
Sample with replacement

$B_1 \neq B_2 \neq B_3 \cdots \neq B_n$



Populations

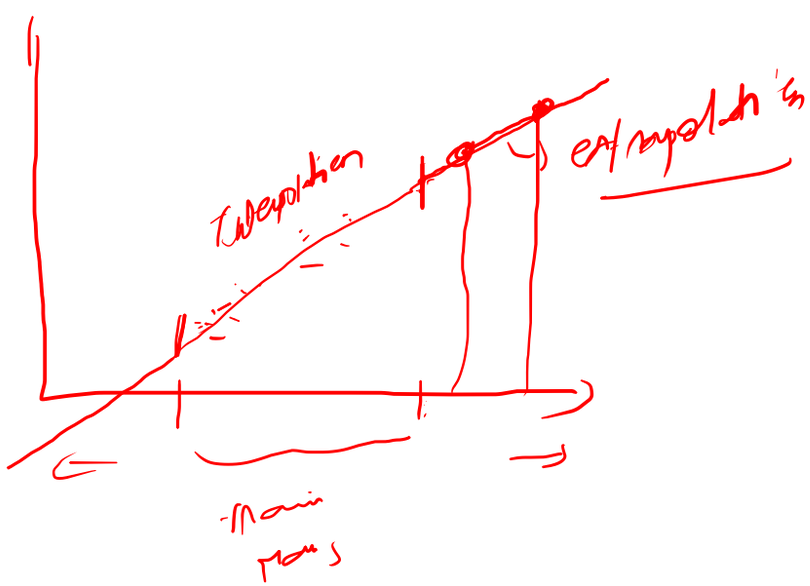
col with 50 inside $= \sqrt{n}$



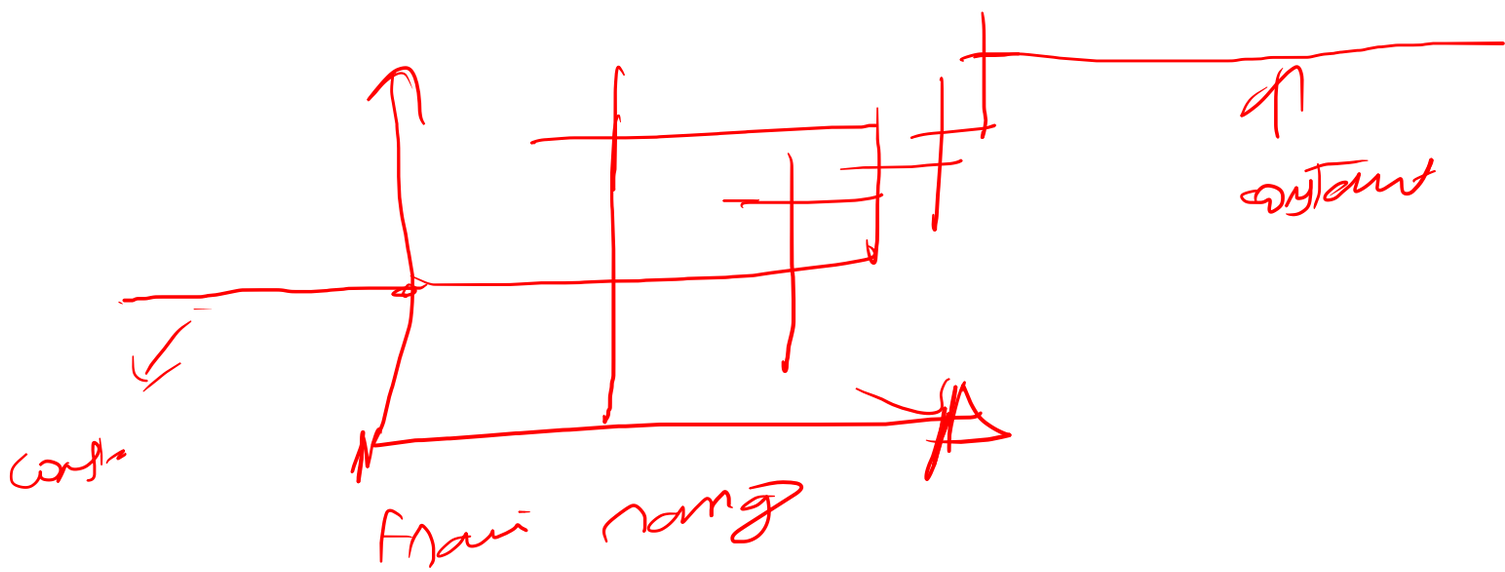

$D_1 \neq D_2 = B_3 \neq \cdots D_n$

Out of bag error

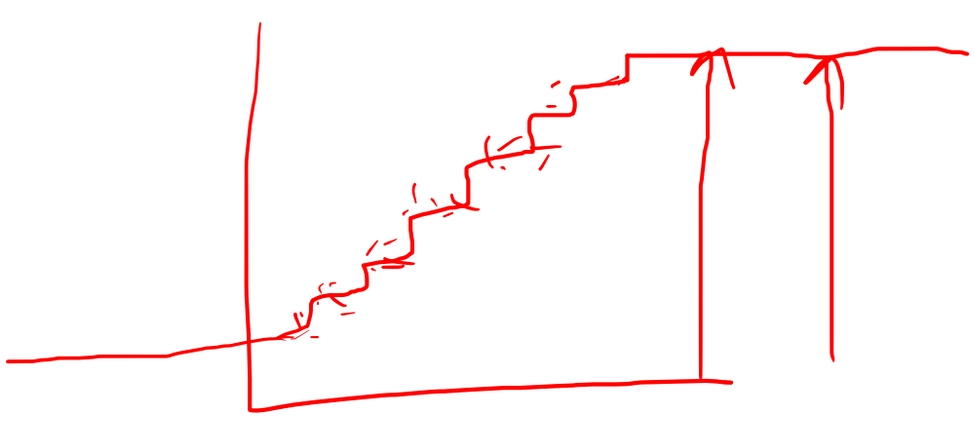
## Linear regression




## Decision regression




Extrapolation in regression
it is not advice to
use

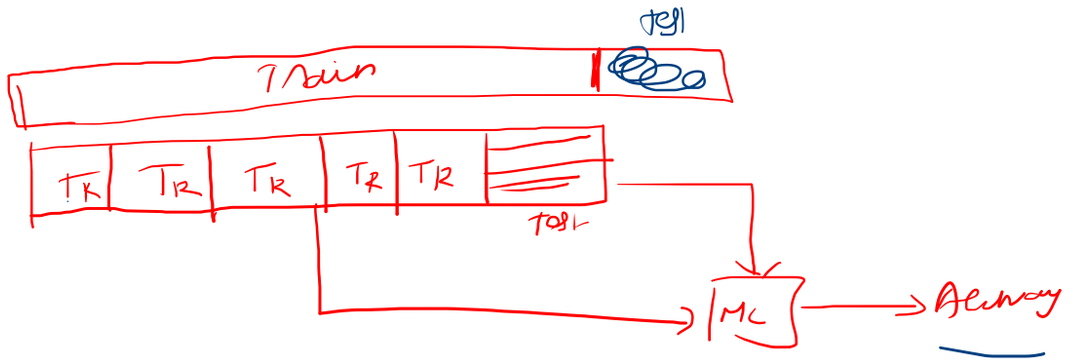


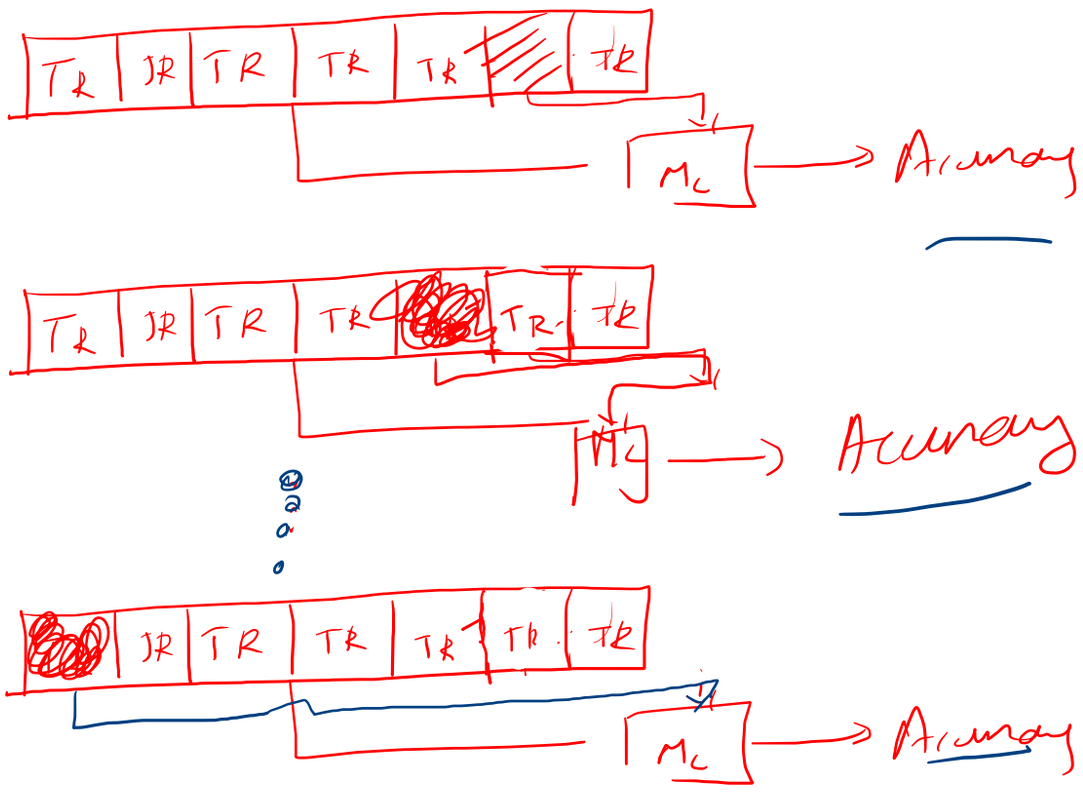
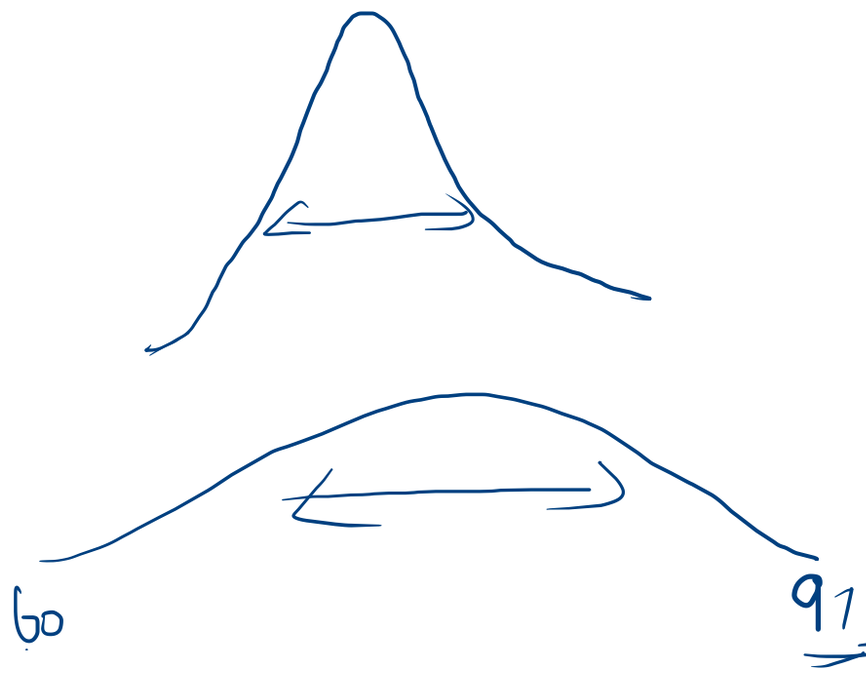
CART

Cross validation [ -> prevention overfitting ] -> detect overfitting

K -> fold cross validation

CV -> cross validation




10 x 1 fold cross validation