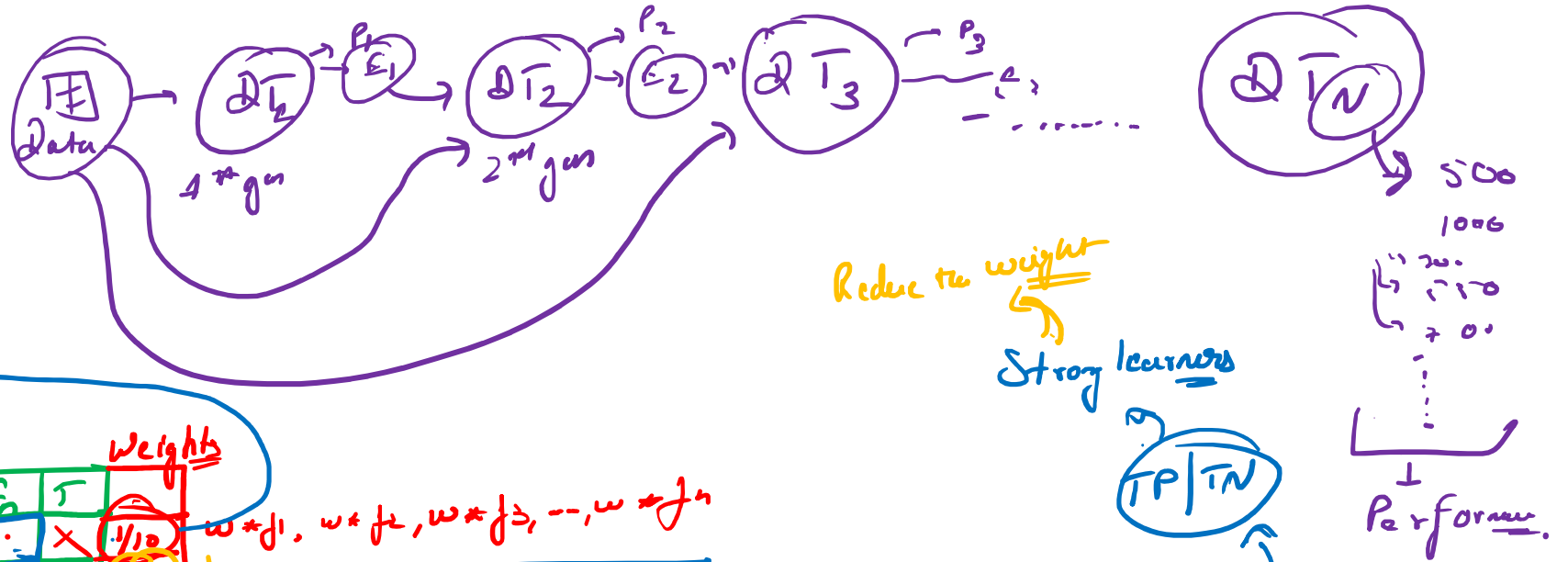
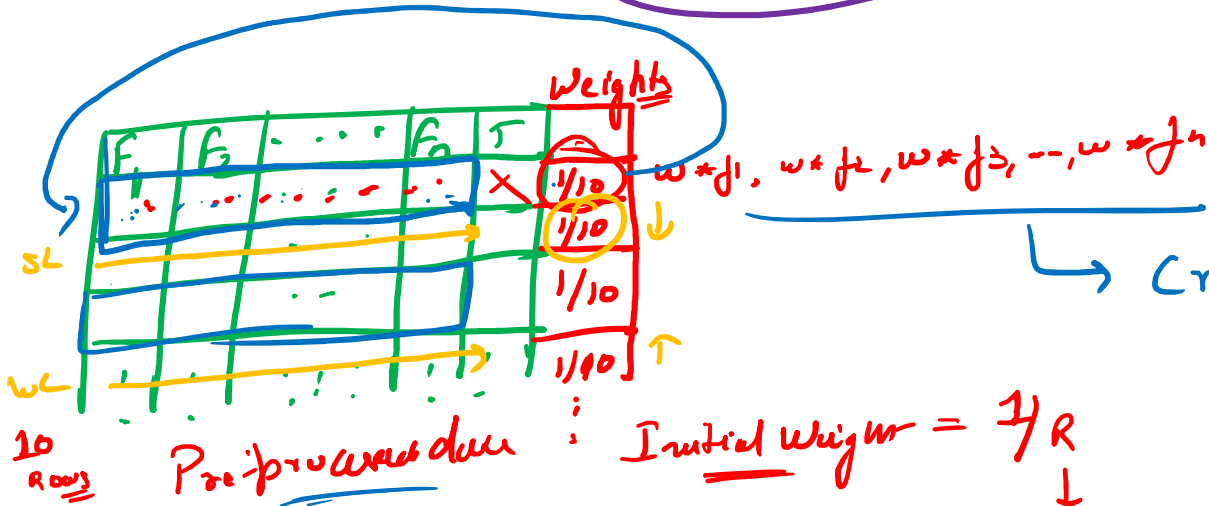


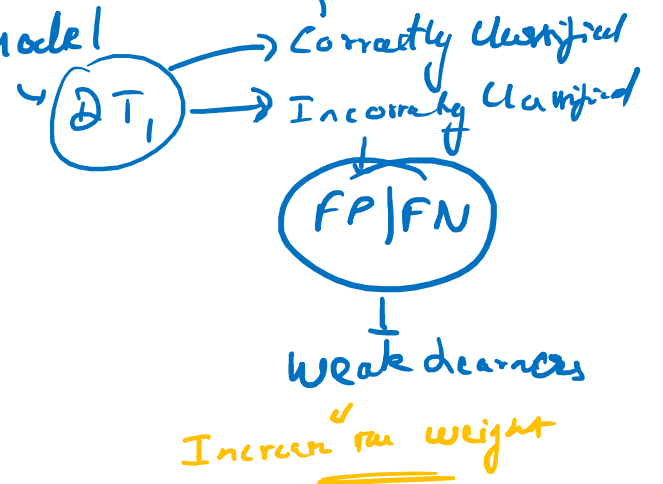
→ Boosting

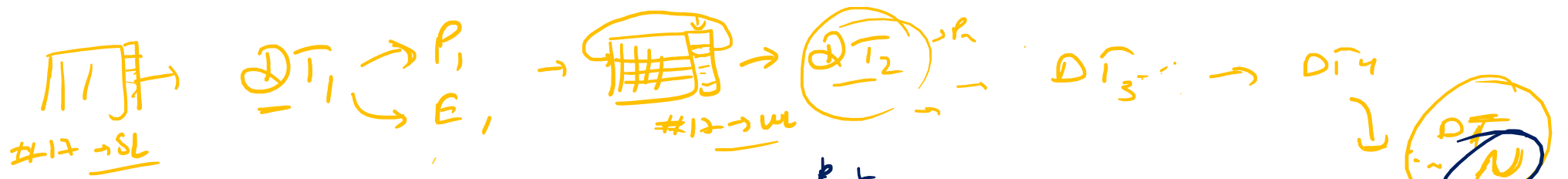


Reduce the weight
Strong learners



Create a D.T. Model





Increase the weight = Old weight $\times e^{p_{\text{err}}}$
 Decrease the weight = Old weight $\times e^{-p_{\text{err}}}$

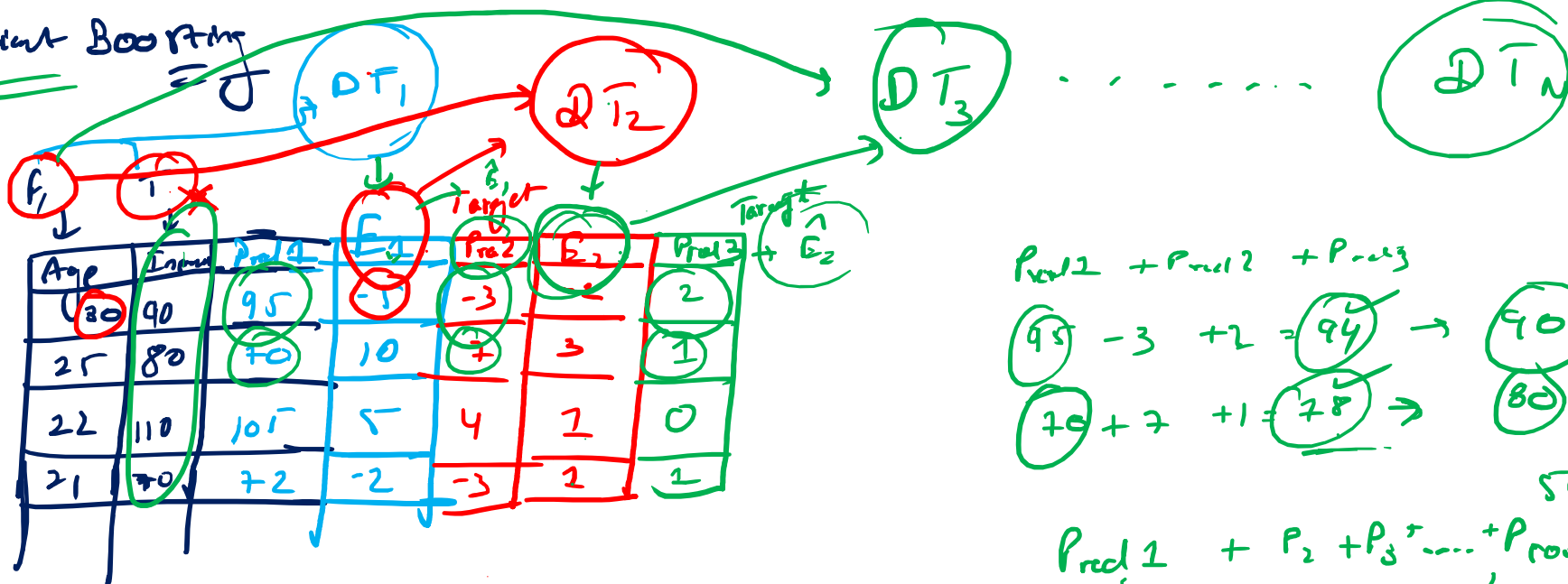
$$P + E = 1$$

Ada Boost (Adaptive Boosting)

$N = 1000$
 \downarrow
 97%
 $N = 100$
 \downarrow
 99%

→ Gradient Boosting

Err² of \hat{y} .



$$Pred1 + Pred2 + Pred3$$

$$95 - 3 + 2 = 94 \rightarrow 90$$

$$70 + 7 + 1 = 78 \rightarrow 80$$

500 Rows

$$Pred1 + P_2 + P_3 + \dots + P_{100}$$

4 lines

↓ (λ)

Add a learning rate

$$Final\ Pred = Pred1 + \lambda \hat{E}_1 + \lambda \hat{E}_2 + \lambda \hat{E}_3 + \dots + \lambda \hat{E}_{100}$$

→ K - fold Cross Validation (CV)

↓
Arbitrary Number

Most used of K in CV = 10

→ Data → Tr + Test

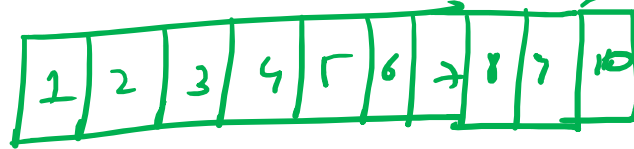
↓ divide the data into
K Parts

↓
Row

1000 Rows

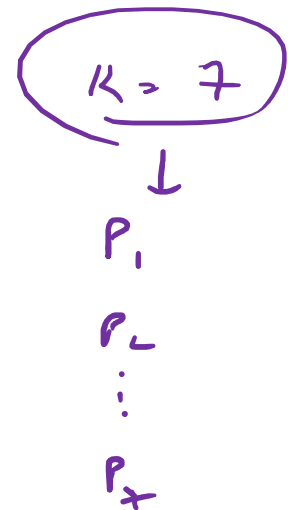
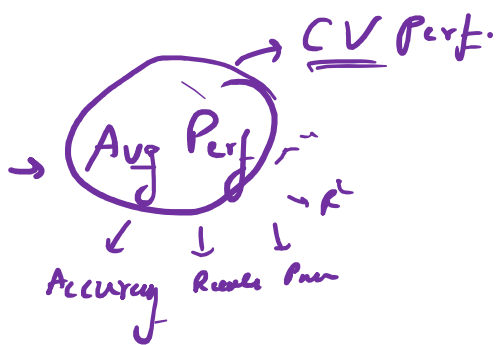
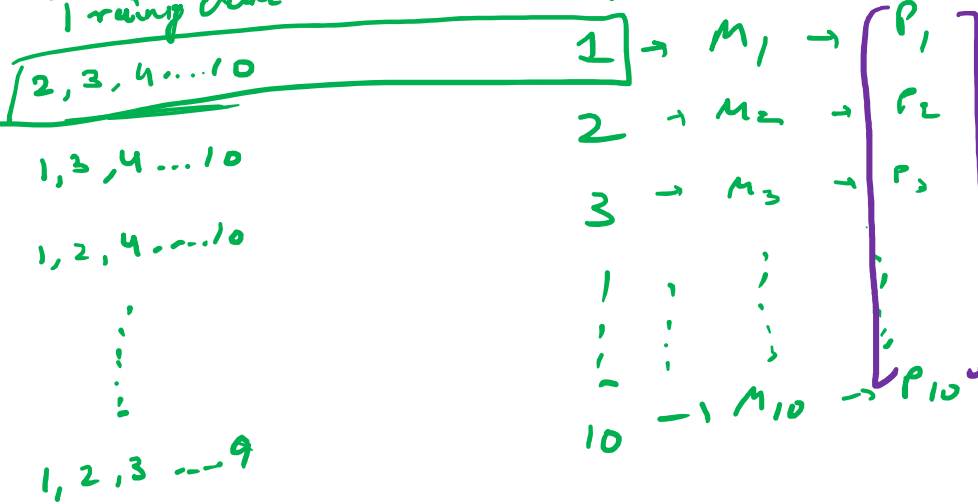
K = 10

↓
 $1000 / 10 = \underline{100 \text{ Rows}}$



Training Data

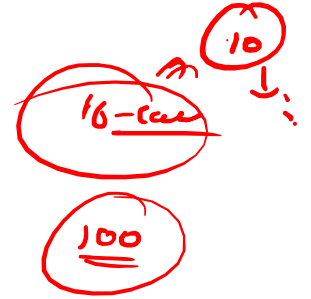
Testing Data



→ Hyperparameters for Ensemble Models

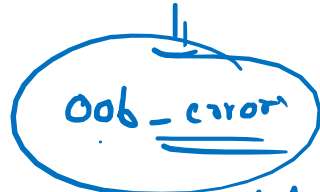
- ① n -estimators \Rightarrow No. of Trees
- ② max-features \Rightarrow sqrt, auto, 0.7 * bagging
- ③ min-sample-leaf \Rightarrow 8, 19, 4
- ④ max-depth = 2, 3, 4, ...

Predictive Power



- ① n -jobs \Rightarrow -1, 1, 2, ... 10 \rightarrow No limit

Avg of 100
OOB Turning Error

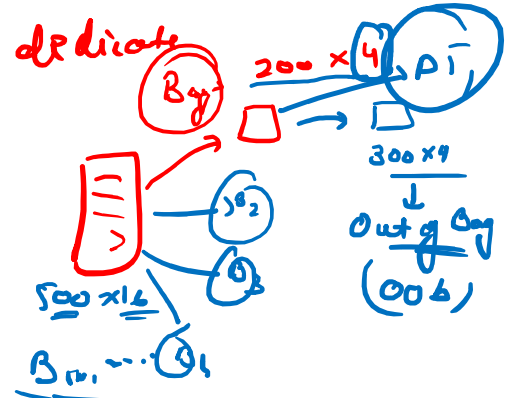
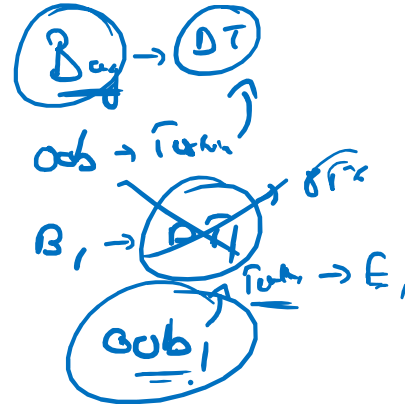


Avg of 100 OOB Arc
OOB-score

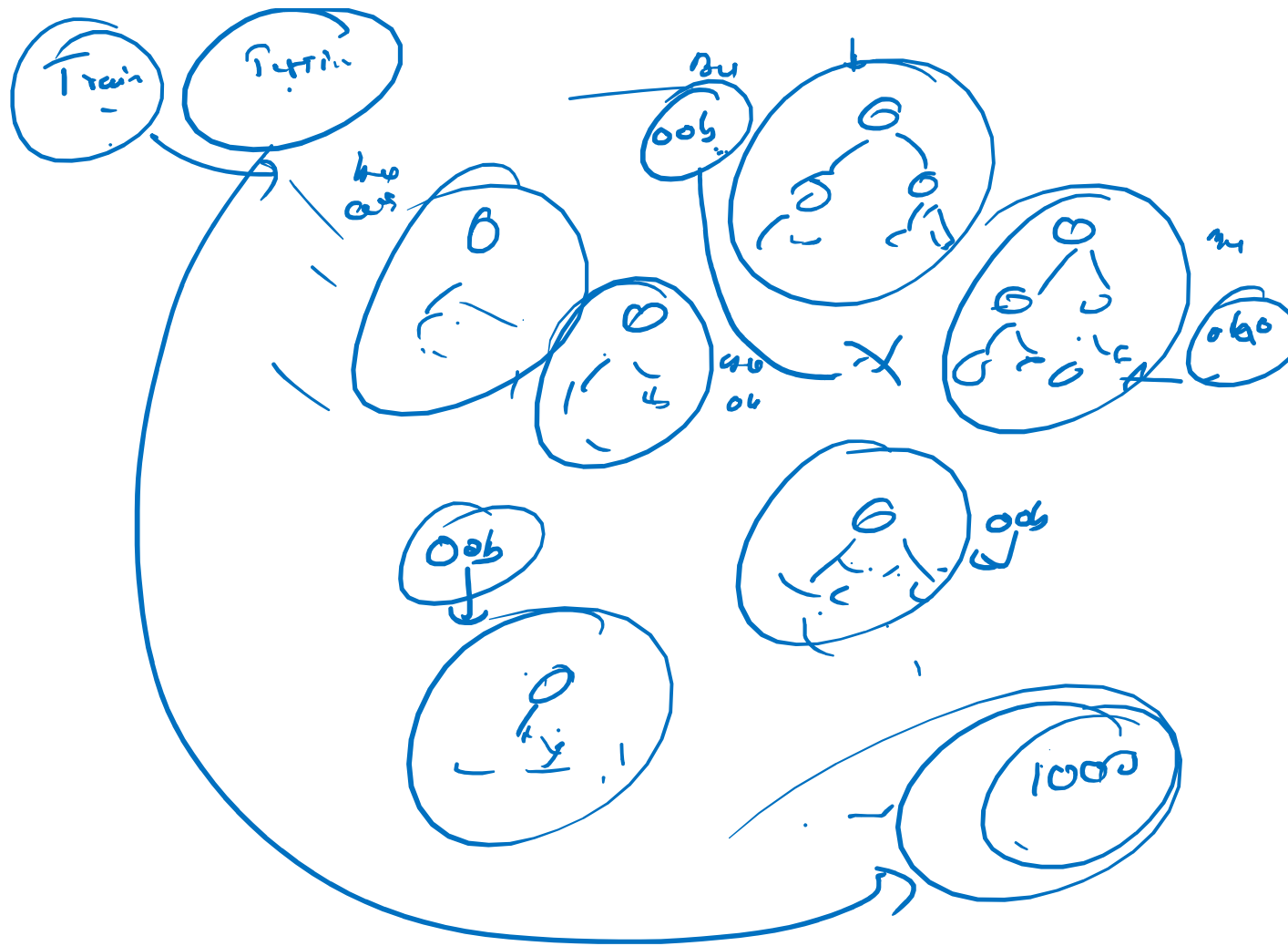
- ② OOB-error \Rightarrow 5%, 1%, 1%
OR

- ③ OOB-score \Rightarrow 90%, 95%, ...

\rightarrow # of CPU cores you want to dedicate



500 bags \Rightarrow 100 OOB
500 OI \Rightarrow 100 Error

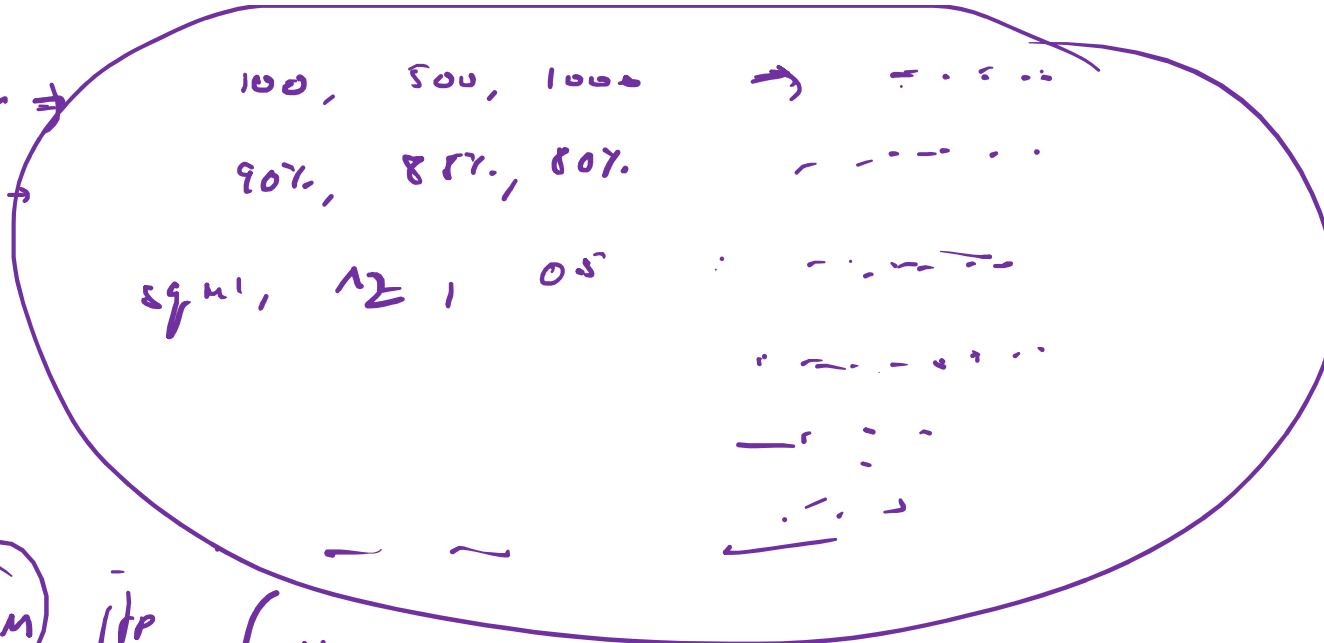


Λ -Criterium \rightarrow

oob.score \rightarrow

max fun \rightarrow

\vdots



MLM

\swarrow 1/p

\swarrow 1/p

\swarrow data

Sklearn \rightarrow

Grid Search CV

\rightarrow O/P \rightarrow The best value of each Parameter

09:37 PM