

Random Forest → is a versatile & widely used
in ML that belongs to the class of ensemble method.
It is type of bagging method, which involves training
many individual model (DT only) & combining their
outputs to make a final prediction.

R Forest → more no of DT

RF :- $\boxed{DT + Bagging + column Sampling}$

Bagging

SVM D - -

KNN D ! , ! , ! , !

RF : DT D -

Random Forest

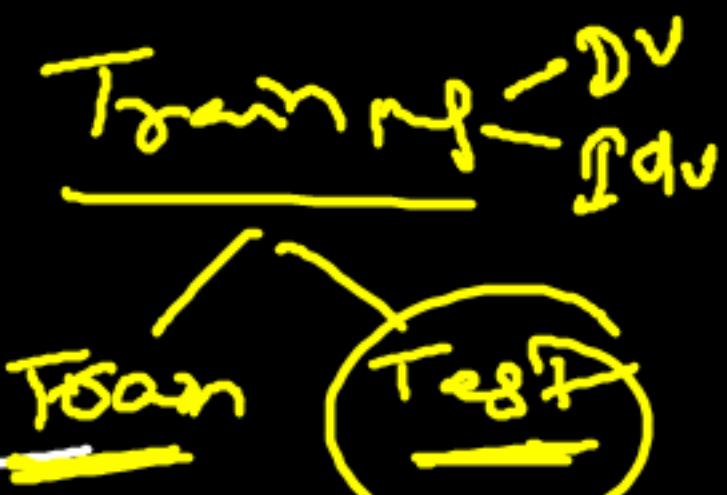
Mute Start Video Security Participants Chat Polls New Share Pause Share Annotate Remote Control Apps More

You are screen sharing

Stop Share

Low Bias & High variance

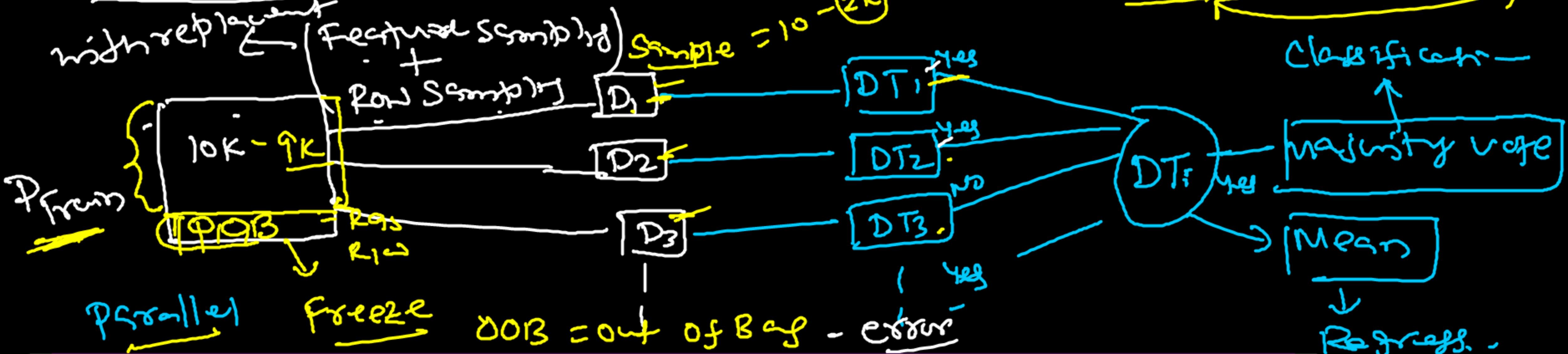
↓
different background



↓ Target - RF

Low Bias & low variance → Objective

with replacement (frequent sample) \Rightarrow Sample = $10^{-2}K$



logist_{RF} = LogistReg()

WRF · fit(X_trn, Y_trn)

y_pred = $(x_{test}) \xrightarrow{g-test}$

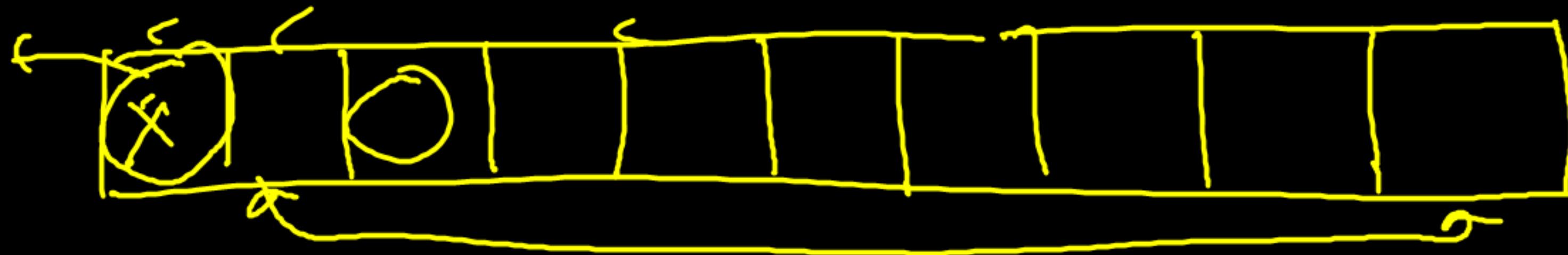
Classification

Majority Vote

Mean

Regress.

Cross validation ✓



Model

Train ✓
Test ↗

$$CV = 1 \quad (2, 3, 4, \dots, 10) = 1 - \text{ignore} = 80\%$$

$$CV = 2 \quad (1, 3, 4, \dots, 10) = 2 = 90\%$$

$$CV = 3 \quad (1, 2, 4, 5, \dots) = 3 = 70\%$$

Train = 1000

Row & column

CAT = DJ, Eng, GKGS, Dust

Model - Test ↗

Final → data leakage

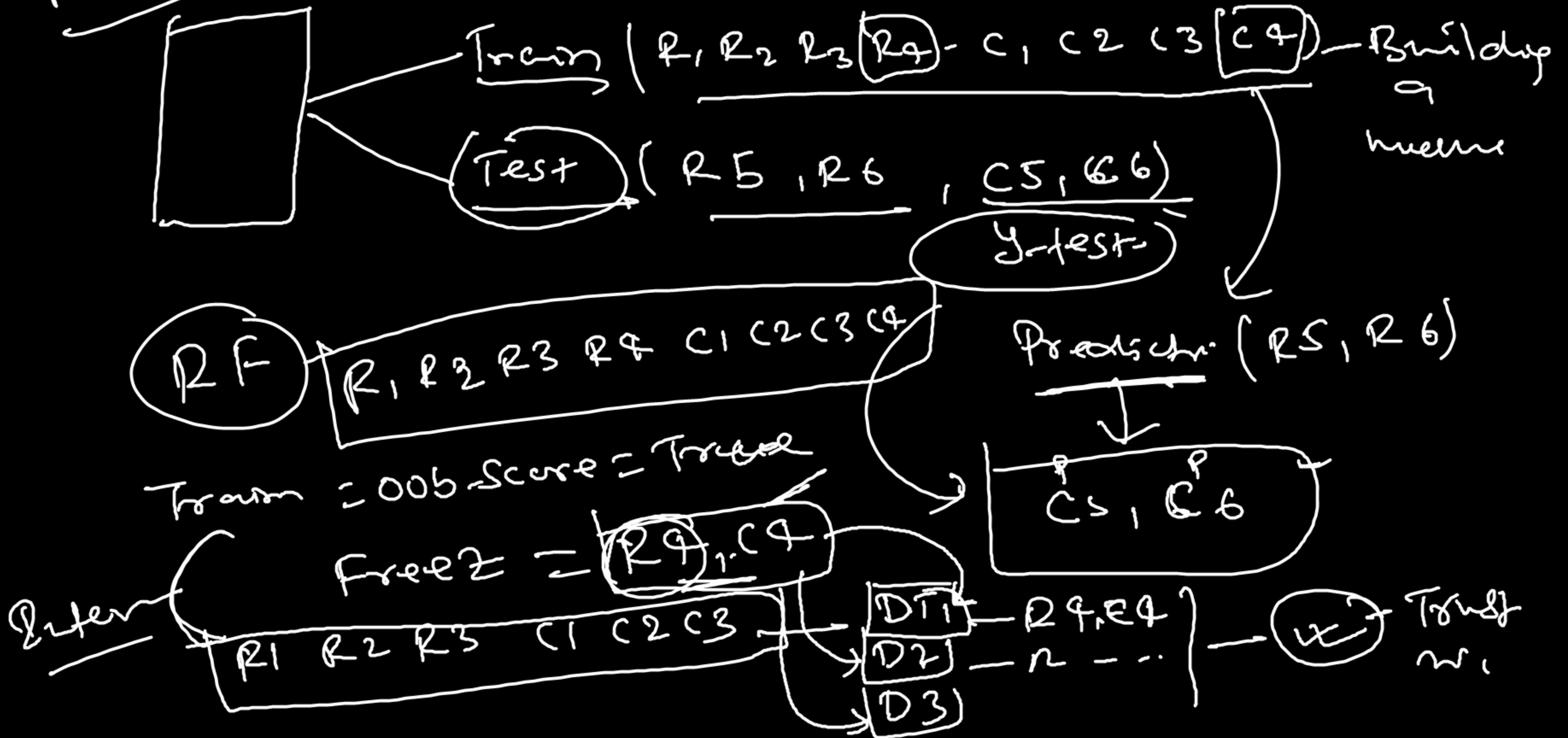
<u>C₁, C₂, C₃</u>	<u>C₄</u>
<u>R₁, R₂, R₃</u>	<u>R₄</u>

Prep ↗

Mean = 85% Train ↗

Max = 90% ↗

Now van schijle



Extremely randomize Tree

↳ Try out a random Sample of possible
Value to determine τ (Tau - Threshold)

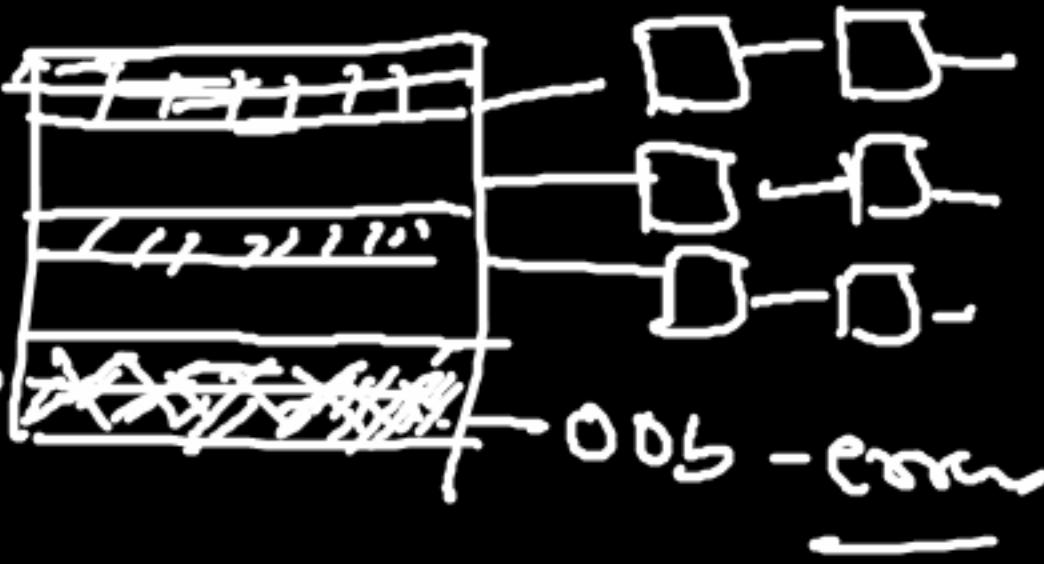
$$RF := [RS + CS + \tau]$$

[005]

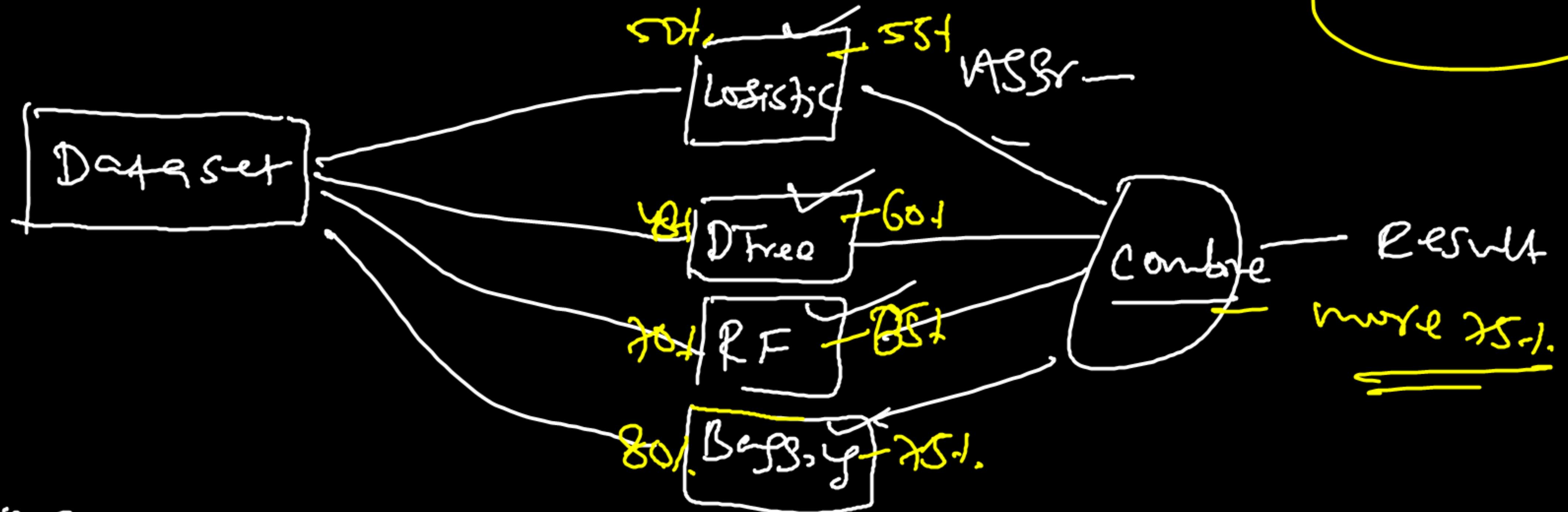
Randomization as a way to reduce variance

τ - Tau - Threshold

Case Study

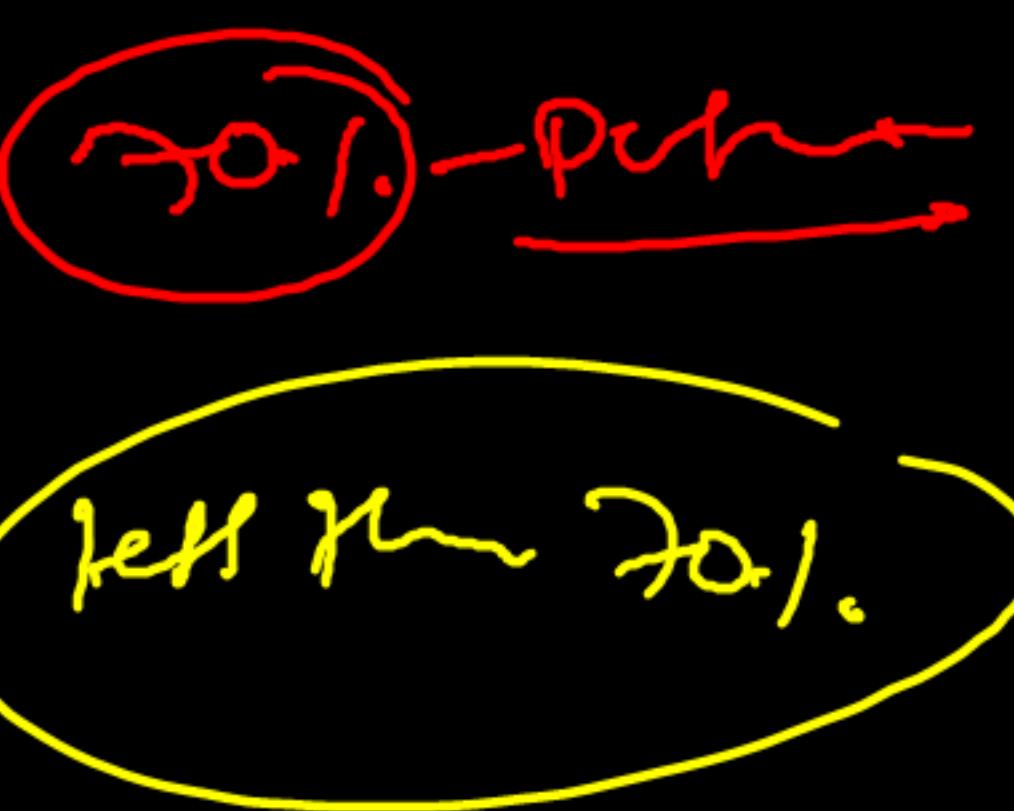


Voting Ensemble - core idea / Intuition



Assumptions :-

- ① All models work independently
- ② All model accuracy union should be 51%. Otherwise Voting will give wrong result.



Architecture

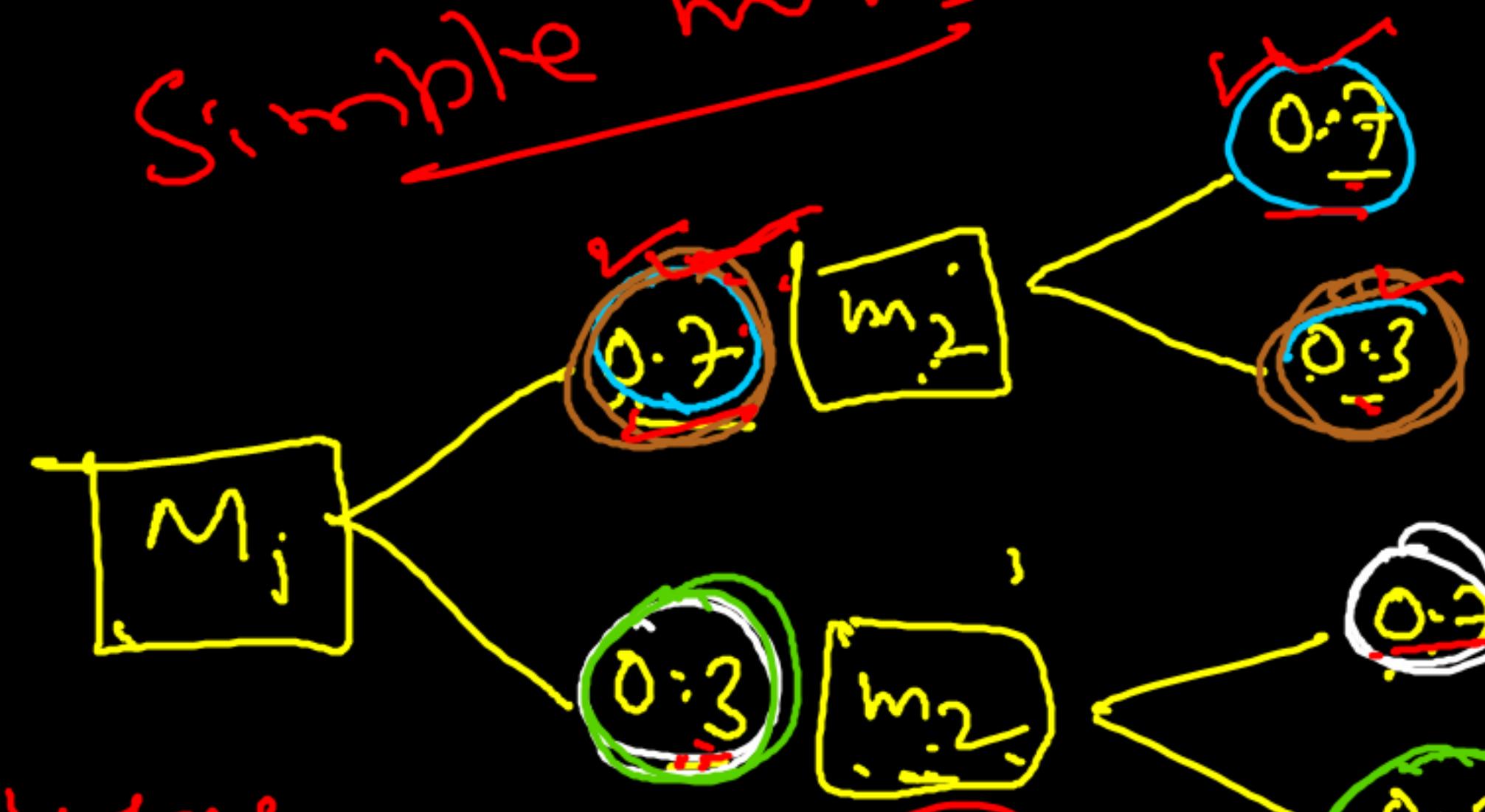
M_1 ✓ 0.7/0.3 ✓

M_2 ✓ 0.7/0.3 ✓

M_3 ✓ 0.7/0.3 ✓

M_4 0.2
0.3

Simple matrix



Voting

correct

$$\{ \begin{matrix} 0.343 \\ 0.147 \\ 0.147 \\ 0.147 \end{matrix} \}$$

$$\sum 0.784 = 78.4 \text{ Accm}$$

0.7
Correct wrong

$$0.7 = 0.343 \leftarrow (0.7, 0.3, 0.7)$$

$$0.3 = 0.147 \leftarrow (0.7, 0.3, 0.3)$$

$$0.7 = 0.147 \leftarrow (0.7, 0.3, 0.3)$$

$$0.3 = 0.063 \times (0.7, 0.3, 0.3)$$

$$0.7 = 0.147 \leftarrow$$

$$0.3 = 0.063 \times$$

$$0.7 = 0.063 \times$$

$$0.3 = 0.027 \times$$

Voting error

-22-1.

Logistic Reg < 0.7
0.3

$$\frac{DT}{T} < 0.65$$

KNN < 0.75
0.25

Wohng

$$\text{COPY} = \begin{cases} ? \\ ? \\ ? \end{cases} \} \text{Summet}$$

