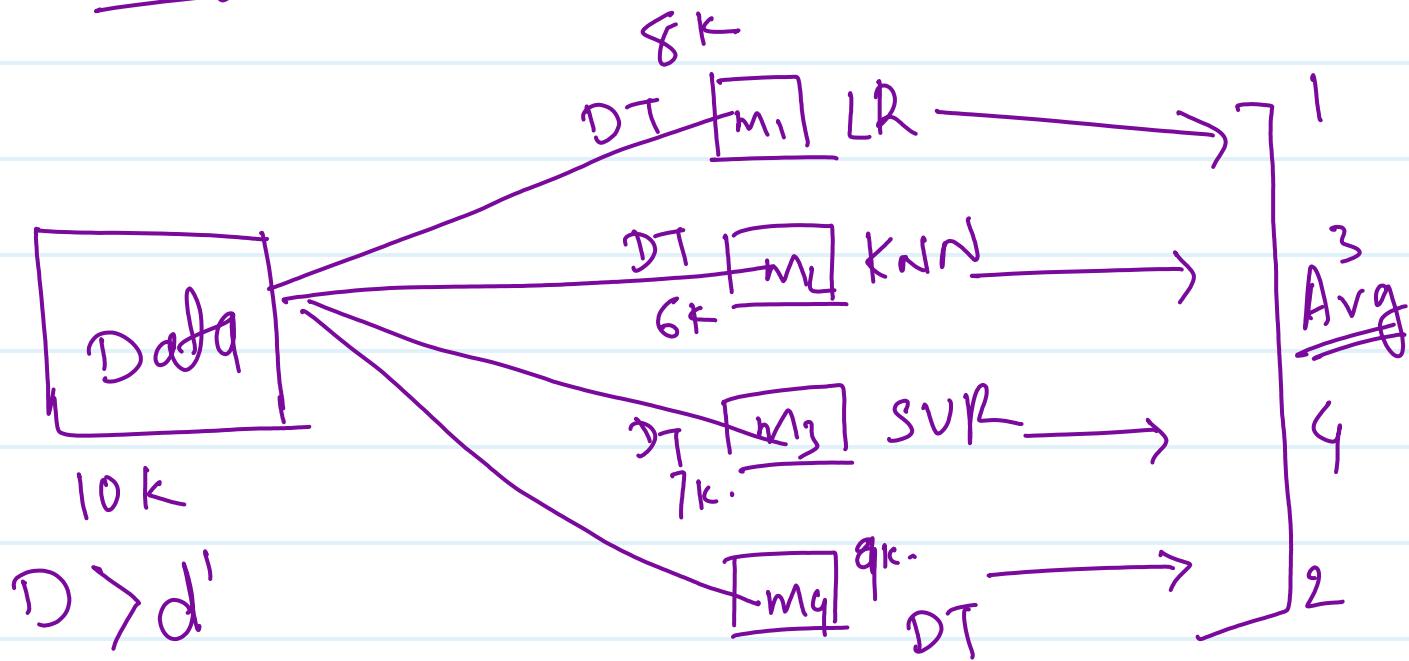


Ensemble Technique.

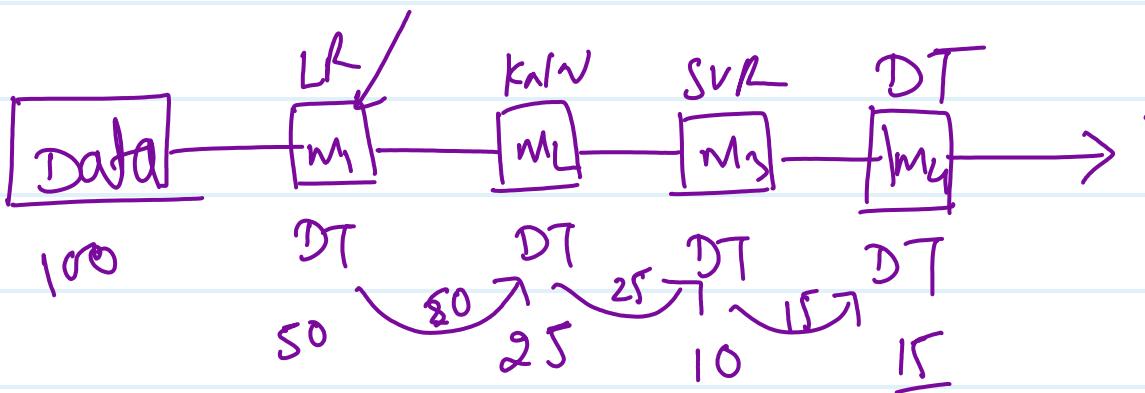
① Bagging

② Boosting.

Bagging



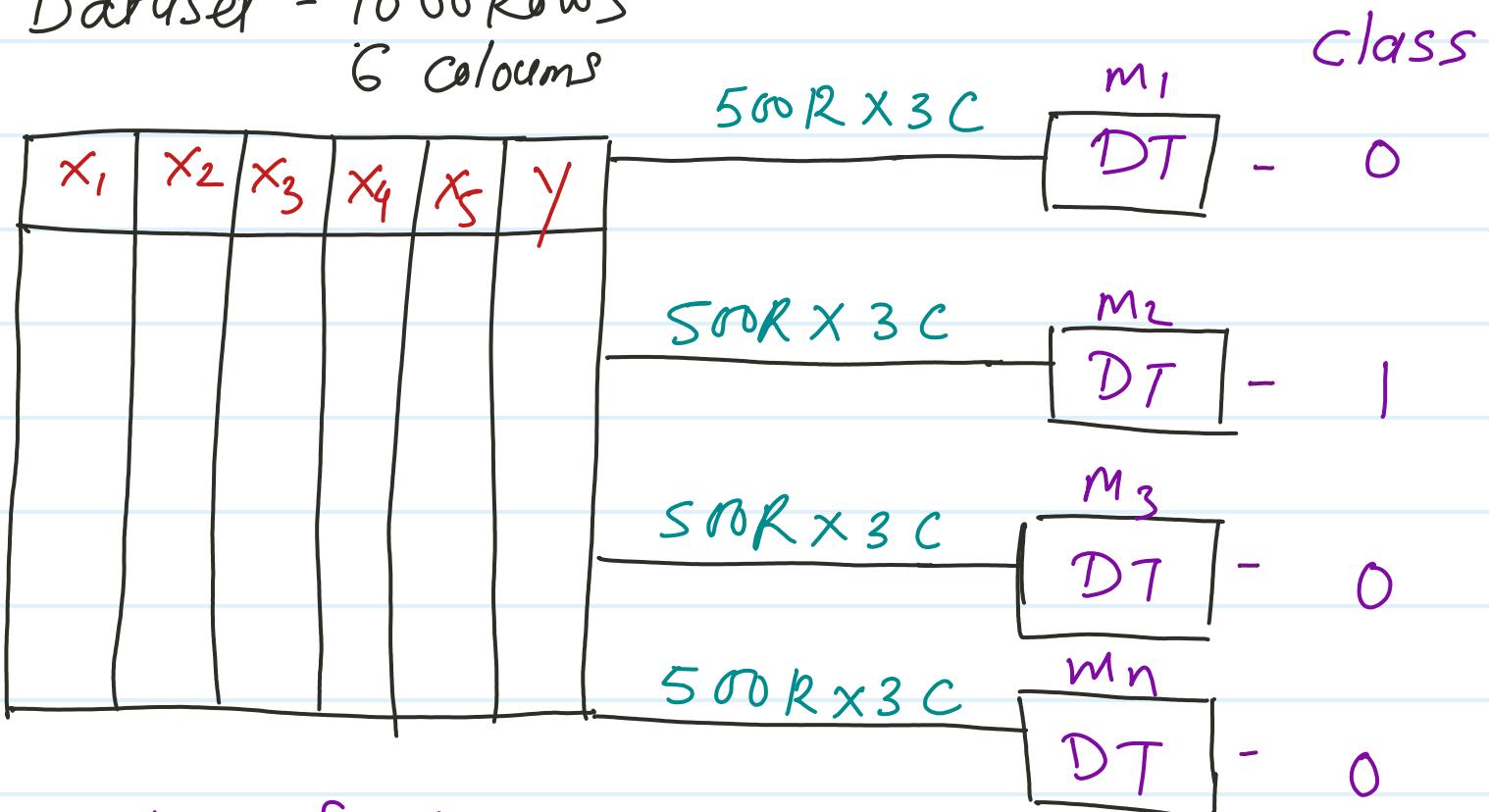
Boosting weak learner



Bagging Method

- * Random Forest classifier
- * Random forest Regressor

Dataset - 1000 Rows
6 columns



classification works on voting based output

entire Dataset - D
subset Dataset - d'

$$D > d'$$

since we are using ensemble technique. Trade off for Bias and variance would be

DT { low Bias
high variance

RF { low Bias
low Variance

Pros -

- ① Random Forest used to make robust model over a decision tree overfitting issue
- ② Random forest widely used in Regression problem due to its capabilities of handle non-linear dataset.

③ Outlier / noise not affected it.

④ Can handle high dimensional data.

⑤ RF does not require feature scaling.

Equation of RF:

$T = \text{num. of DT in Forest}$

$\hat{y}_t(x) = \text{Prediction of the } t^{\text{th}} \text{ DT}$
for input x .

$\hat{Y}_t = \text{final Random Forest prediction}$
for input x .

$$\hat{y}(x) = \frac{1}{T} \sum_{t=1}^T \hat{y}_t(x)$$

We calculate average of all DT.
and it will be our final output.

For hyperparameter tuning, mostly

Randomsearch cv prefer due

to computationd constraint.

