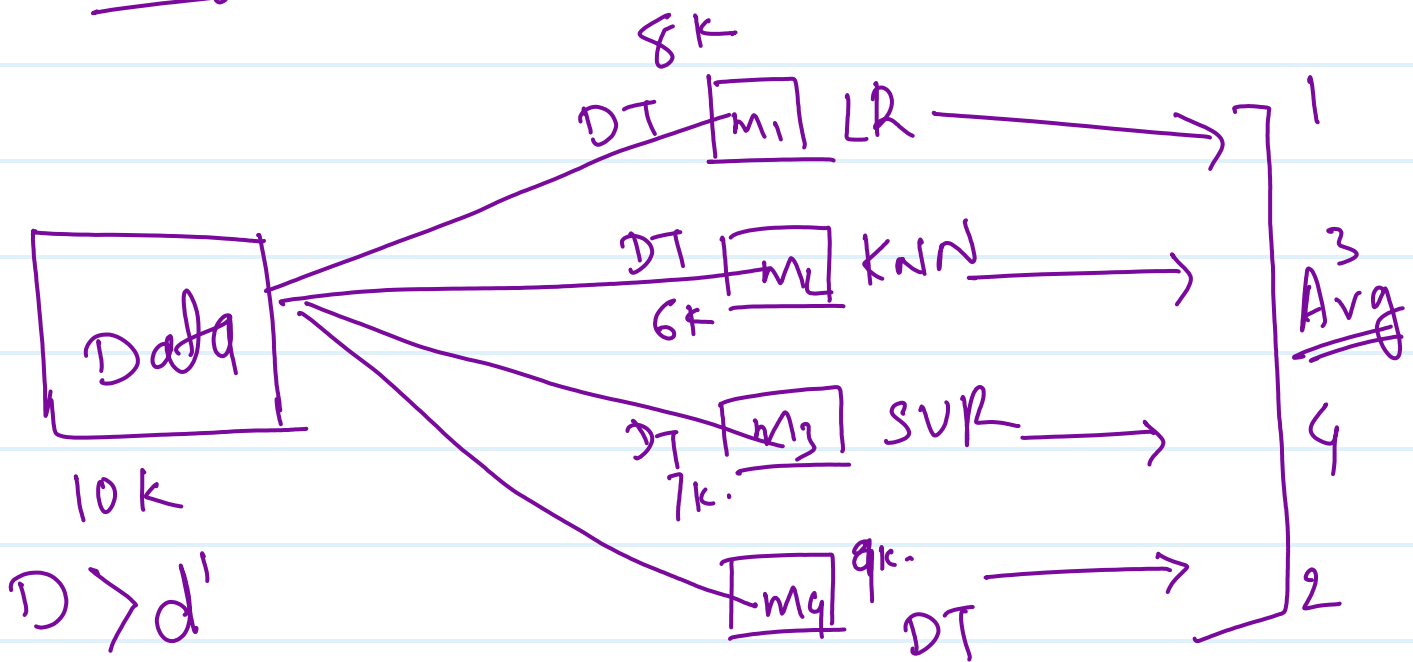


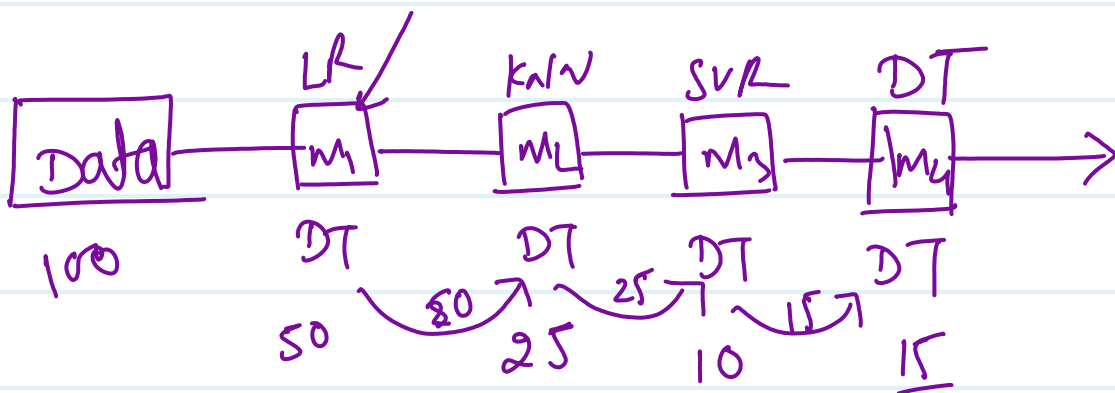
Ensemble Technique.

- ① Bagging
- ② Boosting.

Bagging



Boosting weak learner

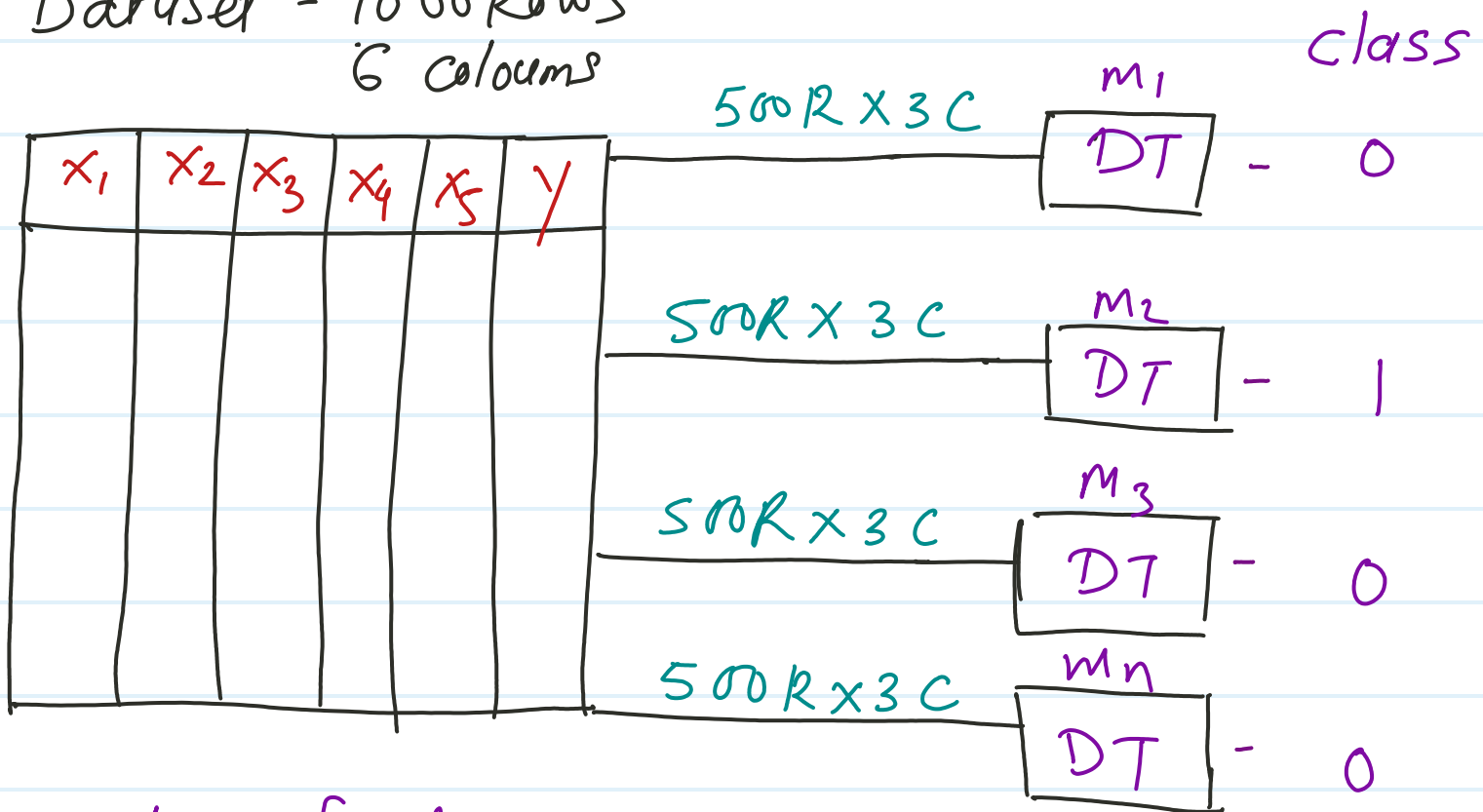


Bagging Method

* 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 required feature scaling.

Equation of RF

T = Num. of DT in Forest

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

\hat{y}_t = 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 computational constraint.

