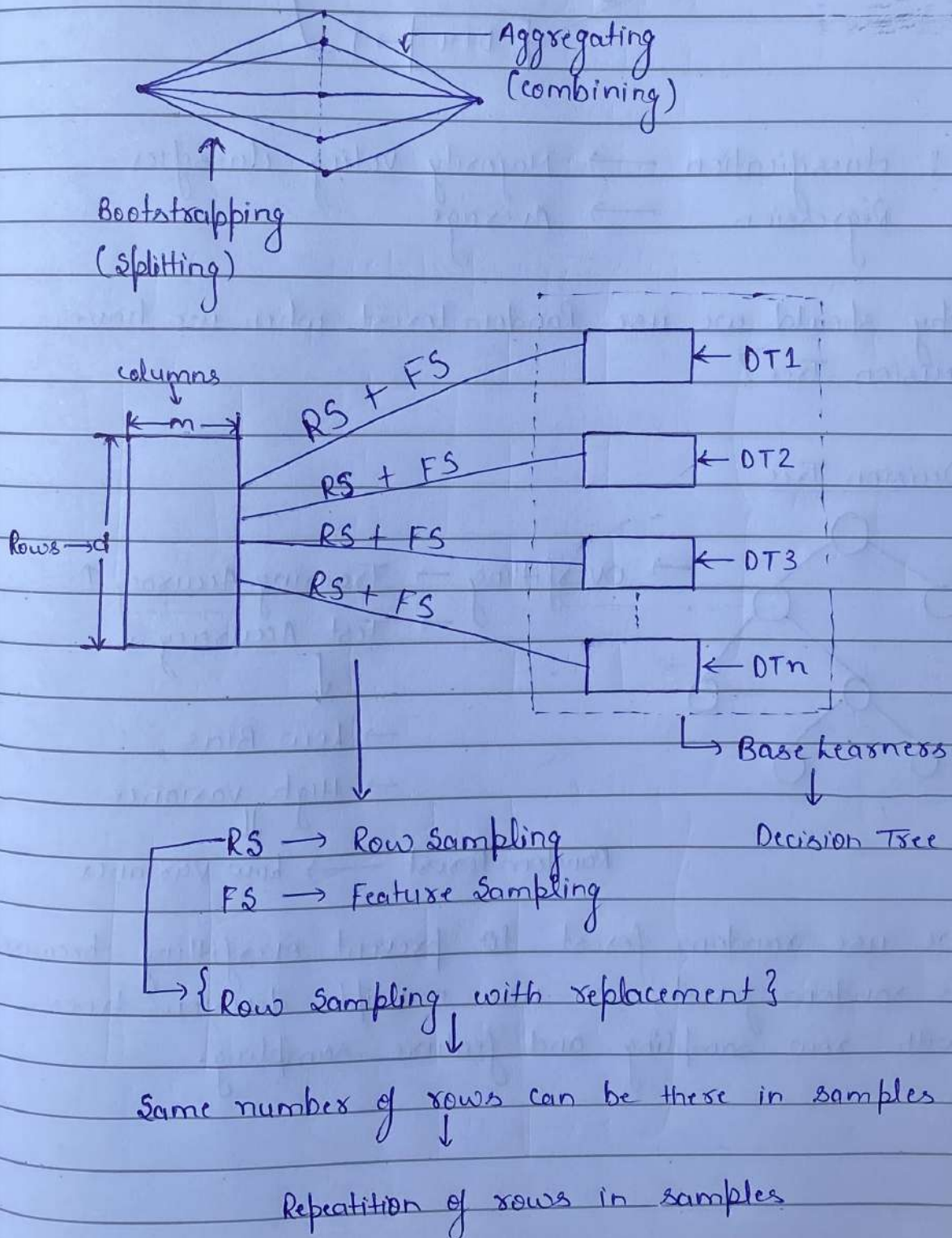
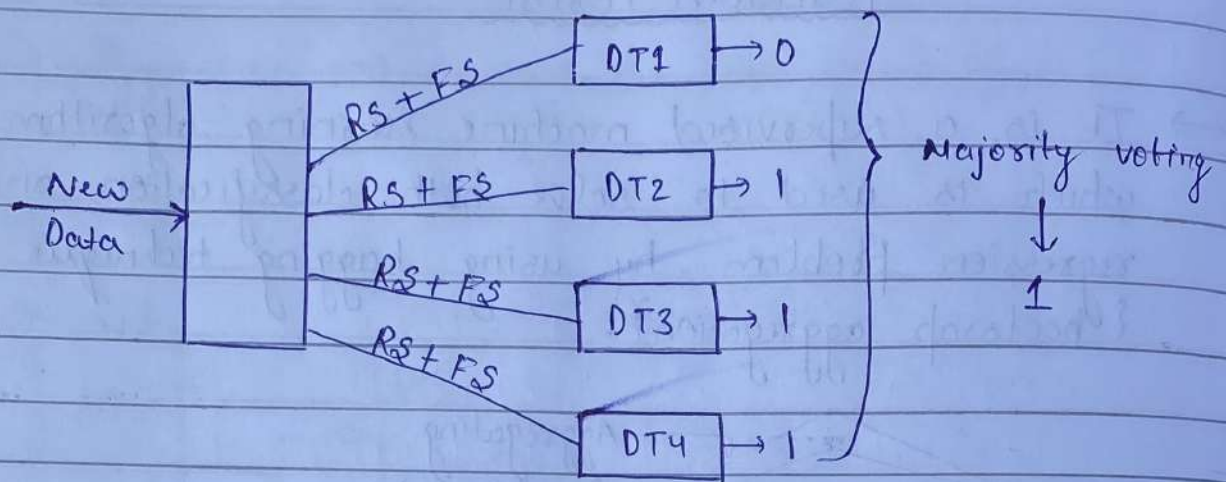


# Random Forest

→ It is a supervised machine learning algorithm which is used to solve both classification and regression problem by using bagging technique {bootstrap aggregation}.



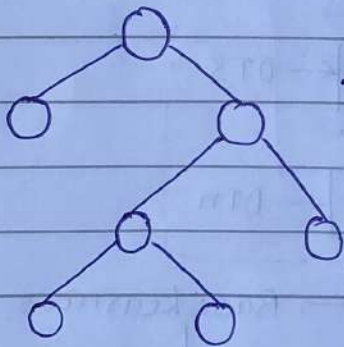


Note : classification  $\rightarrow$  Majority voting classifiers  
 Regression  $\rightarrow$  Average

Q why should we use Random Forest when we have Decision Tree?

$\rightarrow$

Decision Tree



$\rightarrow$  overfitting  $\rightarrow$  Training Accuracy  $\uparrow$   
 $\rightarrow$  Test Accuracy  $\downarrow$

$\downarrow$   
 $\rightarrow$  Low Bias

$\rightarrow$  High variance

$\Downarrow$   
 Random Forest  $\rightarrow$  low variance

We use random forest to prevent overfitting because in random forest, we use multiple decision trees with row sampling and feature sampling.