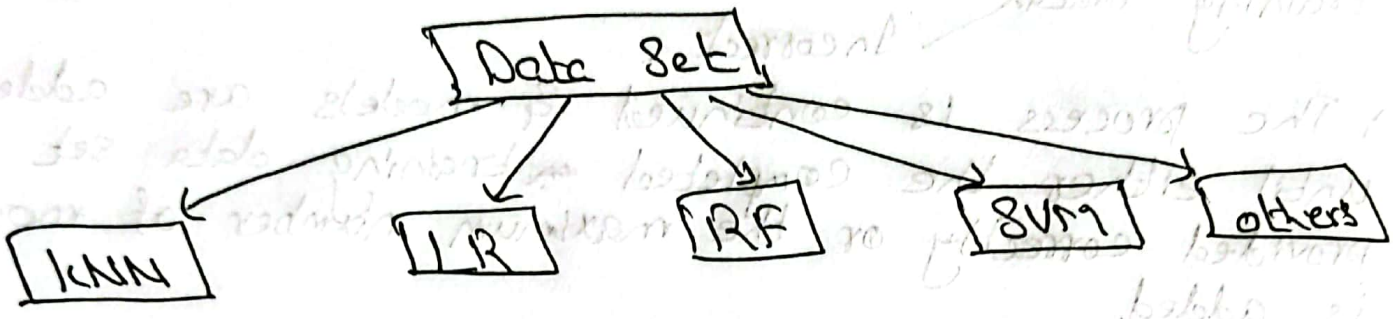
Ensemble Algos
(Lecture 7)

→ Ensemble Learning is consensus based learning algorithm. For example multiple Decision trees or a combination of different algos.

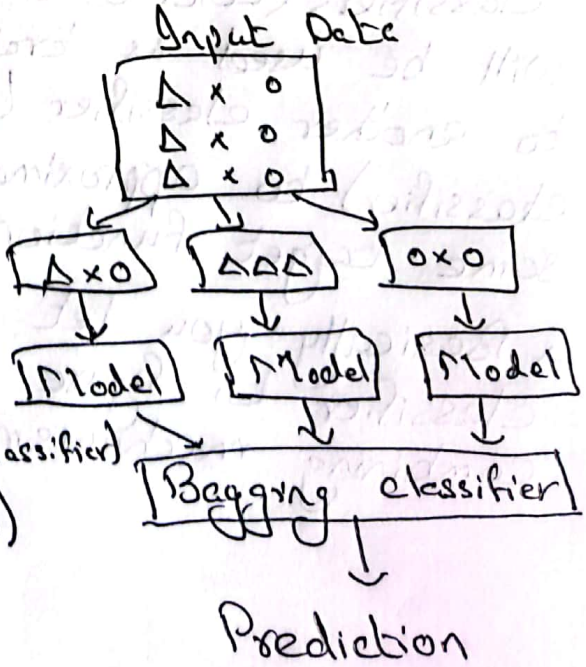
→ This is to overcome an individual weak algo (such as DT's) algo's weakness. For example DT's have a problem of overfitting.



→ Four different techniques Aggregating (bagging):
1) Bootstrap
2) Boosting
3) Stacking
4) Voting

Bootstrap Aggregation :-

→ It decreases variance.
→ Usually applied to DT
→ Special case of model averaging Approach.

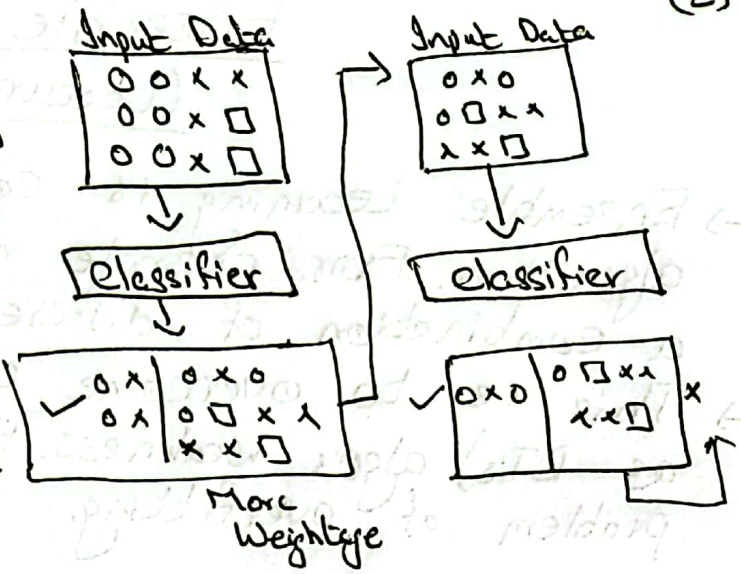


Boosting:-

→ Technique that attempts to build a strong classifier from a number of weak classifiers.

→ Mode by using weak models in series.

→ A model is build from training data
- Correct
- Incorrect

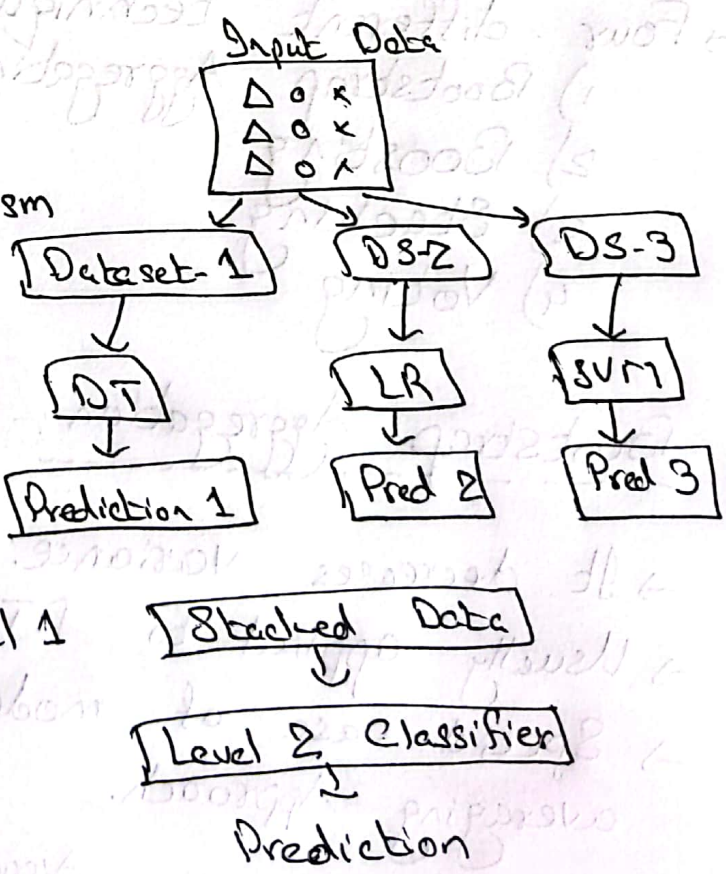


→ The process is continued & models are added until either the completed training data set is provided correctly or the maximum number of models is added.

Stacking, Bleading:-

→ Unlike traditional ensemble learning, the combined mechanism is that the output of classifiers (Level 0 classifiers) will be used as training data to another classifier (Level 1 classifier) to approximate the same target function.

→ Basically you let the Level 1 classifier to figure out the combining mechanism.



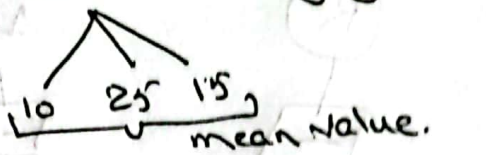
Random Forest:-

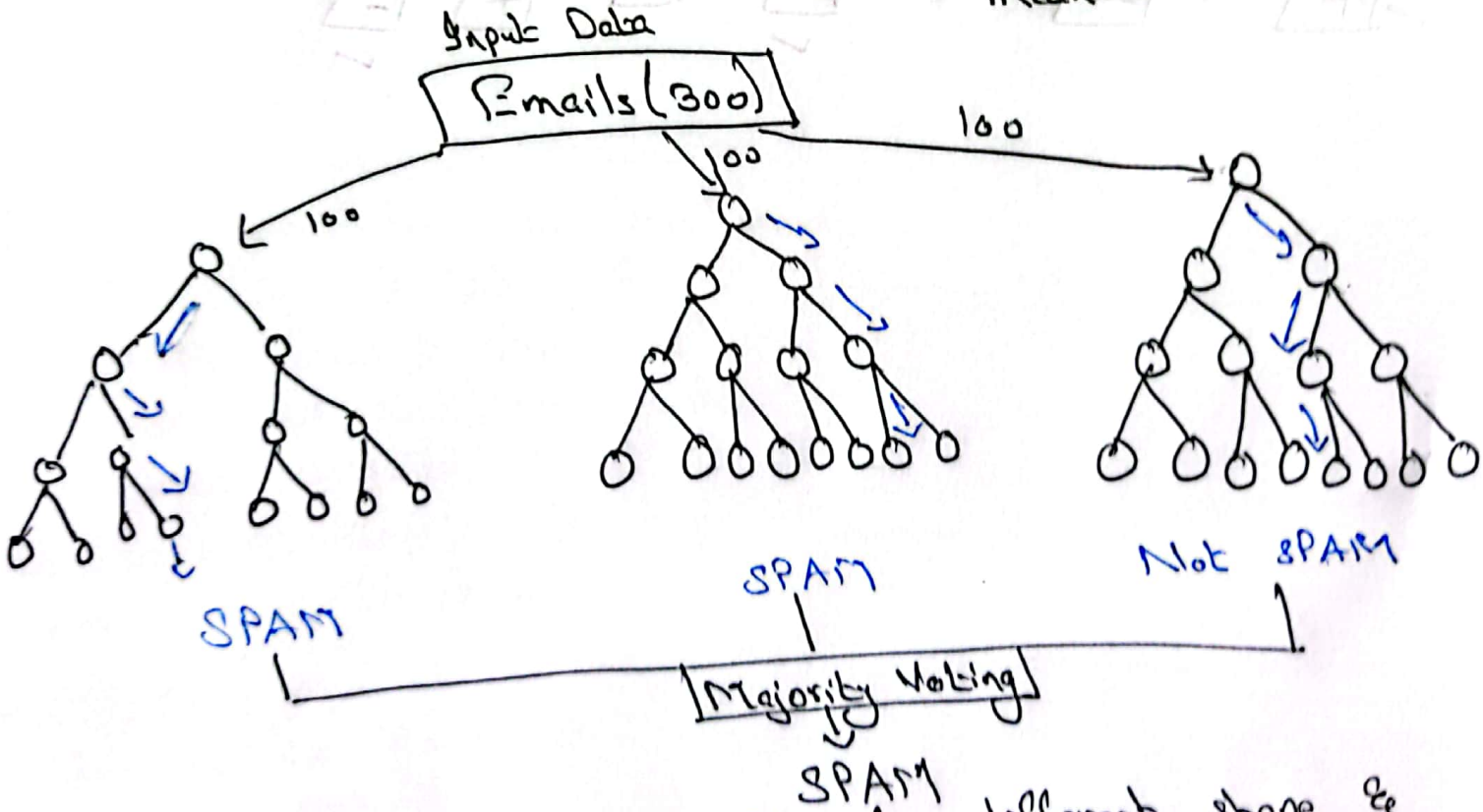
→ RF is an extension of DT to counter shortcomings in DT

→ RF is an ensemble learning technique method, and is used for both classification & regression. However it is more accurate for classification.

→ Steps

1). Create Bootstrap (randomly) Dataset from original data by randomly choosing data (repetition is allowed).
2). Create randomized DTs from bootstrapped dataset.
3). Output of the RF is the class selected by majority of DTs. For regression averaging is employed.





→ RF consists of multiple DTs of different shape & size.

→ Random data leads to random nodes of DTs.

→ 'n' number of Bootstrap samples for 'n' number of DTs.

ID	P	Q	R	Class
20	0	0	1	+
21	1	0	0	−
22	2	0	0	−
23	1	1	2	+

ID	P	Q	R	Class
20	0	0	1	+
22	2	0	0	−
20	0	0	1	+
22	2	0	0	−