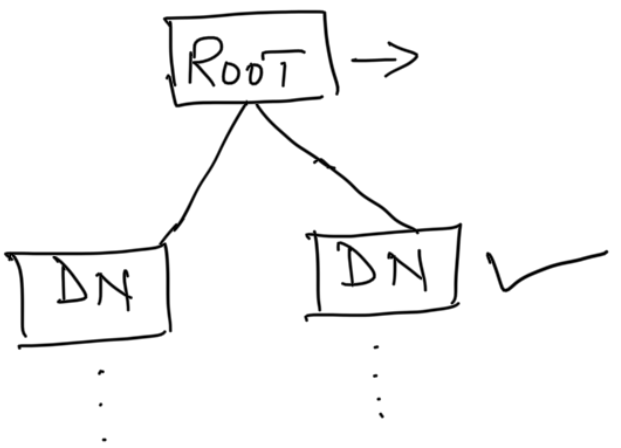
Decision Trees.



A / B / C / D / ....

LN LN ... LN LN ...

Overfit → high chances.

* A complex DT is highly likely to overfit.

How to avoid it??

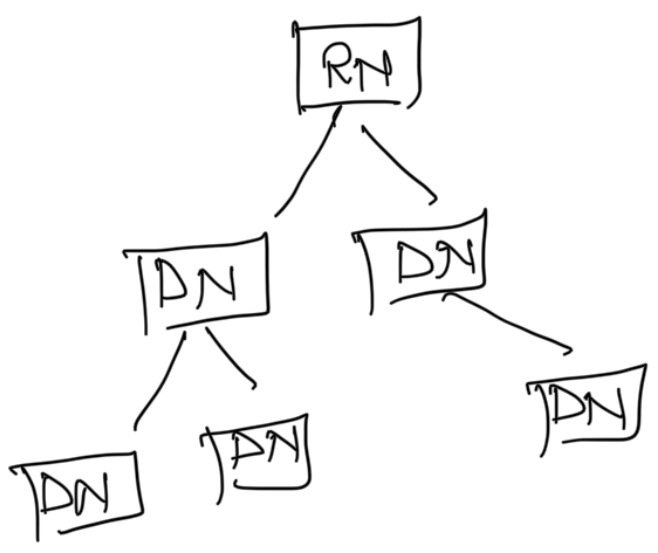
Overfitting → complexity.

complexity → depth

depth → Nodes.

Stop it by limiting the nodes.

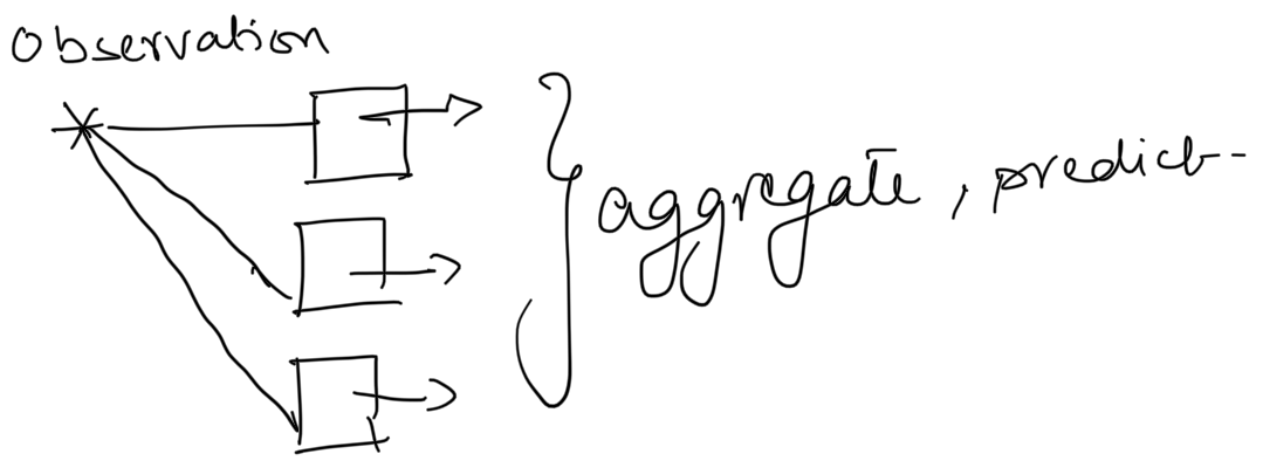
Pruning →



Ensemble learning :-

KBC → audience poll. → better
us
expert →

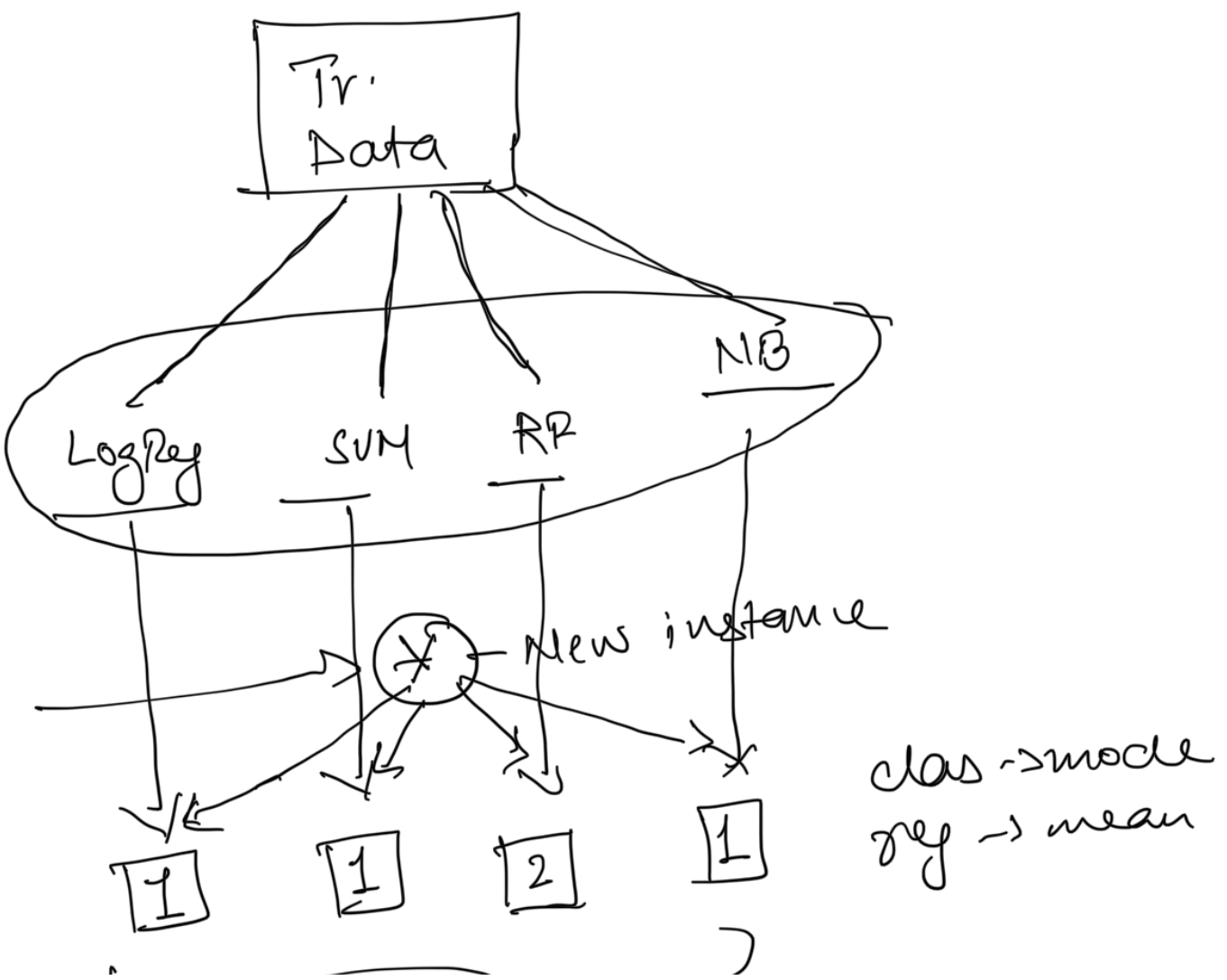
"wisdom of the crowd"



ensemble, ensemble learning.

Types

1) Voting classifiers



voting.

(1)

prize money – 10,000 $

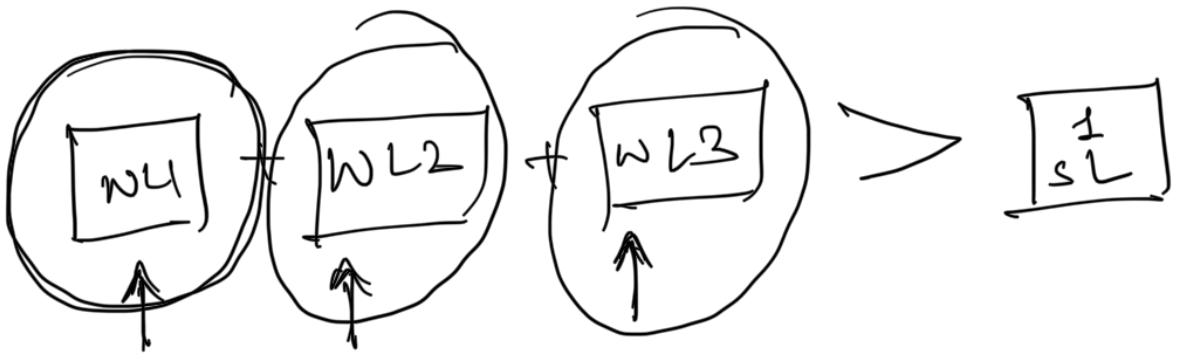
A student.

A: 7 days, less challenge → above average

B: 13 days, more challenge → Top 3

Choice A
Choice B

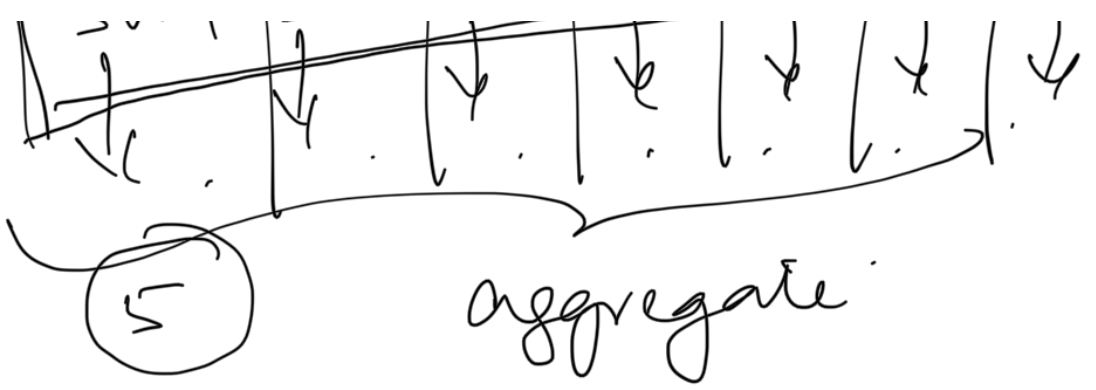
each classifier should be a weak learner.



Weak learner → just better than underfitting

Bagging

DT

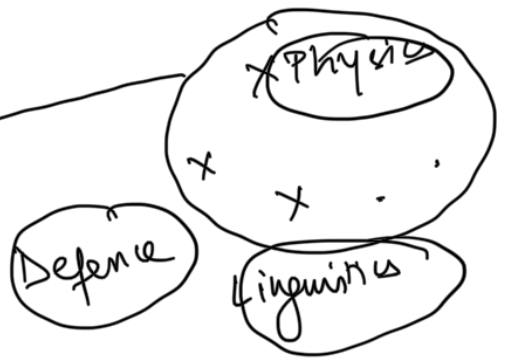


N models; N samples.

DT ← Voting ← Bagging.

diversity → high variance

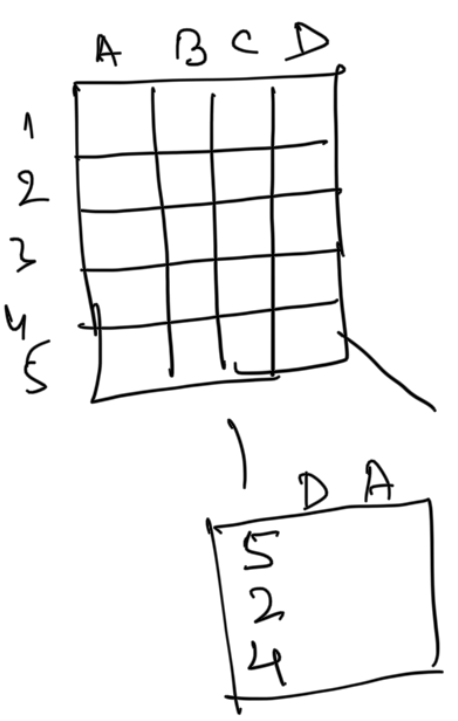
GA



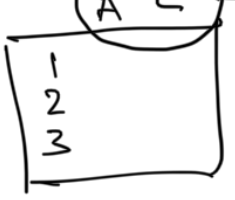
5 members

Random forest.

DT + Bagging



grow the trees,
Not use entropy for identifying best split.
A C → randomly.



Diversity → avoid overfitting

Gradient Boosting

exp	quali	salary	B_1	Res	B_2	Res	B_3	Res
		50	54	-4	-2	-2	-1	-1
		55	54	1	1	0	0	0
		60	56	4	3	1	0	1

MT1

1. Add a Base learner. (B1)
   (exp, quali => salary)
2. Predict and calculate residuals.
3. Add a Base learner (B2)
   (exp, qual => Res)

Regularization rate

$$y = \alpha B_1(x) + \alpha B_2(x) + \alpha B_3(x) + \dots \alpha B_n(x)$$

$n + n/2$

$0 \leq 1$ + + + + + + + + +