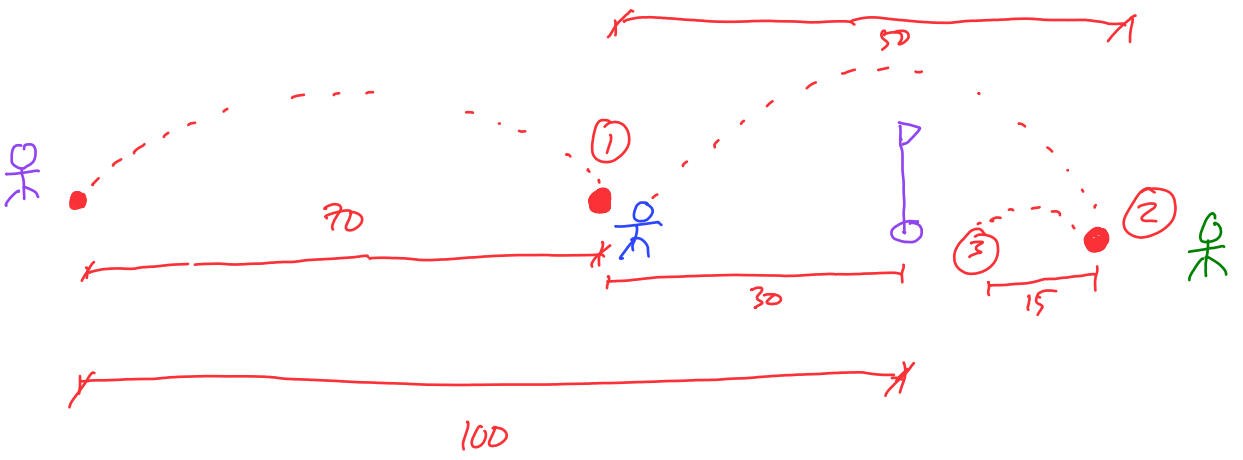
The "key" idea of boosting method

- Train a sequence of weak models where the next model learns from the mistake / error of the current model.

- The final model is a combination of All the model in the sequence



: target : 100

reality : 70 ⇒ error : 100 − 70 = 30

: Target : 30

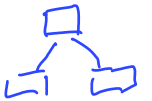
reality : 50 ⇒ error : 30 − 50 = −20

: target : −20 ⇒ error : −20 − (−15) = −5

reality : −15

Data:

	X_1 GPA	X_2 Age	Y (target) Salary
	3.5	60	100
	4.0	29	120
	3.0	49	80

train model 1 (weak model, says a stump 

GPA	Age	Salary	$\hat{Y}_1$ (Model 1 predicts)	ε_1 Model 1's error
3.5	60	100	90	$100 - 90 = 10$
4.0	29	120	150	$120 - 150 = -30$
3.0	49	80	70	$80 - 70 = 10$

train the second model, another stump  on the data where the new target is ε_1

GPA	Age	target ε_1	Model 2 predicts $\hat{\varepsilon}_1$	$\hat{\varepsilon}_2$
3.5	60	10	5	5
4.0	29	-30	-20	-10
3.0	49	10	15	-5

we will fit (third model) on the dataset where the target is ε_2 (error of the second model)

GPA	Age	ε_2 (target)	$\hat{\varepsilon}_2$ (model 3 predicts)
3.5	60	5	5
4.0	29	-10	-8
3.0	49	-5	-6

$\Rightarrow$ we combine the final predictions of the 3 models as:

H = Predictions of Model1 + Predictions of M2 + Pred. of M3

$= \hat{Y}_1 + \hat{\varepsilon}_1 + \hat{\varepsilon}_2$

GPA	Age	Salary (Y)	$\hat{Y}_1$ (Model 1 predicts)	$\hat{\varepsilon}_1$	$\hat{\varepsilon}_2$	Final pred
3.5	60	100	90	5	5	100
4.0	29	120	150	-20	-8	122
3.0	49	80	70	15	-6	79