Example

	age	initial guess	residual
obs 1	21	22	-1
obs 2	22	22	0
obs 3	23	22	1
obs 4	22	22	0
obs 5	21	22	-1

actual value = predicted + residual

what if we had a tree that could predict the residuals made by the initial model?

$\Longrightarrow$ **gradient boosting tree!**

*instead of fitting a bunch of independent trees,
we incrementally improve on our initial guess*

0.01 + 0.01 + 0.01 + $\cdots$ + 0.01
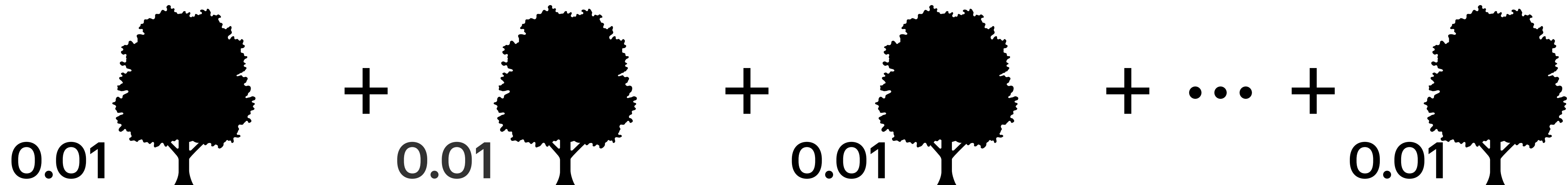
Gradient Boosting Trees

Intuitively:

- Gradient boosting is a method for iteratively building a complex model T by adding simple models.

- Each new simple model added to the ensemble compensates for the weaknesses of the current ensemble:

  ‣ each simple model $T^{(i)}$ we add to our ensemble model T, models the errors of T

  ‣ Thus, with each addition of $T^{(i)}$, the residual is reduced