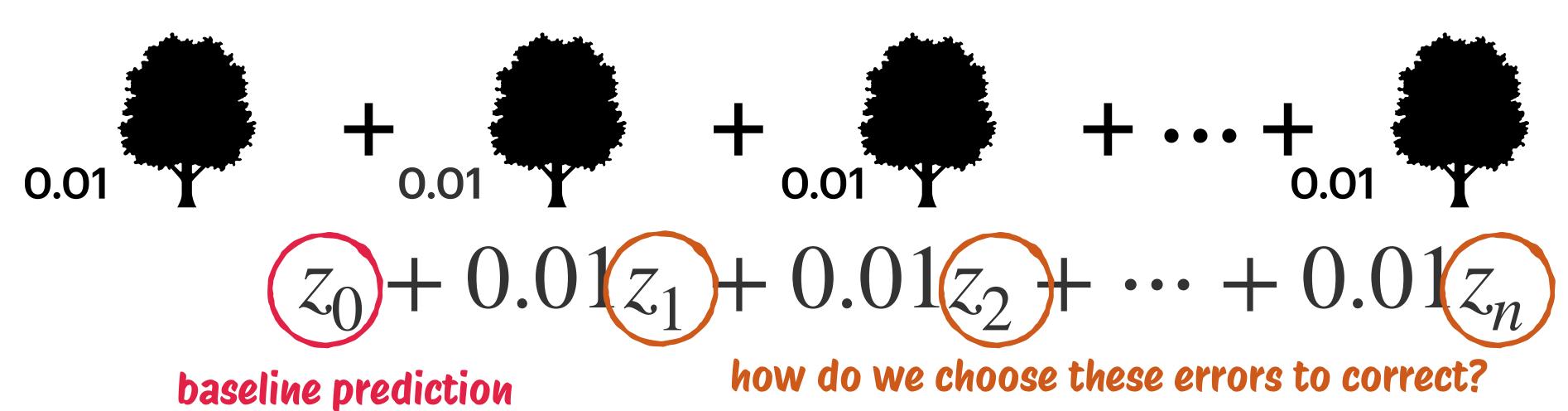
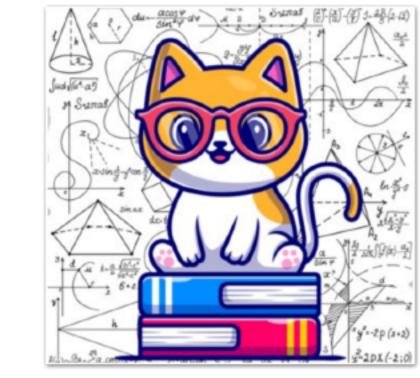
Gradient Boosting Trees: The Algorithm

- 1. Fit a simple model $T^{(0)}$ on the training data $\{(x_1, y_1), \ldots, (x_N, y_N)\}$ Set $T \leftarrow T^{(0)}$ and compute residuals $\{r_1, \ldots, r_N\}$ fort T
- 2. Fit a simple model $T^{(1)}$ to the current **residuals**, i.e. train using $\{(x_1, r_1), \ldots, (x_N, r_N)\}$
- 3. Set $T \leftarrow T + \lambda T^{(1)}$ where λ is the learning rate (usually 0.01 or 0.001)
- 4. Compute residuals, set $r_n \leftarrow r_n \lambda T^{(i)}(x_n), n = 1, ..., N$
- 5. Repeat steps 2-4 until stopping condition is met

Gradient Boosting Trees: The Math





future trees predict error for a regression tree given defined loss function

let
$$F_i$$
 be our predictions $F_i = \sum_{t=0}^l z_t$
$$F_1 = z_0 + z_1$$

$$F_2 = z_0 + z_1 + z_3$$

$$\vdots$$

$$F_i = F_{i-1} + (z_i)$$

$$z_{i} = -\frac{\partial Loss(y, F_{i})}{\partial F_{i}}$$