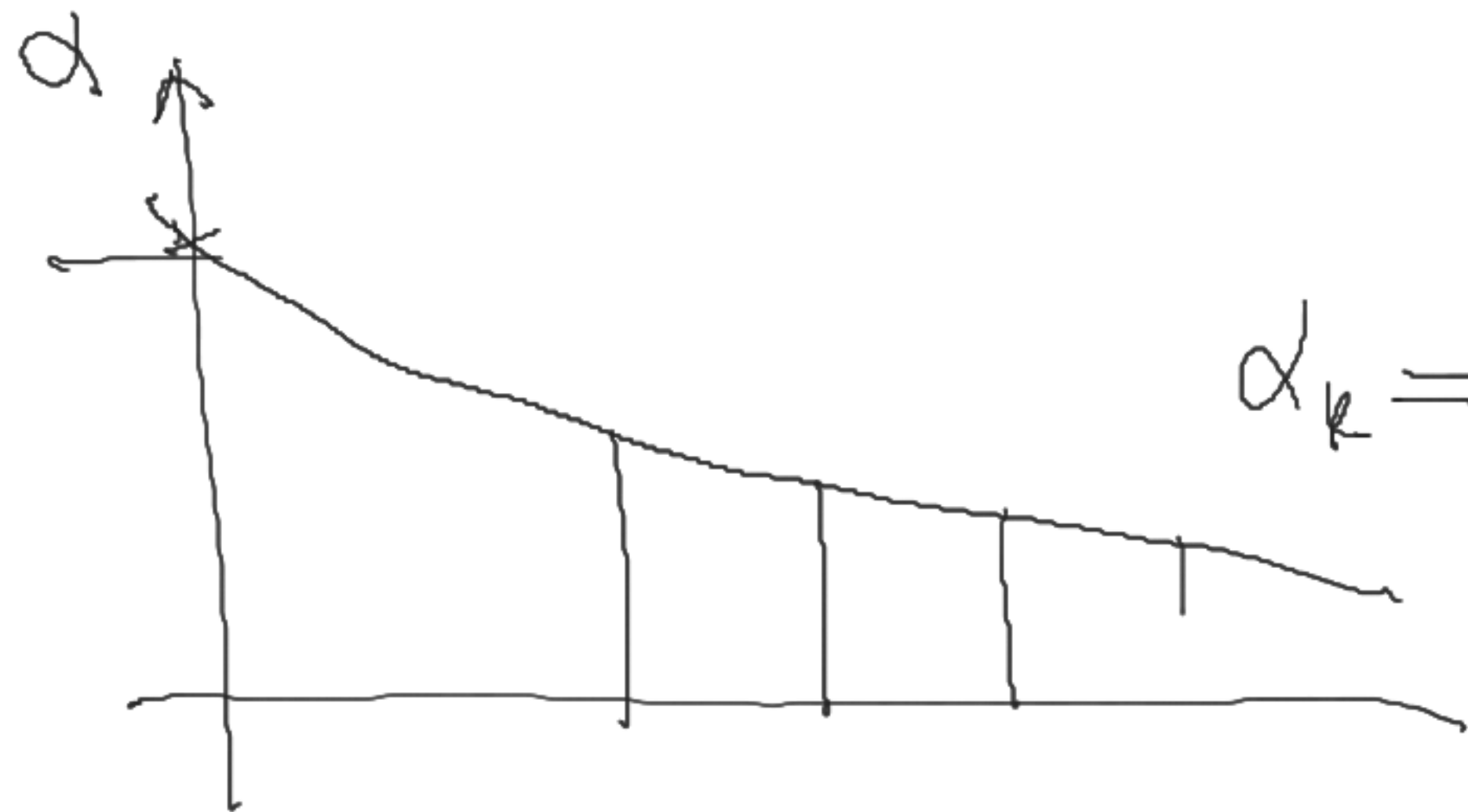


input_size
||
(784,)

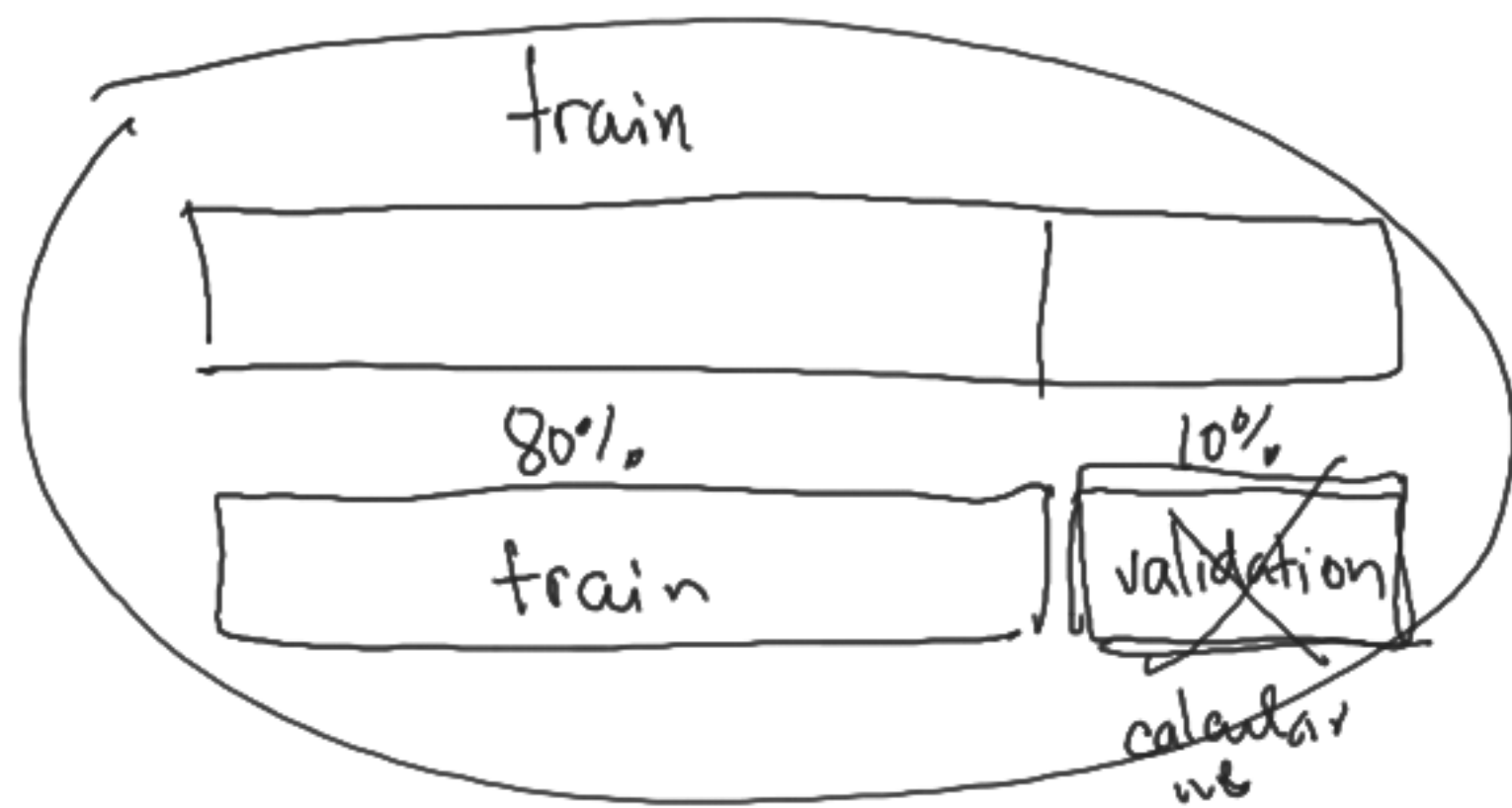
~~b~~
~~60,000~~ x 784

Minibatch GD



$$\alpha_k = \alpha_0 e^{-\text{decay}}$$

iter



desempenho.

