

T-split

α -train

$$\left(\frac{0.85 \times 80,000}{125} \right) \times 10^4 \text{ Epochs}$$

Each Epoch

$$3400$$

$$\begin{aligned} \text{Epoes} &= 5 \\ \text{Days} &= 3 \end{aligned}$$

$$R(c, \text{Day}=4) = 0 \rightarrow \text{stop criterion.}$$

States & Days, cakes.

We need to start with last day as it is near the stop criterion.

↳ Each state may will compare of all possi values of cakes.

↳ Then we need to chk for all choices at that state.

$$R(c, d) = 0.7 d^{-1} \times \sqrt{c}$$

For Dynamics can be constructed from the stop-decision

$$b, T = 4$$

$$R(c, t) = \int [R(x, t) + R(c-x, t+1)]$$

Max Value for reward at $x \in [0, c]$

negative part of $Q = (x, y, z, 0, 0)$

18

good deal after hours of work all
- operations of work at first
designed from stop state
- could be called 'never' the go

Model also at hour or half of
state that to decide if

$$-22.5 \times 2.4 \times 7.0 + (6.0 \times 2.8)$$