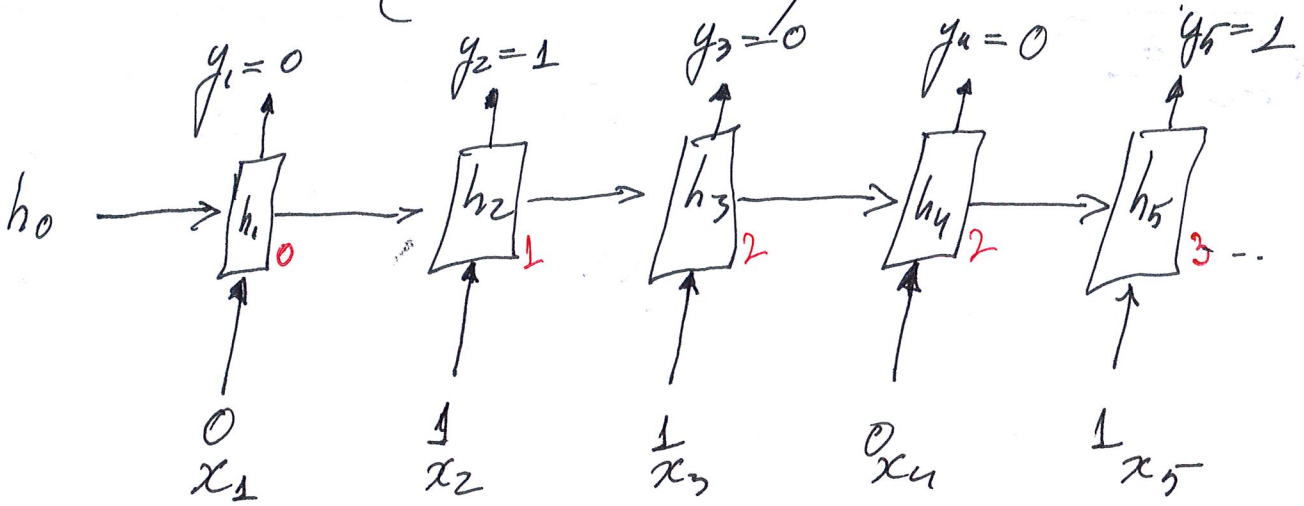
$$P(\underline{\text{laptops}} \mid \text{the students opened their}) \approx \frac{C(\text{the students opened their laptops})}{C(\text{the students opened their})}$$



$$h_t = f(h_{t-1}, x_t) = h_{t-1} + x_t$$

$$y_t = g(h_t) = h_t \,\%\, 2$$

$$h_t = f_W(h_{t-1}, x_t) = f(W_{hh} h_{t-1} + W_{hx} x_t)$$

$$I\, h_{t-1} + I\, x_t$$

$$p(C_1|x) = \frac{e^{z_1/T}}{e^{z_1/T} + e^{z_2/T}}$$

$$p(C_2|x) = \frac{e^{z_2/T}}{e^{z_1/T} + e^{z_2/T}}$$

$z_1 > z_2$

→ sampling

$T = 1 \Rightarrow$ normal softmax

$T \to \infty \Rightarrow p(C_1|x) = 0.5$

$p(C_2|x) = 0.5$

$T \to 0 \Rightarrow p(C_1|x) = 1.0$

$p(C_2|x) = 0.0$

uniform sampling

greedy sampling

auto-regressive generation