. Design an RNN arch. for the
Language modeling task.

$$P(w_i \mid w_1 \cdots w_{i-1})$$

Generate text using this RNN

$y_{ni}$

$h_i$

$w_i$

# Auto regressive Generation



$w_2$

$y_1$

cat

$a$

$h_1$ $\rightarrow$ $h_2$ $\rightarrow$ $h_3$
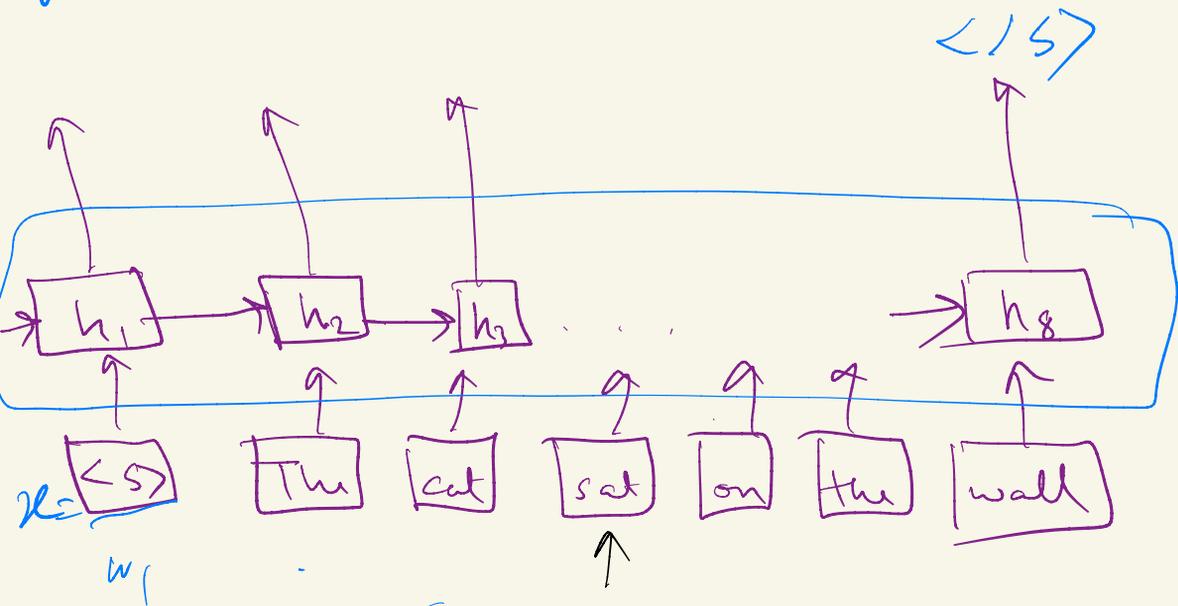
$h_0$

$\langle s \rangle$

The

cat

$w_1$

$w_2 = The$

$w_3 = cat$
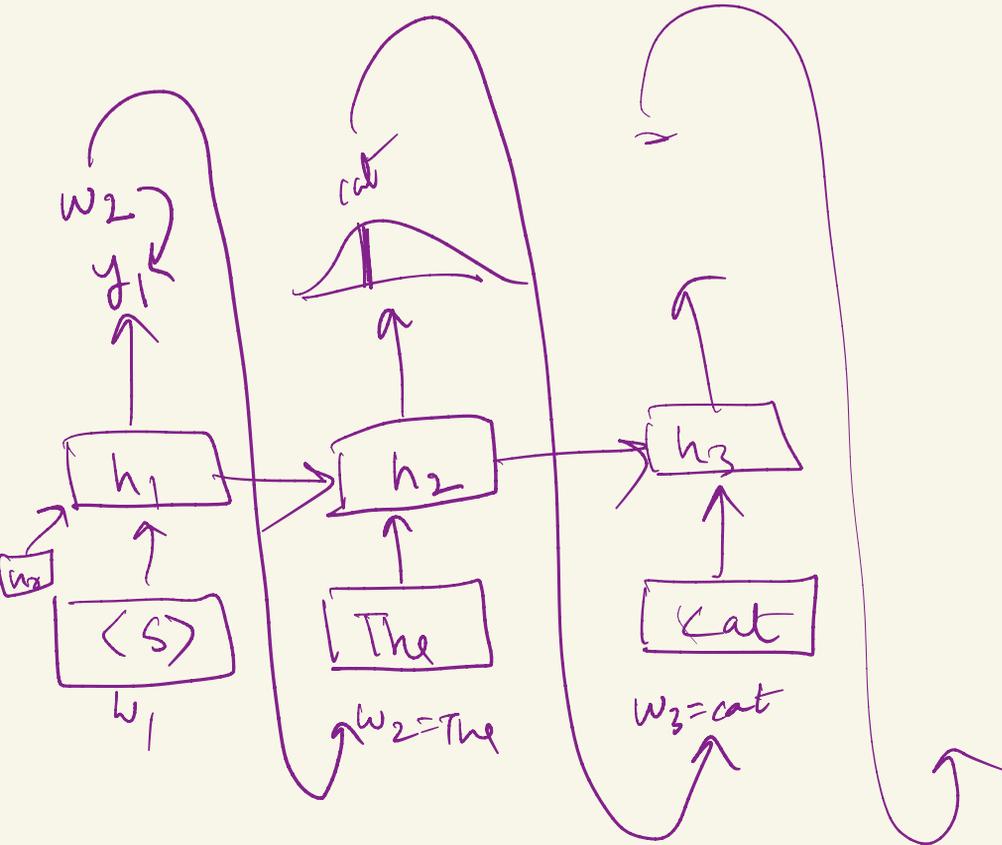
# Train an RNN

gold output at step $t = w_{t+1}$

$L_1 = -\log P(\text{The} | \cdots)$   $L_3 = -\log P(\text{sat} | \cdots)$   " teacher

$L_2 = -\log P(\text{cat} | \cdots)$   forcing "

$y_1^* = \text{The}$   cat   (sat)   —   –   —

$$y_2^d = sat$$
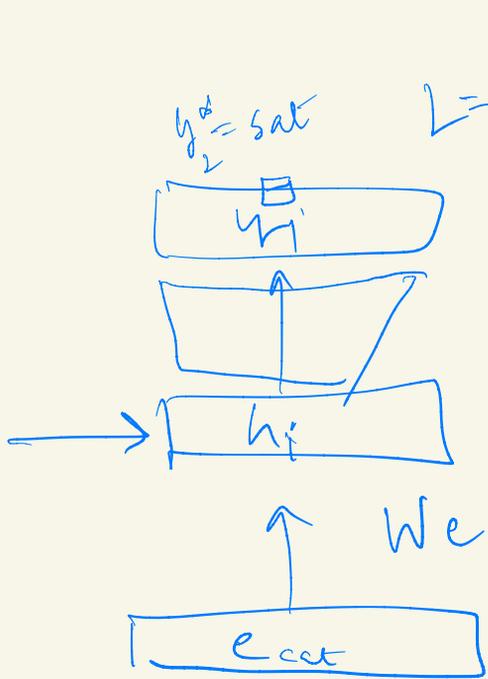
$$L = -\log P(sat \mid ---)$$

$$y_i = softmax(V h_i)$$

$$h_i = f\left(U h_{i-1} + W_e\right)$$

$$h_i$$

$$W_e$$

$$e_{cat}$$

on had high prob.
under the RNN
$p(on \mid the\ cat\ sat)$

$y_1$  $y_2$  $y_3$  $y_n$  $y$

$h_1$  $h_2$  $h_3$  $h_n$  $h_5$

$<s>$  The  cat  sat  on

On

$p$  sofa  wall

$y$

$h_1$  $h_2$  $b_2$  $h_n$

$<s>$  The  Cat  Sat

on