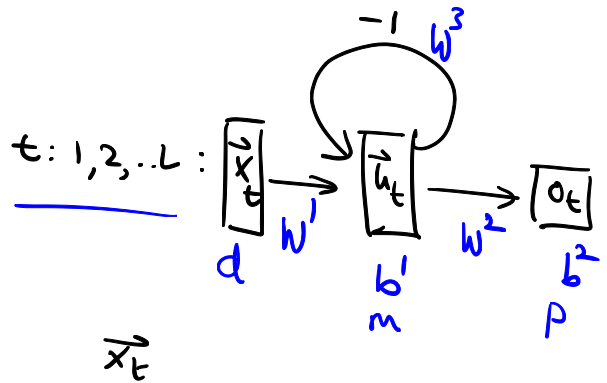
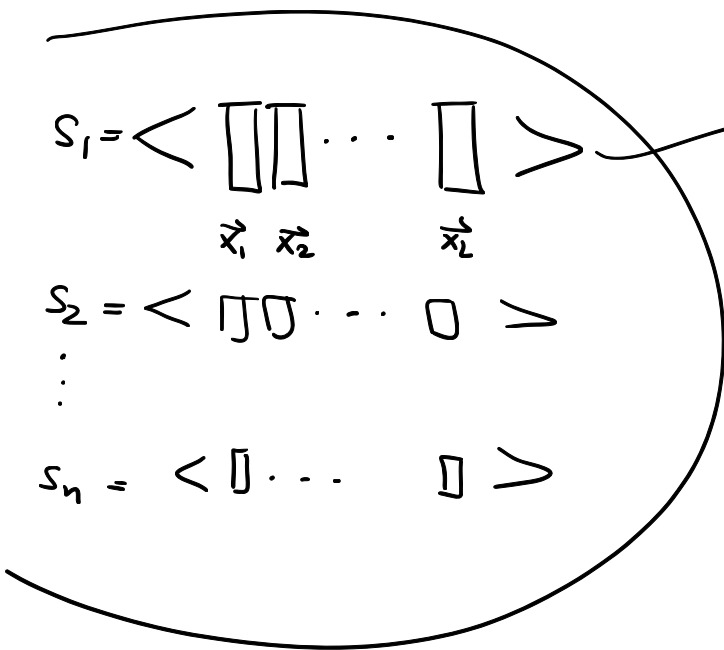


# RNN

→ Recurrent NN



$\vec{net}_{h_t} \equiv$  weighted sum of all inputs

$$= (W^1)^T \vec{x}_t$$

Data

n sequences

L : length of each seq

$b^1$  : hidden bias

$b^2$  : out

$W^1$  :  $d \times m$  - input to hidden

$W^2$  :  $m \times p$  - hidden to output

$W^3$  :  $m \times m$  - hidden to hidden

parameter sharing

















