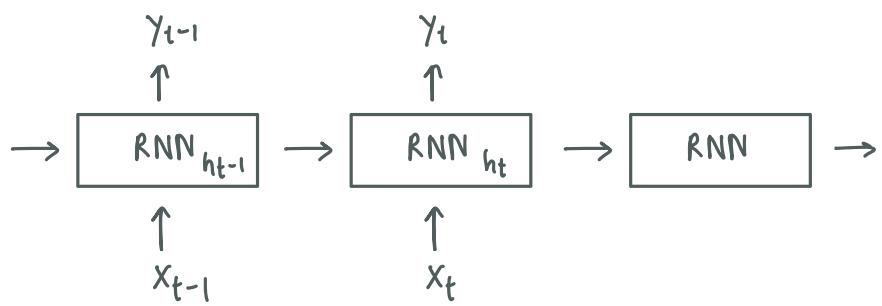
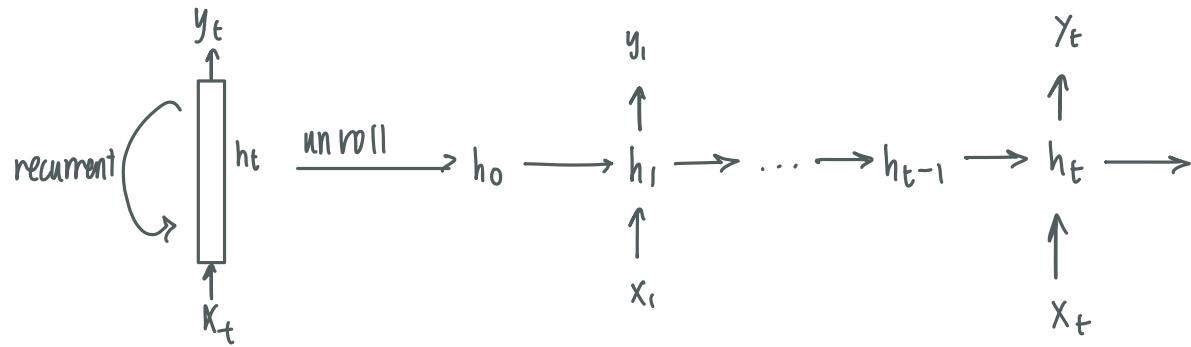


next HW assigned today for PyTorch.

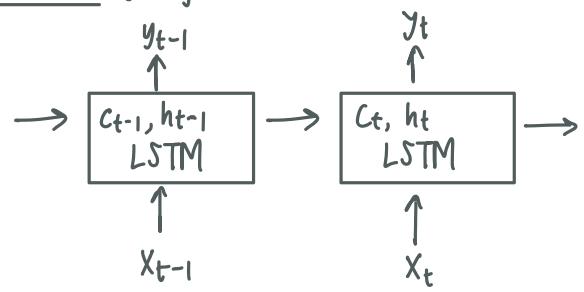
- code for CNN

- compute on Google Colab for GPU

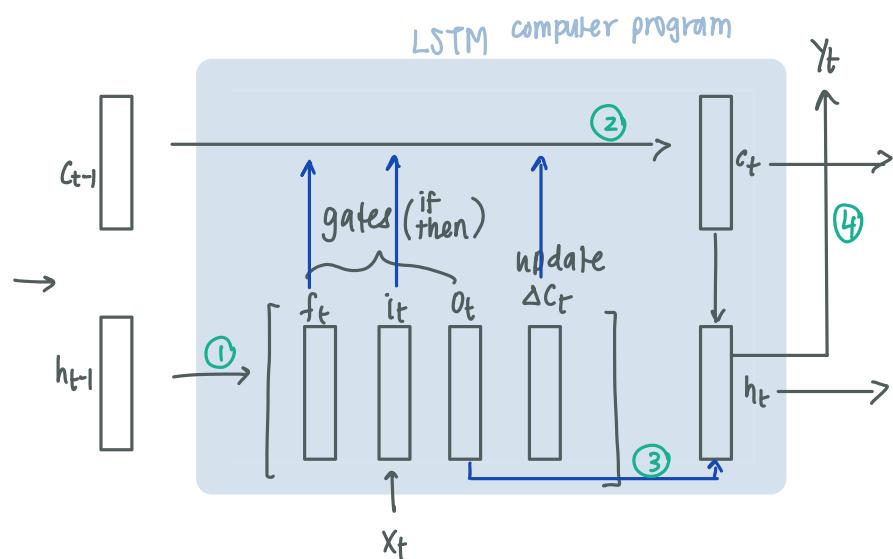
RNN



LSTM (long short term memory)



c stands for (memory) cells



①

$$\left[\underbrace{f_t, i_t, o_t}_{\text{sigmoid}} \underbrace{\Delta c_t}_{\tanh} \right] = \text{Rectify } \begin{matrix} W \\ \downarrow \\ \text{one for each } f_t, i_t, o_t, \Delta c_t \end{matrix}$$

Rectify

$$\begin{matrix} W \\ \downarrow \\ \text{one for each } f_t, i_t, o_t, \Delta c_t \end{matrix}$$

$$\begin{matrix} W_h & W_x \\ \downarrow \\ h_{t-1} \\ x_t \end{matrix}$$

② $c_t = f_t \odot c_{t-1} + i_t \odot \Delta c_f$

③ $h_t = \tanh(o_t \odot c_t)$

④ $y_t = \text{Rectify}(W h_t)$

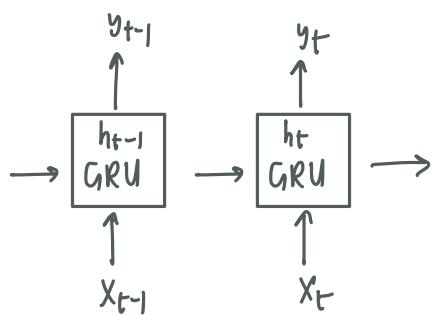
Gated Recurrent Unit (GRU)

① $r_t, z_t = \text{sigmoid}(W \begin{matrix} w_h \\ w_x \end{matrix} h_{t-1})$

↙
one for r_t
one for z_t

remember or not?
↑
 $r_t \odot h_{t-1}$

overall global picture:



h_{t-1} controls initial generation of operations
care must abt h_t

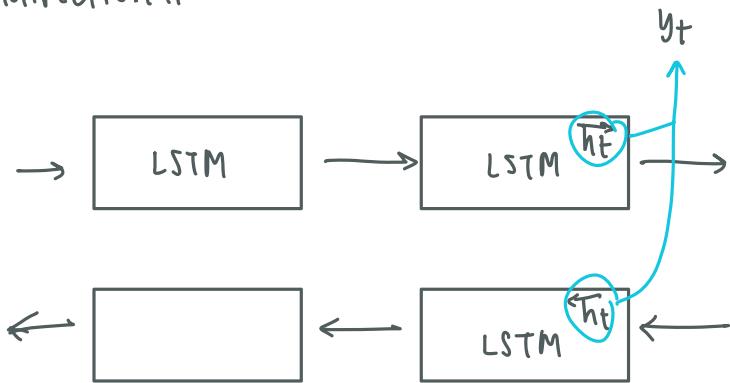
② $\tilde{h}_t = \tanh(W \begin{matrix} w_h \\ w_x \end{matrix} (r_t \odot h_{t-1}) + x_t)$

③ $h_t = z_t \odot \tilde{h}_t + (1 - z_t) \odot h_{t-1}$

↓
replace or not?

④ $y_t = \text{Rectify}(W h_t)$

Bidirectional

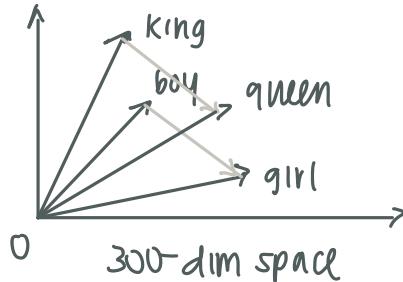


Natural language processing (NLP)

`words` \in `dictionary` budo words

i.e. king queen
boy girl

tokens 6 50,000



arithmetic
queen = king + girl - boy

The diagram consists of two rectangles. A larger rectangle on the left is labeled 'w' at the top center. To its right is a smaller, vertically oriented rectangle labeled 'dr' at the bottom left corner. A horizontal line segment connects the right side of the large rectangle to the left side of the small rectangle.

a word =  (as in classification)

dense \leftarrow (any element can be non-zero)
distributed \leftarrow (distributed representation)

embedding

12
- - -
300

W
300x6000

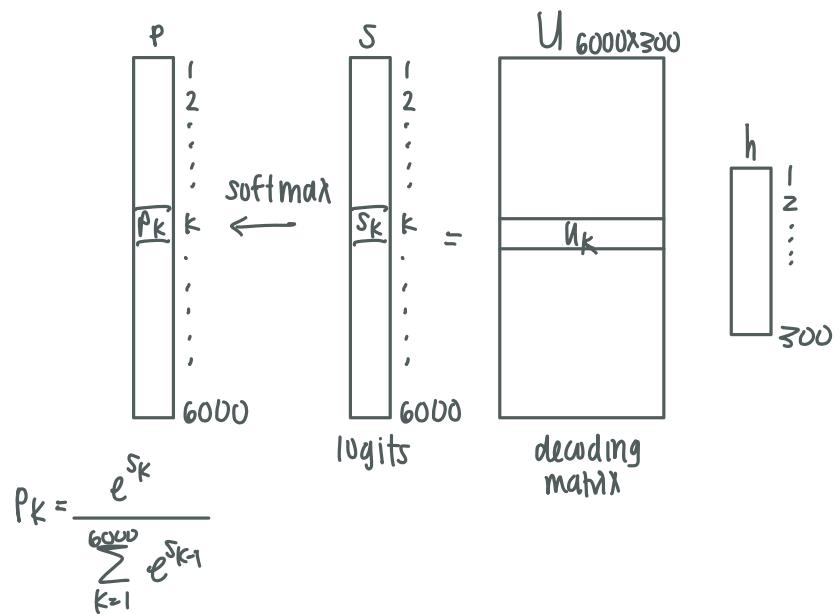
w_1	w_k	w_{kk}			
-------	-------	----------	--	--	--

encoding matrix 300

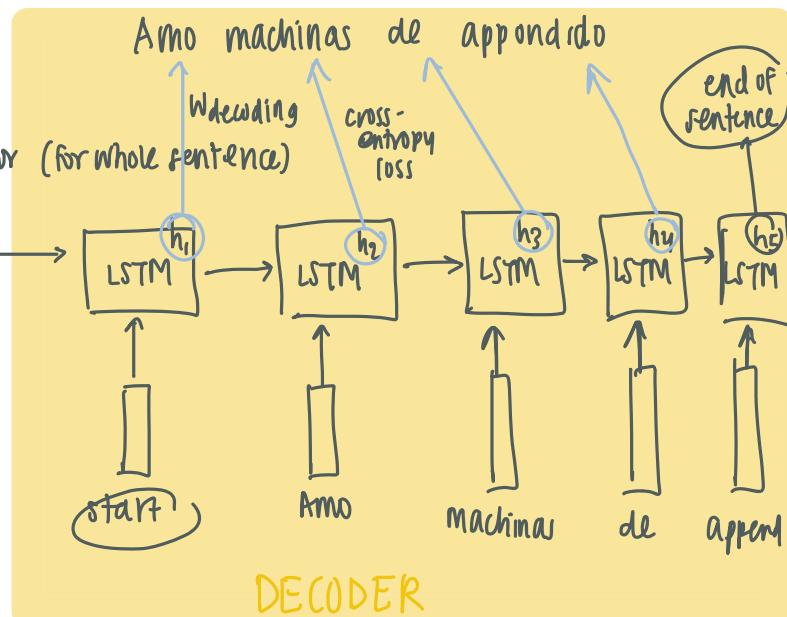
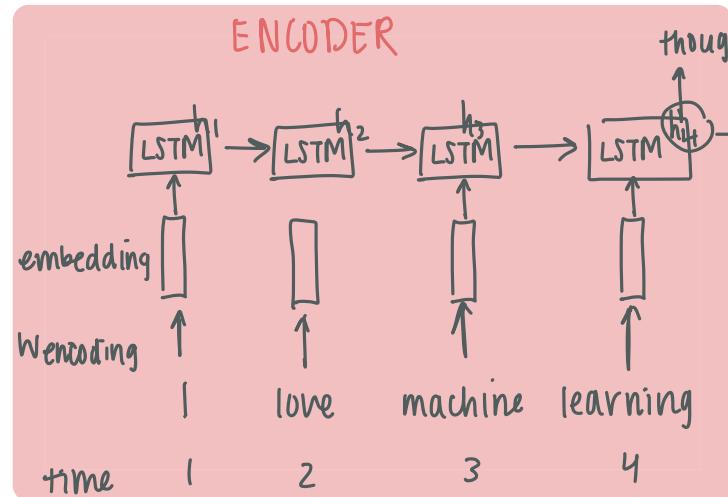
$$= \boxed{} \quad \text{encoding}$$

1
2
3
4
5
6
7
8
9
0

Decoding



Translation



Auto-regressive