Si Zhang
sz324

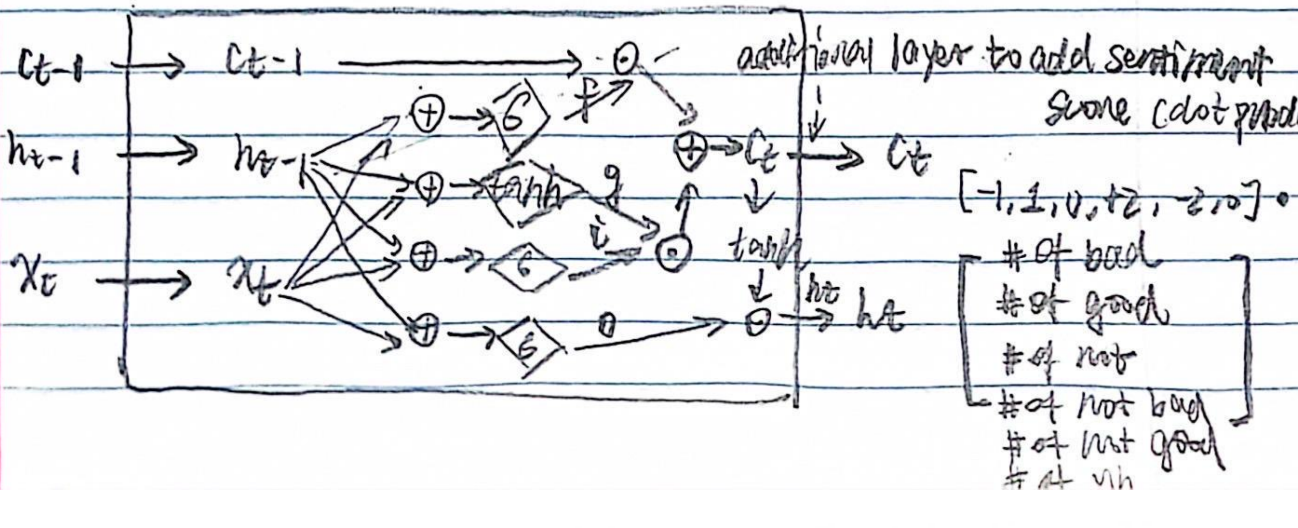
NLP LSTM

1. RNN: $h_t = \sigma(W_{ih}x_t + b_{ih} + W_{hh}h_{t-1} + b_{hh})$

LSTM:
$$i_t = \sigma(W_{ii}x_t + b_{ii} + W_{hi}h_{t-1} + b_{hi})$$
$$f_t = \sigma(W_{if}x_t + b_{if} + W_{hf}h_{t-1} + b_{hf})$$
$$g_t = \tanh(W_{ig}x_t + b_{ig} + W_{hg}h_{t-1} + b_{hg})$$
$$o_t = \sigma(W_{io}x_t + b_{io} + W_{ho}h_{t-1} + b_{ho})$$
$$c_t = f_t \odot c_{t-1} + i_t \odot g_t$$
$$h_t = o_t \odot \tanh(c_t)$$

assumes weights for i_t, f_t and g_t have weights of 0 and very large bias. (i.e. $W_{ii}=0$, $W_{hi}=0$, $W_{if}=0$, $W_{hf}=0$, $W_{ig}=0$, $W_{hg}=0$)
Under this assumption, $i_t \to 1$, $f_t \to 1$, $g_t \to 1$.
$$c_t = 1 \odot c_{t-1} + 1 \odot 1 = c_{t-1} + 1$$
assume C_0 is close to infinity, C_t would be close to infinity as well.
in this case, $\tanh(c_t)$ would be close to 1. ($C_t \to 1$)
$h_t = o_t = \sigma(W_{io}x_t + b_{io} + W_{ho}h_{t-1} + b_{ho})$
make $W_{io} = W_{ih}$, $W_{hh} = W_{ho}$, $b_{ih} = b_{io}$, $b_{hh} = b_{ho}$, then h_t of LSTM under these assumptions would equal to RNN in a very simple form.

2. The input size of the LSTM is 6. ($M=6$), hidden state size could be varied dependent on the training model, assume size of 8 for now. ($K=8$), output size is 1 ($L=1$), which is cumulative sentiment score.



as shown in the graph. I add an additional layer on top of C_t to include the sentiment score calculation. (since h_t is defined by tanh and has range of $(-1,1)$, not situable for sentiment score calculation).

similar to question #1. I assume all the weights of i_t, f_t, g_t and O_t to be 0, bias to be a large number to allow the full activation of ~~under~~ input gate, forget gate, and output gates.

under this assumption, $i_t \to 1$, $f_t \to 1$, $g_t \to 1$, $O_t \to 1$.

$C_t = C_{t-1} + 1$

assume $C_0 = 0$. at time of 0.

then add additional layer defined as Z_t to C_t. ($Z_t = W_{iz} \cdot X_t + bias$)

$\therefore C_t = C_{t-1} + 1 + W_{iz} \cdot X_t + bias$.

Note: $W_{iz} = [-1, 1, 0, +2, -2n]$

$$x_t = \begin{bmatrix} \text{\# of bad} \\ \text{\# of good} \\ \text{\# of not bad} \\ \text{\# of not good} \\ \text{\# of n} \end{bmatrix}$$ (in alphabetical order).

bias set to 0.