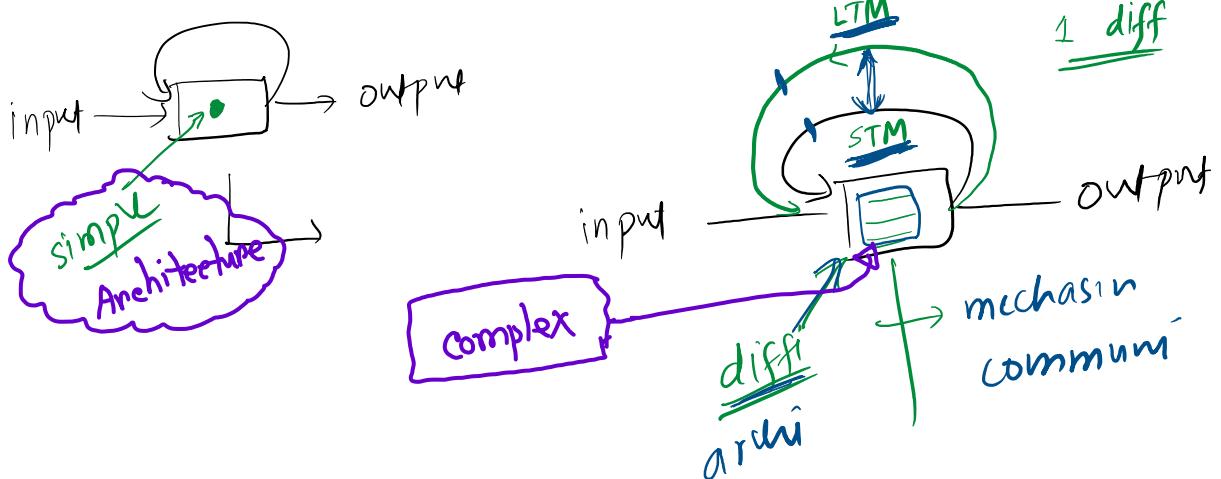
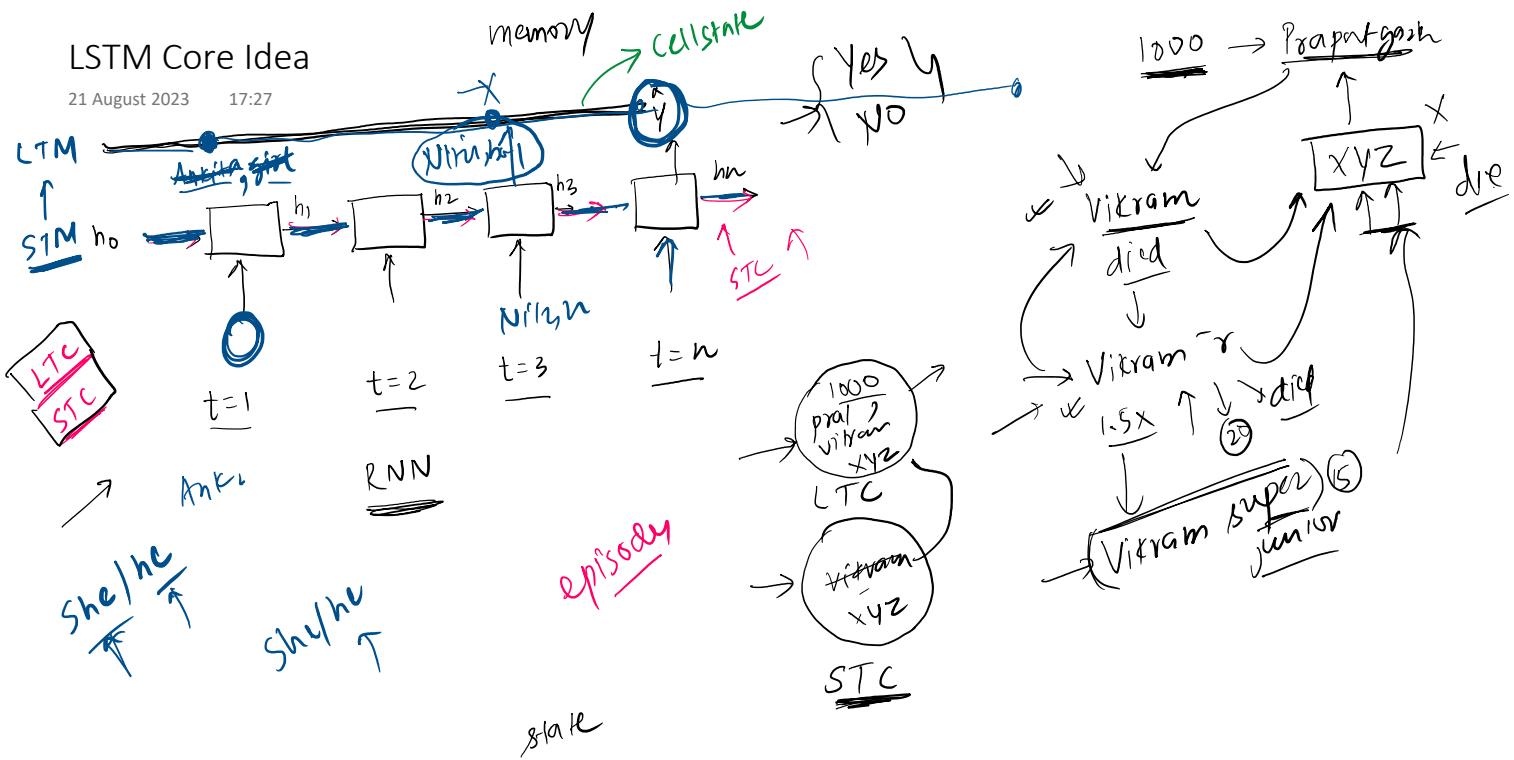


Leeture #61

LSTM Core Idea

21 August 2023 17:27



↪ हाजार वर्ष आगे एकटा पुस्तक कारिबी। एकटा गृह्य चिन प्रत्यक्ष याद फूल हिन विक्रम। Vikram एक बहुत powerful and kind व्यक्ति हिन। पाठेवार व्यक्ति एकटा फूल XYZ एकदिन प्रत्यक्ष आणवार कर्हे। Vikram याकै शिळ याच्या ग्रन्ति तिनि याकै माझा याच डारू ठेवून Vikram एवं काढकाढि फूल फूल हड्ये आणि प्रतिक्षार्थ निपोर्याच्या घावात XYZ एवं काढ माझी फूल। Vikram एवं प्रत्येकोटा हिन Vikram super junior फूल XYZ के मिळेले।

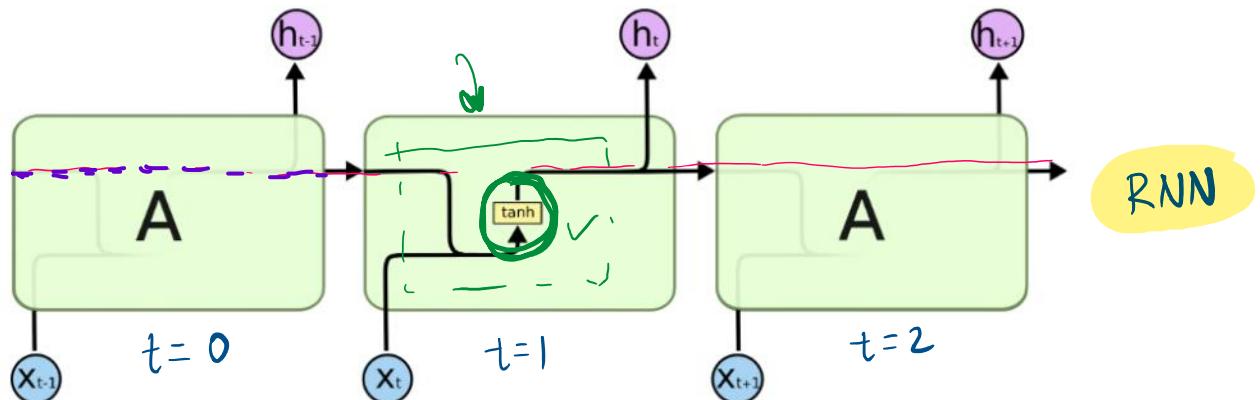
→ गोप्ती कैमन सामग्री?

गोप्तीकृत judge कृपाशमार्य आवडू राहिल्या maintain करूचि, Short term context & long term context. Short term context एवं इटना प्रवाहले judge करूचि आवू long term context असागरे अंतर्गत याकै store कर्हे short term context एवं basis अ long term ए किंवा add वा remove राहेगे गोप्ती judge करूचि। RNN ए प्रत्येक फूलात पाणी, LSTM ए LTM (long term memory), STM (short term memory) थारके। RNN ए कृती STM थारता।

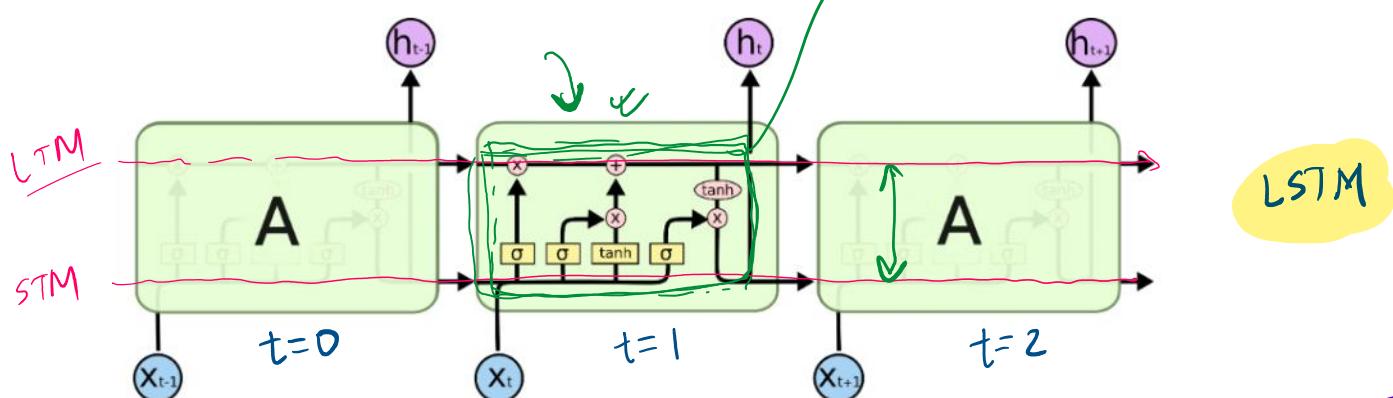
LSTM Architecture

21 August 2023 18:41

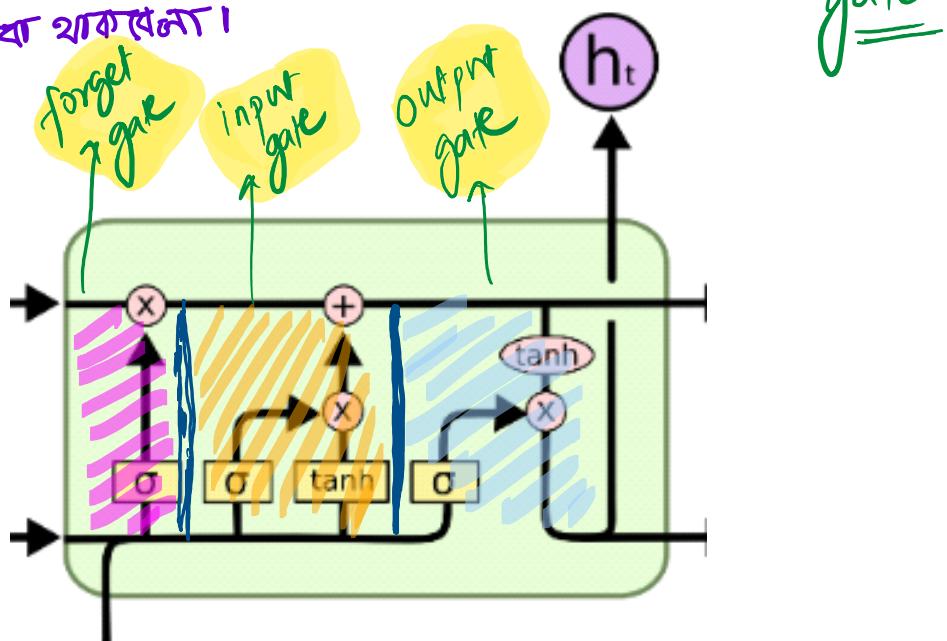
RNN Maintaining only one state.



(LSTM Maintaining two states, LTM, STM)



LSTM ୟ ୩ ସିରନ୍ତେ gate ଥାଏଣୁ ୩ଟେ color ଦିଲ୍ଲେ ଆଶାନ୍ତା କରା (ନିଯମ ଉପରେ) ଏହି
gate ଗୁଣ ଦିଲ୍ଲେ STM, LTM ନିଜଦେଖି ମଧ୍ୟେ communicate କରେବି ଯିବାର ଲୋକେ ଏହିବେ
LTM ଓ କଥା ଏଥାକାରିତା !





Forget Gate: Forget gate current input x_t basis ϕ -গোন context

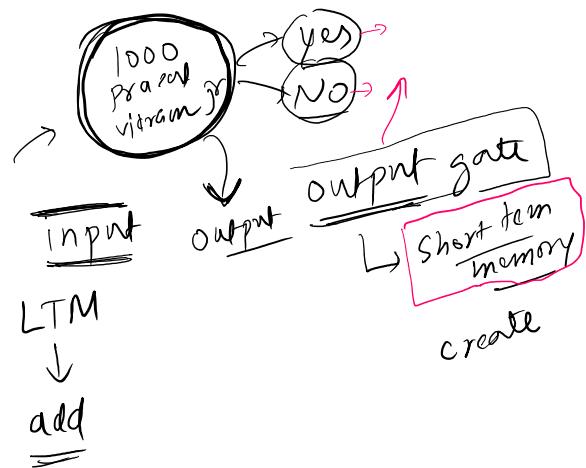
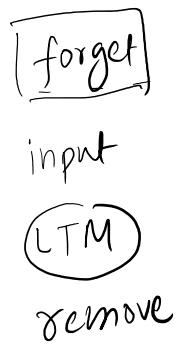
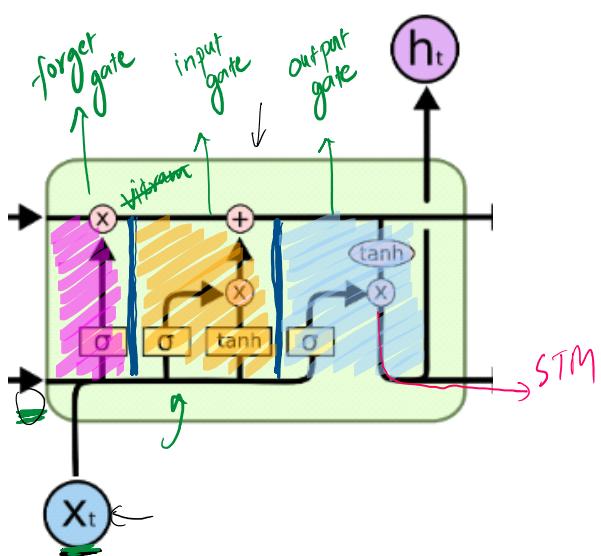
LTM করে remove করে। তুরুতে সামান্য স্টেটিলাম গতির নামুক হবে
vikram এর পর নাম বিক্রম vikram কিছি এর পর নাম বিক্রম জো পর নাম বিক্রম
super এর (এই নাম আশল নামুক)। Forget gate তুরুতে vikram নাম ও পর
vikram জো প্রস্তুত LTS দ্বারা remove করে দেয়।

input gate: Forget gate - vikram জো একটা প্রস্তুত বিক্রম super
হা কে LTM তে যুক্ত করেছে।

output gate: Long term memory এর basis উ output করী হবে
তা decide করে। আবু short term memory create করে।

LSTM Gates

21 August 2023 19:06

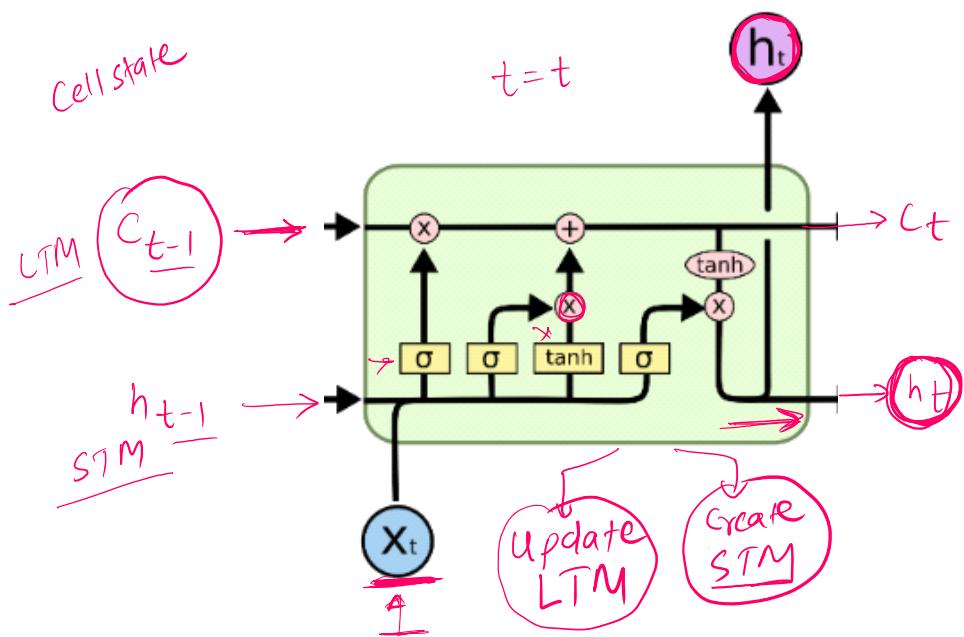


Summary

21 August 2023 19:29

LSTM

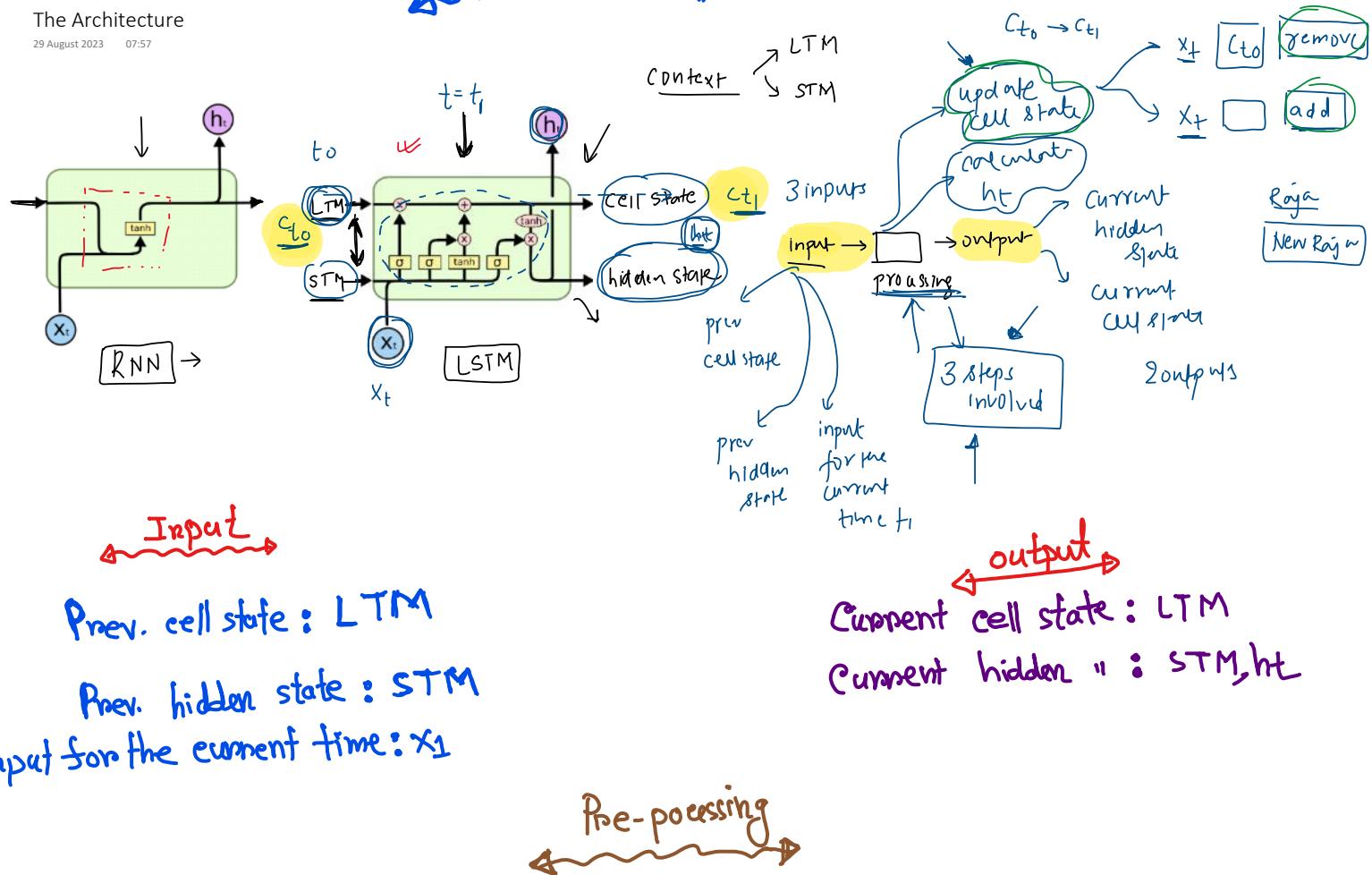
input → $\boxed{\quad}$ → output
↓
process



Lecture-62

The Architecture

29 August 2023 07:57

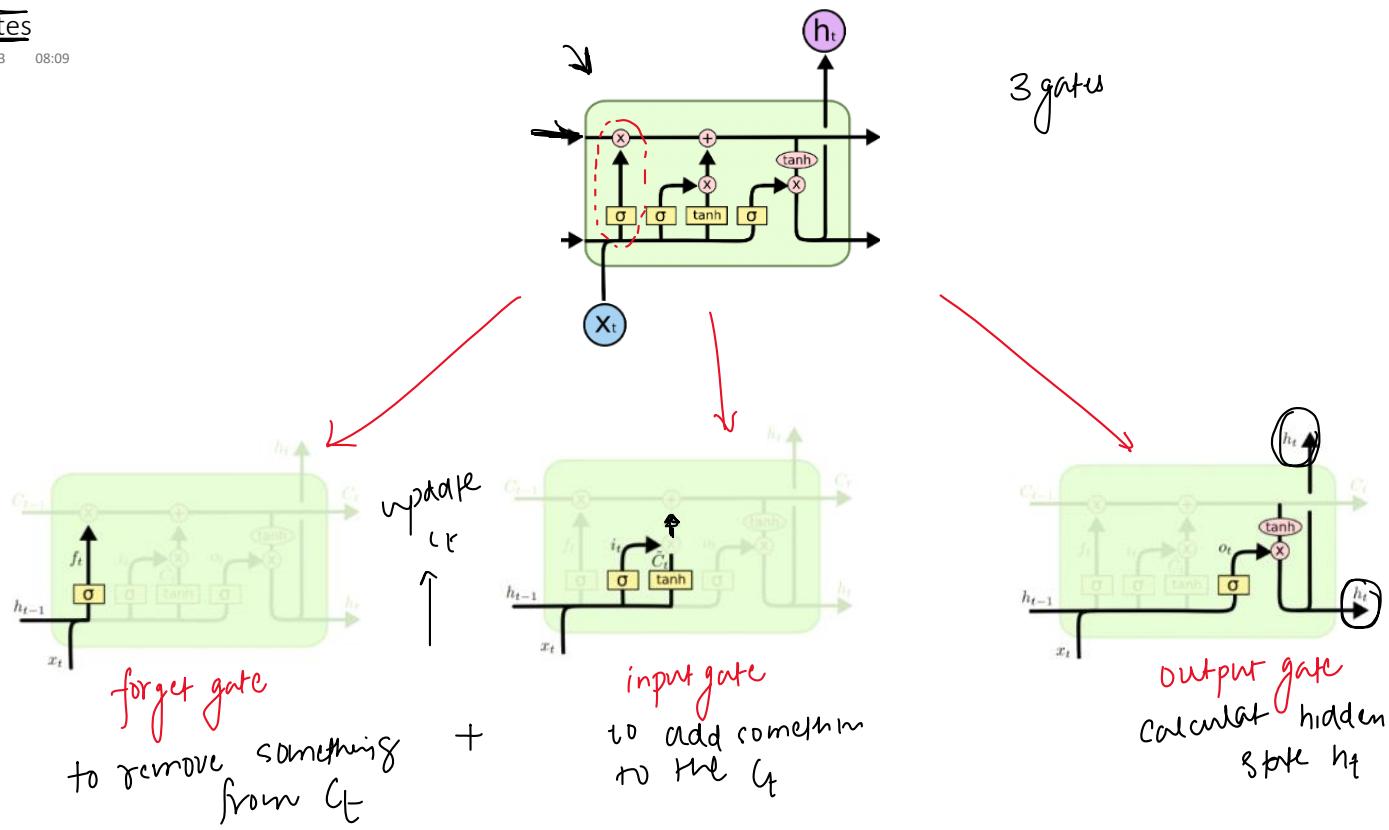


i) Update cell state ($c_{t_0} \sim c_{t_1}$)

ii) Calculate h_t

In update cell state we do two things

- i) Based on current input (x_t) remove something from (c_{t_0})
- ii) " " " " add " " in (c_{t_0})

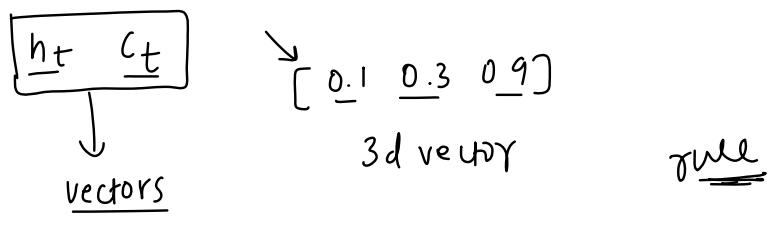
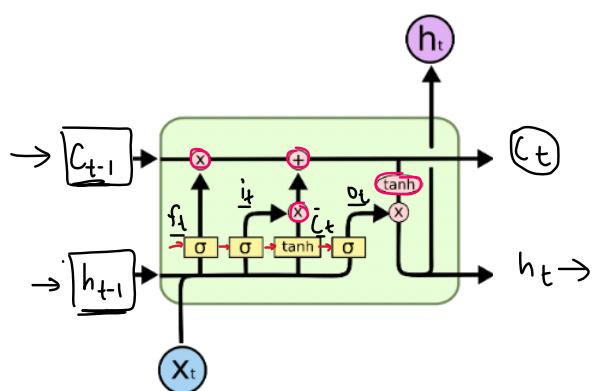


$(\text{forget gate} + \text{input gate}) \approx \text{updating}$

$(\text{Output gate}) \approx \text{calculate hidden state}$

What are C_t and h_t

29 August 2023 08:08



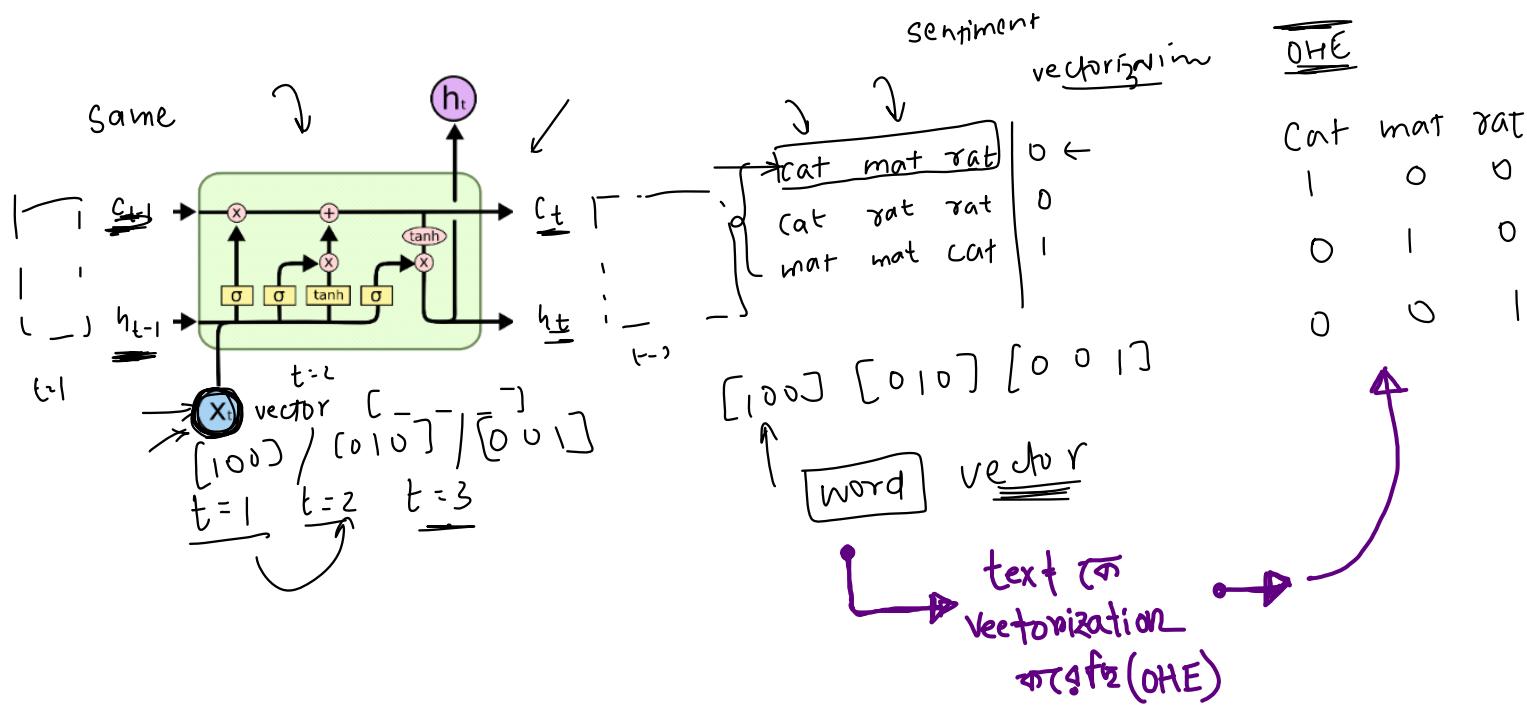
$$\begin{array}{|c|c|} \hline h_t & C_t \\ \hline \end{array} \quad \text{dim equal}$$

$h_t [0.1 \ 0.3 \ 0.9]$
 $C_t [0.55 \ 0.6 \ 0.0]$

Same

→ C_t, h_t रावे वेक्टर

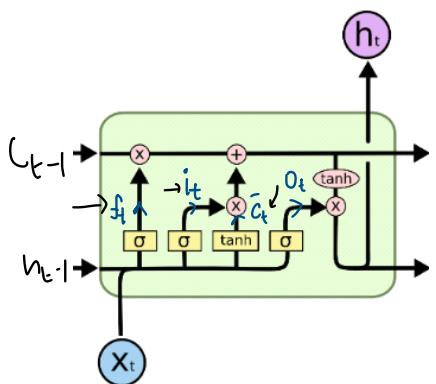
→ असमान C_t and h_t @ dimension एकान्तर हो



** X_t इसके input के t timestep पर
जोकिए word [1 0 0] input हिस्त याहै । **

What are f_t , i_t , o_t and \bar{C}_t

29 August 2023 08:09



The diagram illustrates the structure of a candidate cell state (C_t) at time step t . It consists of three main components:

- Forget gate**: Represented by the vector f_t .
- Input gate**: Represented by the vector i_t .
- Output gate**: Represented by the vector \bar{C}_t .

A bracket labeled "Input" groups i_t and \bar{C}_t . Another bracket labeled "output gate" groups \bar{C}_t and \bar{o}_t . An arrow points from the label "candidate cell state" to the overall structure.

Below the vectors, the word "Vektoren" is written, followed by an arrow pointing to a box containing C_t and h_t . Below the box, the vectors f_t , i_t , \bar{C}_t , and \bar{o}_t are shown again, each underlined.

[x y z]
[]

ଜ୍ଞାନଫୁଟ୍ ଟେ ଗାର୍ ଫଂଗେ ଗେଟ୍ ୯

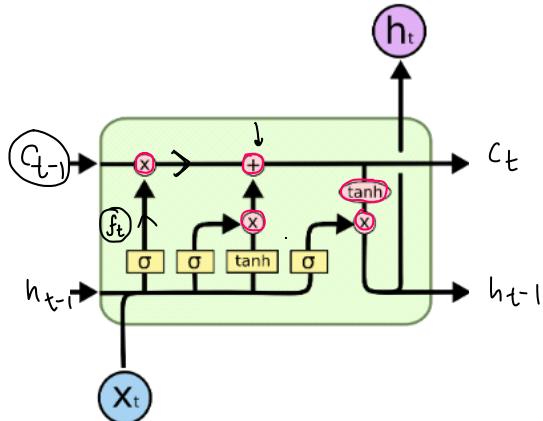
"	i_t	"	input	"	"	
"	C_t	"	"	"	"	(এক Candidate cell state এল)
"	O_t	"	Output	"	"	

ଶବ୍ଦ ଅଧିକାରୀ ମନ୍ତ୍ରୀ ପାଇଁ ଏହାର ଏକ ଅଧିକାରୀ ମନ୍ତ୍ରୀ ପାଇଁ ଏହାର ଏକ ଅଧିକାରୀ ମନ୍ତ୍ରୀ ପାଇଁ

f_t, i_t, \bar{c}_t, o_t ଏହି ।

Pointwise Operations

29 August 2023 18:26



$$\begin{aligned} &\rightarrow \otimes \\ &\rightarrow + \\ &\rightarrow \tanh \end{aligned}$$

$c_{t-1} = \begin{bmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \end{bmatrix}$

$f_t = \begin{bmatrix} 0.26 & 0.34 & 0.53 \end{bmatrix}$

$\text{shape(dim)} \quad \downarrow \quad \text{vector}$

$c_{t-1} \otimes f_t \rightarrow \text{vector} \quad [5 \ 7 \ 9]$

$\rightarrow [4 \ 10 \ 18]$

- \times pointwise multiplication
- $+$ pointwise addition
- \tanh || tanh

$$(c_{t-1} \times f_t) = [4 \ 5 \ 6] \times [1 \ 2 \ 3] = [4 \ 10 \ 18]$$

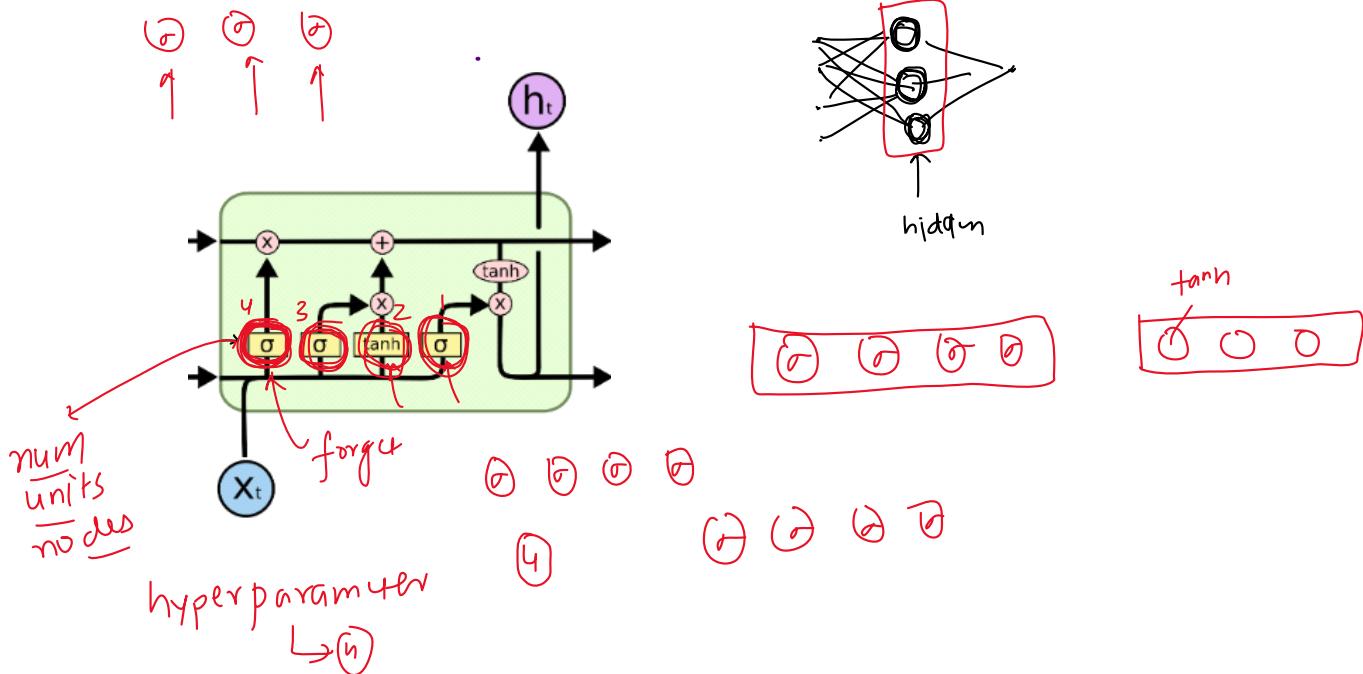
$$(+) = [4 \ 5 \ 6] + [1 \ 2 \ 3] = [5 \ 7 \ 9]$$

$$\tanh(x) = \tanh [4 \ 5 \ 6] = [0.1 \ 0.3 \ 0.9]$$

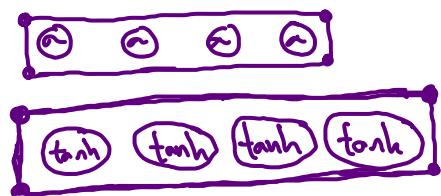
$$\frac{1}{1+e^{-x}}$$

→ Neural Network Layers

29 August 2023 18:34



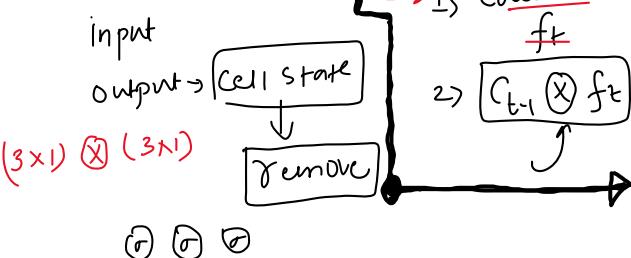
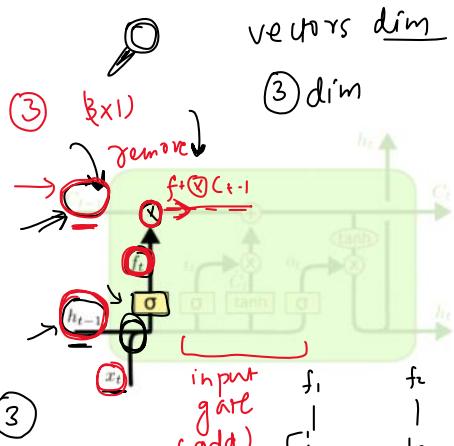
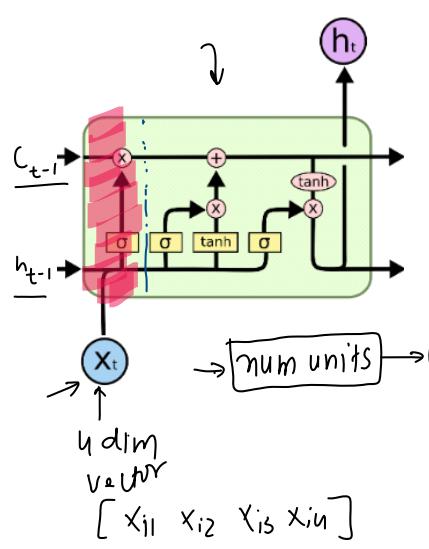
- ④ Yellow box शुल्क रखने वाला नेटवर्क जैसा है वह नेटवर्क जैसा है। आप यह नेटवर्क को " " के लिए node जांचना सकते हैं। यदि node जांचना चाहिए तो क्या करें?
- ⑤ Neural Network के node को शुल्क रखने के लिए आमतर बहुप्रयोगी है। आप यह शुल्क " " के लिए node जांचना सकते हैं। यदि node जांचना चाहिए तो क्या करें?



आप यह node के लिए निम्नलिखित $h_t, e_t, f_t, i_t, \bar{c}_t, o_t$ जैसे डिमेशन रख सकते हैं (4×1)

The Forget Gate

29 August 2023 19:58



$$f_t \rightarrow [f_1 \ f_2 \ f_3] \rightarrow f_t(3)$$

$$f_1 \quad b_1 \quad f_2 \quad b_2 \quad f_3 \quad b_3$$

$$(3 \times 1) \quad (3 \times 1) \quad (7 \times 1) \quad \rightarrow (3 \times 1) + (3 \times 1) = (3 \times 1)$$

$$\rightarrow W_{3 \times 7} \quad (21 \text{ weights}) \quad \frac{21 \text{ wts}}{4 \text{ biases}}$$

forget gate $\leftarrow 0\%$
 $\downarrow 50\%$
 $\downarrow 100\%$

remove long term $\rightarrow [4 \ 5 \ 6]$

$[C_{t-1}] \rightarrow [0 \ 0 \ 0] \otimes [f_t]$

$[1 \ 1 \ 1] \rightarrow [\frac{1}{2} \ \frac{1}{2} \ \frac{1}{2}]$

$[4 \ 5 \ 6] \rightarrow [2 \ 2.5 \ 3]$

$[0 \ 0 \ 0] \rightarrow 100\% \text{ remove}$



$$f_t = \sigma(W_f [h_{t-1}, x_t] + b_f)$$

$$(3 \times 1) \quad (3 \times 7) \quad (7 \times 1) \quad (3 \times 1) \quad (3 \times 1) \rightarrow (3 \times 1)$$

\downarrow weight matrixes

$$f_t \otimes c_{t-1} \rightarrow (3 \times 1)$$

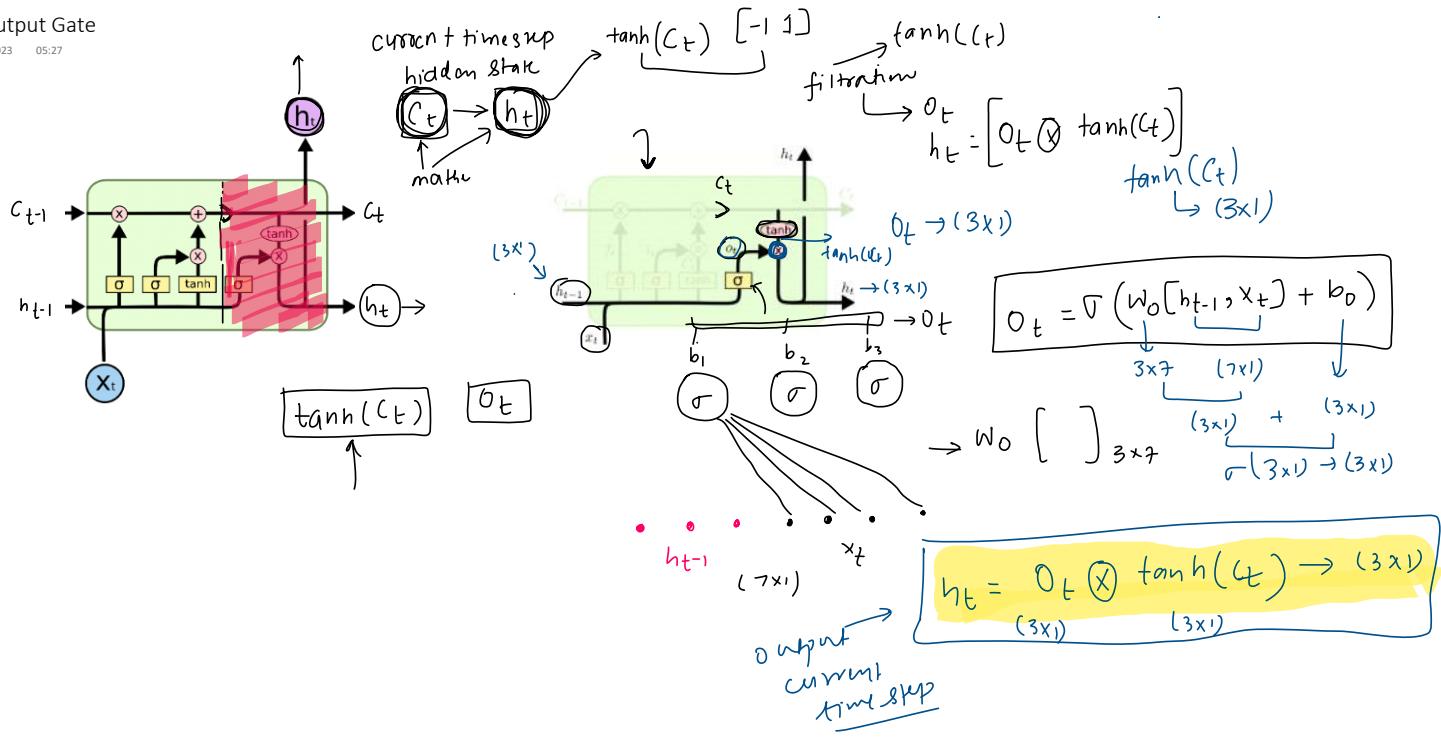
$$(C_{t-1}) \rightarrow \underline{\text{remove}}$$

জন removal এন্টি।
 f_t এ sigmoid আছে
 c_{t-1} output range $(0 \sim 1)$
এখনকারু সংস্করণ
অন্য জাত ধরনের মানের
পুর কর্তৃত অবশ্যই
আমার জন্যে এটি রয়েছে।

সুতরাং, f_t এর পাই power আছে যাতে information
 c_{t-1} এর আসলের দিকে process রয়েছে।

The Output Gate

30 August 2023 05:27

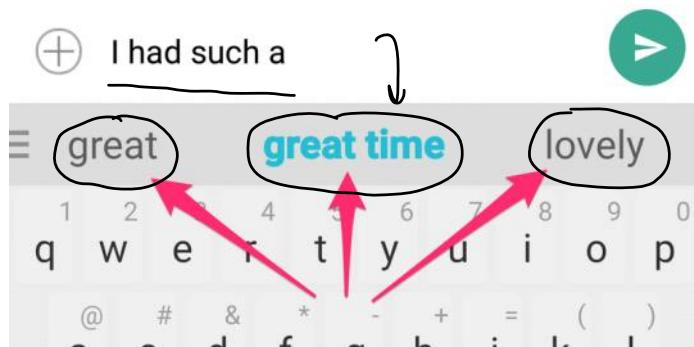


प्राची रात्रि, Current timestep का यह hidden state calculate करना (ht calculate करना)।

Lecture #63

What is a Next Word Predictor

08 September 2023 08:50



code

Eran Brauer
Mobile • 1h ago

Guy Katabi • 8:47 AM
Hi Eran

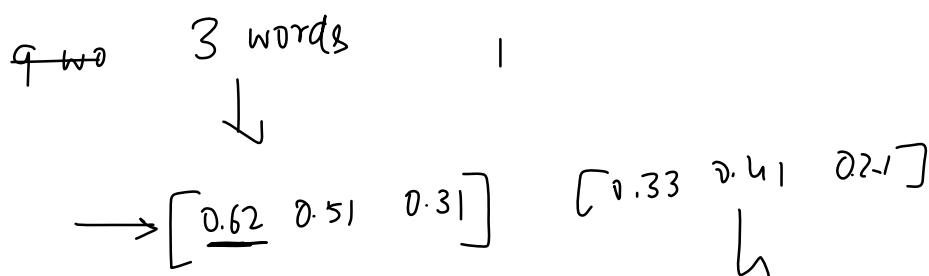
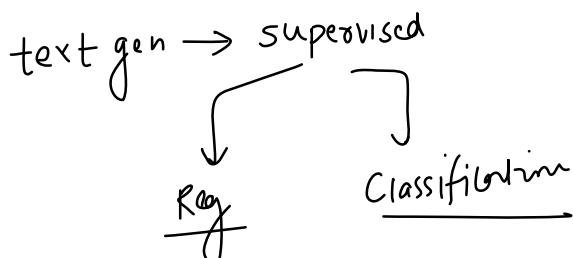
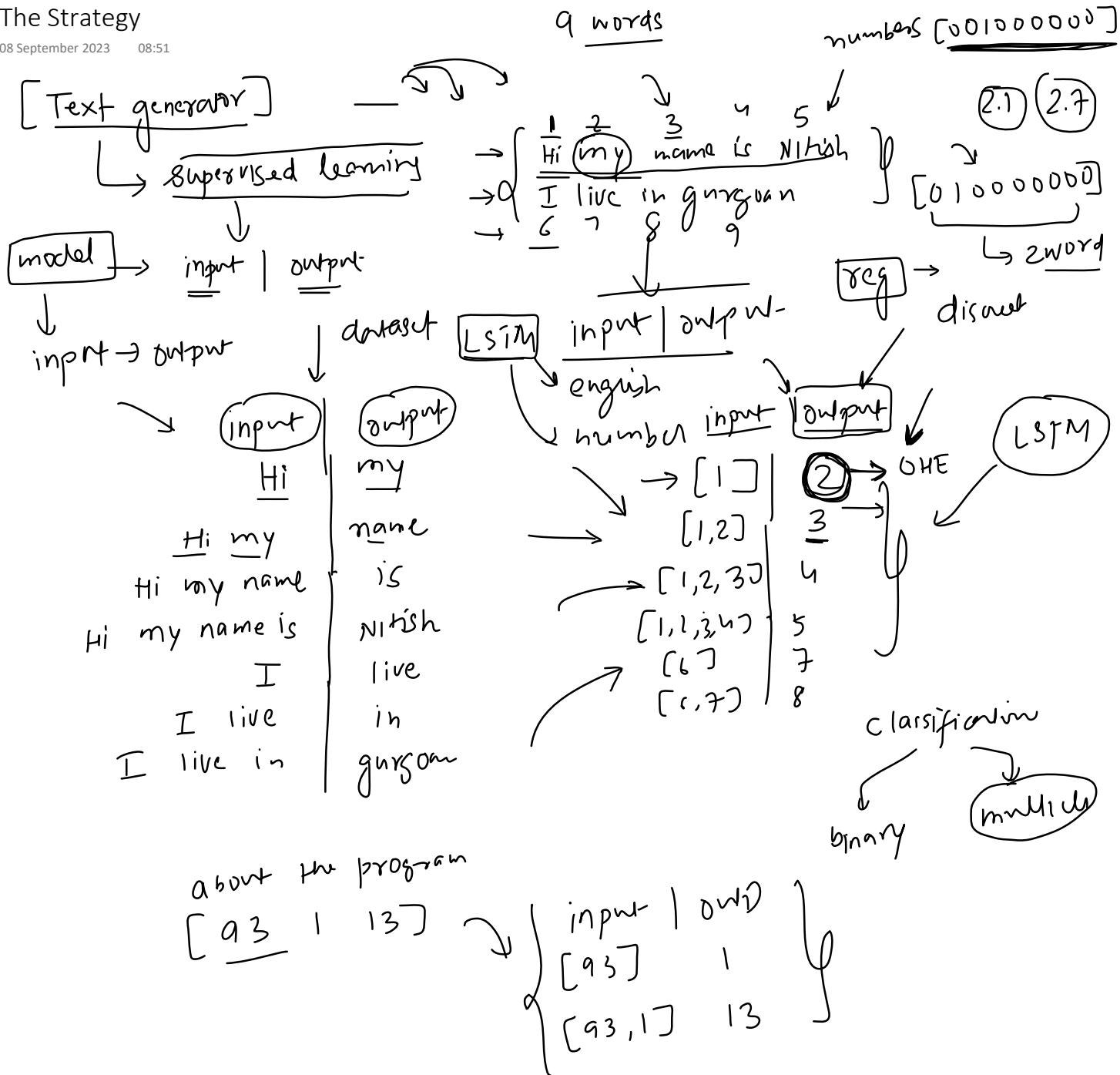
Thanks for reaching out and glad to be in your network.

Image, Video, GIF, Smileys

Send

The Strategy

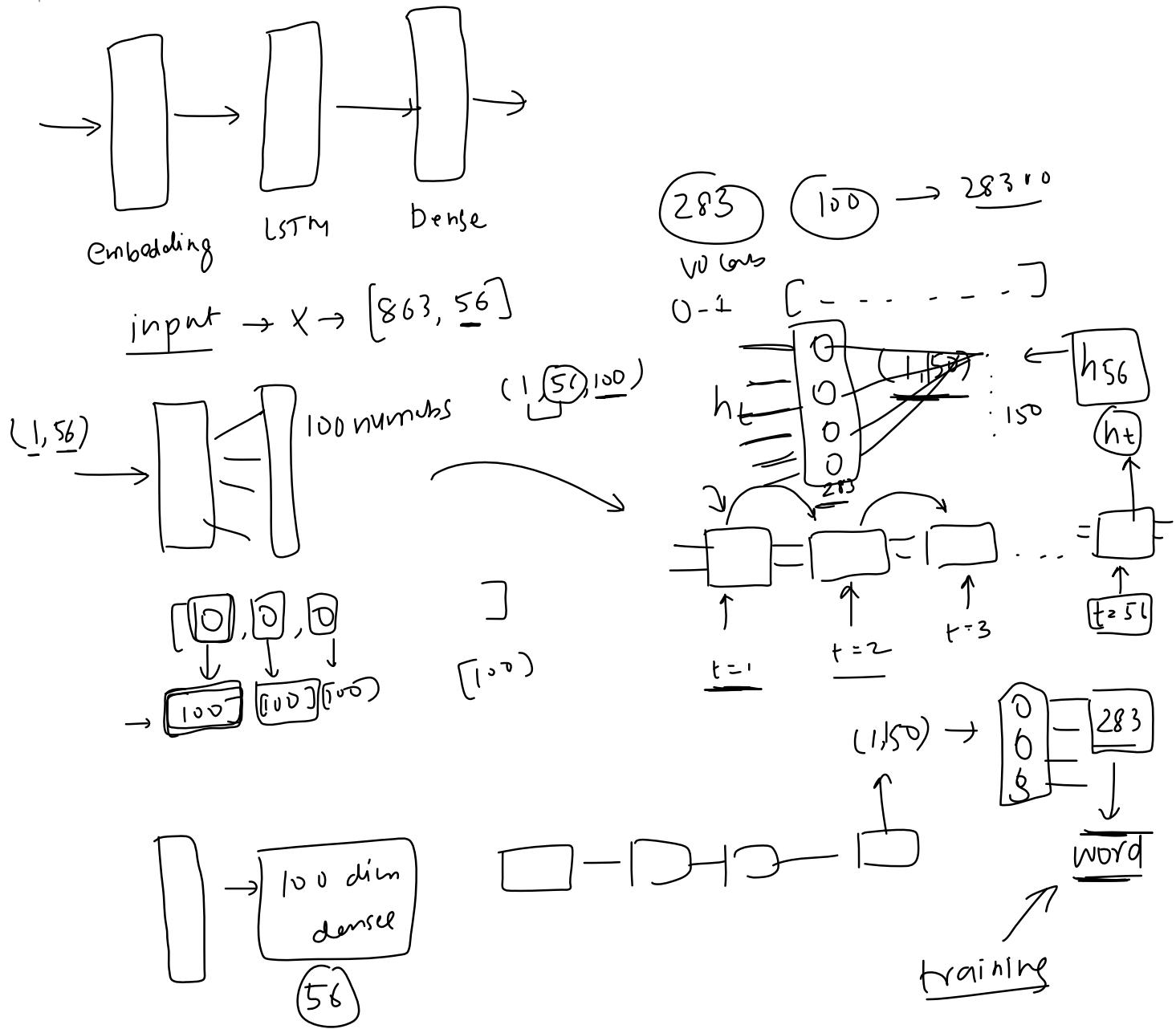
08 September 2023 08:51



[| 0 0] [0 | 0]
↳ first word ↳ second word

The Architecture

08 September 2023 08:55



How to improve performance?

08 September 2023 08:51

