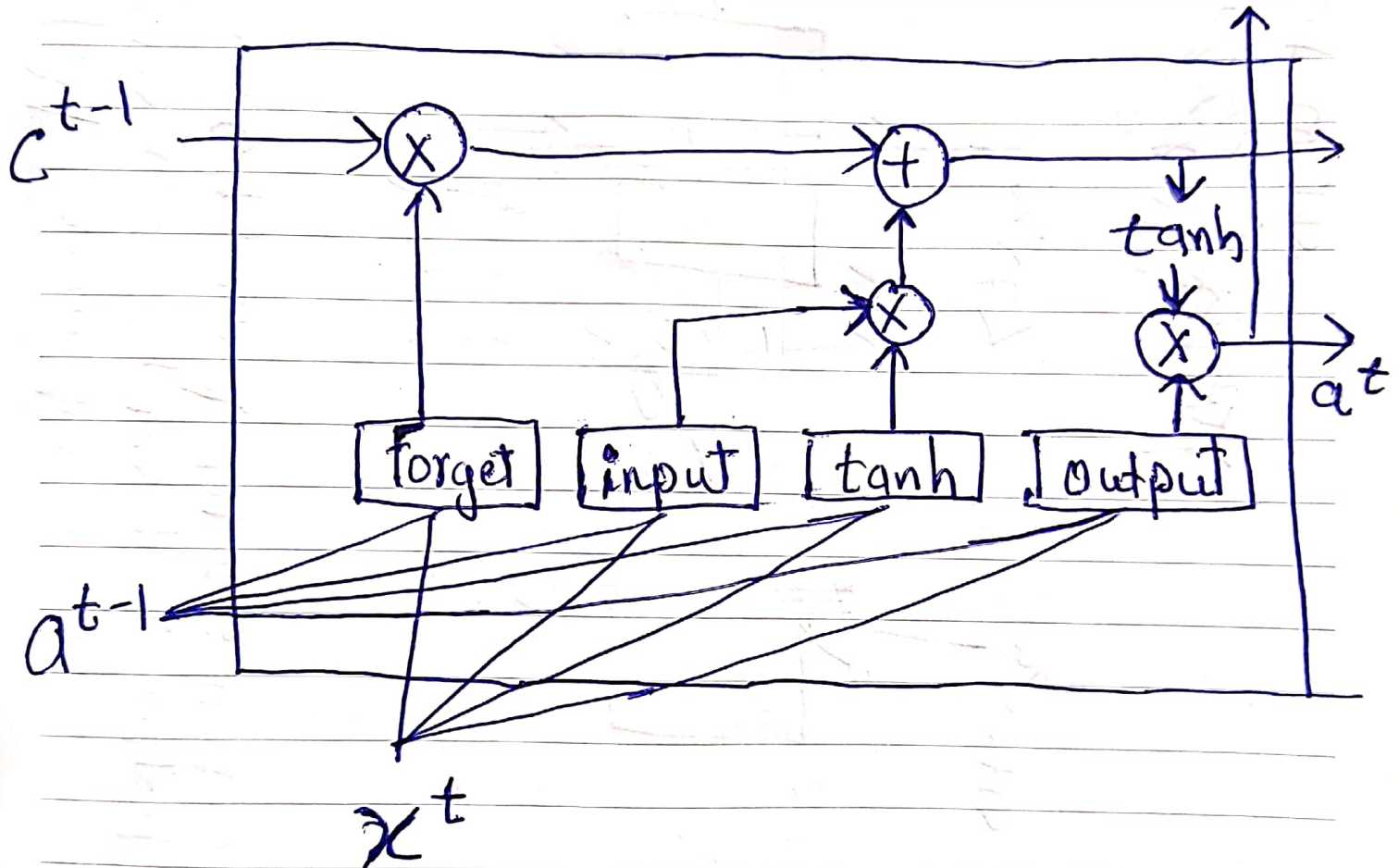


LSTM

LSTM has 3 gate and output gate

\therefore it has 4 weight matrices.



At times we encounter situations that are difficult to handle,
but with the right attitude and confidence, the most difficult journey can be completed easily.

24

शनि SAT

LSTM parameter counts

lets assume,

$$\text{input dim} = 64, \text{ num units} = 32$$

∴ num_units + input_dim :

$$\text{concat} [a^{t-1}, x^t], +1 \text{ bias}$$

there are 4 NN layers (4 gates :

$$W_{\text{forget}}, W_{\text{input}}, W_{\text{output}}, W_{\text{cell}}$$

25

रवि SUN

∴ NO. of Parameters =

$$[(\text{num units} + \text{input-dim} + 1) * \text{num units}] * 4$$

$$= [(32 + 64 + 1) * 32] * 4$$

$$= 12,416$$

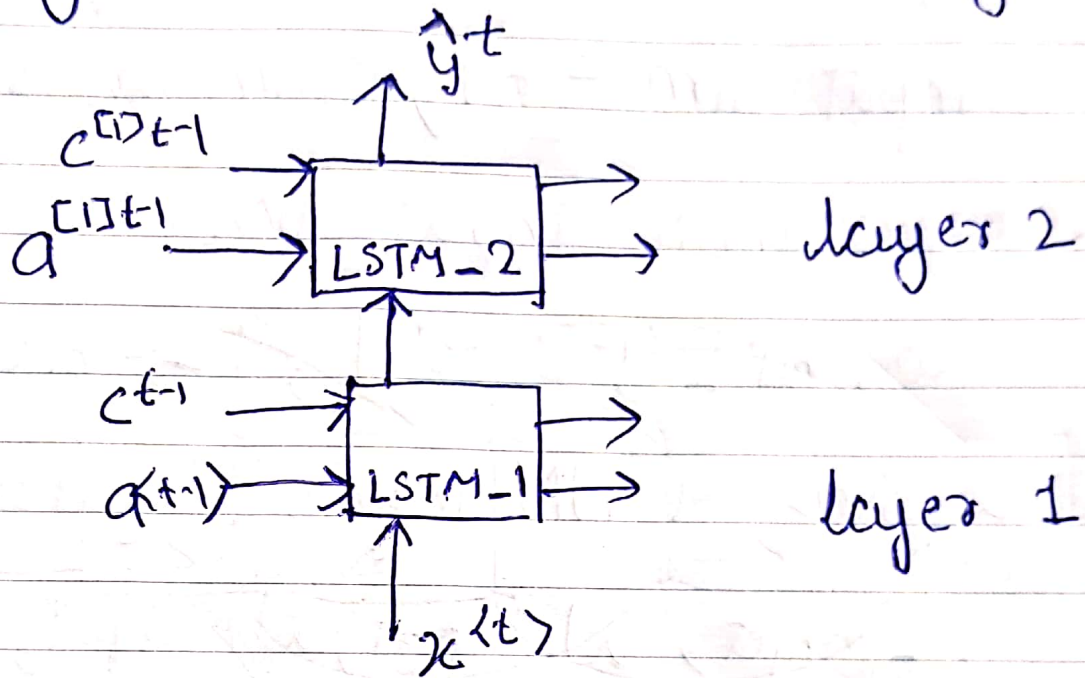
P.5: num units = NO. of units in hidden gates/activation units

कई बार हमें ऐसी परिस्थितियों का सामना करना पड़ता है जो बहुत कठिन लगती हैं, पर सही नज़रिये और आत्मविश्वास के सहारे कठिन से कठिन सफर भी आसानी से तय किया जा सकता है।

Multilayer LSTM parameter counts

शुक्र FRI

Lets say we have 2 LSTM layers,
then,



~~If layer 1 we can count~~ first, we will count parameters in layer 1, then, will pass output units/activation units/num units to next layer as inputs, After that, will count parameter of layer 2 in same way.

22

गुरु THU

let's assume, input dim = 64

activation units / output units / hidden units /
or num units = 32for layer 1

$$[(\text{num units} + \text{input units} + 1) * \text{num units}] * 4$$

$$= [(32 + 64 + 1) * 32] * 4 = 12416$$

for layer 2now input dim = 32, \therefore

$$[(32 + 32 + 1) * 32] * 4 = 8320$$

*** Note:

activation unit = num unit = output unit
= hidden unit = gate unit.

meaning is same Just name is diff.

हमारी सबसे बड़ी पूंजी यह है कि किसी की आँखों में हमारे लिए आँसू हो,
और हमारा सबसे बड़ा अपराध यह है कि किसी की आँखों में हमारी वजह से आँसू हो।

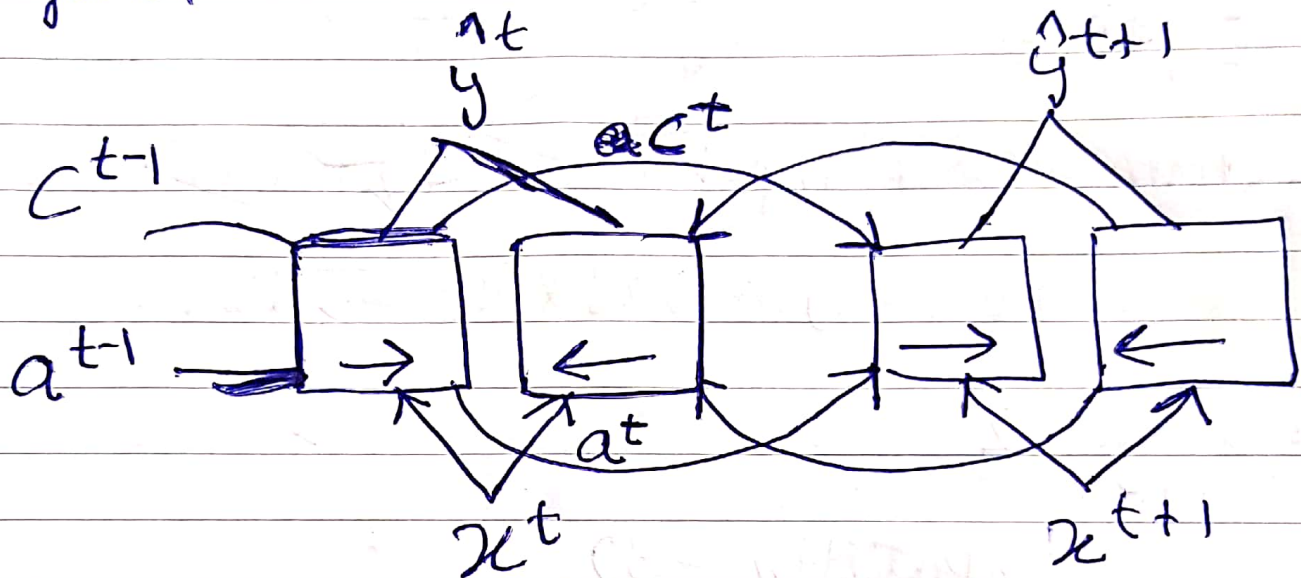
Bidirectional LSTM Parameters

21

बुध WED

Count

Formula is Same, we just multiply it by 2 as it has bi directional flow.



Input dim = 64 , Num-unit/activation unit = 32

$$2 \times [(inputdim + numunit + 1) \times numunit] \times 4$$

$$= 2 \times [(64 + 32 + 1) \times 32] \times 4$$

$$= 24832$$

At times, the best way to solve the problem is to change your route and approach.

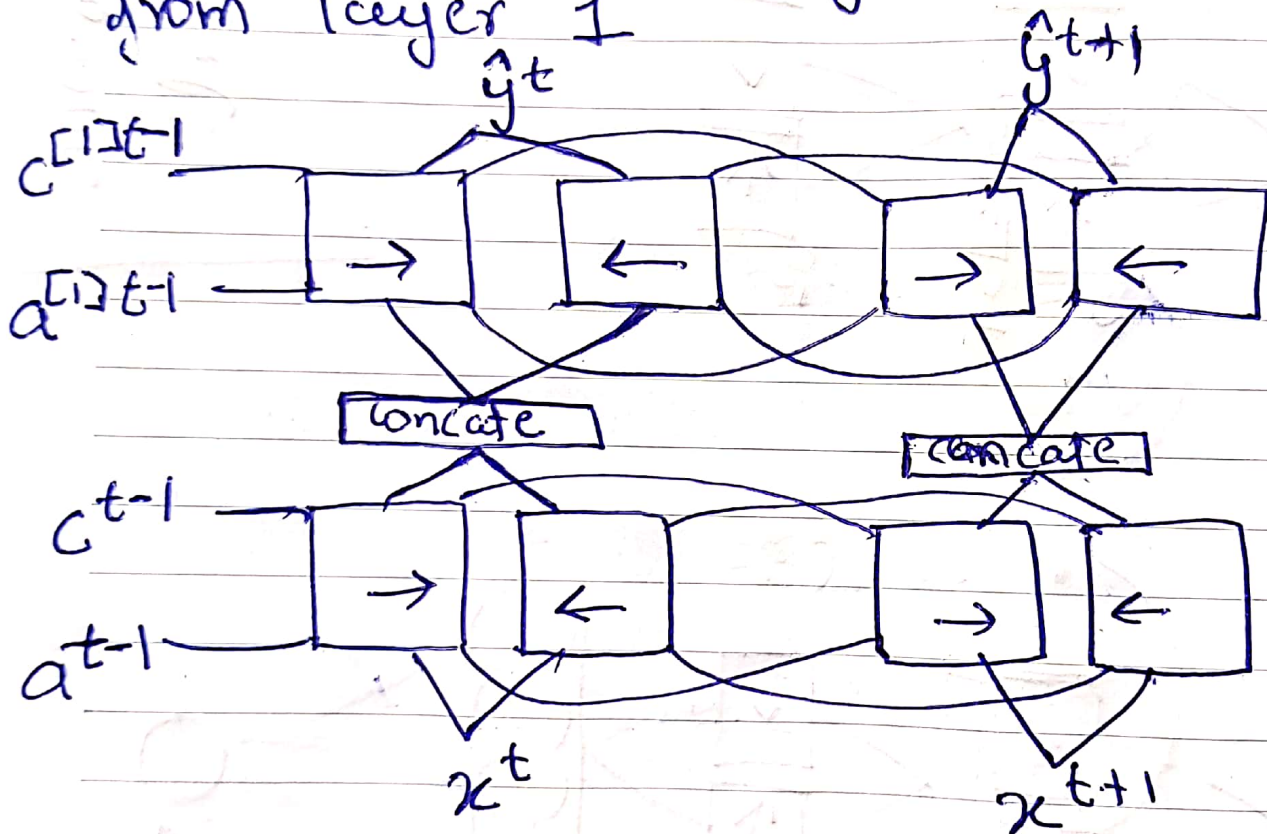
20

मंगल TUE

multilayer Bidirectional LSTM

Parameters Counts.

here, idea is, input dim of 2nd layer will be concatenation of both num units from layer 1



activation unit/num unit from 1st layer will be 32-32 from both

∴ On concatenation $\text{Concat}(32, 32) = 64$

∴ Input dim of 2nd layer = 64

$$= 2 * [(64 + 32 + 1) * 32] * 4$$

$$= 24832$$

कभी-कभी अपना रास्ता और दृष्टिकोण बदल देना ही मुश्किलों का सही हल होता है।

GRU Parameter Counts 14 बुध WED

GRU RNN have 2 gates :

update gate & relevance gate,
and one output cell.

∴ 3 weight matrices.

∴ Every thing is SAME as ~~LSTM~~
LSTM's formulas. ~~Just~~ Just

instead of multiply it by 4,

multiply ~~it~~ it by 3, as it has
3 weight matrices, rest, everything
is same in order to count
parameters.

for LSTM

of Parameters =

$$[(\text{num units} + \text{input dims} + 1) * \text{num units}] * 4$$

for GRU

of Parameters =

$$[(\text{num units} + \text{input dims} + 1) * \text{num units}] * 3$$

ing-3D

ing-3D