Sub :

Day ID: 18321046

Time : Date : / /

Ans. to the Ques. No. 3

Update gate: Update gate is responsible for determining and maintaing the amounts of past information that would be passed to the future.

$$z_t = \sigma(W_z * x_t + U_z * h_{t-1})$$

Here, x_t - input vector z_t = updat gate

W_z, U_z - weight matrices

h_{t-1} - previous hidden state.

reset gate: Reset gate is used for determining the amount of past information that is needed to forget.

$$r_t = \sigma(W_r * x_t + U_r * h_{t-1})$$

r_t - reset gate

x_t - input

W_r, U_r - weight matrices

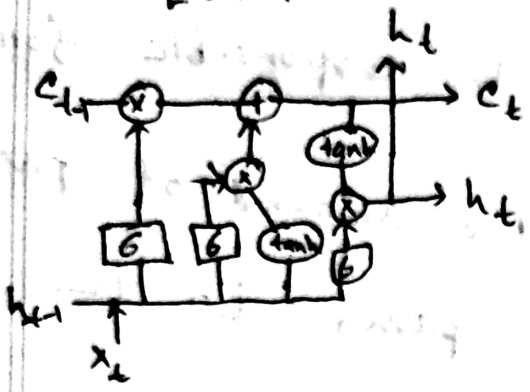
h_{t-1} - previous hidden state.

Sub :

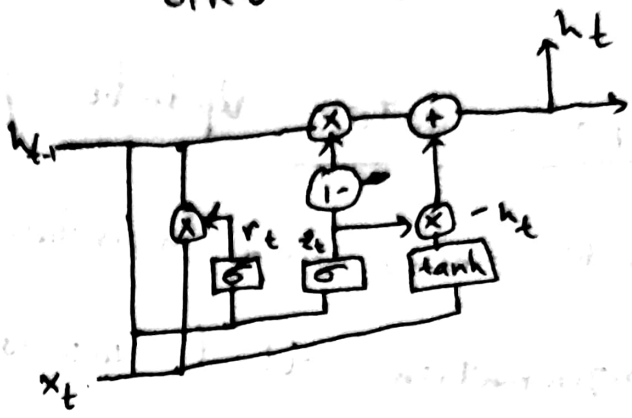
Day

Time : Date : / /

LSTM



GRU



Here, GRU and LSTM both have x_t and h_{t-1} inputs. And both do the calculation and pass values along. But the main difference in the workflow of LSTM and GRU is that GRU's don't need the cell layer to pass values along. In GRU it has two gates but lstm has three gates with less memory unit, GRU do the calculations within each iteration insures that the h_t values beign passed along either retain a high amount of old information or are jump-started with a high amount of new information.

Sub :

Day

Time : Date : / /

The reasons why GRU is faster ~~than~~ in compare to RNN are —

- In GRU there are two gates (update gate, reset gate) but ~~if~~ LSTM have three gates (forget, input, output). ~~that~~ GRU has less memory cell and can work fast, as forget gate in LSTM control the prodedure and make the system slow.
- GRU has less gates, so the system is less complex with respect to LSTM.
- GRU don't have cell state unlike LSTM. So It makes the work easier and faster.
- GRU have almost same architecture as Vanila RNN though vanila RNN is simple and use one function which take much time than GRV.