

Question-3

In GRU, the update gate and reset gate decide what info should be passed to the output.

Update gate helps to determine how much the past info needs to be passed along the future.

Reset gate is used from the model to decide how much of the past info to forget.

LSTM

VS

GRU

has two gate — update & reset
update & reset — input, output,
forget.

more complex — less complex

good for large dataset — good for small dataset

Does not expose — exposes
the complete history of the memory
memory of hidden layer & hidden layer completely

Why GRU is faster than RNN

GRU has only two gates reset & update gate. It uses less training parameters & therefore uses less memory, executes faster & trains faster. GRU are simpler & thus easier to modify. On the other hand, RNN have feedback loops in the recurrent layers which lets them maintain information in memory over time. But it can be difficult to train RNN to solve problems that require long term temporal dependencies which causes the gradient of the loss function. And if make any error it takes long time for back propagation.

That's why GRU is faster than RNN.

Why GRU is faster than RNN

GRU has only two gates reset & update gate. It uses less training parameters & therefore uses less memory, executes faster & trains faster. GRU are simpler & thus easier to modify. On the other hand, RNN have feedback loops in the recurrent layers which lets them maintain information in memory over time. But it can be difficult to train RNN to solve problems that require long term temporal dependencies which causes the gradient of the loss function. And if make any error it takes long time for back propagation. That's why GRU is faster than RNN.