NLP ASSIGNMENT 3

Q1) RNN has two components - one function that gives output and another function that generates another input that passes through another function for output. Thus, there is processing and memory at same time. Basically, this has given to fundamental blocks of LSTM and GRU architectures.

Q2) BPTT is an algorithm to train weights based on gradients.

Q3) during back propagation, gradients often get reduced continuously and sometimes even reach very small values leading to vanishing gradients. Then weight updation cannot take place and convergence to optimum does not take place. Sometimes, gradients also increase continuously during back propagation leading to exploding gradients and this leads to very large weight updates and divergence from optimum.

Q4) LSTM is an RNN with output giving units and feedback connectors.

Q5) GRU is an RNN with a forget gate and no output gate.

Q6) in peephole LSTM, cell state can also be seen.

Q7) BRNN has two RNNs - one forward and one backward to include learning from even forward layers, moving along backward direction.

Q8) LSTM has 3 gates - input gate, forget gate, output gate. Forget gate has sigmoid function, input gate has point by point multiplication of sigmoid and tanh functions and output gate has point by point multiplication of sigmoid and tanh functions.

Q9) biLSTM is bidirectional LSTM containing two LSTMs, one in forward and other in backward direction.

Q10) biGRU is bidirectional GRU containing two GRUs, one in forward and other in backward direction.