DL ASSIGNMENT 8

Q1) stateless RNN initialises states separately while stateful RNN uses previous batch’s output states.

Q2) vectorisation of text is very important via encoding and decoding.

Q3) padding and truncation handles variable length input sequences. decoder handles variable length output sequences.

Q4) beam search follows best nodes at particular search. Set beam width and decode.

Q5) attention mechanism helps to understand which words to focus more which further helps in understanding the context of a text.

Q6) residual connections are most important in transformer architecture since they retain position related information.

Q7) when number of classes is large, sampled softmax eases computation.