NLP ASSIGNMENT 5

Q1) seq2seq models take text as input and give text as output. In common architecture, there is an encoder that encodes text and a decoder that decodes into text.

Q2) problems with vanilla RNN include vanishing and exploding gradients, cannot have deep networks or stockings, slow training etc.

Q3) high gradients are replaced with small gradients to prevent exploding gradient problem.

Q4) attention mechanism aims to understand the relative importance of words and their context.

Q5) CRFs are used to predict sequences.

Q6) self attention relates different positions in sequence to understand inputs better.

Q7) bandana attention does both aligning and translation simultaneously. It is called additive attention as it does does a linear combination of both encoder and decoder.

Q8) language model estimates the probabilility distribution of various language units.

Q9) multihead mechanism can run through attention mechanism several times in parallel. It is a module for attention mechanisms.

Q10) BLUE is an algorithm that shows how good a machine translation is using a score.