1.what are sequence to sequence model ?

Ans : sequence to sequence model that takes a sequence of items words, letters, time series, etc and outputs another sequence of items.In the case of Neural Machine Translation, the input is a series of words, and the output is the translated series of words.

2.what are the problems with vanilla RNNs ?

Ans : The vanishing or exploding gradient problem. The vanishing and/or exploding gradient problems are regularly experienced with regards to RNNs,Exploding gradients,RNNs cannot be stacked up,Slow and Complex training procedures,Difficult to process longer sequences.

3.what is gradient Clipping ?

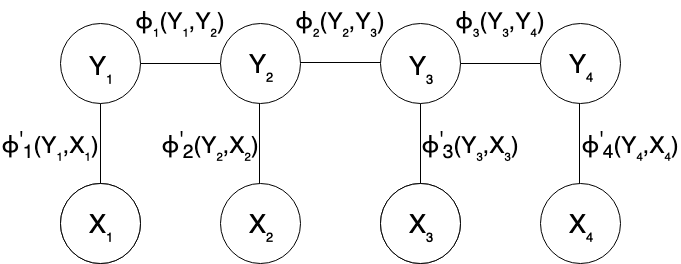
Ans : Gradient clipping is a technique to prevent exploding gradients in very deep networks, usually in recurrent neural networks.

4.Explain Attention mechanism.

Ans : Attention mechanism was to permit the decoder to utilize the most relevant parts of the input sequence in a flexible manner by a weighted combination of all of the encoded input vector with the most relevant vectors being attributed the highest weights.Attention is an interface connecting the encoder and decoder that provides the decoder with information from every encoder hidden state

5.Explain the conditional random fields(CRFs)

Ans : Conditional Random Fields is a class of discriminative models best suited to prediction tasks where contextual information or state of the neighbors affect the current prediction.



6.Explain self-attention.

Ans : Self Attention is also called intra Attention attention mechanism relating different positions of a single sequence in order to compute a representation of the same sequence. It has been shown to be very useful in machine reading, abstractive summarization, or image description generation.the self-attention mechanism allows the inputs to interact with each other (“self”) and find out who they should pay more attention to (“attention”).

7.what is Bahdanau Attention.?

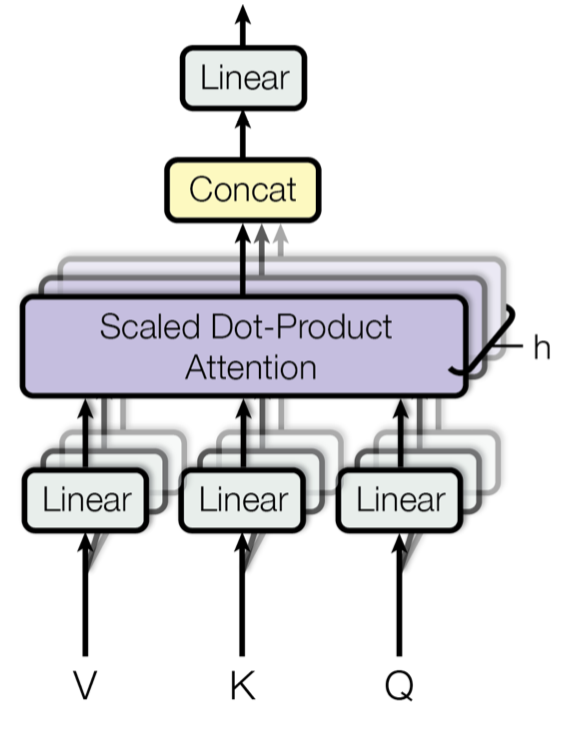
Ans : Bahdanau attention mechanism that learns to align and translate jointly. It is also known as Additive attention as it performs a linear combination of encoder states and the decoder states.

8.what is a Language model ?

Ans : Language model is use of various statistical and probabilistic techniques to determine the probability of a given sequence of words occurring in a sentence. Language models analyze bodies of text data to provide a basis for their word predictions.

9.what is multi head attention ?

Ans : Multi-head Attention is a module for attention mechanisms which runs through an attention mechanism several times in parallel. The independent attention outputs are then concatenated and linearly transformed into the expected dimension.



10.what is Bilingual Evaluation understudy(BLEU)?

Ans :Bilingual evaluation understudy is a score for comparing a candidate translation of text to one or more reference translations.