1. What are Sequence-to-sequence models?

A Seq2Seq model is a 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 Problem with Vanilla RNNs?

A major issue with the vanilla RNN is that they suffers from vanishing/exploding gradients similarly to issues with deep feedforward networks. At each timestep, the hidden state ht is multiplied by W, at the last timestep, the value of ht is effectively multiplied by W .

3. What is Gradient clipping?

Gradient Clipping is a method where the error derivative is changed or clipped to a threshold during backward propagation through the network, and using the clipped gradients to update the weights.

4. Explain Attention mechanism

An attention model, also known as an attention mechanism, is an input processing technique of neural networks. This mechanism helps neural networks solve complicated tasks by dividing them into smaller areas of attention and processing them sequentially

5. Explain Conditional random fields (CRFs)

Conditional random fields (CRFs) are a class of statistical modeling methods often applied in pattern recognition and machine learning and used for structured prediction. Whereas a classifier predicts a label for a single sample without considering "neighbouring" samples, a CRF can take context into account.

6. Explain self-attention

Self Attention, also called intra Attention, is an 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.

7. What is Bahdanau Attention?

The Bahdanau attention was proposed to address the performance bottleneck of conventional encoder-decoder architectures, achieving significant improvements over the conventional approach. In this tutorial, you will discover the Bahdanau attention mechanism for neural machine translation.

8. What is a Language Model?

Language modeling (LM) is the 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?

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)

BLEU (BiLingual Evaluation Understudy) is a metric for automatically evaluating machine-translated text. The BLEU score is a number between zero and one that measures the similarity of the machine-translated text to a set of high quality reference translations.