1. What are Sequence-to-sequence models?

**Ans : Sequence to Sequence (often abbreviated to seq2seq) models is a special class of Recurrent Neural Network architectures that we typically use (but not restricted) to solve complex Language problems like Machine Translation, Question Answering, creating Chatbots, Text Summarization, etc.**

1. What are the Problem with Vanilla RNNs?

**Ans : The most common issues with RNNS are gradient vanishing and exploding problems. The gradients refer to the errors made as the neural network trains. If the gradients start to explode, the neural network will become unstable and unable to learn from training data.**

1. What is Gradient clipping?

**Ans : Gradient clipping is a technique to prevent exploding gradients in very deep networks, usually in recurrent neural networks. A neural network is a learning algorithm, also called neural network or neural net, that uses a network of functions to understand and translate data input into a specific output.**

1. Explain Attention mechanism

**Ans : A neural network is considered to be an effort to mimic human brain actions in a simplified manner. Attention Mechanism is also an attempt to implement the same action of selectively concentrating on a few relevant things, while ignoring others in deep neural networks.**

1. Explain Conditional random fields (CRFs)

**Ans : 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 "neighboring" samples, a CRF can take context into account.**

1. Explain self-attention

**Ans : 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.**

1. What is Bahdanau Attention?

**Ans : The Bahdanau attention was proposed to address the performance bottleneck of conventional encoder-decoder architectures, achieving significant improvements over the conventional approach.**

1. What is a Language Model?

**Ans : 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. They are used in natural language processing (NLP) applications, particularly ones that generate text as an output.**

1. 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. ... Intuitively, multiple attention heads allows for attending to parts of the sequence differently (e.g. longer-term dependencies versus shorter-term dependencies).**

1. What is Bilingual Evaluation Understudy (BLEU)

**Ans : 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.**