**NLP- ASSIGNMENT-5**

1. **What are Sequence-to-sequence models?**

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?**

However, RNNs suffer from the problem of vanishing gradients, which hampers learning of long data sequences. The gradients carry information used in the RNN parameter update and when the gradient becomes smaller and smaller, the parameter updates become insignificant which means no real learning is done.

1. **What is Gradient clipping?**

Gradient clipping is a technique to prevent exploding gradients in very deep networks, usually in recurrent neural networks. ... With gradient clipping, pre-determined gradient threshold be introduced, and then gradients norms that exceed this threshold are scaled down to match the norm.

1. **Explain Attention mechanism**

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)**

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.

1. **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.

1. **What is Bahdanau Attention?**

Bahdanau attention mechanism proposed an 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.

1. **What is a Language Model?**

A statistical language model is a probability distribution over sequences of words. Given such a sequence, say of length m, it assigns a probability {\displaystyle P(w\_{1},\ldots ,w\_{m})}P(w\_{1},\ldots ,w\_{m}) to the whole sequence.

1. **What is Multi-Head Attention?**

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).**

BLEU is an algorithm for evaluating the quality of text which has been machine-translated from one natural language to another.