**ADITYA AMIN**

**ASSIGN : NLP-03**

1. Explain the basic architecture of RNN cell.

The basic architecture of a Recurrent Neural Network (RNN) cell consists of a recurrent connection that allows the network to process sequential data. The cell's architecture can vary, but the most common type of RNN cell is the Long Short-Term Memory (LSTM) cell.

1. Explain Backpropagation through time (BPTT)

Backpropagation Through Time (BPTT) is a learning algorithm used to train recurrent neural networks (RNNs) on sequential data. It is an extension of the backpropagation algorithm that is commonly used for training feedforward neural networks.

1. Explain Vanishing and exploding gradients

Vanishing and exploding gradients are common issues that can occur during the training of deep neural networks, particularly recurrent neural networks (RNNs). These problems affect the backpropagation algorithm, where gradients are calculated and propagated backward through the network to update the model's parameters.

1. Explain Long short-term memory (LSTM)

Long Short-Term Memory (LSTM) is a type of recurrent neural network (RNN) architecture designed to effectively capture and process long-term dependencies in sequential data. It was introduced to overcome the limitations of traditional RNNs, such as the vanishing/exploding gradient problem and the difficulty in retaining information over long sequences.

1. Explain Gated recurrent unit (GRU)

Gated Recurrent Unit (GRU) is a type of recurrent neural network (RNN) architecture that addresses some of the limitations of traditional RNNs, such as the vanishing gradient problem and difficulty in capturing long-term dependencies. GRU is similar to LSTM but has a simpler structure with fewer gating mechanisms.

1. Explain Peephole LSTM.

Peephole LSTM is an extension of the Long Short-Term Memory (LSTM) architecture that incorporates peephole connections to the LSTM cell. It enhances the standard LSTM by allowing the cell to directly influence the gating mechanism using information from the cell's own memory.

1. Bidirectional RNNs

Bidirectional Recurrent Neural Networks (RNNs) are a type of RNN architecture that processes sequential data in both forward and backward directions. It combines two separate RNNs, one operating in the forward direction and the other in the backward direction, to capture information from both past and future contexts.

1. Explain the gates of LSTM with equations.

Forget Gate:

The forget gate determines how much of the previous cell state should be forgotten or retained. It takes the previous cell state C\_{t-1} and the current input X\_t as inputs, and outputs a forget gate activation f\_t between 0 and 1.

The equation for the forget gate is:

f\_t = sigmoid(W\_f · [h\_{t-1}, X\_t] + b\_f)

Input Gate:

The input gate controls how much new information is added to the cell state. It takes the previous cell state C\_{t-1} and the current input X\_t as inputs, and outputs an input gate activation i\_t between 0 and 1.

The equation for the input gate is:

i\_t = sigmoid(W\_i · [h\_{t-1}, X\_t] + b\_i)

Cell State Update:

The cell state update step involves computing a candidate cell state C\_tilde\_t based on the current input X\_t and the previous hidden state h\_{t-1}. The candidate cell state is then combined with the forget gate output and the input gate output to update the cell state C\_t.

The equation for the candidate cell state is:

C\_tilde\_t = tanh(W\_c · [h\_{t-1}, X\_t] + b\_c)

1. Explain BiLSTM\

BiLSTM (Bidirectional Long Short-Term Memory) is a variant of the Long Short-Term Memory (LSTM) architecture that incorporates bidirectional processing. It consists of two LSTM networks, one processing the input sequence in the forward direction and the other processing it in the backward direction. The outputs of both directions are combined to obtain the final output sequence.

1. Explain BiGRU

BiGRU (Bidirectional Gated Recurrent Unit) is a variant of the Gated Recurrent Unit (GRU) architecture that incorporates bidirectional processing. It consists of two GRU networks, one processing the input sequence in the forward direction and the other processing it in the backward direction. The outputs of both directions are combined to obtain the final output sequence.