**Assignment\_3**

1. Explain the basic architecture of RNN cell.

**Ans: Recurrent Neural Network: The fundamental feature of a Recurrent Neural Network (RNN) is that the network contains at least one feed-back connection, so the activations can flow round in a loop.**

**There are two main architectures that are used in almost every application of recurrent neural networks: long-short term memory (LSTM) (Hochreiter and Schmidhuber, 1997) and gated recurrent unit (GRU) (Cho et al., 2014). Both of these use every time step to calculate an output and to update the internal state.**

1. Explain Backpropagation through time (BPTT)

**Ans: Backpropagation Through Time, or BPTT, is the application of the Backpropagation training algorithm to recurrent neural network applied to sequence data like a time series. A recurrent neural network is shown one input each timestep and predicts one output. Conceptually, BPTT works by unrolling all input timesteps**.

1. Explain Vanishing and exploding gradients

**Ans: The exploding gradient is** **the inverse of the vanishing gradient and occurs when large error gradients accumulate, resulting in extremely large updates to neural network model weights during training**. **As a result, the model is unstable and incapable of learning from your training data.**

**The term vanishing gradient refers to the fact that** **in a feedforward network (FFN) the backpropagated error signal typically decreases (or increases) exponentially as a function of the distance from the final layer**.

1. Explain Long short-term memory (LSTM)

**Ans: Long Short-Term Memory (LSTM) networks are a type of recurrent neural network capable of learning order dependence in sequence prediction problems. This is a behavior required in complex problem domains like machine translation, speech recognition, and more. LSTMs are a complex area of deep learning**.

1. Explain Gated recurrent unit (GRU)

**Ans: Gated recurrent units (GRUs) are a gating mechanism in recurrent neural networks. The GRU is like a long short-term memory (LSTM) with a forget gate, but has fewer parameters than LSTM, as it lacks an output gate.**

1. Explain Peephole LSTM

**Ans: One popular LSTM variant, is adding “peephole connections.” This means that we let the gate layers look at the cell state. In this peephole connection we can see that all the gates are having an input along with the cell state.**

1. Bidirectional RNNs

**Ans: Bidirectional long-short term memory(Bidirectional LSTM) is** **the process of making any neural network o have the sequence information in both directions backwards (future to past) or forward(past to future)**

1. Explain the gates of LSTM with equations.
2. Explain BiLSTM

**Ans: A Bidirectional LSTM, or biLSTM, is** **a sequence processing model that consists of two LSTMs: one taking the input in a forward direction, and the other in a backwards direction**.

1. Explain BiGRU

**Ans: A Bidirectional GRU, or BiGRU, is a sequence processing model that consists of two GRUs. one taking the input in a forward direction, and the other in a backwards direction. It is a bidirectional recurrent neural network with only the input and forget gates**.