1. Explain the basic architecture of RNN cell.

**Ans : There are two main architectures that are used in almost every application of recurrent neural networks: long-short term memory (LSTM) and gated recurrent unit (GRU). 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 training algorithm used to update weights in recurrent neural networks like LSTMs.**

1. Explain Vanishing and exploding gradients

**Ans : In a network of n hidden layers, n derivatives will be multiplied together. If the derivatives are large then the gradient will increase exponentially as we propagate down the model until they eventually explode, and this is what we call the problem of exploding gradient .**

1. Explain Long short-term memory (LSTM)

**Ans : Long short-term memory (LSTM) is an artificial recurrent neural network (RNN) architecture used in the field of deep learning. LSTM networks are well-suited to classifying, processing and making predictions based on time series data, since there can be lags of unknown duration between important events in a time series.**

1. Explain Gated recurrent unit (GRU)

**Ans : 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 : Peephole connections refer to a modification to the basic LSTM architecture. Surprisingly, LSTM augmented by “peephole connections” from its internal cells to its multiplicative gates can learn the fine distinction between sequences of spikes separated by either 50 or 49 discrete time steps.**

1. Bidirectional RNNs

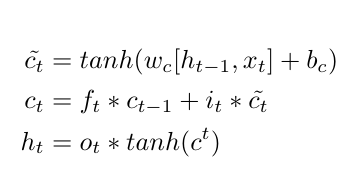
**Ans : Bidirectional recurrent neural networks(RNN) are really just putting two independent RNNs together. The input sequence is fed in normal time order for one network, and in reverse time order for another.**

1. Explain the gates of LSTM with equations.

**Ans : First equation is for Input Gate which tells us that what new information we’re going to store in the cell state(that we will see below).**

**Second is for the forget gate which tells the information to throw away from the cell state.**

**Third one is for the output gate which is used to provide the activation to the final output of the lstm block at timestamp ‘t’.**



1. 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.**