1.Explain the basic architecture of RNN cell.

Ans : A recurrent neural network (RNN) is a special kind of artificial neural network that permits continuing information related to past knowledge by utilizing a special kind of looped architecture. They are employed in many areas regarding data with sequences, such as predicting the next word of a sentence.there are various type of architecture.

Many-to-Many Architecture. Many-to-One Architecture. One-to-One Architecture. One-to-Many Architecture.RNN cell is a

anything that has a state and performs some operation that takes a matrix of inputs.

2.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.BPTT works all input timesteps.

3.Explain vanishing and exploding gradients.

Ans : Exploding gradient occurs when the derivatives or slope will get larger and larger as we go backward with every layer during backpropagation. This situation is the exact opposite of the vanishing gradients. This problem happens because of weights, not because of the activation function.

4.Explain Long short term memory(LSTM)

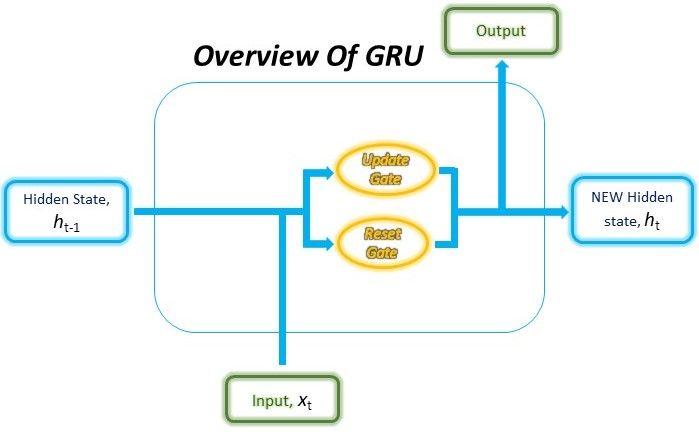
Ans : LSTM stands for Long-Short Term Memory. LSTM is a type of recurrent neural network but is better than traditional recurrent neural networks in terms of memory. Having a good hold over memorizing certain patterns LSTMs perform fairly better.LSTMs use a series of 'gates' which control how the information in a sequence of data comes into, is stored in and leaves the network. There are three gates in a typical LSTM. forget gate, input gate and output gate. These gates can be thought of as filters and are each their own neural network.

5.Explain gated recurrent unit(GRU)

Ans : A gated recurrent unit (GRU) is part of a specific model of recurrent neural network that intends to use connections through a sequence of nodes to perform machine learning tasks associated with memory and clustering

There are two gates:

Reset gate,update gate



Take the input current input and the previous hidden state as vectors. For each gate, calculate the parameterized current input and previously hidden state vectors by performing element-wise multiplication (Hadamard Product) between the vector and the respective weights for each gate.

6.Explain peephole LSTM.

Ans : peephole connection that all the gates are having an input along with the cell state.peephole connections provide feedback from the cell to the gates, allowing the gates to carry out their operations as a function of both the incoming inputs and the previous state of the cell.LSTM has connections it is capable of processing the entire sequence of data, apart from single data points such as images.

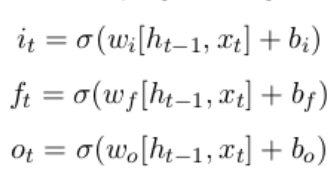
7.Bidirectional RNNs

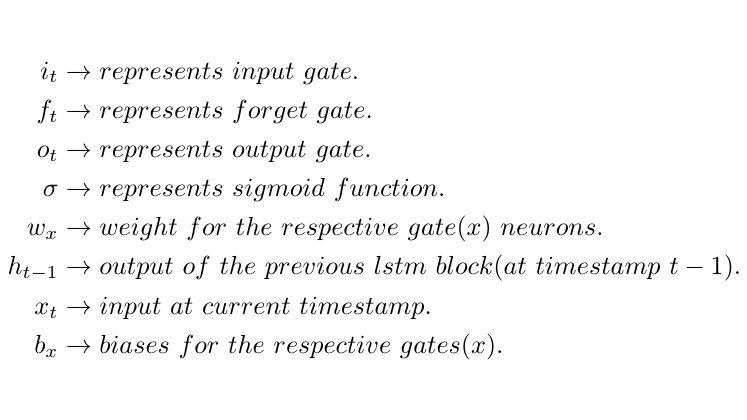
Ans: A Bidirectional RNN is a combination of two RNNs training the network in opposite directions, one from the beginning to the end of a sequence, and the other, from the end to the beginning of a sequence.

8.Explain the gates of LSTM with equations.

Ans : There are three different gates in LSTM cell:

forget gate, input gate, output gate.





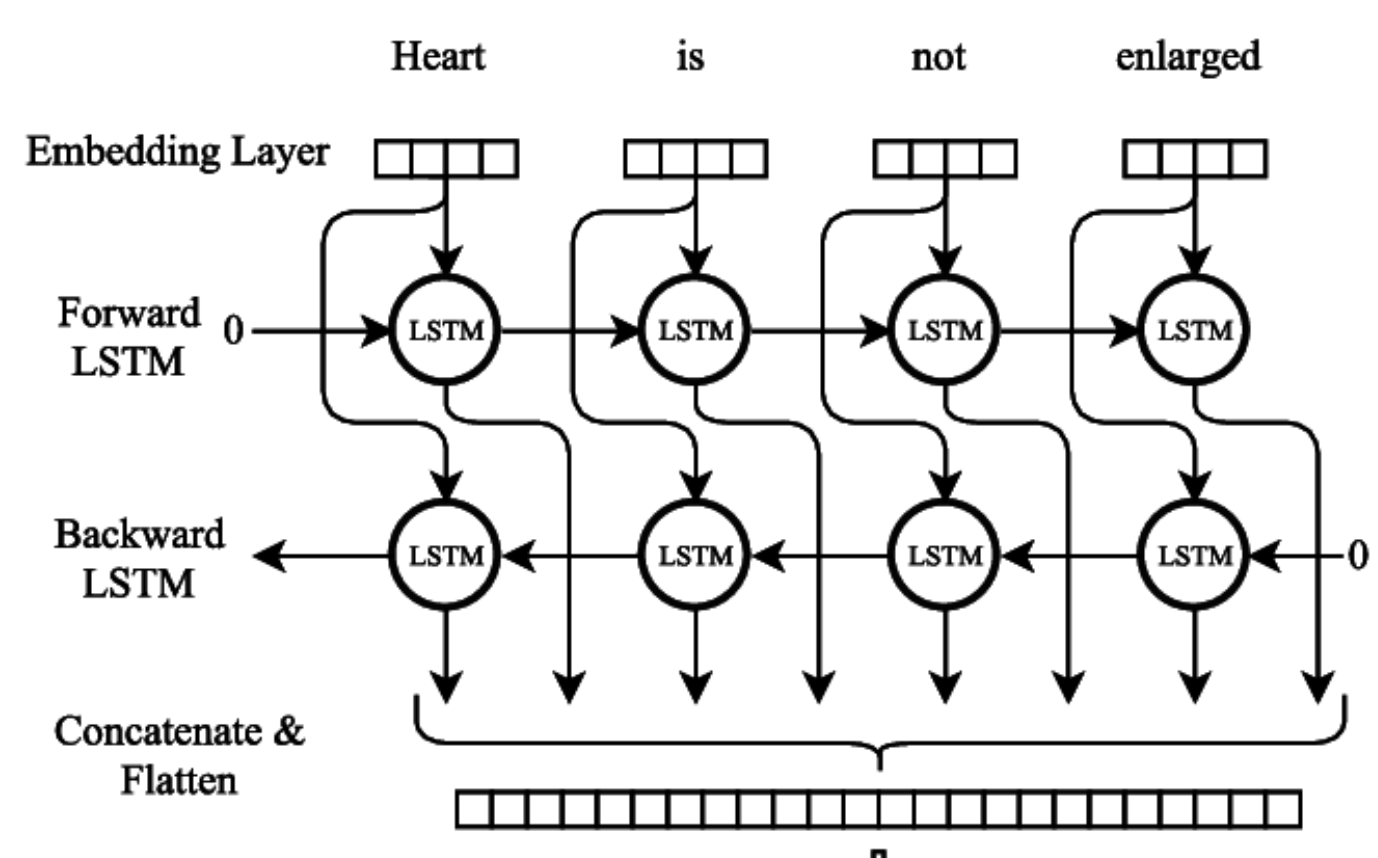
First equation is for Input Gate which tells us that what new information we’re going to store in the cell state.

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’.

9.Explain BiLSTM.

Ans : Bidirectional Long Short-term Memory. Bidirectional LSTM 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.BiLSTM calculates the input sequence from the opposite direction to a forward hidden sequence and a backward hidden sequence . The encoded vector is formed by the concatenation of the final forward and backward outputs where is the output sequence of the first hidden layer.



10.Explain BiGRU

Ans :

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.

GRUs are very similar to Long Short Term Memory(LSTM). Just like LSTM, GRU uses gates to control the flow of information. They are relatively new as compared to LSTM. This is the reason they offer some improvement over LSTM and have simpler architecture.