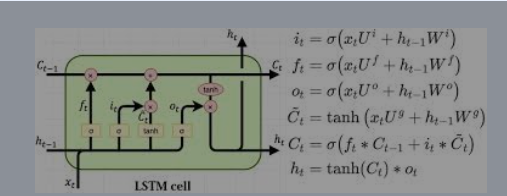
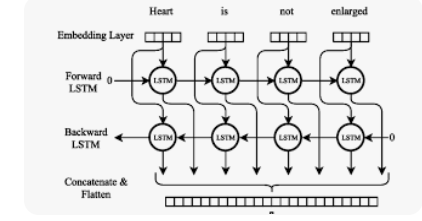
1. Explain the basic architecture of RNN cell.  
   All RNN are in the form of a chain of repeating modules of a neural network. In standard RNNs, this repeating module will have a very simple structure, such as a single tanh layer. LSTMs also have a chain-like structure, but the repeating module is a bit different structure.
2. Explain Backpropagation through time (BPTT)  
   Recurrent Neural Networks are those networks that deal with sequential data. They predict outputs using not only the current inputs but also by taking into consideration those that occurred before it. In other words, the current output depends on current output as well as a memory element (which takes into account the past inputs).
3. Explain Vanishing and exploding gradients  
   As the backpropagation algorithm advances downwards(or backward) from the output layer towards the input layer, the gradients often get smaller and smaller and approach zero which eventually leaves the weights of the initial or lower layers nearly unchanged. As a result, the gradient descent never converges to the optimum. This is known as the vanishing gradients problem.  
   The gradients keep on getting larger and larger as the backpropagation algorithm progresses. This, in turn, causes very large weight updates and causes the gradient descent to diverge. This is known as the exploding gradients problem.
4. Explain Long short-term memory (LSTM)  
   LSTM stands for long short-term memory networks, used in the field of Deep Learning. It is a variety of recurrent neural networks (RNNs) that are capable of learning long-term dependencies, especially in sequence prediction problems.
5. Explain Gated recurrent unit (GRU)  
   Gated recurrent units (GRUs) are a gating mechanism in recurrent neural networks, introduced in 2014 by Kyunghyun Cho et al. 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.
6. Explain Peephole LSTM  
   One popular LSTM variant, introduced by Gers & Schmidhuber (2000), 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.
7. Bidirectional RNNs  
   Bidirectional RNN ( BRNN ) duplicates the RNN processing chain so that inputs are processed in both forward and reverse time order. This allows a BRNN to look at future context as well. Two common variants of RNN include GRU and LSTM . LSTM does better than RNN in capturing long-term dependencies.
8. Explain the gates of LSTM with equations.  
   
9. Explain BiLSTM  
   
10. Explain BiGRU

