## Question 1

Question 1. A boolean function having n=2 variables have 2^(2^n) = 16 functions where 14 are linearly separable only 2 of them are not they are (XOR and XNOR). So for n =3 variable there will be 2^(2^3)= 2^8 = 256 functions. There are 152 functions are not linearly separable. Can you find a boolean function in terms of a,b,c 3 boolean variables to form any function that is linearly separable. You may think of feature map. Can you find the generic formula for any number of N.

N = 2 linearly not separable 2
N = 3 linearly not separable 152
N=4 ??
N=5 ??

Implement the following boolean logic with perceptron network.

$$F(A,B,C) = AB' + BA' + BC' + CB'$$

Where A,B,C are inputs and F is the output. You can assume each neuron with threshold 2. Is it linearly separable?? Explain please.

Add file

## Question 2

Write the core difference of RNN ,Peephole-LSTM and GRU? What is the basic difference of LSTM forget gate and GRU's reset gate. Show the matrix workflow of the LSTM.

Add file

## Question 3

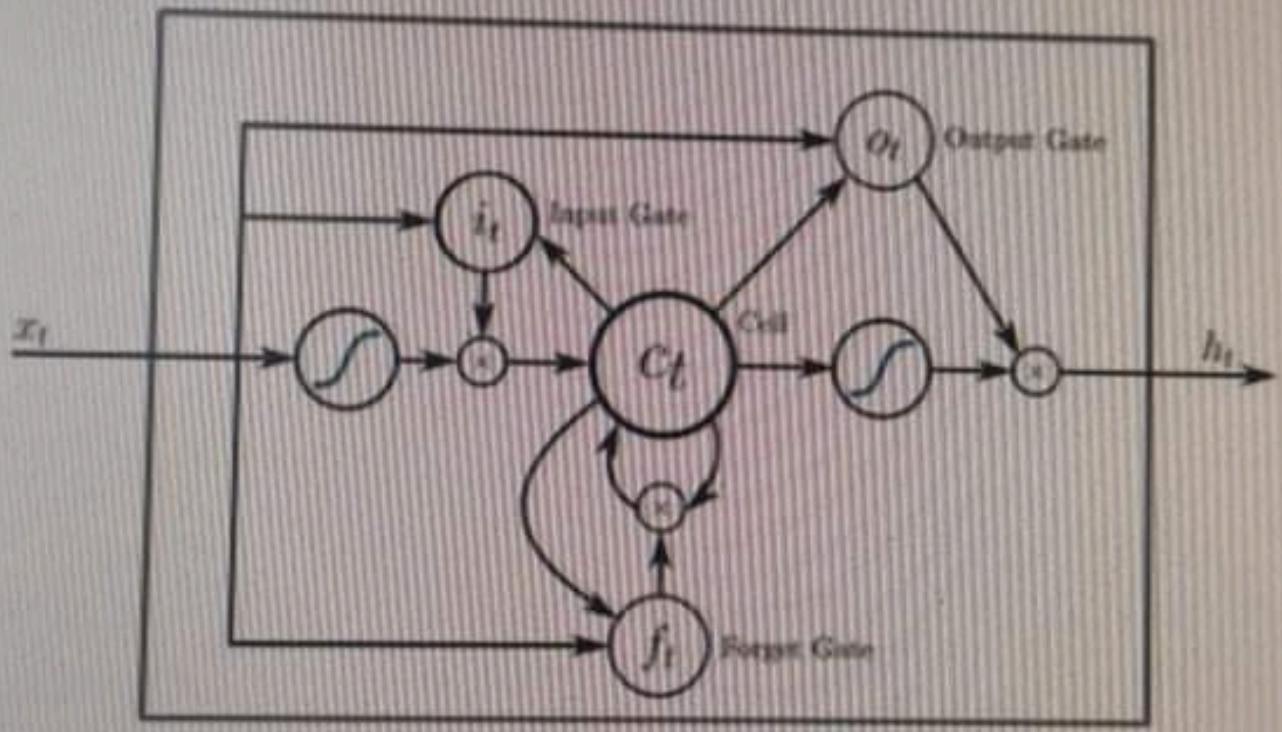
What is update gate and reset gate of GRU(Gated Recurrent Unit)? What is the difference of the work flow of LSTM and GRU? Show your reasons why GRU is faster in compare to RNN??

Add file

## Question 4

Q 1. What is vanishing gradient problem? How does LSTM solve the vanishing gradient

Q 1. What is vanishing gradient problem? How does LSTM solve the vanishing gradient problem? Explain your answer with help of cell gates.




LSTM cell Structure

Question 2:

Major difference of RNN, Peephole-LSTM and GRU is standard RNN suffer from vanishing and exploding gradient problems. LSTMs deals with these problems by introducing new gates such as input and forget gates which allows for a good control over the gradient flow and enable better preservation of long range dependencies. By using increasing number of repeating layer in LSTM long range dependency in RNN is resolved. We can let the gate layers look at the cell state by adding peephole connections. RNNs donot have a cell state, they only have hidden states and those hidden states serves as the memory for RNNs.

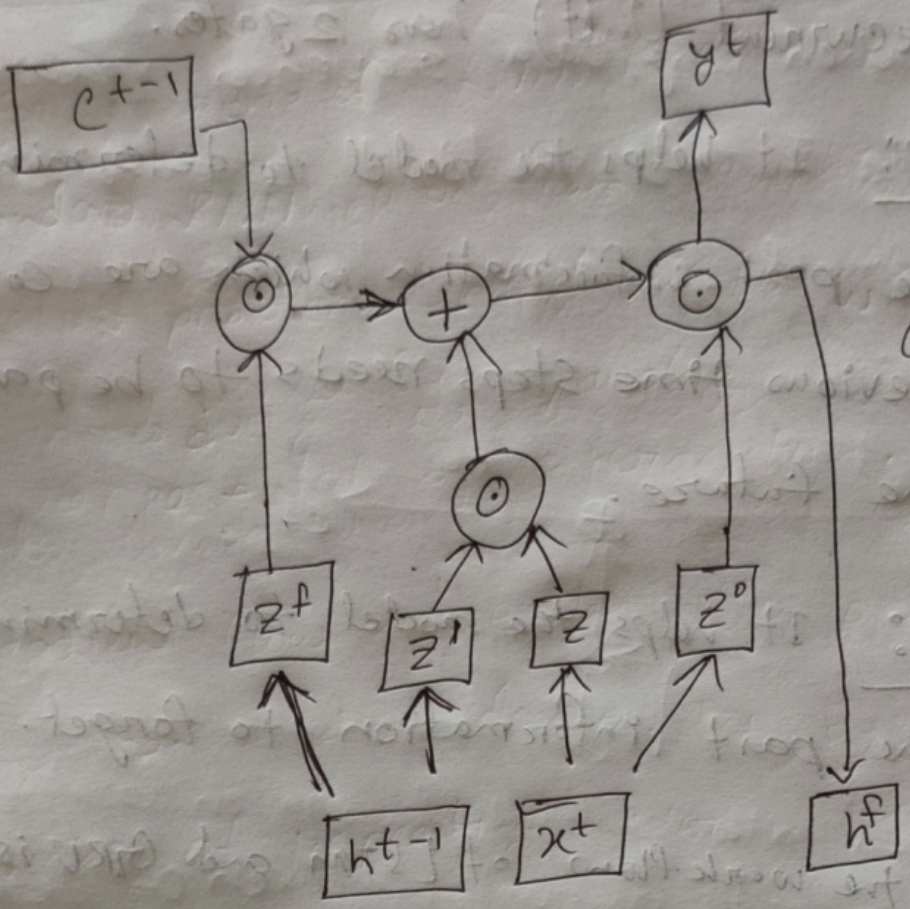
GRU is easy to modify and doesn't require memory units so training speed is faster then

LSTM. LSTM has three doors, whereas, GRU has two doors. Moreover, LSTM has two activation function $Q_1 Q_1$ and $Q_2 Q_2$ while GRU has only one activation function.

The basic difference of LSTM forget gate and GRU's reset gate is forget gate controls what is kept vs forgotten from the previous cell state. It will decide how much information from the previous state should kept and forget remaining.

On contrast, output gate controls controls which parts of the cell are output to the hidden state, It will determine what the next hidden state will be.

# Matrix workflow of LSTM:




$$c^t = z^f \odot c^{t-1} + z^i \odot z$$

$$h^t = z^o \odot \tanh(c^t)$$

$$y^t = \sigma(Wh')$$