Name: Enrolment No:	UPES UNIVERSITY WITH A PURPOSE

UNIVERSITY OF PETROLEUM AND ENERGY STUDIES
Online End Semester Examination, May 2021

Course: Neural Networks. **Semester: VI**
Program: B. Tech. AIML **Time 03 hrs.**
Course Code: CSAI 3001 **Max. Marks: 100**

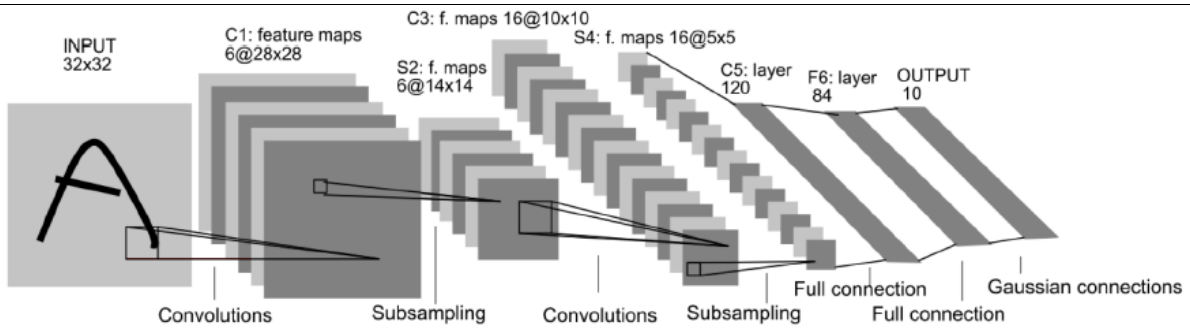
SECTION A

1. **Each Question will carry 5 Marks**
2. **More then one answers may be correct in given choices**
3. **Instruction: Complete the statement / Select the correct answer(s)**

S. No.	**Question**	**CO**
Q 1	Which of the following is/are forms of regularization in neural networks. A. Weight Decay B. L1 Regularization C. L2 Regularization D. Dropout	**CO1**
Q2	Suppose when you are training your convolutional neural network, you find that the training loss just doesn't go down after initialization. What could you try to fix this problem? A. Change the network architecture B. Change learning rates C. Ensure training data is being read correctly D. Find a better model E. Normalize the inputs to the network F. Add a regularization term	**CO2**
Q3	Neural networks A. Optimize a convex objective function B. Can only be trained with stochastic gradient descent C. Can use a mix of different activation functions D. Can be made to perform well even when the number of parameters/weights is much greater than the number of data points.	**CO2**

Q4	Which of the following are true A. The backpropagation algorithm is free from weight initialization B. backpropagation algorithm help to find the learning rate. C. For a saddle point slops in the orthogonal direction is zero. D. The convergence of the NN trained with backpropagation algorithm can not be guaranties.	**CO2**
Q5	Which of the following are true A. Hyperparameters of a model can be selected based on the performance of the validation set. B. The performance evaluation for early stopping is calculated on the test set. C. Larger batch size reduce the variance in the gradient estimation of the SGD. D. Momentum based optimization algorithms converge the NN faster then SGD.	**CO4**
Q6	Which of the following statements are true for an input of dimension H x W x C (stride=1,padding=0) A. 1 x 1 convolution operation reduces H,W,C B. 1 x 1 convolution operation can reduces C but not H,W. C. Pooling can reduce H,W but not C D. Pooling can reduce H,W , C	**CO4**
SECTION B **1. Each question will carry 10 marks** **2. Instruction: Write short / brief notes**		
Q 7	If $\sigma(x) = \frac{1}{1+e^{-x}}$ Prove that $\frac{\partial \sigma}{\partial x} = \sigma(x)(1-\sigma(x))$	**CO3**
Q 8		**CO4**

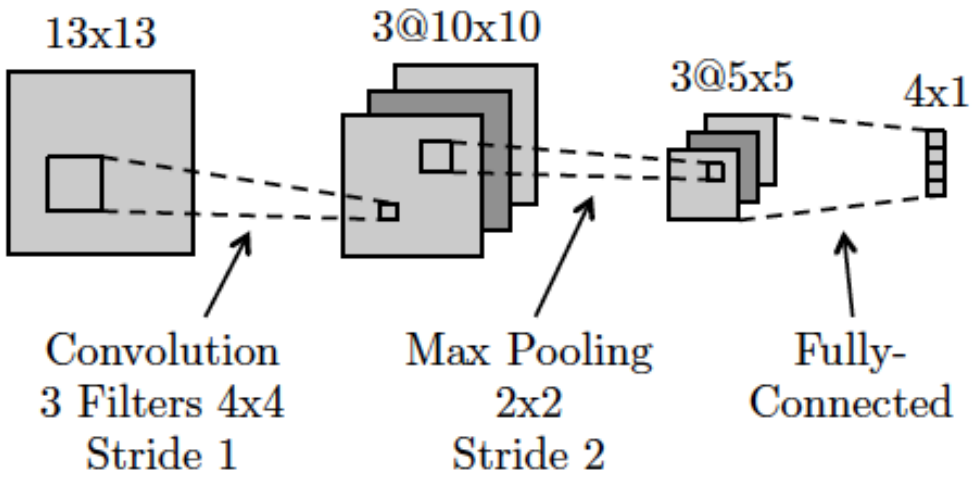
	`` class MyNet(nn.Module): ... def forward(self, x): self.h1 = self.downconv1(x) self.h2 = self.downconv2(self.h1) self.h3 = self.rfconv(self.h2) self.h4 = self.upconv1(torch.cat([self.h3, self.h2], 1)) self.h5 = self.upconv2(self.h4) self.out = self.finalconv(torch.cat([self.h5, x], 1)) return self.out ``` The methods downconv1, rfconv, etc. implement convolution layers. Draw the diagram of the NN architecture implemented in the code given above. .	
Q 9	Devise the Neural Style Transfer algorithm. OR Write a short not on "How to train a GAN?".	**CO5**
Q 10	Write a short note on the importance of the weight initialization in the neural network. Briefly Discuss the various weight initialization techniques.	**CO3**
Q 11	Suppose you want to redesign the AlexNet architecture to reduce the number of arithmetic operations required for each backprop update. 1. Would you try to cut down on the number of weights, units, or connections? Justify your answer.(5) 2. Would you modify the convolution layers or the fully connected layers? Justify your answer.(5)	**CO4**
Section C 1. **Each Question carries 20 Marks.** 2. **Instruction: Write long answer.**		
Q12	Here is the historical LeNet Convolutional Neural Network architecture of Yann LeCun et al. for digit classification that we've discussed in class. Here, the INPUT layer takes in a 32x32 image, and the OUTPUT layer produces 10 outputs. The notation 6@28x28 means 6 matrices of size 28x28.	**CO3**



1. Given that the input size is 32x32, and the Layer 1 size is 28x28, what's the size of the convolutional filter in the first layer (i.e. how many inputs is each neuron connected to)?(5)
2. How many independent parameters (weight and bias) are in layer C1?((5)
3. How many independent parameters (weight and bias) are in layer C3? (5)
4. How many independent parameters (weight and bias) are in layer F6? (5)

OR

Below is a diagram of a small convolutional neural network that converts a 13x13 image into 4 output values. The network has the following layers/operations from input to output: convolution with 3 fillters, max pooling, ReLu, and finally a fully-connected layer. For this network we will not be using any bias/offset parameters (b). Please answer the following questions about this network.



1. How many weights in the convolutional layer do we need to learn?(5)
2. How many ReLu operations are performed on the forward pass? (5)
3. How many weights do we need to learn for the entire network? (5)

	4. What is the disadvantage of a fully-connected neural network compared to a convolutional neural network with the same size layers? (5)	