National University of Computer and Emerging Sciences, Lahore Campus

Course Name:	Deep Learning	Course Code:	CS5102
Program:	MS(DS)	Semester:	Fall 2018
Duration:	180 Minutes	Total Marks:	100
Paper Date:	26 Dec 18	Weight	40
Section:	ALL	Page(s):	11
Exam Type:	Final		

Student : Name:______________________________

Roll No._________________

Section:_______

Instructions

1: Please show all your work. Please write answers on the question paper. You can attach extra sheet to the exam if needed.

2: This is an open book exam.

3: All questions are explained clearly. If you find some question's statement ambiguous, you can make reasonable assumptions as long as you state them clearly.

4: All problems carry equal marks.

Good luck!

Problem 1: The following training patterns are given:

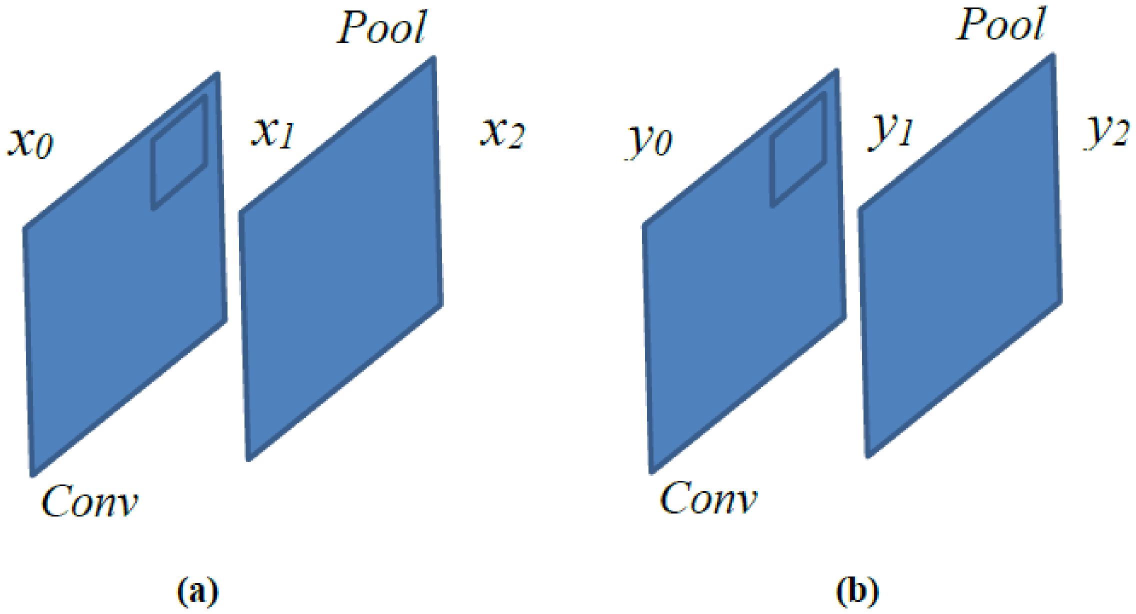
x such that [-3,-2] ϵ C_1 , [0,1] ϵ C_2 and [3,2] ϵ C_1

a) Plot the data points.

b) You have a single neuron with threshold activation and a single input *x* to classify this data. Can we achieve 100% accuracy? Find the weight that gives you the best performance.

c) Can you transform *x* such that you can classify this data with a single neuron? What input transformation $y = f(x)$ makes the data linearly separable?

Problem 2:



(a) **(b)**

A CNN is being used for a classification task, and one part of the CNN is shown above in (a). The input image is passed through a convolutional layer followed by a max-pooling layer to produce outputs x_1 and x_2 respectively.

The same CNN is passed another input $y_0 = cx_0$ where $c > 0$ as shown in (b).

(a): How is y_1, the output of convolutional layer with new input, related to x_1? Justify your answer. In the light of this, what can you say about effect of scaling a convolutional layer's input by a positive constant?

(b): How is y_2, the output of pooling layer with new input, related to x_2? Justify your answer. In the light of this, what can you say about effect of scaling a max-pooling layer's input by a positive constant?

(c) : As a result of (a) and (b), can you say that a CNN containing convolutional and pooling layers only will have no effect on scaling the input by c > 0.

(d) : What if the input is scaled by c < 0? Will y_2 and x_2 still maintain the same relationship? If not, which layer will change?

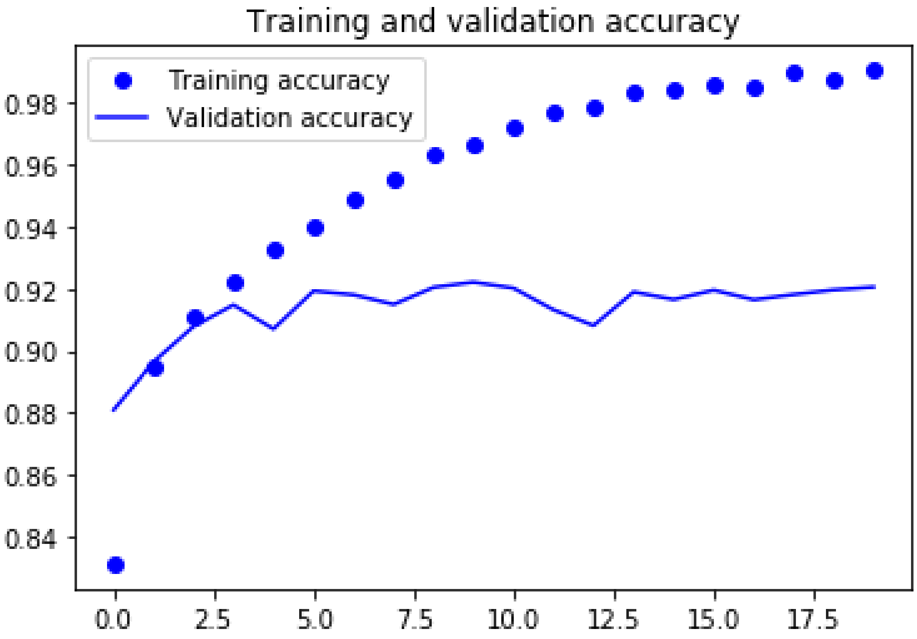
Problem 3:

(a): In the generator/discriminator two player game discussed in class, we intend to make a modification.

We have a pre-trained MNIST classifier (discriminator) which can give us the probability of input character *not being* 6 ($p(xi)$ is not equal to 6). We need to use the generator to create new images of character 6 which look like actual 6 images.

Write the generator-discriminator adversarial cost function for this problem. Also write the generator and discriminator updates.

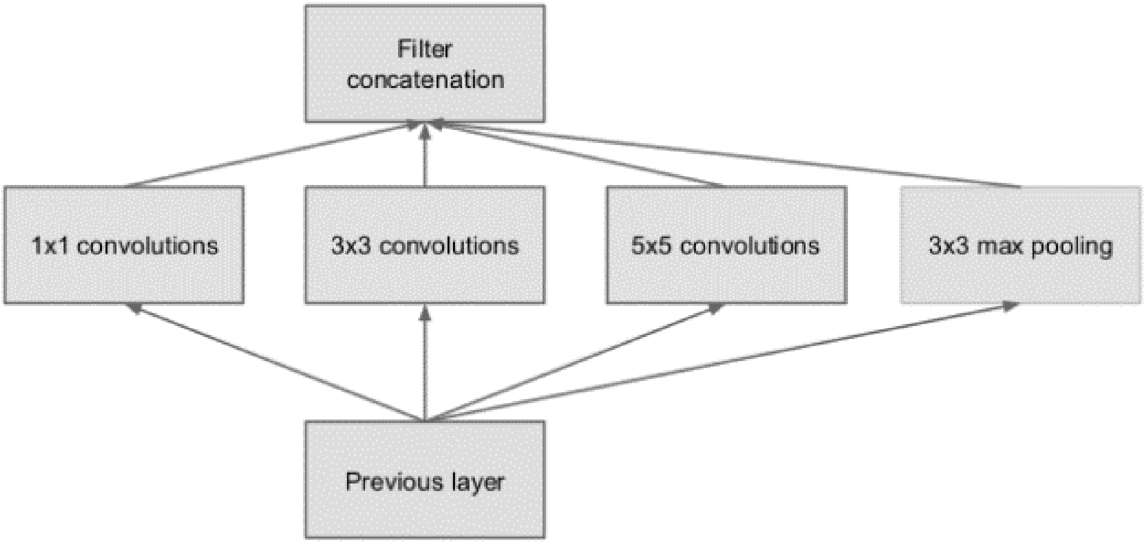
(b): For a given neural network, training produces the following plot.



What is most likely wrong with this training?

Problem 4:

A single inception module (shown below) is given an input of size (128 x 128 x 32). Each convolution layer has 32 filters in it.



(a): Find the values of P (zero padding), S(stride) and F(filter size) and K(number of filters) for each layer so that each layer produces an output of (64 x 64 x 32).

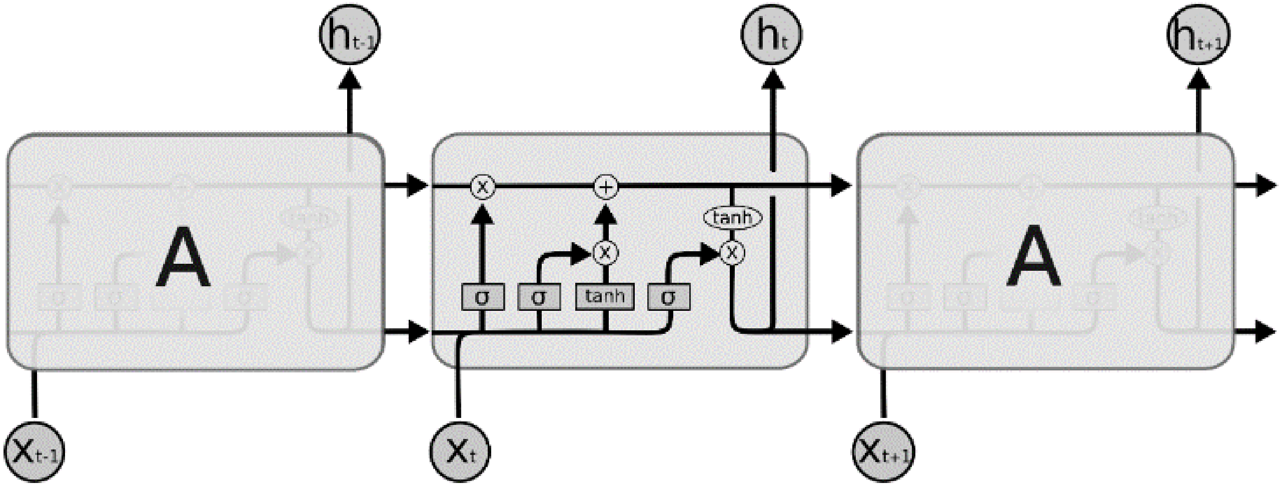
(b): What is the size of the total output after filter concatenation? How can we reduce the depth of the inception module output if we require an input of (64 x 64 x 24) in next stage?

Problem 5: In a video processing problem, you intend to predict the next scene type (action, drama, conversation, scene change) in a movie, based on the previous video frames.

The classification depends not just on the current frame, but also the movie type (T), cast likelihood(C) and previous scene (P).

(a): Out of VGG, GoogleNet, ResNet, RNN and LSTM, what type of convolutional neural network will you use to model this problem and why? Draw a rough diagram of where each input will go (T, C, P, x_i) in your model.

(b): For the LSTM shown below, modify the machine for a problem with memory delay. Specifically, change the LSTM machine such that current input x_i does not have an effect on current output but only affects the output after *n* instances.


h_{t-1}
h_t
h_{t+1}
A
σ
σ
tanh
σ
tanh
A
X_{t-1}
X_t
X_{t+1}