

IT549: Deep Learning

2nd In-Semester Exam (22nd March 2024)

DA-IICT, Gandhinagar

Time Duration: 30 Minutes

Total Marks: 27

Q.1 What do you mean by data-normalization? Why is it required for training Machine Learning/Deep Learning models? (3 points)

Q.2 What do you mean by model regularization? Why is it required? How do you know whether the Deep learning model is overfitted or well-fitted? (2 points)

Q.3 Why is stochastic gradient descent(SGD) better than standard gradient descent? Write down the algorithm (pseudocode or steps) for SGD. (3 points)

Q.4 What do you mean by homogenous ensemble learning? How do you perform it for deep neural networks? (2 points)

Q.5 What do you mean by the parameter sharing concept in a convolutional neural network (CNN) or recurrent neural network? (2 points).

Q.6 What do you mean by skip connection in DL? What are its advantages and disadvantages? (2 points)

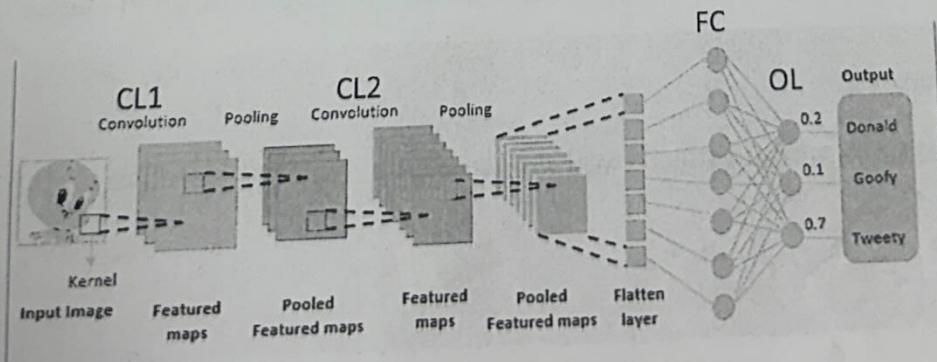
Q.7 One marketing company approach you for developing an ML/DL/AI model to estimate or predict the customer preference of their 100 products. The company has a huge dataset of customer information and products they brought.

What are the (at least two) questions you will ask the marketing company? Why? (2 points)

what will you look for in data before modelling a deep learning model? (1 marks)

Q.8 Compute the no. of parameters (CL1, CL2, FC-OL) in the following CNN architecture for classification. Which activation function is in the output layer, and which loss will you use? (3 points)

The input image size is 16x16 (grayscale). The kernel/filter size is 5x5 on the CNN layer. There are 10 filters in the CL1 layer and 20 in the CL2 layer. The convolution type is "same". Polling is 2x2. Stride is one everywhere.



Q.9 Given the following matrix, compute the performance measures, e.g. sensitivity and specificity for the 'Like' class from the following confusion matrix. (2 points)

		Predicted		
	classes	Like	Neutral	Dislike
Actual	Like	200	25	100
	Neutral	50	400	100
	Dislike	40	50	100

500 - Pred
- 100 - Wrong -

100 Q.10 Five class classifier deep learning models predicted 100 examples as class "C1" out of 500, which was the wrong prediction. There are 800 examples of "C1" class in the test set. The total no. of examples is 5000 in the test set. Compute the precision and recall. (2 points)

Q.11 (3 points)

Answer True or False. 0.5 points for the correct answer and -0.25 for the wrong answer, 0 points if you do not attempt it.

- (a) The node-dropout technique reduces the number of parameters in the network.

- (b) In the vanilla CNN model, If the input image size is increased, number of parameters will increase in CNN layers. _____
- (c) In ADAM optimization, the update step size of each parameter is inversely proportional to first-order gradient movement (e.g. error gradient). _____
- (d) A standard recurrent neural network (RNN or LSTM) requires the same sequence length for each observation. _____
- (e) Batch normalization has four learnable parameters per neuron in Dense ANN. _____
- (f) An increase in model capacity will reduce model bias but increase model variance. _____