DHARMSINH DESAI UNIVERSITY, NADIAD
FACULTY OF TECHNOLOGY
THIRD SESSIONAL
SUBJECT: Machine Learning (CE622)

Examination	**: B.Tech Semester VI**	**Seat No.**	**: 103**
Date	**: 17/03/2025**	**Day**	**: Monday**
Time	**: 2:30 PM to 3:45 PM**	**Max. Marks**	**: 36**

INSTRUCTIONS:
1. Figures to the right indicate maximum marks for that question.
2. The symbols used carry their usual meanings.
3. Assume suitable data, if required & mention them clearly.
4. Draw neat sketches wherever necessary.

Q.1 Do as directed. **[12]**

CO4 N (a) What problem can occur in the training of multilayered feedforward network if the learning rate is very high? [2]

CO4 U (b) State true or false and justify your answer, "The gradient descent algorithm always gives optimal solution." [2]

CO3 R (c) Which machine learning / deep learning model is best suited to develop automatic speech recognition system? [2]

CO3 U (d) State whether the following statements are true or false: [1]
i) 1x1 filter can be used to create a linear projection of a stack of feature maps.
ii) 1x1 filter can be used to increase the number of feature maps in a model.

CO3 R (e) Write the equations for backpropagation through the convolutional and input layers in a CNN. Explain the terms used in the equations. [2]

CO3 N (f) Consider an Image of 3x3 and a filter of size 3x3. Compute the output of the max pooling layer for the network given in figure 1.1. For the convolution layer take stride size =1 and padding size =1. For max pooling take filter size = 2 and stride size=1. [3]

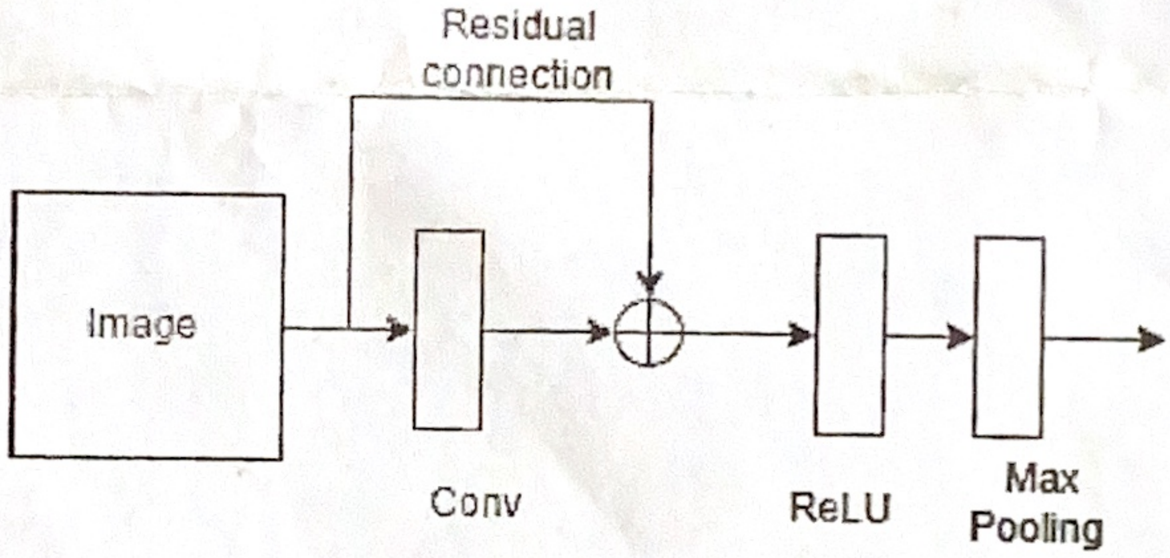

Figure 1.1

Image

0.5	0.3	0.2
0.3	0.2	0.4
0.4	0.4	0.6

Filter

1	0	-1
1	0	-1
1	0	-1

Q.2 Attempt *Any TWO* from the following questions. **[12]**

CO3 N (a) Describe the error back propagation algorithm to train Multilayered Feedforward Neural Network. Derive the formula to update output layer and hidden layer weights. [6]

CO3 N (b) Describe the problem of part of speech tagging of the given English sentence. Also, describe the technique to solve the problem. [6]

CO3 N (c) What is generative AI? Which are the prominent models of generative AI? What are the applications of generative AI? [6]

Q.3 Attempt the following questions. **[12]**

CO3 C (a) Consider the following CNN classifier. Input image is represented as WxHxC where W and H represent Width and Height of the image and C represents the number of channels. [6]

INPUT(256×256×3)—>CONV(F=9,K=64)—>CONV(F=5,K=128)---→POOL(F=2,S=2) —>BATCHNORM()------>CONV(F=3,K=128)---->CONV(F=5,K=64) →POOL(F=2,S=2) —>FC(N=512)--->DropO(0.5)---->FC(N=128)----- > FC(N=10).

In the given network, CONV denotes a convolutional layer with K filters each of size FxF and each having stride and padding parameters 1 and 0 respectively. POOL indicates an FxF pooling layer with padding 0 and stride size S. FC stands for a fully-connected layer with N neurons. BATCHNORM stands for batch normalization layer. DropO stands for dropout layer. For each layer, **calculate the number of weights, number of biases and the size of the associated feature maps.**

CO3 N (b) i) Explain in detail how the **Inception architecture** differs from traditional CNNs using appropriate example. [6]

ii) **In a traditional CNN architecture,** an input image of size 7×7×512 produces an output of size 1×1×4096 after a convolution operation. Identify the size and number of filters used. Is it possible to replace a fully connected layer by a convolution layer? If yes, explain in brief how?

OR

CO3 C (a) i) Assume the following documents. [6]

D1	Italy is world champion.
D2	Germany and Italy played each other in Semifinal.
D3	Germany was in the semifinal.
D4	Germany won the semifinal.

Consider (is, was, in, the, each, other, and) as stop words. Create TF-IDF feature matrix for all the unique words in given dataset. Show all the necessary calculations.

ii) Explain how the CBOW (Continuous Bag of Words) architecture converts words into vector format with an example.

CO3 N (b) Describe the working of depthwise separable convolutions and how they contribute to the efficiency of the MobileNet architecture. Provide an example to illustrate the working of depthwise separable convolutions. [6]