VIT Vidyalankar Institute of Technology
(Accredited A+ by NAAC)
(Autonomous Institute Affiliated to University of Mumbai)

Mid Semester Examination

Branch	Date	Sem.	Roll No. / Exam Seat No.	Subject	Student's Signature	Junior Supervisor's Name and Sign
CMPN	16/04	VIII	MSE-3	DL		

Question No.	A	B	C	D	E	F	G	H	Total	Total out of (20 /30 / 40)
1										
2										
3										
4										

Examiners Signature	Student's Sign (After receiving the assessed answer sheet)

Q 1A) An autoencoder is a type of ANN used for unsupervised learning. It is designed to learn efficient representation of Input data, used for the purpose of dimensionality reduction, feature learning or data denoising.

The basic architecture of an autoencoder consists of 3 main component An Encoder, a decoder and a bottleneck layer.

Encoding

$$z = W_1 x + b_1$$

Where x is the input data, W_1 is the encoding weight matrix, b_1 is the encoding bais vector, z is the code.

Decoding:

$$x' = W_2 z + b_2$$

Where W_2 is the decoding weight matrix, b_2 is the decoding bais vector & x' is the reconstructed data.

The goal of the Autoencoder is to minimize the reconstruction Error between X and x'

Autoencoder has various architectures including stacked autoencoder and has powerful tools for feature learning and representation learning in deep learning.

Q1B) In Backpropagation once the gradients of the loss with respect to the weights are computed, the weights are updated to minimize the loss
The delta rule specifies how much each weight should be adjusted based on the gradient of the loss with respect to that weight.
The updated weights are calculated as

$$\Delta w_{ji} = \eta \delta_j o_i$$

$$\delta_j = o_j(1-o_j)(t_j-o_j)$$ if j is an output unit

$$\delta_j = o_j(1-o_j)\sum_k \delta_k w_{kj}$$ If J is an hidden unit.

Where η is a constant called the learing rate

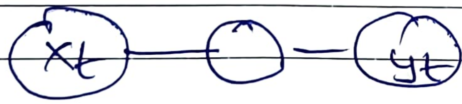
t_j is the correct teacher output for unit j

δ_i is the error measure for unit i.

Q. A recurrent neural network is a type of AI network which uses sequential data or time series data. The basic architecture of RNN consists of Input x_t output y_t & at each time step we unfold the network for k time steps to get the output at time step k+1.

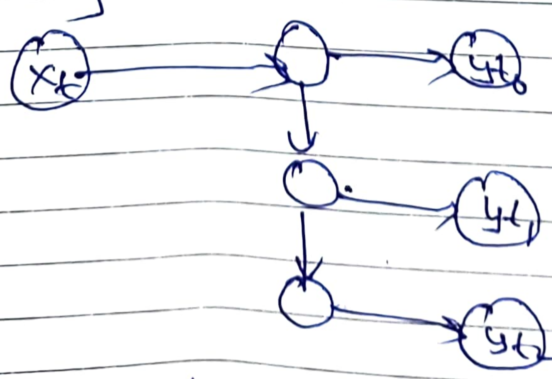
different types of RNN with varying architecture are

(1) One to one



there is single (x_t, y_t) pair.

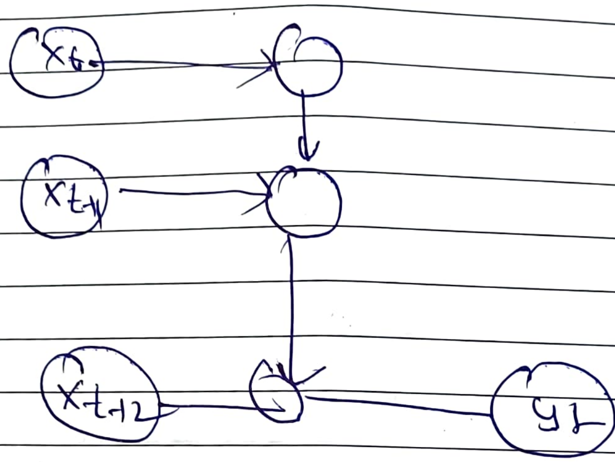
(2) One to many



a single i/p at x_t can produce multiple o/p eg (y_{t0}, y_{t1}, y_{t2})

music generation is the example

3) Many to one



In this case many inputs from different time steps produce a single output example (x_t, x_{t+1}, x_{t+2}) can produce a single o/p y_t such n/w are employed in sentimental analysis or emotional detection

Q2

A) In Long Short Term memory (LSTM) cell there are 3 main gates Input gate, the forget gate and the output gate.

(1) Input Gate:- Controls the flow of new information into the cell state
(2) It takes the input x_i and passes them through the sigmoid activation function
(3) The output of the sigmoid function determine how much of the new information should be added to the cell state

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i)$$

(2) Forget gate:
The Forget gate controls the flow of information from the previous cell state C_{t-1} to the current cell state C_t
It decides which cell state should be retained and which should be discarded

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$$

(3) Output gate.
The output gate ~~controls~~ controls the flow of information from the current cell state C_t to the current hidden state h_t.
It determine which parts of the cell state should be exposed as the output

of LSTM cell.
mathematically the output gate

$$o_t = \sigma(W_o [h_{t-1}, x_t] + b_o).$$

Q2 B) In CNN the stride parameter determine the step size at which the CN filter moves across the input data or feature maps. It controls how much the filter shift over the input data during the convolution operation.

The filter moves horizontally and vertically across the i/p data by the specified number of pixels. This result in downe sampling or reduction in the spatial dimension of the o/p feature mape depending on the size of the input data & the stride value

Example

Suppose we have input image with 5x5 dimension & 3x3 filter with a stride 2

$$\begin{bmatrix} 1, & 2, & 3, & 4 & 5 \\ 6 & 7 & 8 & 9 & 10 \\ 11 & 12 & 13 & 14 & 15 \\ 16 & 17 & 18 & 19 & 20 \\ 21 & 22 & 23 & 24 & 25 \end{bmatrix}$$

Input

~~Kerne~~ ~~Kernel~~

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

Kernel

with stride 2, the

$$\begin{bmatrix} 4, & 6 & 8 \\ 17 & 19 & 11 \\ 24 & 26 & 28 \\ 44 & 46 & 48 \end{bmatrix}$$

Q3(A) ~~Input~~

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 10 \\ 11 & 12 & 13 & 14 & 15 \\ 16 & 17 & 18 & 19 & 20 \\ 21 & 22 & 23 & 24 & 25 \end{bmatrix}$$

Input

$$\begin{bmatrix} 0.1 & 0.2 & 0.3 \\ 0.1 & 0.2 & 0.3 \\ 0.1 & 0.2 & 0.3 \end{bmatrix}$$

Kernel.

Convolution operation across the input matrix element-wise multiplication & Sum the result

Feature map

$$\begin{bmatrix} 4.6 & 7.9 & 11.2 \\ 14.1 & 17.4 & 20.7 \\ 23.6 & 26.9 & 30.2 \end{bmatrix}$$

Now we apply Relu activation function to each activation ~~function~~. feature map

$$\begin{bmatrix} 4.6 & 7.9 & 11.2 \\ 14.1 & 17.4 & 20.9 \\ 23.6 & 26.9 & 30.2 \end{bmatrix}$$

Q3B LeNet is a CNN that Yann LeCun introduced in 1989.

LeNet is a common term for LeNet-5

Feature of LeNet-5
Every convolution layer include three parts: Convolution, pooling and nonlinear activation functions

The average pooling layer is used for sub sampling

tanh is used as the activation function
Using multi-layered perception or fully connected layers as the last classifier
It consists of 7 layers. The first layer consists of an input image with the 32 x 32. It is convolved with 6 filters of size 5x5 resulting in dimension of 28 x 28 x 6. The second layer is pooling operation which filters size 2x2 and stride 2. Hence the resulting image dimension will be 14 x 14 x 6.
The third layer also involves in a convolution operation with 16 filters of size 5x5 followed by 4th pooling layer & filter size of 2x2 & stride = 2
The resulting dimension is 5x5x16

Summary of Lenet 5 - Architecture

	Layer	Feature map	Size	Kernel size	stride	Activation
	Image	1	32x32			
1	Convolution	6	28x28	5x5	1	tanh
2	Average pooling	6	14x14	2x2	2	tanh
3	Convolution	16	10x10	5x5	1	tanh
4	Avg pooling	16	5x5	2x2	2	tanh
5	Convolution	120	1x1	5x5	1	tanh
6	FC	-	84			tanh
o/p	FC	-	10			softmax