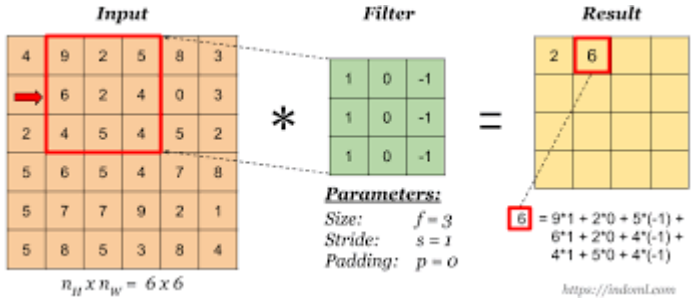
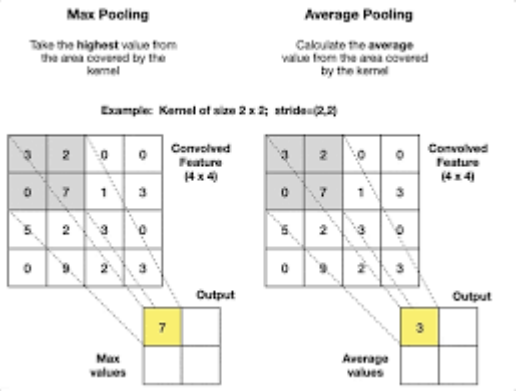
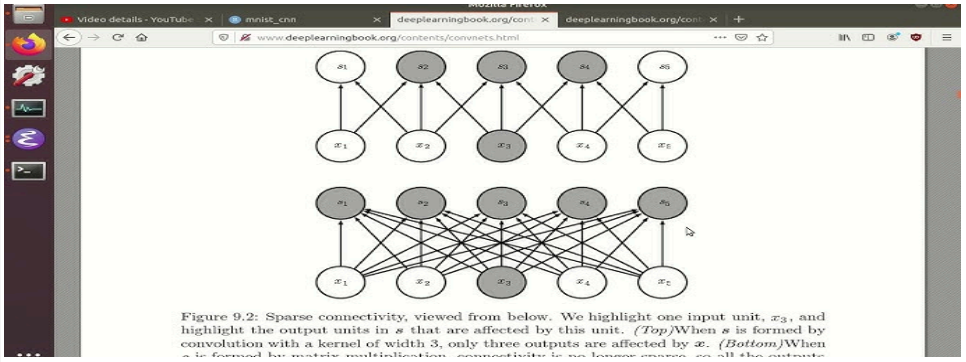
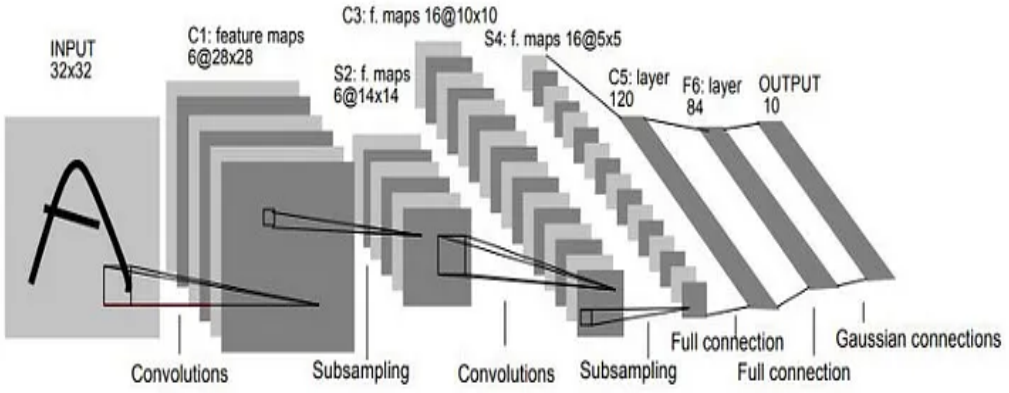
USN | | | | | | | | | | |

Internal Assessment Test 3 Answer Sheme – December 2024

Sub:	DEEP LEARNING ANSWER SCHEME			Sub Code:	21CS743	Branch:	AInDS
Date:	17/12/2024	Duration:	90 minutes	Max Marks: 50	Sem	VII	OBE

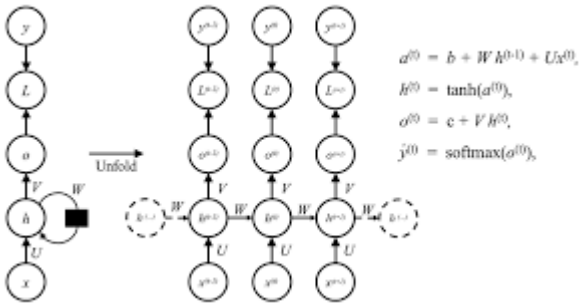
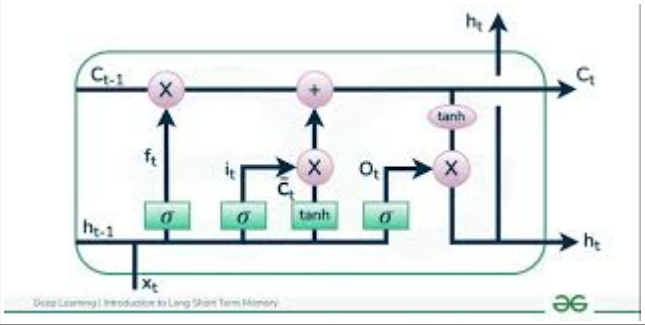
		Answer any FIVE Questions	MARKS	CO	RBT
1		**What is convolution Network? Explain about Convolution operation and Pooling** A convolutional neural network (CNN) is a type of artificial neural network used primarily for image recognition and processing, due to its ability to recognize patterns in images **(2 Marks)** The convolution operation, the main part of the CNN, applies specific filters or kernel functions to a selected region of the image to detect local features. In other words, by convolution, it is possible to focus on a specific feature of the image at a time by applying specific filters. **(3 marks)**  **(2 Marks)** Pooling is a technique used in Convolutional Neural Networks (CNNs) to downsample the spatial dimensions of the input feature maps, reducing the amount of computation and parameters in the network while retaining important information. (**3 marks**)  (**2 Marks**)	10	CO3	L2

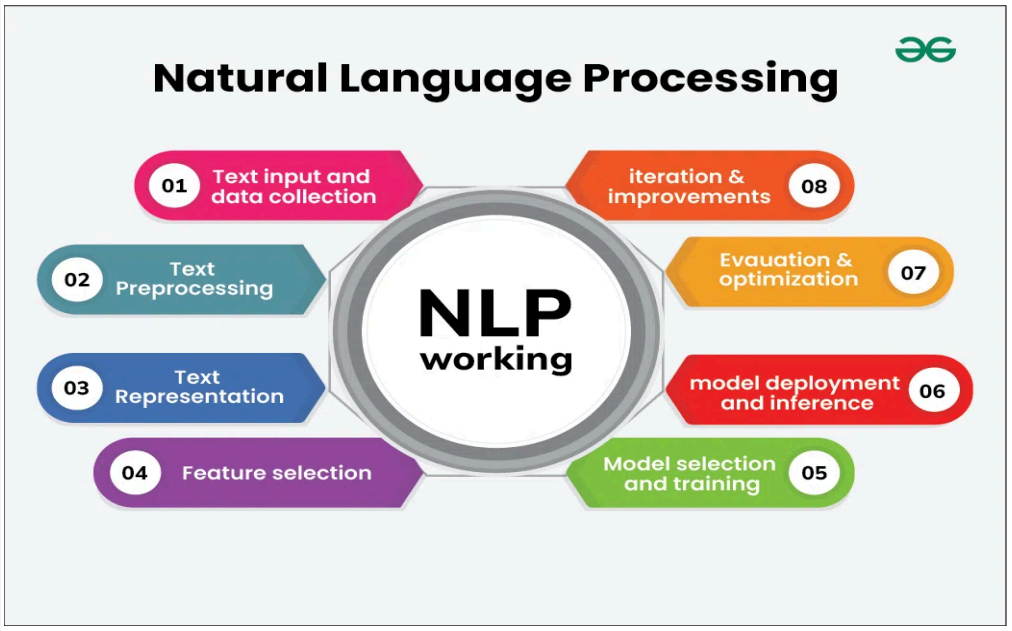
2	a	**Explain Sparse Interactions and Parameter Sharing in convolutional network** Parameter sharing refers to using the same parameter for more than one function in a model. In a traditional neural net, each element of the weight matrixis used exactly once when computing the output of a layer. It is multiplied byone element of the input and then never revisited **(3 marks)** Sparse Interactions operation enables the extraction of hierarchical features from input data. Sparse Interactions: Sparse interactions within convolution optimize computation by focusing on essential connections, reducing computational complexity **(2 Marks)**  Figure 9.2: Sparse connectivity, viewed from below. We highlight one input unit, x_3, and highlight the output units in s that are affected by this unit. *(Top)*When s is formed by convolution with a kernel of width 3, only three outputs are affected by x. *(Bottom)*When	5	CO3	L3
	b	**List Examples of data types with different dimensionalities and number of Channels** The data used with a convolutional network usually consists of severalchannels, each channel being the observation of a different quantity at somepoint in space or time. **(2 Marks)** Gradients obtained by back-propagation may then be used with any general-purpose gradient-based techniques to train an RNN | Dimensions	Single channel	Multichannel | |---|---|---| | 1D | **Raw audio** (single amplitude value per time point) | **Skeleton animatin data** (orientation of each joint) | | 2D | **Audio spectrogram** (one FFT coefficient per time point per frequency) | **Color image** (RGB triplet per (x,y) tuple) | | 3D | **CT scan** (one value per (x,y,z) tuple) | **Color video** (one RGB triplet per (x,y) tuple per time instant) | **(3 Marks)**	5	CO3	

3		**Explain about LeNet in detail.** The LeNet-5 signifies CNN's emergence and outlines its core components. However, it was not popular at the time due to a lack of hardware, especially GPU (Graphics Process Unit, a specialised electronic circuit designed to change memory to accelerate the creation of images during a buffer intended for output to a show device) and alternative algorithms, like SVM, which could perform effects similar to or even better than those of the LeNet. **(1 Mark)** **Features of LeNet-5**Every convolutional layer includes three parts: convolution, pooling, and nonlinear activation functionsUsing convolution to extract spatial features (Convolution was called receptive fields originally)**The average pooling layer** is used for subsampling.**tanh**' is used as the activation functionsing **Multi-Layered Perceptron** or **Fully Connected Layers** as the last classifierThe sparse connection between layers reduces the complexity of computation (4 Mark)  (	10	C03	L2

Layer		Feature Map	Size	Kernel Size	Stride	Activation
Input	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	Average Pooling	16	5x5	2x2	2	tanh
5	Convolution	120	1x1	5x5	1	tanh
6	FC	-	84	-	-	tanh
Output	FC	-	10	-	-	softmax

(5 Marks)

4		Discuss about RNN and how to compute the gradient in RNN. A recurrent neural network (RNN) is a deep learning model that is trained to process and convert a sequential data input into a specific sequential data output. **(2 marks)**  $a^{(t)} = b + W h^{(t-1)} + U x^{(t)},$ $h^{(t)} = \tanh(a^{(t)}),$ $o^{(t)} = c + V h^{(t)},$ $\hat{y}^{(t)} = \text{softmax}(o^{(t)}),$ **(3 Marks)** Gradients obtained by back-propagation may then be used with any general-purpose gradient-based techniques to train an RNN **(5 marks)** $\nabla_c L = \sum_t \left(\frac{\partial o^{(t)}}{\partial c}\right)^\top \nabla_{o^{(t)}} L = \sum_t \nabla_{o^{(t)}} L$ $\nabla_b L = \sum_t \left(\frac{\partial h^{(t)}}{\partial b^{(t)}}\right)^\top \nabla_{h^{(t)}} L = \sum_t \text{diag}\left(1 - \left(h^{(t)}\right)^2\right) \nabla_{h^{(t)}} L$ $\nabla_V L = \sum_t \sum_i \left(\frac{\partial L}{\partial o_i^{(t)}}\right) \nabla_V o_i^{(t)} = \sum_t (\nabla_{o^{(t)}} L)\, h^{(t)\top}$ $\nabla_W L = \sum_t \sum_i \left(\frac{\partial L}{\partial h_i^{(t)}}\right) \nabla_W h_i^{(t)}$ $= \sum_t \text{diag}\left(1 - \left(h^{(t)}\right)^2\right) (\nabla_{h^{(t)}} L)\, h^{(t-1)\top}$ $\nabla_U L = \sum_t \sum_i \left(\frac{\partial L}{\partial h_i^{(t)}}\right) \nabla_U h_i^{(t)}$ $= \sum_t \text{diag}\left(1 - \left(h^{(t)}\right)^2\right) (\nabla_{h^{(t)}} L)\, x^{(t)\top}$	10	C03	L2
5	a	**Explain LSTM with Block Diagram.** Long Short-Term Memory, is a type of recurrent neural network (RNN) that uses gates to capture both short-term and long-term memory. LSTMs are designed to process and retain information over multiple time steps. They are widely used in deep learning and are ideal for sequence prediction tasks **(2 Marks)**  **(3 Marks)**	5	C03	L2

	b	**Discuss about the components of the Conventional neural Network layer.** The convolutional layer is the core building block of a CNN, and it is where the majority of computation occurs. It requires a few components, which are input data, a filter and a feature map. Let's assume that the input will be a color image, which is made up of a matrix of pixels in 3D. **(3 Marks)** Diagram **(2 marks)**	5	C03	
6		**Explain about Natural Language Processing** NLP powers many applications that use language, such as text translation, voice recognition, text summarization, and chatbots. You may have used some of these applications yourself, such as voice-operated GPS systems, digital assistants, speech-to-text software, and customer service bots **(2 Marks)**  **(3 Marks)** Working of Natural Language Processing (NLP) (**5 Marks)** **Text Input and Data Collection** **Text Preprocessing** **Text Representation** **Feature Extraction** **Model Selection and Training** **Model Deployment and Inference** **Evaluation and Optimization** **Iteration and Improvement**	10	C05	L2

CI **CCI** **HOD**