Question Paper

Exam Date & Time: 02-May-2024 (02:30 PM - 05:30 PM)

MANIPAL ACADEMY OF HIGHER EDUCATION

SIXTH SEMESTER B.TECH END SEMESTER EXAMINATIONS, MAY 2024

DEEP LEARNING [CSE 3271]

Marks: 50 **Duration: 180 mins.**

Answer all the questions.

Instructions to Candidates: Answer ALL questions Missing data may be suitably assumed

1)	A)	What are the advantages of deep learning models over machine learning models? With a neat, labelled graph, illustrate how different deep neural networks improve their performance compared to machine learning models with the increasing amount of data.	(4)
	B)	What is the significance of gradient based optimization techniques in computation graph of a deep neural network? Consider the Sigmoid function whose inputs are given by the output of a linear model $w_1 * x_1 + w_2 * x_2$. Draw the computation graph and mark all the gradients.	(3)
	C)	Consider a feed forward network with n inputs and m outputs. What is the size of model parameters, input and output? A 4-input neuron has weights 1, 2, 3 and 4. The transfer function is linear with the constant of proportionality being equal to 2. The inputs are 4, 10, 5 and 20 respectively. What is the output?	(3)
2)	A)	What are the limitations of feed-forward neural networks? Show how convolution neural networks overcome the limitation? For the following cases, explain how you would arrive at the output of same size as input by applying convolution operation. Justify your answers pictorially. i. Input image of size 7*7 and filter size of 3 * 3 ii. Input image of size 7*7 and filter size of 5 * 5	(4)
	B)	In the context of linear regression, what is the objective function J(w) that is minimized when applying L2 regularization ? Derive the expression and show how does L2 regularization help in preventing overfitting in linear regression models?	(3)
	C)	List and explain the use-cases of pretrained models in transfer learning.	(3)
3)	A)	Illustrate how dropout differ from L1 or L2 regularization techniques. Are there any benefits of using dropout regularization compared to L1 or L2 regularization? Explain in detail. Suppose you have a neural network with an input layer of 500 neurons, two hidden layers of 200 neurons each, and an output layer of 10 neurons. If dropout regularization is applied to the hidden layers with a dropout rate of 0.3, how many neurons in total are expected to be dropped out during a single forward pass?	(4)
	B)	How can you determine the appropriate level of data augmentation for a given dataset and model architecture? You have a dataset of 1000 images, and you want to apply random rotation as a data augmentation technique. If you rotate each image by a random angle between -10 and 10 degrees, how many possible unique rotations are there for each image? How many additional augmented samples are added now ?	(3)

	C)	How can you determine the maximum number of epochs for a given problem with early stopping enabled? Explain with a neat, labelled graph marking the early stopping checkpoint.	(3)
4)		Describe how parameter sharing works within the context of RNNs, highlighting its role in capturing temporal dependencies. Provide examples to demonstrate the impact of parameter sharing on the efficiency and effectiveness of RNN training.	(4)
	A)		
	B)	Analyse the mathematical formulation of the LSTM cell state update equations. How do the input gate, forget gate, and output gate interact to regulate the flow of information through the network, and how does this contribute to the model's ability to learn complex temporal patterns?	(3)
	C)	List the potential drawbacks of using a traditional autoencoder.	(3)
5)		Devise a strategy leveraging autoencoders to mitigate image noise effectively, ensuring a high-quality restoration process	(4)
	A)		
	B)	Explain the role of the encoder and decoder in a Variational Autoencoder, emphasizing their functions in the generative process.	(3)
	C)	Compare the number of learnable parameters in RNN and LSTM	(3)

-----End-----