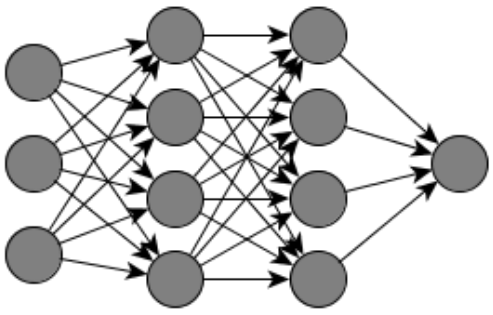
M.Sc. Computer Science
MCSE304: Deep Learning
Unique Paper Code: 223412304
Semester III
OBE Examination December 2021
Year of Admission: 2020

Time: Three Hours **Max. Marks: 70**

Note: Answer any 4 questions. All questions carry equal marks.

1.	a) Consider a dataset of computerized tomography (CT) scans. You want to build a prediction model for automated diagnosis the lower back pain (LBP) using this dataset. Will it be feasible to use transfer learning for this task? Justify your answer. Briefly explain the purpose of spatial pyramid pooling (SPP) in CNN? b) Explain the role of VALID padding in convolutional layer. How does it differ from SAME padding? For the dataset in part (a), which padding would you prefer for automated diagnosis of LBP? c) You validated the LBP diagnostic model built in part (a) on an independent cohort of CT scans. It is observed that the model did not perform well on this cohort. What could be the possible reason for the same?
2.	a) Why do we use autoencoders when there are already powerful techniques PCA (Principal Component Analysis) for dimension reduction? b) Consider a 3-layer neural network. What problem arises when you initialize the weights to be zero for all the layers? How can you avoid that problem? c) Consider the following neural network, comprising four layers, called layer 1, layer2, layer3, and layer4. $W^{[i]}$ and $b^{[i]}$ denote the weights and biases of layer i.  Determine the dimensions of following arrays: • $W^{[1]}$

<table>
<tr><td></td><td>

$b^{[1]}$
$W^{[2]}$
$b^{[2]}$
$W^{[3]}$
$b^{[3]}$
$W^{[4]}$

</td></tr>
<tr><td>3.</td><td>a) Consider a dataset of COVID-19 with five attributes like name, age, temperature, BR (breathing rate), and HBR (heartbeat rate). If there are some irregularities in the values corresponding to BR and HBR attributes, enumerate all possible problems encountered in model design. How regularization helps in resolving those problems?
b) What do you understand by learning rate in a neural network model? What happens if the learning rate is too high or too low?</td></tr>
<tr><td>4.</td><td>a) Consider a task of predicting financial risk using a dataset having five independent variables (<code>Location_Score</code>, <code>Internal_Audit_Score</code>, <code>External_Audit_Score</code>, <code>Fin_Score</code>, and <code>Loss_score</code>) and one binary dependent variable <code>IsUnderRisk</code>. The confusion matrix generated for the test set is as follows:

<table>
<tr><td></td><td>Risk</td><td>No-Risk</td></tr>
<tr><td>Risk</td><td>500</td><td>300</td></tr>
<tr><td>No-Risk</td><td>200</td><td>1100</td></tr>
</table>

Compute the metrics accuracy, recall, precision, and F1-score. If the above dataset also includes some past risk results, which model is most appropriate in that scenario?

b) Explain orthogonalization, evaluation metrics, train/dev/test set with given Cat Classification Table.

<table>
<tr><td>Classifier</td><td>Accuracy</td><td>Running time</td></tr>
<tr><td>A</td><td>91%</td><td>83ms</td></tr>
<tr><td>B</td><td>95%</td><td>95ms</td></tr>
<tr><td>C</td><td>92.5%</td><td>86ms</td></tr>
</table>
</td></tr>
<tr><td>5.</td><td>a) Explain the process of computing gradients using backpropagation function?
b) Illustrate the use of sequence to sequence (Seg2Seq) model with the help of a suitable example.</td></tr>
<tr><td>6.</td><td>a) An input image has been converted into a matrix of size 32×32 along with a filter of size 5×5 with a Stride of 1. Determine the size of the convoluted matrix with a padding amount of 2.
b) Explain stochastic gradient descent, momentum, and adaptive sub-gradient method with a suitable case study.</td></tr>
</table>