SRM Institute of Science and Technology
Faculty of Engineering and Technology
School of Computing

Mode of Exam
OFFLINE

DEPARTMENT OF COMPUTATIONAL INTELLIGENCE

SRM Nagar, Kattankulathur – 603203, Chengalpattu District, Tamilnadu

Academic Year: AY2023-24-ODD set C

Test: CLAT-2 **Date: 10/10/2023**

Course Code & Title: 18AIC301J, Deep Learning Techniques. **Duration: 90 minutes**

Year & Sem: 3rd year, Vth sem **Max. Marks: 50**

Course Articulation Matrix: ***(to be placed)***

Course Learning Rationale (CLR):	The purpose of learning this course is to:		Learning			Program Learning Outcomes (PLO)													
CLR-1 :	Illustrate the basic concepts of deep learning		1	2	3	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
CLR-2 :	Gain knowledge in Optimization algorithms and dimensionality reduction		Level of Thinking (Bloom)	Expected Proficiency (%)	Expected Attainment (%)	Engineering Knowledge	Problem Analysis	Design & Development	Analysis, Design, Research	Modern Tool Usage	Society & Culture	Environment & Sustainability	Ethics	Individual & Team Work	Communication	Project Mgt. & Finance	Life Long Learning	PSO - 1	PSO - 2	PSO - 3
CLR-3 :	Develop a broad understanding of word2vec models and Convolution Neural Network models																			
CLR-4 :	Acquire knowledge in Transfer learning and Sequential Models																			
CLR-5 :	Implement the attention mechanism and advanced deep learning models																			
Course Learning Outcomes (CLO):	At the end of this course, learners will be able to:																			
CLO-1 :	Understand various deep learning models to solve real world problems		2	85	75	L	H	-	-	-	-	-	-	M	-	-	-	H	M	M
CLO-2 :	Compare the optimization algorithms and high dimensional data using reduction techniques		2	80	70	M	M	-	-	-	-	-	-	M	-	-	-	M	L	M
CLO-3 :	Implement word2vec models and Convolution Neural Network models		3	85	75	M	H	-	-	-	-	-	-	H	-	-	-	H	M	M
CLO-4 :	Apply RNN and transfer learning to real world scenarios		3	85	80	M	H	-	-	-	-	-	-	H	-	-	-	H	M	H
CLO-5 :	Use deep learning models to solve real-world applications		3	80	70	M	H	-	-	-	-	-	-	H	-	-	-	L	L	L

Part - A
(20 x 1 = 20 Marks)

Instructions: Answer all

Q. No	Answer with choice variable	Marks	BL	CO	PO	PI Code
1	(d)All of the above	1	1	2	1	1.6.1*
2	*(a)Learning rate is adaptive	1	2	2	1	1.6.1
3	*(b)Rectified Linear Unit	1	1	2	1	1.6.1
4	*D) All of the above	1	2	3	1	1.6.1
5	*(b) Dropout	1	2	2	1	1.6.1
6	*C) 218 x 218 x 5	1	3	3	2	2.4.1
7	*(c)Resnet	1	1	3	1	1.6.1
8	*C) PCA	1	1	3	1	1.6.1
9	*(c)3	1	3	3	2	1.7.1
10	*(d)Pooling	1	1	2	1	1.6.1
	Part – B **(2x 5 = 10 Marks ; 2x 10 = 20 Marks)**					
11	Unit2: SLIDE 5, DLclass5.pdf	5	2	2	1	1.6.1

<table>
<tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>12</td><td>Unit2: SLIDE 14,DLclass1.pdf</td><td>5</td><td>1</td><td>3</td><td>1</td><td>1.6.1</td></tr>
<tr><td>13</td><td>[(32-3+2*2)/1]+1=34*34*8</td><td>10</td><td>2</td><td>2</td><td>1</td><td>1.6.1</td></tr>
<tr><td>14</td><td>Unit3:DLclass5.pdf</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>15</td><td>Unit2: DLclass2.pdf</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>16</td><td>Unit2: DLclass3.pdf</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>17</td><td>
<table>
<tr><th>Layer</th><th>Activation map dimensions</th><th>Number of weights</th><th>Number of biases</th></tr>
<tr><td>INPUT</td><td>$128 \times 128 \times 3$</td><td>0</td><td>0</td></tr>
<tr><td>CONV-9-32</td><td></td><td></td><td></td></tr>
<tr><td>POOL-2</td><td></td><td></td><td></td></tr>
<tr><td>CONV-5-64</td><td></td><td></td><td></td></tr>
<tr><td>POOL-2</td><td></td><td></td><td></td></tr>
<tr><td>CONV-5-64</td><td></td><td></td><td></td></tr>
<tr><td>POOL-2</td><td></td><td></td><td></td></tr>
<tr><td>FC-3</td><td></td><td></td><td></td></tr>
</table>
Solution: Successively:
• $120 \times 120 \times 32$ and $32 \times (9 \times 9 \times 3 + 1)$
• $60 \times 60 \times 32$ and 0
• $56 \times 56 \times 64$ and $64 \times (5 \times 5 \times 32 + 1)$
• $28 \times 28 \times 64$ and 0
• $24 \times 24 \times 64$ and $64 \times (5 \times 5 \times 64 + 1)$
• $12 \times 12 \times 64$ and 0
• 3 and $3 \times (12 \times 12 \times 64 + 1)$
</td><td>10</td><td>3</td><td>3</td><td>2</td><td>1.7.1</td></tr>
<tr><td>18</td><td>Unit3: DLclass4.pdf</td><td></td><td></td><td></td><td></td><td></td></tr>
</table>

Question Paper Setter

Approved by Audit Professor/ Course Coordinator

*** Performance Indicators are available separately for Computer Science and Engineering in AICTE examination reforms policy.**

Course Outcome (CO) and Bloom's level (BL) Coverage in Questions

SRM Institute of Science and Technology
Faculty of Engineering and Technology
School of Computing

Mode of Exam
OFFLINE

DEPARTMENT OF COMPUTATIONAL INTELLIGENCE

SRM Nagar, Kattankulathur – 603203, Chengalpattu District, Tamilnadu

Academic Year: AY 23-24-ODD **SET D**

Test: CLAT-2 **Date: 10/10/2023**

Course Code & Title: 18AIC301J, Deep Learning Techniques. **Duration: 90 minutes**

Year & Sem: 3rd year, Vth sem **Max. Marks: 50**

Course Articulation Matrix: ***(to be placed)***

Course Learning Rationale (CLR):	The purpose of learning this course is to:		Learning				Program Learning Outcomes (PLO)														
CLR-1 :	Illustrate the basic concepts of deep learning		1	2	3		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
CLR-2 :	Gain knowledge in Optimization algorithms and dimensionality reduction		Level of Thinking (Bloom)	Expected Proficiency (%)	Expected Attainment (%)		Engineering Knowledge	Problem Analysis	Design & Development	Analysis, Design, Research	Modern Tool Usage	Society & Culture	Environment & Sustainability	Ethics	Individual & Team Work	Communication	Project Mgt. & Finance	Life Long Learning	PSO - 1	PSO - 2	PSO - 3
CLR-3 :	Develop a broad understanding of word2vec models and Convolution Neural Network models																				
CLR-4 :	Acquire knowledge in Transfer learning and Sequential Models																				
CLR-5 :	Implement the attention mechanism and advanced deep learning models																				
Course Learning Outcomes (CLO):	At the end of this course, learners will be able to:																				
CLO-1 :	Understand various deep learning models to solve real world problems		2	85	75		L	H	-	-	-	-	-	-	M	-	-	-	H	M	M
CLO-2 :	Compare the optimization algorithms and high dimensional data using reduction techniques		2	80	70		M	M	-	-	-	-	-	-	M	-	-	-	M	L	M
CLO-3 :	Implement word2vec models and Convolution Neural Network models		3	85	75		M	H	-	-	-	-	-	-	H	-	-	-	H	M	M
CLO-4 :	Apply RNN and transfer learning to real world scenarios		3	85	80		M	H	-	-	-	-	-	-	H	-	-	-	H	M	H
CLO-5 :	Use deep learning models to solve real-world applications		3	80	70		M	H	-	-	-	-	-	-	H	-	-	-	L	L	L

Part - A
(10 x1 = 20 Marks)

Instructions: Answer all

Q. No	Question	Marks	BL	CO	PO	PI Code
1	CNN is mostly used when there is a) structured Data b) **unstructured Data** c)Both a and b d)none of the above	1	1	2	1	1.6.1*
2	Which neural network has only one hidden layer between the input and output? a) shallow neural network b) Deep neural network c)**Feed-forward neural networks** d) Recurrent neural networks	1	1	2	1	1.6.1
3	Which of the following SGD variants is based on both momentum and adaptive learning? a) RMSprop. b) Adagrad. **c) Adam** d)Nestrov	1	2	2	1	1.6.1

4	Which of the following functions can be used as an activation function in the output layer if we wish to predict the probabilities of n classes (p1, p2,...,pk) such that sum of p over all n equals to 1? **(a) Softmax** (b) ReLu (c) Sigmoid (d) Tanh	**1**	**3**	**2**	**1**	**1.6.1**
5	Which of the following is NOT a solution for Dying ReLU? a) Batch normalization b) Dropout c) Leaking ReLU **d)Low learning rate**	**1**	**1**	**2**	**1**	**1.6.1**
6	Suppose an input to AVG-Pooling layer is given above. The pooling size of neurons in the layer is (3, 3). 3 4 5 4 5 6 5 6 7 What would be the output of this Pooling layer? a) 3 **b) 5** c) 5.5 d) 7	**1**	**3**	**3**	**2**	**2.4.1**
7	Which of the following techniques perform similar operation as a dropout in a neural network? **a) bagging** b) Boosting c)Stacking d)None of these	**1**	**2**	**3**	**1**	**1.6.1**
8	The input image has been converted into a matrix of size 28*28 and a kernel/filter of size 7*7 with a stride of 1. What will be the size of the convoluted matrix? a)20*20 b)21*21 **c)22*22** d)25*25	**1**	**1**	**3**	**1**	**1.6.1**
9	The number of nodes in the input layer is 10 and the hidden layer is 5. The maximum number of connections from the input layer to the hidden layer are- **a)50** b) more than 50 c)less than 50 d)It is an arbitrary value	**1**	**1**	**3**	**1**	**1.6.1**
10	How many input nodes are required to process a colored image of 28X28? **a)28 X 28 X 3** b) 28 X 28 X 1 c) 56 X 56 X 1 d) 56 X 56 X 3	**1**	**1**	**3**	**2**	**2.4.1**

SRM Institute of Science and Technology
Faculty of Engineering and Technology
School of Computing

Mode of Exam
OFFLINE

DEPARTMENT OF COMPUTATIONAL INTELLIGENCE

SRM Nagar, Kattankulathur – 603203, Chengalpattu District, Tamilnadu

Academic Year: AY 23-24-ODD **SET D**

Course Code & Title: 18AIC301J, Deep Learning Techniques **Date: 10/10/2023**

Part – B (4x 5 = 20 Marks)
Answer all the questions

11	Explain how the following concepts help to overcome overfitting. (a)L2 regularization (b)Data augmentation Unit-2 DL class 3- ppt	**5**	**2**	**2**	**1**	**1.6.1**
12	What is dimensionality deduction? Illustrate the concept of SVD. Unit -2 DL class 5- page 7	**5**	**2**	**2**	**1**	**2.1.3**
13	Outline the concepts of one hop representation with an example. Unit 3 ppt- page 2	**5**	**2**	**3**	**1**	**1.6.1**
14	Explain any two word-representations used in NLP? Detail the techniques with an example each. Unit 3 ppt- page 5	**5**	**2**	**3**	**1**	**1.6.1**
	Part – C (2x 10 = 20 Marks) **Answer all the questions**					
15	Define Autoencoders. Explain relation between PCA and Autoencoders. Unit 2 DL class 5 - page 2	**10**	**3**	**2**	**1**	**1.6.1**
	OR					
16	What are the limitations of vennila gradient decent algorithm? Explain the optimization algorithms to overcome that. Unit 2- DL class 2- page 2	**10**	**3**	**2**	**1**	**1.6.1**
17	Calculate CNN operation using below Using 5*5 input data calculate using kernel filter 3 3 3 1 0 0 0 1 3 1 3 1 2 2 3 2 0 0 2 2 2 0 0 0 1 Kernel filter 3*3 0 1 2 2 2 0 0 1 2 Use input to calculate below operation in CNN 1)stride 2)padding 3)max pooling 4)avg pooling Convolution ans: 12 9 10 1 14 9 3 6 1 Stride = 1 If stride= 1, padding is not required Max pooling 14	**10**	**3**	**3**	**2**	**2.4.1**

	Avg pooling 8.66					
	OR					
18	Explain the intuition behind GoogleNet Unit 3- DL class 5- page-10	10	3	3	1	1.6.1

SRM Institute of Science and Technology
College of Engineering and Technology
School of Computing

SET C

DEPARTMENT OF COMPUTATIONAL INTELLIGENCE

SRM Nagar, Kattankulathur – 603203, Chengalpattu District, Tamilnadu

Academic Year: AY2023-24-ODD

Test: CLAT-2
Course Code & Title: 18AIC301J, Deep Learning Techniques.
Year & Sem: 3rd year, Vth sem

Date: 10/10/2023
Duration: 90 minutes
Max. Marks: 50

Course Articulation Matrix: ***(to be placed)***

Course Learning Rationale (CLR):	The purpose of learning this course is to:		Learning			Program Learning Outcomes (PLO)														
CLR-1 :	Illustrate the basic concepts of deep learning		1	2	3		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
CLR-2 :	Gain knowledge in Optimization algorithms and dimensionality reduction		Level of Thinking (Bloom)	Expected Proficiency (%)	Expected Attainment (%)		Engineering Knowledge	Problem Analysis	Design & Development	Analysis, Design, Research	Modern Tool Usage	Society & Culture	Environment & Sustainability	Ethics	Individual & Team Work	Communication	Project Mgt. & Finance	Life Long Learning	PSO - 1	PSO - 2	PSO – 3
CLR-3 :	Develop a broad understanding of word2vec models and Convolution Neural Network models																				
CLR-4 :	Acquire knowledge in Transfer learning and Sequential Models																				
CLR-5 :	Implement the attention mechanism and advanced deep learning models																				
Course Learning Outcomes (CLO):	At the end of this course, learners will be able to:																				
CLO-1 :	Understand various deep learning models to solve real world problems		2	85	75		L	H	-	-	-	-	-	-	M	-	-	-	H	M	M
CLO-2 :	Compare the optimization algorithms and high dimensional data using reduction techniques		2	80	70		M	M	-	-	-	-	-	-	M	-	-	-	M	L	M
CLO-3 :	Implement word2vec models and Convolution Neural Network models		3	85	75		M	H	-	-	-	-	-	-	H	-	-	-	H	M	M
CLO-4 :	Apply RNN and transfer learning to real world scenarios		3	85	80		M	H	-	-	-	-	-	-	H	-	-	-	H	M	H
CLO-5 :	Use deep learning models to solve real-world applications		3	80	70		M	H	-	-	-	-	-	-	H	-	-	-	L	L	L

Part - A
(10 x1 = 15 Marks)

Instructions: Answer all

Q. No	Question	Marks	BL	CO	PO	PI Code
1	The technique which is used to predict the target word from the context____ A. CBOW B. CNN **C. Skip Gram Model** D. LSTM	1	1	2	1	1.6.1*
2	Which of the following layers is NOT a part of	1	2	2	1	1.6.1

	CNN? A. Convolutional Layer B. Pooling Layer C. Code Layer D. Fully connected Lay Which of the following layers is NOT a part of CNN? A. Convolutional Layer					

	B. Pooling Layer C. Code Layer D. Fully connected Laye Which of the following layers is NOT a part of CNN? A. Convolutional Layer B. Pooling Layer **C. Code Layer** D. Fully connected Laye					
3	**When uses PCA?** (a)When my data is small and with a few features (b)Everytime before uses a Machine Learning algorithm (c)When I have a overfit case **(d)You want to find latent features and reduce dimensionality**	**1**	**1**	**2**	**1**	**1.7.1**
4	Which of the following model is best suited for sequential data? A. Convolutional Neural Networks (ConvNets) B. Capsule Neural Networks (CapsNets) C. RNN (Recurrent Neural Network) **D. Autoencoder**	**1**	**2**	**2**	**1**	**1.6.1**
5	Which of the following is FALSE about Dropout? A. Dropout is a learnable parameter in the network B. Dropout increases the accuracy and performance of the model C. Dropout introduces sparsity in the network D. Dropout makes training process noisy	**1**	**1**	**3**	**1**	**1.6.1**
6	Data Augmentation helps in: **A. Reducing overfitting** B. Increasing generalization capacity of the network C. Generating data from data D. All of the abov	**1**	**1**	**3**	**1**	**1.6.1**
7	_____ computes the average of the elements present in the region of feature map covered by the filter. A. Max Pooling	**1**	**1**	**3**	**1**	**1.6.1**

	B. Average Pooling C. Global pooling D. None of these					
8	Which of the following is a hyperparameter in a neural network? A. Activation Function **B. Learning Rate** C. Momentum D. All of the above	**1**	**3**	**3**	**2**	**1.7.1**
9	Using too many epochs while training a network may lead to: A. High training time **B. Overfitting** C. Unnecessary time wastage D. All of the above	**1**	**1**	**3**	**1**	**1.6.1**
10	You have an input volume that is 63x63x16, and convolve it with 32 filters that are each 7x7, and stride of 1. You want to use a "same" convolution. What is the padding? (a)1 (b)2 **(c)3** (d)7	**1**	**3**	**3**	**2**	**1.7.1**
	Part – B **(4x 5 = 20 Marks)**					
11	Differentiate Skip Gram model and continuous bag of model with example. Skip Gram Model explanation 1 Marks Example 1 continuous bag of model with example 2 Example 1	**5**	**1**	**3**	**1**	**1.6.1**
12	Discuss about the limitations of gradient descent algorithm. 1. **Choice of learning rate:** The choice of learning rate is crucial for the convergence of gradient descent and its variants. Choosing a learning rate that is too large can lead to oscillations or overshooting while choosing a learning rate that is too small can lead to slow convergence or getting stuck in local minima. 2. **Sensitivity to initialization:** Gradient descent and its variants can be sensitive to the initialization of the model's parameters, which can affect the convergence and the quality of the solution.	**5**	**1**	**2**	**1**	**1.6.1**

	3. **Time-consuming:** Gradient descent and its variants can be time-consuming, especially when dealing with large datasets and high-dimensional models. The convergence speed can also vary depending on the variant used and the specific problem. 4. **Local optima:** Gradient descent and its variants can converge to a local minimum instead of the global minimum of the cost function, especially in non-convex problems. This can affect the quality of the solution, and techniques like random initialization and multiple restarts may be used to mitigate this issue.					
13	Discuss about the Nesterov and Momentum Based Gradient decedent with Pseudocode.	**5**	**1**	**2**	**1**	**1.7.1**
14	Illustrate the use of convolution operation in CNN with example.	**5**	**1**	**3**	**1**	**1.6.1**
	Part – C **(2x 10 = 20 Marks)**					
15	Explain the one hot representation and Distributed word representation with example. Illustrate the challenges of both representations. https://www.geeksforgeeks.org/ml-one-hot-encoding-of-datasets-in-python/	**10**	**2**	**2**	**1**	**1.6.1**
16	(OR) Illustrate the need of optimization in deep learning and discuss about the optimization techniques.	**10**	**2**	**2**	**2**	**1.6.1**
17	Explain the Alexnet Model in detail with its architecture diagram.	**10**	**3**	**3**	**2**	**1.7.1**
18	(OR) Discuss about the ResNet architecture with its diagram and list out the pros by comparing with other architecture	**10**	**3**	**3**	**2**	**1.6.1**

***Performance Indicators are available separately for Computer Science and Engineering in AICTE examination reforms policy.**

Course Outcome (CO) and Bloom's level (BL) Coverage in Questions

SRM Institute of Science and Technology
Faculty of Engineering and Technology
School of Computing

Mode of Exam
OFFLINE

DEPARTMENT OF COMPUTATIONAL INTELLIGENCE

SRM Nagar, Kattankulathur – 603203, Chengalpattu District, Tamilnadu

Academic Year: AY 23-24-ODD **SET**

Test: CLAT-2 **Date: 10/10/2023**

Course Code & Title: 18AIC301J, Deep Learning Techniques. **Duration: 90 minutes**

Year & Sem: 3rd year, Vth sem **Max. Marks: 50**

Course Articulation Matrix: ***(to be placed)***

Course Learning Rationale (CLR):	The purpose of learning this course is to:	Learning			Program Learning Outcomes (PLO)														
CLR-1 :	Illustrate the basic concepts of deep learning	1	2	3	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
CLR-2 :	Gain knowledge in Optimization algorithms and dimensionality reduction	Level of Thinking (Bloom)	Expected Proficiency (%)	Expected Attainment (%)	Engineering Knowledge	Problem Analysis	Design & Development	Analysis, Design, Research	Modern Tool Usage	Society & Culture	Environment & Sustainability	Ethics	Individual & Team Work	Communication	Project Mgt. & Finance	Life Long Learning	PSO - 1	PSO - 2	PSO - 3
CLR-3 :	Develop a broad understanding of word2vec models and Convolution Neural Network models																		
CLR-4 :	Acquire knowledge in Transfer learning and Sequential Models																		
CLR-5 :	Implement the attention mechanism and advanced deep learning models																		
Course Learning Outcomes (CLO):	At the end of this course, learners will be able to:																		
CLO-1 :	Understand various deep learning models to solve real world problems	2	85	75	L	H	-	-	-	-	-	-	M	-	-	-	H	M	M
CLO-2 :	Compare the optimization algorithms and high dimensional data using reduction techniques	2	80	70	M	M	-	-	-	-	-	-	M	-	-	-	M	L	M
CLO-3 :	Implement word2vec models and Convolution Neural Network models	3	85	75	M	H	-	-	-	-	-	-	H	-	-	-	H	M	M
CLO-4 :	Apply RNN and transfer learning to real world scenarios	3	85	80	M	H	-	-	-	-	-	-	H	-	-	-	H	M	H
CLO-5 :	Use deep learning models to solve real-world applications	3	80	70	M	H	-	-	-	-	-	-	H	-	-	-	L	L	L

Part - A
(10 x1 = 10 Marks)

Instructions: Answer all

Q. No	Question	Marks	BL	CO	PO	PI Code
1	Which of the following is FALSE about CNN? A. We must flatten the output before feeding it to a fully connected layer B. There can be only one fully connected layer in CNN C. We can use as many convolutional and pooling layers in CNN D. None of the above	1	1	3	1	1.6.1*
2	Filter of size 3X3 is rotated over input matrix of size 4X4 (stride=1). What will be the size of output matrix after applying zero padding? A. 4X4 B. 3X3 C. 2X2 D. 1X1	1	3	2	1	1.6.1
3	Which of the following SGD variants is based on both momentum and adaptive learning? A. RMSprop. B. Adagrad.	1	2	2	1	1.6.1

	C. Adam D. Nestrov					
4	Which of the following functions can be used as an activation function in the output layer if we wish to predict the probabilities of n classes (p1, p2,...,pk) such that sum of p over all n equals to 1? A. Softmax B. ReLu C. Sigmoid D. Tanh	**1**	**2**	**2**	**1**	**1.6.1**
5	Which of the following is a hyperparameter in a neural network? A. Activation Function B. Learning Rate C. Momentum D. All of the above	**1**	**1**	**2**	**1**	**1.6.1**
6	Which of the following is NOT a way to increase generalization in auto encoder? A. Use larger code size B. Use L1 and L2 regularization C. Add some random noise to the input D. Limit the number of nodes in the hidden layers	**1**	**3**	**2**	**2**	**2.4.1**
7	How many numbers of convolution layers in ZFNet architecture? A. 3 B. 4 C. 5 D. 6	**1**	**2**	**3**	**1**	**1.6.1**
8	Which of the following is TRUE about Softmax and Sigmoid function? A. Softmax is usually used for hidden layers and sigmoid for outer layers B. Sigmoid is usually used for hidden layers and softmax for outer layers C. Softmax function is usually used for binary classification problem D. All of the above	**1**	**1**	**3**	**1**	**1.6.1**
9	How many numbers of inception modules available in GoogleNet architecture? A. 5 B. 7 C. 9 D. 3	**1**	**1**	**3**	**1**	**1.6.1**
10	After training a neural network, you observe a large gap between the training accuracy (100%) and the test accuracy (42%). Which of the following methods is commonly used to reduce this gap? A. Generative Adversarial Networks B. Dropout C. Sigmoid activation D. RMSprop optimizer	**1**	**3**	**3**	**2**	**2.4.1**

SRM Institute of Science and Technology
Faculty of Engineering and Technology
School of Computing

Mode of Exam
OFFLINE

DEPARTMENT OF COMPUTATIONAL INTELLIGENCE

SRM Nagar, Kattankulathur – 603203, Chengalpattu District, Tamilnadu

Academic Year: AY 23-24-ODD **SET A**

Course Code & Title: 18AIC301J, Deep Learning Techniques **Date: 10/10/2023**

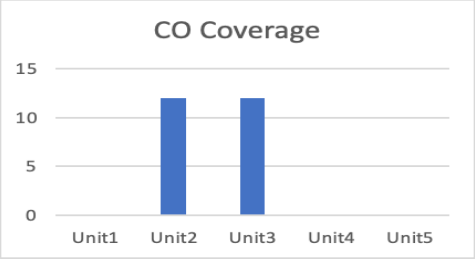
Part – B (4x 5 = 20 Marks)
Answer all the questions

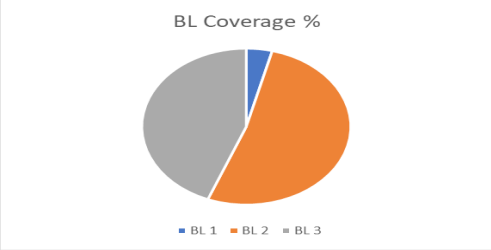
<table>
<tr><td>11</td><td>Explain about Nesterov accelerated gradient decent and RMSProp optimizers in details.</td><td>5</td><td>2</td><td>2</td><td>1</td><td>1.6.1</td></tr>
<tr><td>12</td><td>Explain about dimensionality reduction using Auto encoders with an example.</td><td>5</td><td>2</td><td>2</td><td>1</td><td>2.1.3</td></tr>
<tr><td>13</td><td>Explain the concept of skip-gram model with an example.</td><td>5</td><td>2</td><td>3</td><td>1</td><td>1.6.1</td></tr>
<tr><td>14</td><td>You come up with a CNN classifier for 3 classes. For each layer, calculate the output size and number of parameters associated with each layer.
The
<table>
<tr><th>Layer</th><th>Activation map size</th><th>No of parameters</th></tr>
<tr><td>input</td><td>128*128*3</td><td>0</td></tr>
<tr><td>Conv -5-16</td><td></td><td></td></tr>
<tr><td>MaxPool-2</td><td></td><td></td></tr>
<tr><td>Conv-3-32</td><td></td><td></td></tr>
<tr><td>MaxPool-2</td><td></td><td></td></tr>
<tr><td>Conv-3-64</td><td></td><td></td></tr>
<tr><td>MaxPool-2</td><td></td><td></td></tr>
<tr><td>FC-1(128)</td><td></td><td></td></tr>
<tr><td>FC-2(25)</td><td></td><td></td></tr>
</table>
notation follows the convention:

• CONV-K-N denotes a convolutional layer with N filters, each of them of size K x K. Padding and stride parameters are always 0 and 1 respectively.
• POOL-K indicates a K x K pooling layer with stride K and padding 0.
FC(X) indicates X number of neurons.</td><td>5</td><td>3</td><td>3</td><td>1</td><td>1.6.1</td></tr>
<tr><td colspan="7">Part – C (2x 10 = 20 Marks)
Answer all the questions</td></tr>
<tr><td>15</td><td>(i) Write in detail about principal component analysis with an example.
(ii) What are the limitations of vennila gradient decent algorithm?</td><td>7+3</td><td>2</td><td>2</td><td>1</td><td>1.6.1</td></tr>
<tr><td colspan="7">OR</td></tr>
<tr><td>16</td><td>(i) Explain the techniques to overcome over fitting in deep neural network model.
(ii) Compare PCA and Auto encoders</td><td>7+3</td><td>2</td><td>2</td><td>1</td><td>1.6.1</td></tr>
<tr><td>17</td><td>Consider a CNN with following layers stacked 1-convolution, 1-activation(ReLu), 1-Avg pooling, 1-flattening and 1-dense layer with 3 neurons and softmax function. Show the processing and output of each layer for the following input data.
<table>
<tr><td>3</td><td>3</td><td>3</td><td>1</td><td>0</td></tr>
<tr><td>0</td><td>0</td><td>1</td><td>3</td><td>1</td></tr>
<tr><td>3</td><td>1</td><td>2</td><td>2</td><td>3</td></tr>
</table></td><td>10</td><td>3</td><td>3</td><td>2</td><td>2.4.1</td></tr>
</table>

	2 0 0 2 2 2 0 0 0 1 Kernel filter 3*3 0 1 2 -1 2 0 0 -1 2					
	OR					
18	Write in detail about AlexNet and Google Net with their architectures and highlight the difference between them.	10	2	3	1	1.6.1

***Performance Indicators are available separately for Computer Science and Engineering in AICTE examination reforms policy. Course Outcome (CO) and Bloom's level (BL) Coverage in Questions**





QP Setter / Course Coordinator

Approved by the Audit Professor