Question Paper Name:	5324 Neural Networks and Deep Learning 30th June 2019 Shift 1
Subject Name:	Neural Networks and Deep Learning
Creation Date:	2019-06-30 13:01:46
Duration:	180
Total Marks:	100
Display Marks:	Yes

Neural Networks and Deep Learning

Group Number :	1
Group Id :	489994230
Group Maximum Duration :	0
Group Minimum Duration :	120
Revisit allowed for view? :	No
Revisit allowed for edit? :	No
Break time:	0
Group Marks:	100

Neural Networks and Deep Learning

Section Id :	489994286
Section Number :	1
Section type :	Online
Mandatory or Optional:	Mandatory
Number of Questions:	53
Number of Questions to be attempted:	53
Section Marks:	100
Display Number Panel:	Yes
Group All Questions:	No

Sub-Section Number:	1
Sub-Section Id:	489994312
Question Shuffling Allowed :	Yes

Question Number : 1 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical
Correct Marks : 1 Wrong Marks : 0

What is the secret of success of deep learning?

a. Vanishing gradient problem does not exist anymore
b. Number of weights associated with the network have become in the order of billions
c. Brute computing power using GPUs have become manifold
d. Machine intelligence has almost become equivalent to human intelligence

---

Options :
1. 1

3. 3

4. 4

Question Number : 2 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical
Correct Marks : 1 Wrong Marks : 0

When is the cell said to be fired?

a. If potential of the cell body reaches a steady threshold value
b. If there is impulse reaction
c. During upbeat of heart
d. None of the above

Options :
1. 1
2. 2
3. 3
4. 4

Question Number : 3 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical
Correct Marks : 1 Wrong Marks : 0

Who developed the first learning machine in which connection strengths could be adapted automatically?

a. McCulloch-pits
b. Marvin Minsky
c. Hopfield
d. Rosenblatt

Options :
1. 1
2. 2
3. 3
4. 4

Question Number : 4 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical
Correct Marks : 1 Wrong Marks : 0

Who proposed the first perceptron model in 1958?

a. McCulloch-pits
b. Marvin Minsky
c. Hopfield
d. Rosenblatt

Options :
1. 1
2. 2
3. 3
4. 4

Question Number : 5 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical
Correct Marks : 1 Wrong Marks : 0

5. sign of weight indicates?
   a. Excitatory input
   b. Inhibitory input
   c. Can be either excitatory or inhibitory as such
   d. None of the Above

Options :
1. 1
2. 2
3. 3
4. 4

Question Number : 6 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical
Correct Marks : 1 Wrong Marks : 0

Competitive learning net is used for?
   a. Pattern grouping
   b. Pattern storage
   c. Pattern grouping or storage
   d. None of the above

Options :
1. 1
2. 2
3. 3
4. 4

Question Number : 7 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical
Correct Marks : 1 Wrong Marks : 0

Which loss function out of the following is best for image classification (one class out of many) using deep neural network?

   a. Categorical cross entropy
   b. Mean squared error loss
   c. Binary cross entropy
   d. Hinge loss

Options :
1. 1
2. 2
3. 3
4. 4

Question Number : 8 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical
Correct Marks : 1 Wrong Marks : 0

The problem that a CNN with 50 layers performs worse than a CNN with 30 players on both training and test data can be effectively resolved by using

   a. Regularization
   b. Validation set
   c. Resnet architecture
   d. Softmax activation

1. 1
2. 2
3. 3
4. 4

Question Number : 9 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical
Correct Marks : 1 Wrong Marks : 0

In Delta Rule for mean squared error minimization, weights are adjusted proportional to

a. the changes in the output vector
b. the difference between desired output and actual output
c. the difference between input and output of the multi-layered perceptron
d. None of the above

Options :
1. 1
2. 2
3. 3
4. 4

Question Number : 10 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical
Correct Marks : 1 Wrong Marks : 0

What is a "hidden layer" of a neural network?

a. The last layer of neurons, which is hidden from the inputs.
b. One of the middle layers, which aren't directly connected to either inputs or outputs.
c. A group of neurons with zero weights, which have effectively hidden themselves from the rest of the network.
d. A layer of neurons that's in control of turning inputs on or off, which affects how training is accomplished.

Options :
1. 1
2. 2
3. 3
4. 4

Question Number : 11 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical
Correct Marks : 1 Wrong Marks : 0

Given variables: (i) the present input; (ii) the previous cell state; (iii) the previous output; (iv) the previous hidden state; the response of a traditional LSTM depends on:

a. (i),(ii),(iii)
b. (i),(ii),(iii),(iv)
c. (i),(ii),(iv)
d. (i),(iii),(iv)

Options :
1. 1
2. 2
3. 3

Question Number : 12 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical
Correct Marks : 1 Wrong Marks : 0

Which of the following is correct regarding Gated Recurrent Units (GRUs)?

a. It combines the forget and input gates into a single "update gate."
b. It merges the cell state and hidden state
c. It takes the previous hidden state multiplied by reset gate as an input.
d. All of these

Options :

1. 1
2. 2
3. 3
4. 4

Question Number : 13 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical
Correct Marks : 1 Wrong Marks : 0

Please mark the correct weight update law for RBM. Here, W stands for weights of RBM, b stands for hidden unit bias vector of RBM, c stands for visible layer bias vector of RBM and $\eta = 1$ is the learning rate.

a. $W = W + \eta(h(x^{(t)})x^{(t)} - h(\tilde{x}^{(t)})\tilde{x}^{(t)}); b = b + \eta(h(x^{(t)}) - h(\tilde{x}^{(t)}); c = c + (x^{(t)} - \tilde{x}^{(t)}))$
b. $W = W - \eta(h(x^{(t)})x^{(t)} - h(\tilde{x}^{(t)})\tilde{x}^{(t)}); b = b + \eta(h(x^{(t)}) - h(\tilde{x}^{(t)}); c = c - \eta(x^{(t)} - \tilde{x}^{(t)}))$
c. $W = W + \eta(h(x^{(t)})x^{(t)} - h(\tilde{x}^{(t)})\tilde{x}^{(t)}); b = b - \eta(h(x^{(t)}) - h(\tilde{x}^{(t)}); c = c + \eta(x^{(t)} - \tilde{x}^{(t)}))$
d. $W = W - \eta(h(x^{(t)})x^{(t)} - h(\tilde{x}^{(t)})\tilde{x}^{(t)}); b = b - \eta(h(x^{(t)}) - h(\tilde{x}^{(t)}); c = c - (x^{(t)} - \tilde{x}^{(t)}))$

Options :

1. 1
2. 2
3. 3
4. 4

Question Number : 14 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical
Correct Marks : 1 Wrong Marks : 0

Which deep network was used for the image classification for the first time in image-net challenge?

a. VGG
b. Googlenet
c. ZFnet
d. Alexnet

Options :

1. 1
2. 2
3. 3

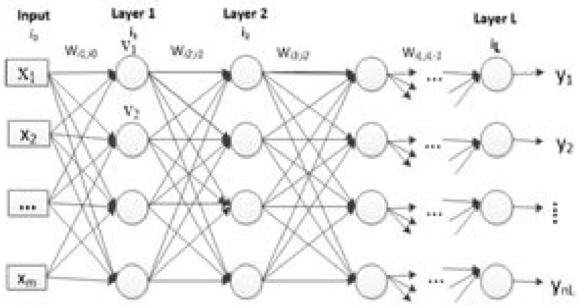
Question Number : 15 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

Which deep network was used for the image classification for the first time in image-net challenge?

a. VGG
b. Googlenet
c. ZFnet
d. Alexnet

In the following figure, please mark the correct expression for the back-error propagated for the 1st layer



a. $\delta_{i_1} = (v_{i_1})(1 - v_{i_1}) \sum_{i_2=1}^{n_2} \delta_{i_2} W_{i_2 i_1}$

b. $\delta_{i_1} = (v_{i_1})(1 - v_{i_1}) \sum_{i_2=1}^{n_2} \sum_{i_3=1}^{n_3} \delta_{i_3} \delta_{i_2} W_{i_2 i_1}$

c. $\delta_{i_1} = (v_{i_1}) \sum_{i_2=1}^{n_2} \sum_{i_3=1}^{n_3} \delta_{i_3} \delta_{i_2} W_{i_2 i_1}$

d. $\delta_{i_1} = (v_{i_1})(1 - v_{i_1}) \sum_{i_2=1}^{n_2} \sum_{i_3=1}^{n_3} \delta_{i_3} \delta_{i_2} W_{i_2 i_1} W_{i_3 i_2}$

Options :

1. 1
2. 2
3. 3
4. 4

Question Number : 16 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

Adding a momentum rate in backpropagation algorithm helps in

a. Faster convergence to nearest local minimum
b. Avoids local minima
c. Ensures global convergence
d. Prevents over-fitting

Options :

1. 1
2. 2
3. 3
4. 4

Question Number : 17 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

What is the output of the following Python program?

```
i = 0
while i < 5:
    print(i),
    i += 1
    if i == 3:
        break
    else:
        print(0),
```

a. 0 0 1 0 2 0
b. 0 0 1 0 2
c. 0 1 2 0
d. None of these

Options :

1. 1
2. 2
3. 3
4. 4

Question Number : 18 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

What will be displayed by the following python code?

```
def f(value,
values):
    v = 1
    values[0] = 44
t = 3
v = [1, 2, 3]
f(t, v)
print(t, v[0])
```

a. (1, 1)
b. (1, 44)
c. (3, 1)
d. (3, 44)

Options :

1. 1
2. 2
3. 3
4. 4

Question Number : 19 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 1 Wrong Marks : 0

ht vector associated with a Kohonen lattice is updated as

$$w_\gamma = w_\gamma + \eta_w h_\gamma (x - w_\gamma)$$

The learning rate is

a. A fixed value less than 1
b. $\eta_w = \eta_i (\frac{\eta_f}{\eta_i})^{\frac{t}{t_{max}}}$
c. Increasing monotonically with time
d. Always a negative value

Options :

1. 1
2. 2
3. 3
4. 4

Question Number : 20 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical
Correct Marks : 1 Wrong Marks : 0

The feed-forward network (FFN) is different from the feedback network as

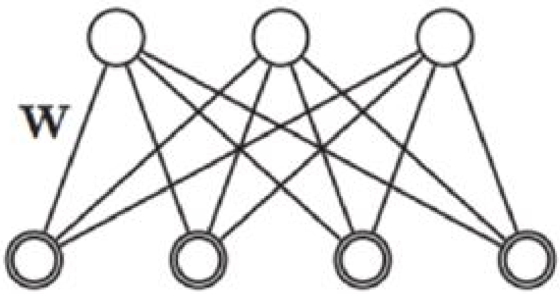
a. FFN has no memory
b. FFN response is dynamical in nature
c. FFN allows connection of a neuron to any one from any layer
d. FFN can become unstable

Options :

1. 1
2. 2
3. 3
4. 4

Question Number : 21 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical
Correct Marks : 1 Wrong Marks : 0

Which of the following options is most appropriate for the network shown below?



a. Bayesian net, Boltzmann machine
b. Undirected graph, Boltzmann machine
c. Directed graph, restricted Boltzmann machine
d. Undirected graph, restricted Boltzmann machine

Options :

1. 1
2. 2
3. 3

Question Number : 22 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical
Correct Marks : 1 Wrong Marks : 0

An RGB image of size 32x32 is the input to an 8x8 CNN layer of 16 conv filters with stride 2. What is the size of each filter (tensor)?

a. (8, 8, 16)
b. (16, 16, 2)
c. (8, 8, 3)
d. (16, 16, 8)

Options :
1. 1
2. 2
3. 3
4. 4

Question Number : 23 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical
Correct Marks : 1 Wrong Marks : 0

What does the Inception Net use to reduce the feature depth while preserving the spatial dimensions of features?

a. Pooling
b. Feature concatenation
c. 1x1 convolutions
d. Dropouts

Options :
1. 1
2. 2
3. 3
4. 4

Question Number : 24 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical
Correct Marks : 1 Wrong Marks : 0

Fully connected layers are not preferred in deep CNNs because
a. They are computationally expensive
b. They consume more memory
c. They are incompatible with convolutional layers
d. They have too many trainable parameters

Options :
1. 1
2. 2
3. 3
4. 4

Question Number : 25 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical
Correct Marks : 2 Wrong Marks : 0

Write the derivative of ReLU activation function f with input $x_i \in \mathbb{R}, \forall i = 0,1,\ldots,N$. Here, $exp(x)$ is exponential function and $\delta(x)$ is the Dirac delta function

a. 1
b. $\frac{1}{1+exp(-x_i)}$
c. x_i
d. $\int_{-\infty}^{x_i} \delta(x)dx$

Options :
1. 1
2. 2
3. 3
4. 4

Question Number : 26 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical
Correct Marks : 2 Wrong Marks : 0

Random variables x_1, x_2 and x_3 take discrete values in set $\{0,1\}$. Given $p(x_1 = 1) = 3/4$; $p(x_2 = 0) = 1/2$; and x_1 is independent of x_2. Use $(x_1, x_2) \in \begin{bmatrix} 00 \\ 01 \\ 10 \\ 11 \end{bmatrix} and p(x_1, x_2) = p(x_2, x_3) = \begin{pmatrix} 0.125 \\ 0.375 \\ 0.250 \\ 0.250 \end{pmatrix}$ to find the value of $p(x_1 = 0, x_2 = 0, x_3 = 1)$

a. $1/64$
b. $1/32$
c. $3/32$
d. $1/16$

Options :
1. 1
2. 2
3. 3
4. 4

Question Number : 27 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical
Correct Marks : 2 Wrong Marks : 0

...ted to the following linear model

$$Aw = y, A\epsilon\mathbb{R}^{n\times m}, y\epsilon\mathbb{R}^{n}, w\epsilon\mathbb{R}^{m}$$

Given $n \ll m$,the minimum norm solution for the parameter vector w is

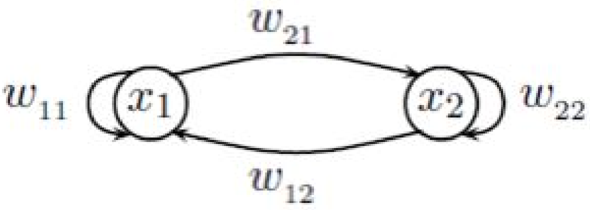
a. $w = (A^T A)A^T y^{-1}$
b. $w = A^T(A^{-1}A^T)y$
c. $w = A^T(AA^T)^{-1}y$
d. $w = (A^T A)^{-1}A^T y$

Options :

1. 1
2. 2
3. 3
4. 4

Question Number : 28 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical
Correct Marks : 2 Wrong Marks : 0

In the following recurrent network, the activation functions are linear. The instantaneous cost function is $E(t+1) = \frac{1}{2}\left(x_1^d(t+1) - x_1(t+1)\right)^2 + \frac{1}{2}\left(x_2^d(t+1) - x_2(t+1)\right)^2$. The weight update equation for w_{12} is



a. $w_{12} \leftarrow w_{12} + \eta(x_1^d(t+1) - x_1(t+1))\frac{\partial x_1(t+1)}{\partial w_{12}} + \eta(x_2^d(t+1) - x_2(t+1))\frac{\partial x_2(t+1)}{\partial w_{12}}$ where, $\frac{\partial x_1(t+1)}{\partial w_{12}} = w_{11}\frac{\partial x_1(t)}{\partial w_{12}} + w_{12}\frac{\partial x_2(t)}{\partial w_{12}} + x_2(t)$ and $\frac{\partial x_2(t+1)}{\partial w_{12}} = w_{21}\frac{\partial x_1(t)}{\partial w_{12}} + w_{22}\frac{\partial x_2(t)}{\partial w_{12}}$

b. $w_{12} \leftarrow w_{12} + \eta(x_1^d(t+1) - x_1(t+1))\frac{\partial x_1(t+1)}{\partial w_{12}} + \eta(x_2^d(t+1) - x_2(t+1))\frac{\partial x_2(t+1)}{\partial w_{12}}$ where, $\frac{\partial x_1(t+1)}{\partial w_{12}} = w_{11}\frac{\partial x_1(t)}{\partial w_{12}} + w_{12}\frac{\partial x_2(t)}{\partial w_{12}}$ and $\frac{\partial x_2(t+1)}{\partial w_{12}} = w_{21}\frac{\partial x_1(t)}{\partial w_{12}} + w_{22}\frac{\partial x_2(t)}{\partial w_{12}}$

c. $w_{12} \leftarrow w_{12} + \eta(x_1^d(t+1) - x_1(t+1))\frac{\partial x_1(t+1)}{\partial w_{12}}$

where, $\frac{\partial x_1(t+1)}{\partial w_{12}} = w_{11}\frac{\partial x_1(t)}{\partial w_{12}} + w_{12}\frac{\partial x_2(t)}{\partial w_{12}}$

d. $w_{12} \leftarrow w_{12} + \eta(x_2^d(t+1) - x_2(t+1))\frac{\partial x_2(t+1)}{\partial w_{12}}$

where $\frac{\partial x_2(t+1)}{\partial w_{12}} = w_{21}\frac{\partial x_1(t)}{\partial w_{12}} + w_{22}\frac{\partial x_2(t)}{\partial w_{12}}$

Options :

1. 1
2. 2
3. 3
4. 4

Question Number : 29 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical
Correct Marks : 2 Wrong Marks : 0

A nonlinear function is given as $y = x \sin x\,; 0 < x < 1$. Representation of this function using 1-D KSOM based function approximation model consisting of three neurons whose weights are 0.25, 0.5 and 0.75 respectively is

$$y = \frac{\Sigma h_\gamma (y_\gamma + A_\gamma (x - w_\gamma))}{\Sigma h_\gamma}$$

Neighborhood function value is 1 for the winner and nearest neighbor is 0.25. The Jacobian associated with each neuron are

a. $A_\gamma = 0.06$ at $w_\gamma = 0.25$; $A_\gamma = 0.81$ at $w_\gamma = 0.5$; $A_\gamma = 0.9$ at $w_\gamma = 0.75$
b. $A_\gamma = 0.49$ at $w_\gamma = 0.25$; $A_\gamma = 0.92$ at $w_\gamma = 0.5$; $A_\gamma = 1.2$ at $w_\gamma = 0.75$
c. $A_\gamma = 0.06$ at $w_\gamma = 0.25$; $A_\gamma = 0.24$ at $w_\gamma = 0.5$; $A_\gamma = 0.5$ at $w_\gamma = 0.75$
d. $A_\gamma = 0.41$ at $w_\gamma = 0.25$; $A_\gamma = 0.31$ at $w_\gamma = 0.5$; $A_\gamma = 2.1$ at $w_\gamma = 0.75$

Options :
1. 1
2. 2
3. 3
4. 4

Question Number : 30 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical
Correct Marks : 2 Wrong Marks : 0

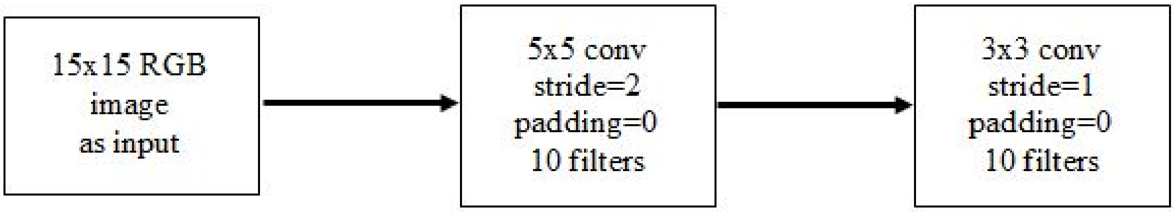
The cost function to update the joint angle vector θ in kinematic control of a robot arm using KSOM based network is

a. $E = \frac{1}{2} \| \theta_0^{out} - \frac{1}{s}\Sigma_\gamma h_\gamma \left(\theta_\gamma + A_\gamma (v_0 - w_\gamma)\right) \|^2; s = \Sigma_\gamma h_\gamma$
b. $E = \frac{1}{2} \| \Delta\theta^{out} - \frac{1}{s}\Sigma_\gamma h_\gamma A_\gamma \Delta v \|^2; s = \Sigma_\gamma h_\gamma$
c. $E = \frac{1}{2} (y^d - y)^2$
d. $E = \frac{1}{2} \| y^d - y \|^2$

Options :
1. 1
2. 2
3. 3
4. 4

Question Number : 31 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical
Correct Marks : 2 Wrong Marks : 0

...lowing figure, the output of the last convolution layer is



a. 15x15x10
b. 6x6x10
c. 4x4x10
d. None of the above

Options :
1. 1
2. 2
3. 3
4. 4

Question Number : 32 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical
Correct Marks : 2 Wrong Marks : 0

The memory requirement of the network given in Q31 is close to

a. 1800 x 4 bytes
b. 1500 x 4 bytes
c. 1200 x 4 bytes
d. 900 x 4 bytes

Options :
1. 1
2. 2
3. 3
4. 4

Question Number : 33 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical
Correct Marks : 2 Wrong Marks : 0

Which of the following can be best used to reduce under fitting in a deep network?

a. Sigmoid non linearity at lower layers
b. Auxiliary outputs at lower layers
c. Max pooling at lower layers
d. 1x1 conv at lower layers

Options :
1. 1
2. 2
3. 3
4. 4

Question Number : 34 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical
Correct Marks : 2 Wrong Marks : 0

...t to a convolutional layer with 11x11 filters with stride 1 is an image of size 15x15. In order to preserve the size of the input in the output, padding of size P along the border can be used, where

a. P=2
b. P=3
c. P=4
d. P=5

Options :

1. 1
2. 2
3. 3
4. 4

Question Number : 35 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 2 Wrong Marks : 0

A 3-input neuron has weights 3, 6 and 9. The transfer function is linear with the constant of proportionality being equal to 4. The inputs are 2, 3 and 5 respectively. The output will be:

a. 238
b. 276
c. 119
d. 123

Options :

1. 1
2. 2
3. 3
4. 4

Question Number : 36 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 2 Wrong Marks : 0

In the following python code using a sequential model from keras library, what is the total number of weights to be trained (including the biases)?

```
Model. Add (Dense (24, input dim= 100, activation='relu'))
Model. Add (Dense (2, activation='softmax'))
```

a. 2448
b. 2402
c. 2472
d. 2474

Options :

1. 1
2. 2
3. 3
4. 4

Question Number : 37 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 2 Wrong Marks : 0

...he output of the following Python program?

```
For i in range (4):
    print(i+1),
    if(i==1):
        break
    else:
        print(i),
    For j in range (1):
        print('z'),
```

a. 1 0 z 2
b. 1 0 z 2 z
c. 1 z 2 z
d. None of these

Options :
1. 1
2. 2
3. 3
4. 4

Question Number : 38 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical
Correct Marks : 2 Wrong Marks : 0

What will be displayed by the following Python code?

```
class adden:
    def __init__(self, a):
        self. a = a
    def add (self, b):
        self.a += 1
        return self. a+b
func = adden (4)
print (func. add (4))
```

a. 4
b. 5
c. 10
d. 9

Options :
1. 1
2. 2
3. 3
4. 4

Question Number : 39 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical
Correct Marks : 2 Wrong Marks : 0

...f the following statements are correct?

Statement 1: LSTM and GRU both can capt... dependencies.

Statement 2: LSTMs are faster to train than GRUs because the former involve lesser parameters.

a. 1 only
b. 2 only
c. Both
d. None

Options :

1. 1
2. 2
3. 3
4. 4

Question Number : 40 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 2 Wrong Marks : 0

Assume a 4×3 RBM and a training vector $x^{(t)}$. Assuming a random number generator generates uniform random vectors $U_1 = \begin{pmatrix} 0.27 \\ 0.62 \\ 0.41 \end{pmatrix}$, $U_2 = \begin{pmatrix} 0.03 \\ 0.68 \\ 0.52 \\ 0.29 \end{pmatrix}$. Using these random vectors in the same order for one iteration of training of a 4×3 RBM given the training vector $x^{(t)}$. Assume initial weights and biases are zero vectors. Given $x^{(t)} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$. The reconstructed hidden activation $h(\tilde{x}^{(t)})$ and the reconstruction error $\frac{\|x^{(t)} - \tilde{x}^{(t)}\|^2}{2}$ are given by

a. $\begin{pmatrix} 0.5 \\ 0.5 \\ 0.5 \end{pmatrix}$ and 0.5 sq units

b. $\begin{pmatrix} 0.5 \\ 1 \\ 0.5 \end{pmatrix}$ and 1 sq units

c. $\begin{pmatrix} 1 \\ 0.5 \\ 1 \end{pmatrix}$ and 1.5 sq units

d. $\begin{pmatrix} 0.5 \\ 1 \\ 1 \end{pmatrix}$ and 2 sq units

Options :

1. 1
2. 2
3. 3
4. 4

Question Number : 41 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 2 Wrong Marks : 0

A convolution layer has 200 3x3 filters with no padding and stride size=1. Its input size is 30x30x10. What is the approximate computational complexity (number of multiplication operations) of this layer? (here M=1,000,000)

a. 16 M
b. 14 M
c. 6 M
d. 2 M

Options :

1. 1
2. 2
3. 3
4. 4

Question Number : 42 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 2 Wrong Marks : 0

While evaluating a NN model for disease detection (binary classification) in a plant, it is found that out of 20 diseased samples, it correctly detects only 10 of them, and out of 100 healthy samples, it wrongly detects 5 as diseased. What is the precision of the model?

a. 95.0%
b. 66.7%
c. 87.5%
d. 50.0%

Options :

1. 1
2. 2
3. 3
4. 4

Question Number : 43 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 2 Wrong Marks : 0

For the following system:

$\dot{x} = y$

$\dot{y} = -kx + y$

Taking the Lyapunov function $V(x,y) = \frac{1}{2}kx^2 + \frac{1}{2}y^2$, the system is

a. Stable
b. Asymptotically stable
c. Unstable
d. None of the above

Options :

1. 1
2. 2
3. 3
4. 4

Question Number : 44 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 2 Wrong Marks : 0

A CNN layer has ten FxF conv filters with stride 1 and padding 5 along the border. Its input and output both have size (28, 28, 10). What is the value of F?

a. 1
b. 5
c. 11
d. 28

Options :

1. 1
2. 2
3. 3
4. 4

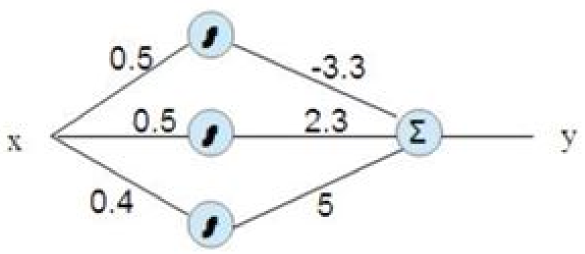
Sub-Section Number:	3
Sub-Section Id:	489994314
Question Shuffling Allowed :	Yes

Question Number : 45 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 4 Wrong Marks : 0

The following figure is a two layered network whose hidden neurons are sigmoid ally activated while the output neuron is linear i.e. output is summation of all inputs to this neuron. The input to the neuron x is not known but the output is $y = 1.5$. You need to search the input space starting with $x = 0.5$ using gradient descent algorithm. Take $\eta = 1$;

$$x(t+1) = x(t) - \eta \frac{dE}{dx};\ E = \frac{1}{2}(y^d - y)^2$$



The value of x after two iterations is

a. 0.62

b. 1.04

c. 0.02

d. 0.25

Options :

1. 1
2. 2
3. 3
4. 4

Question Number : 46 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical
Correct Marks : 4 Wrong Marks : 0

A process is described by $\begin{bmatrix} x_1(t+1) \\ x_2(t+1) \end{bmatrix} = \begin{bmatrix} 0.6 & -1 \\ 0 & 0.6 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$ where matrix parameters are unknown. Weights are updated using real-time recurrent learning. Take $u(0) = x_1(0) = x_2(0) = 0.5$ and $\eta = 1$. Assume initial parameter values to be $W_1 = \begin{bmatrix} 0.2 & -0.6 \\ 0 & 0.5 \end{bmatrix}; W_2 = \begin{bmatrix} 0 \\ 0.1 \end{bmatrix}$. Assume initial values of gradients to be zeros. The parameter values after one iteration are

a. $W_1 = \begin{bmatrix} -0.2 & -0.6 \\ -0.25 & 0.25 \end{bmatrix}; W_2 = \begin{bmatrix} 0 \\ -0.35 \end{bmatrix}$

b. $W_1 = \begin{bmatrix} 0.2 & 0.6 \\ -0.25 & -0.25 \end{bmatrix}; W_2 = \begin{bmatrix} 0 \\ 0.35 \end{bmatrix}$

c. $W_1 = \begin{bmatrix} -0.2 & 0.6 \\ 0.25 & -0.25 \end{bmatrix}; W_2 = \begin{bmatrix} 0 \\ -0.35 \end{bmatrix}$

d. $W_1 = \begin{bmatrix} 0.2 & -0.6 \\ 0.25 & 0.25 \end{bmatrix}; W_2 = \begin{bmatrix} 0 \\ 0.35 \end{bmatrix}$

Options :

1. 1
2. 2
3. 3
4. 4

Question Number : 47 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical
Correct Marks : 4 Wrong Marks : 0

... a $(4 \times 3 \times 2)$ Deep Neural Network and a training vector $v_1 = [1010]$.

Assume a random number generator generates uniform random vectors $U_1 = \begin{pmatrix} 0.27 \\ 0.62 \\ 0.41 \end{pmatrix}$,

$U_2 = \begin{pmatrix} 0.03 \\ 0.68 \\ 0.52 \\ 0.29 \end{pmatrix}$. Use these random vectors in the same order for one iteration of training of a $(4 \times 3$ RBM) using $v_1 = [1\ 0\ 1\ 0]$. The weights and biases of the 4×3 RBM after one iteration assuming initial weights and biases are zero vectors are [Use learning rate $\eta = 1$ in the update rule]

a. $W = \begin{pmatrix} 0 & 0 & 0.5 & -0.5 \\ 0 & 0.1 & 0.5 & -0.5 \\ 0 & 0 & -0.5 & 0.5 \end{pmatrix}, b = \begin{pmatrix} 0.1 \\ 0.1 \\ 0.1 \end{pmatrix}, c = \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}$

b. $W = \begin{pmatrix} 0 & 0 & 0.5 & -0.5 \\ 0 & 0 & 0.5 & -0.5 \\ 0 & 0 & 0.5 & -0.5 \end{pmatrix}, b = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, c = \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}$

c. $W = \begin{pmatrix} 0.5 & -0.5 & 0 & 0 \\ 0.5 & -0.5 & 0 & 0 \\ 0.5 & -0.5 & 0 & 0 \end{pmatrix}, b = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, c = \begin{pmatrix} 0.5 \\ 0 \\ 1 \\ -1 \end{pmatrix}$

d. $W = \begin{pmatrix} 0 & 0 & 0.5 & -0.5 \\ 0 & 0.1 & 0.5 & -0.5 \\ 0 & 0 & -0.5 & 0.5 \end{pmatrix}, b = \begin{pmatrix} 0.1 \\ 0.1 \\ 0.1 \end{pmatrix}, c = \begin{pmatrix} 0 \\ 0.5 \\ 1 \\ -1 \end{pmatrix}$

Options :

1. 1
2. 2
3. 3
4. 4

Question Number : 48 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical

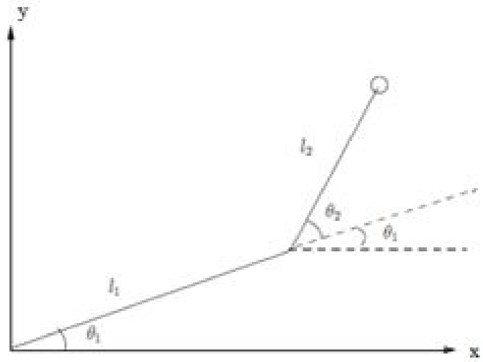
Correct Marks : 4 Wrong Marks : 0

Forward kinematics of a two-link planar robot is given as

$$x = l_1 \cos\theta_1 + l_2 \cos(\theta_1 + \theta_2)$$

$$y = l_1 \sin\theta_1 + l_2 \sin(\theta_1 + \theta_2)$$

where θ_1, θ_2 are link 1 and 2 angles as shown in the figure.



The corresponding inverse kinematics is learned using a 2-D KSOM based network of a lattice size 10×10. For every KSOM neuron, the associated linear model is:

$$\theta(u) = \theta_\lambda + A_\lambda (u - w_\lambda)$$

where $u = \begin{bmatrix} x \\ y \end{bmatrix}$; w_λ is the cluster center of the λ^{th} neuron. Given $l_1 = l_2 = 1m$.

After training, $u = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0.5 \\ 1.0 \end{bmatrix}$ is presented to the network to find out the corresponding joint angle vector $\theta = \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix}$. The winning neuron cluster center for this input is $w_{\lambda_w} = \begin{bmatrix} 0.42 \\ 0.92 \end{bmatrix}$. Ideal values for $\theta_{\lambda_w}; A_{\lambda_w}$ associated with this winning neuron are

a. $\theta_{\lambda_w} = \begin{pmatrix} 0.10 \\ 2.08 \end{pmatrix}; A = \begin{pmatrix} -0.92 & -0.82 \\ 0.42 & -0.58 \end{pmatrix}$

b. $\theta_{\lambda_w} = \begin{pmatrix} -2.18 \\ 2.08 \end{pmatrix}; A = \begin{pmatrix} 0.92 & 0.10 \\ 0.42 & 0.99 \end{pmatrix}$

c. $\theta_{\lambda_w} = \begin{pmatrix} 0.56 \\ 0.82 \end{pmatrix}; A = \begin{pmatrix} -0.82 & -0.90 \\ 0.56 & -0.44 \end{pmatrix}$

d. $\theta_{\lambda_w} = \begin{pmatrix} 0.96 \\ 2.08 \end{pmatrix}; A = \begin{pmatrix} -0.92 & -0.10 \\ -0.42 & -0.99 \end{pmatrix}$

Options :

1. 1
2. 2
3. 3
4. 4

Question Number : 49 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 4 Wrong Marks : 0

A robot with single Sonar while avoiding an obstacle recorded following data. Please note that the robot either goes straight or turns based on the sonar readings which are estimates of the distance of the obstacle up-front. When both velocities are equal, the robot moves straight while unequal velocities lead to left or right turn.

Sonar S	Left Wheel Velocity, v_l	Right Wheel Velocity, v_r
40cm	8cm/sec	8cm/sec
20cm	4cm/sec	0cm/sec

A radial basis function network is designed that can mimic this behavior where sonar data is input and velocities are output. Take two centers as 0.4 m, and 0.2 m. The radial function is given as $e^{-x^2/2\sigma^2}$ where $\sigma = 0.2$ for both the centers. There is a bias at the output node. The parameters w_1, w_2 and bias θ for predicting right wheel velocity are

a. $w_1 = 7.18, w_2 = -2.97$ and bias $\theta = 2.62$
b. $w_1 = 11.56, w_2 = -8.76$ and bias $\theta = 1.75$
c. $w_1 = 11.41, w_2 = -8.92$ and bias $\theta = 1.0$
d. $w_1 = 12.0, w_2 = -8.3$ and bias $\theta = 2.0$

Options :
1. 1
2. 2
3. 3
4. 4

Question Number : 50 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical
Correct Marks : 4 Wrong Marks : 0

ling salesman problem (TSP) asks the following question: Given a list of cities and the distances between each pair of cities, what is the shortest possible route that visits each city exactly once and returns to the origin city? This problem can be solved using a Kohonen SOM 1-d lattice. Given following five city problem:

City	X coordinate	Y Coordinate
A	50	250
B	30	30
C	200	200
D	100	60
E	80	150

Select a 1-d lattice with five neurons with weights (30, 20), (20,30), (0,0), (40,60) and (60,40) respectively. Obviously neuron 2 and 5 are neighbours of neuron 1 while neuron 4 and 1 are neighbours of neuron 5. On convergence, each neuron will represent one city coordinate while TS will travel from neuron 1 to neuron 5 via neurons 2, 3 and 4 respectively before returning to the city represented by neuron 1. All neurons participate in weight update. Distance measures for the winning neuron, 1st immediate neighbour, and 2nd immediate neighbour are 1, 0.5 and 0.1 respectively. Learning rate is fixed at 1.0. Updated weights of neurons as the first city coordinates are presented to this lattice

a. $w_1 = (50,250), w_2 = (30,30), w_3 = (200,200), w_4 = (100,60), w_5 = (80,150)$
b. $w_1 = (30,20), w_2 = (20,30), w_3 = (0,0), w_4 = (40,60), w_5 = (60,40)$
c. $w_1 = (40,135), w_2 = (23,52), w_3 = (5,25), w_4 = (45,205), w_5 = (50,250)$
d. $w_1 = (32,43), w_2 = (23,52), w_3 = (25,125), w_4 = (50,250), w_5 = (55,145)$

Options :

1. 1
2. 2
3. 3
4. 4

Question Number : 51 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical

Correct Marks : 4 Wrong Marks : 0

[illegible] connected output layer of a classification network has 3 neurons. It is trained using the loss function $\mathcal{L} = -\sum_i t_i \log v_i$, for $i = 0$ [illegible] [illegible] as x_k w[illegible] the input nodes, and its output as $v_i = softmax(h_i)$, where $h_i = \sum_k w_{ki} x_k$ and w_{ki} are neuron weights. For a sample, having $i = 1$ as the true output, the gradient of the loss function with respect to the weight w_{k0} will be

a. 0
b. $\frac{exp(h_0)x_k}{\sum_j exp(h_j)}$
c. $\frac{exp(h_1)x_k}{\sum_j exp(h_j)}$
d. $\frac{exp(h_0)exp(h_1)x_k}{\left(\sum_j exp(h_j)\right)^2}$

Options :
1. 1
2. 2
3. 3
4. 4

Question Number : 52 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical
Correct Marks : 4 Wrong Marks : 0

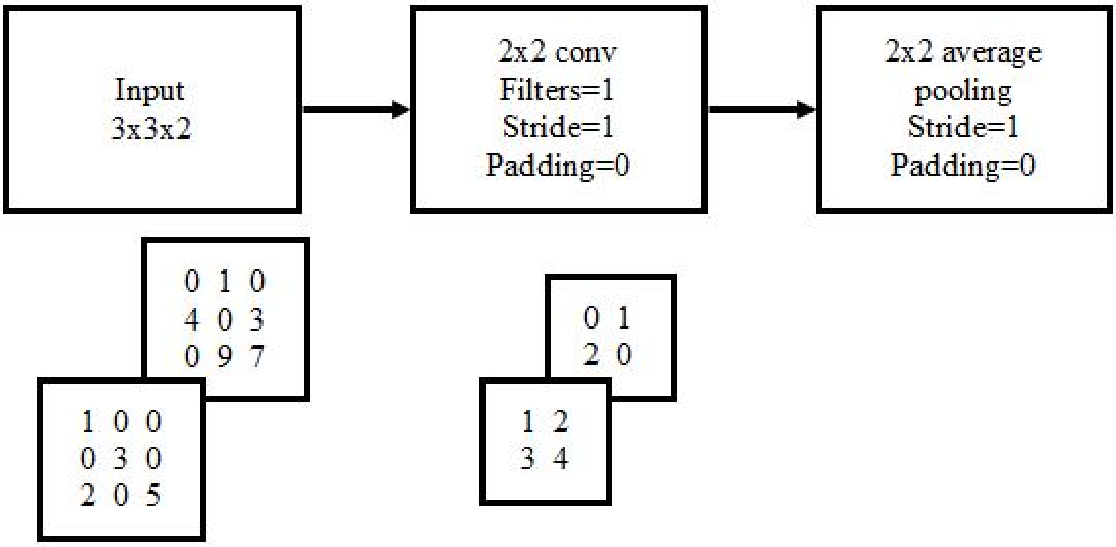
The effective receptive field of a stack of L layers of FxF convolution filters, with stride S (< F) at each layer, is same as that of one MxM convolutional layer. Here, M is given by

a. $M = FS^{L-1} + (F - S)(1 + S + \ldots + S^{L-2})$
b. $M = FS^{L-1} + (F - S)(L - 1)$
c. $M = FS^{L} + (F - 1)(L - 1)$
d. None of the above

Options :
1. 1
2. 2
3. 3
4. 4

Question Number : 53 Question Type : MCQ Option Shuffling : No Display Question Number : Yes Single Line Question Option : No Option Orientation : Vertical
Correct Marks : 4 Wrong Marks : 0

…ollowing CNN, the values of the input and the filter are specified. The final output will be:



a. [21.75]
b. [7.5, 14.25]
c. [10.875]
d. [15.0, 14.25]

Options :
1. 1
2. 2
3. 3
4. 4