Id	1
Question	Back propagation is a learning technique that adjusts weights in the neural network by propagating weight changes.
A	Forward from source to sink
B	Backward from sink to source
C	Forward from source to hidden nodes
D	Backward from sink to hidden nodes
Answer	
Marks	1.5
Unit	I

Id	2
Question	Identify the following activation function : φ(V) = Z + (1/ 1 + exp (− x * V + Y)), Z, X, Y are parameters
A	Step function
B	Ramp function
C	Sigmoid function
D	Gaussian function
Answer	
Marks	1.5
Unit	II

Id	3
Question	A neuron with 3 inputs has the weight vector [0.2 -0.1 0.1]^T and a bias θ = 0. If the input vector is X = [0.2 0.4 0.2]^T then the total input to the neuron is:
A	0.2
B	1
C	0.02
D	-1.0
Answer	
Marks	1.5
Unit	I

Id	4
Question	Which of the following neural networks uses supervised learning? (A) Multilayer perceptron (B) Self organizing feature map (C) Hopfield network
A	(A) only
B	(B) only
C	(A) and (B) only
D	(A) and (C) only
Answer	
Marks	1.5
Unit	II

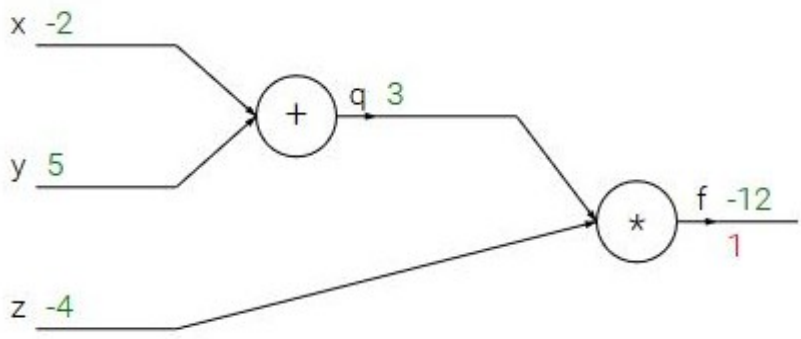| Id	5
Question	In Delta Rule for error minimization
A	weights are adjusted w.r.to change in the output
B	weights are adjusted w.r.to difference between desired output and actual output
C	weights are adjusted w.r.to difference between input and output
D	none of the above
Answer	
Marks	1.5
Unit	I

Id	6
Question	Which of the following can be used for clustering of data ?
A	Single layer perception
B	Multilayer perception
C	Self organizing map
D	Radial basis function
Answer	
Marks	1.5
Unit	I

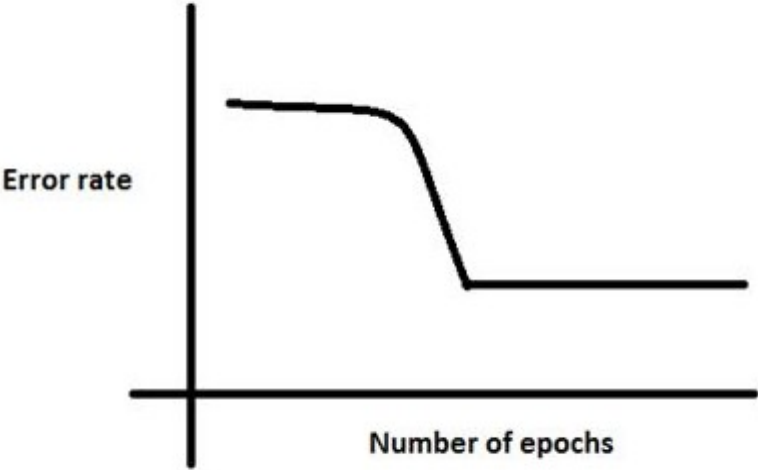| Id	7
Question	A network is created when we multiple neurons stack together. Let us take an example of a neural network simulating an XNOR function. You can see that the last neuron takes input from two neurons before it. The activation function for all the neurons is given by: $$f(x) = \begin{cases} 0, & for\ x < 0 \\ 1, & for\ x \geq 0 \end{cases}$$ Suppose X1 is 0 and X2 is 1, what will be the output for the above neural network?
A	0
B	1
C	2
D	3
Answer	
Marks	1.5
Unit	I

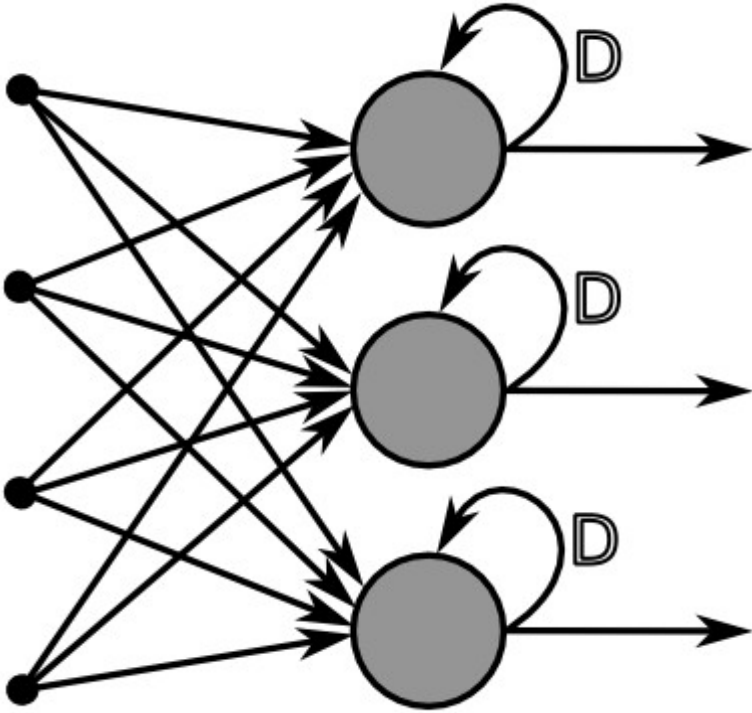| Id	8
Question	In a neural network, knowing the weight and bias of each neuron is the most important step. If you can somehow get the correct value of weight and bias for each neuron, you can approximate any function. What would be the best way to approach this?
A	Assign random values
B	Search every possible combination of weights and biases till you get the best value
C	Iteratively check that after assigning a value how far you are from the best values, and slightly change the assigned values values to make them better
D	None of these
Answer	
Marks	1.5
Unit	I

Id	9
Question	What are the steps for using a gradient descent algorithm? 1. Calculate error between the actual value and the predicted value 2. Reiterate until you find the best weights of network 3. Pass an input through the network and get values from output layer 4. Initialize random weight and bias 5. Go to each neurons which contributes to the error and change its respective values to reduce the error
A	1, 2, 3, 4, 5
B	5, 4, 3, 2, 1
C	3, 2, 1, 5, 4
D	4, 3, 1, 5, 2
Answer	
Marks	1.5
Unit	II

Id	10
Question	**Suppose you have inputs as x, y, and z with values -2, 5, and -4 respectively. You have a neuron 'q' and neuron 'f' with functions:** **q = x + y** **f = q * z** **Graphical representation of the functions is as follows:** **What is the gradient of F with respect to x, y, and z?**
A	(-3,4,4)
B	(4,4,3)
C	(-4,-4,3)
D	(3,-4,-4)
Answer	
Marks	1.5
Unit	I

Id	11
Question	Which of the following techniques perform similar operations as dropout in a neural network?
A	Bagging
B	Boosting
C	Stacking
D	None of these
Answer	
Marks	1.5
Unit	II

Id	12
Question	Which of the following gives non-linearity to a neural network?
A	Stochastic Gradient Descent
B	Rectified Linear Unit
C	Convolution function
D	None of the above
Answer	
Marks	1.5
Unit	III

Id	13
Question	In training a neural network, you notice that the loss does not decrease in the few starting epochs. **The reasons for this could be:** 1.The learning is rate is low 2.Regularization parameter is high 3.Stuck at local minima **What according to you are the probable reasons?**
A	1 and 2
B	2 and 3
C	1 and 3
D	Any of these
Answer	
Marks	1.5
Unit	I

Id	14
Question	**You are building a neural network where it gets input from the previous layer as well as from itself.** **Which of the following architecture has feedback connections?**
A	Recurrent Neural network
B	Convolutional Neural Network
C	Restricted Boltzmann Machine
D	None of these
Answer	
Marks	1.5
Unit	II

Id	15
Question	**What is the sequence of the following tasks in a perceptron?** 1. **Initialize weights of perceptron randomly** 2. **Go to the next batch of datase** 3. **If the prediction does not match the output, change the weights** 4. **For a sample input, compute an output**
A	1, 2, 3, 4
B	4, 3, 2, 1
C	3, 1, 2, 4
D	1, 4, 3, 2
Answer	
Marks	1.5
Unit	I

Id	16
Question	Suppose that you have to minimize the cost function by changing the parameters. Which of the following technique could be used for this?
A	Exhaustive Search
B	Random Search
C	Bayesian Optimization
D	Any of these
Answer	
Marks	1.5
Unit	I

Id	17
Question	In which neural net architecture, does weight sharing occur?
A	Convolutional neural Network
B	Recurrent Neural Network
C	Fully Connected Neural Network
D	Both A and B
Answer	
Marks	1.5
Unit	II

Id	18
Question	Batch Normalization is helpful because
A	It normalizes (changes) all the input before sending it to the next layer
B	It returns back the normalized mean and standard deviation of weights
C	It is a very efficient backpropagation technique
D	None of these
Answer	
Marks	1.5
Unit	III

Id	19
Question	Instead of trying to achieve absolute zero error, we set a metric called bayes error which is the error we hope to achieve. What could be the reason for using bayes error?
A	Input variables may not contain complete information about the output variable
B	System (that creates input-output mapping) may be stochastic
C	Limited training data
D	All the above
Answer	
Marks	1.5
Unit	III

Id	20
Question	In a neural network, which of the following techniques is used to deal with overfitting?
A	Dropout
B	Regularization
C	Batch Normalization
D	All of these
Answer	
Marks	1.5
Unit	II

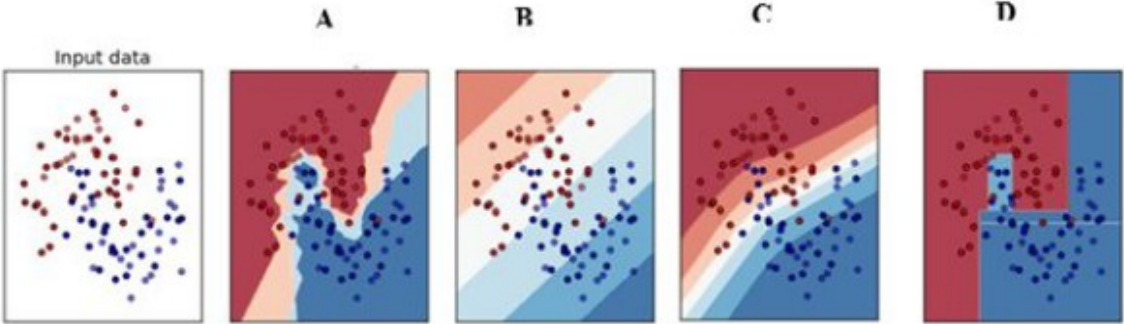| Id	21
Question	What is a dead unit in a neural network?
A	A unit which doesn't update during training by any of its neighbour
B	A unit which does not respond completely to any of the training patterns
C	The unit which produces the biggest sum-squared error
D	None of these
Answer	
Marks	1.5
Unit	I

Id	22
Question	Which of the following statement is the best description of early stopping?
A	Train the network until a local minimum in the error function is reached
B	Simulate the network on a test dataset after every epoch of training. Stop training when the generalization error starts to increase
C	Add a momentum term to the weight update in the Generalized Delta Rule, so that training converges more quickly
D	A faster version of backpropagation, such as the "Quickprop' algorithm
Answer	
Marks	1.5
Unit	III

Id	23
Question	For a classification task, instead of random weight initializations in a neural network, we set all the weights to zero. Which of the following statements is true?
A	There will not be any problem and the neural network will train properly
B	The neural network will train but all the neurons will end up recognizing the same thing
C	The neural network will not train as there is no net gradient change
D	None of these
Answer	
Marks	1.5
Unit	I

Id	24
Question	**There is a plateau at the start. This is happening because the neural network gets stuck at local minima before going on to global minima.** Error rate Number of epochs **To avoid this, which of the following strategy should work?**
A	Increase the number of parameters, as the network would not get stuck at local minima
B	Decrease the learning rate by 10 times at the start and then use momentum
C	Jitter the learning rate, i.e. change the learning rate for a few epochs
D	None of these
Answer	
Marks	1.5
Unit	III

Id	25
Question	**For an image recognition problem (recognizing a cat in a photo), which architecture of neural network would be better suited to solve the problem?**
A	Multi Layer Perceptron
B	Convolutional Neural Network
C	Recurrent Neural network
D	Perceptron
Answer	
Marks	1.5
Unit	II

Id	26
Question	 **Which of the following is a decision boundary of Neural Network?**
A	**B**
B	**A**
C	**D**
D	**All of these**
Answer	
Marks	1.5
Unit	I

Id	27
Question	**What are the factors to select the depth of neural network?** 1. Type of neural network (eg. MLP, CNN etc) 2. Input data 3. Computation power, i.e. Hardware capabilities and software capabilities 4. Learning Rate 5. The output function to map
A	1, 2, 4, 5
B	2, 3, 4, 5
C	1, 3, 4, 5
D	All of these
Answer	
Marks	1.5
Unit	II

Id	28
Question	Consider the scenario. The problem you are trying to solve has a small amount of data. Fortunately, you have a pre-trained neural network that was trained on a similar problem. Which of the following methodologies would you choose to make use of this pre-trained network?
A	Re-train the model for the new dataset
B	Assess on every layer how the model performs and only select a few of them
C	Fine tune the last couple of layers only
D	Freeze all the layers except the last, re-train the last layer
Answer	
Marks	1.5
Unit	III

Id	29
Question	Artificial intelligence is--------------
A	It uses machine-learning techniques. Here program can learn From past experience and adapt themselves to new situations
B	Computational procedure that takes some value as input and produces some value as output.
C	Science of making machines performs tasks that would require intelligence when performed by humans
D	None of these
Answer	
Marks	1.5
Unit	I

Id	30
Question	Evolutionary computation is-----------
A	Combining different types of method or information
B	Approach to the design of learning algorithms that is structured along the lines of the theory of evolution.
C	Decision support systems that contain an information base filled with the knowledge of an expert formulated in terms of if-then rules.
D	None of these
Answer	
Marks	1.5
Unit	I

Id	31
Question	Search space is-------
A	The large set of candidate solutions possible for a problem
B	The information stored in a database that can be, retrieved with a single query.
C	Worth of the output of a machine learning program that makes it understandable for humans
D	None of these
Answer	
Marks	1.5
Unit	III

Id	32
Question	Perceptron is_____
A	General class of approaches to a problem.
B	Performing several computations simultaneously
C	Structures in a database those are statistically relevant
D	Simple forerunner of modern neural networks, without hidden layers
Answer	
Marks	1.5
Unit	I

Id	33
Question	Which of the following are universal approximators?
A	Kernel SVM
B	Neural Networks
C	Boosted Decision Trees
D	All of the above
Answer	
Marks	1.5
Unit	III

Id	34
Question	In which of the following applications can we use deep learning to solve the problem?
A	Protein structure prediction
B	Prediction of chemical reactions
C	Detection of exotic particles
D	All of these
Answer	
Marks	1.5
Unit	III

Id	35
Question	Which of the following statements is true when you use 1 X 1 convolutions in a CNN?
A	It can help in dimensionality reduction
B	It can be used for feature pooling
C	It suffers less overfiting due to small kernel size
D	All of the above
Answer	
Marks	1.5
Unit	III

Id	36
Question	Statement 1: It is possible to train a network well by initializing all the weights as 0. Statement 2: It is possible to train a network well by initializing biases as 0. Which of the statements given above is true?
A	Statement 1 is true while Statement 2 is false
B	Statement 2 is true while statement 1 is false
C	Both statements are true
D	Both statements are false
Answer	
Marks	1.5
Unit	III

Id	37
Question	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	Less than 50
C	More than 50
D	It is an arbitrary value
Answer	
Marks	1.5
Unit	I

Id	38
Question	The input image has been converted into a matrix of size 28 X 28 and a kernel/filter of size 7 X 7 with a stride of 1. What will be the size of the convoluted matrix?
A	22 X 22
B	21 X 21
C	28 X 28
D	7 X 7
Answer	
Marks	1.5
Unit	III

Id	39
Question	In a simple MLP model with 8 neurons in the input layer, 5 neurons in the hidden layer and 1 neuron in the output layer. What is the size of the weight matrices between hidden output layer and input hidden layer?
A	[1 X 5] , [5 X 8]
B	[8 X 5] , [1 X 5]
C	[8 X 5] , [5 X 1]
D	[5 x 1] , [8 X 5]
Answer	
Marks	1.5
Unit	I

Id	40
Question	Assume a simple MLP model with 3 neurons and inputs= 1,2,3. The weights to the input neurons are 4,5 and 6 respectively. Assume the activation function is a linear constant value of 3. What will be the output ?
A	32
B	643
C	96
D	48
Answer	
Marks	1.5
Unit	I

Id	41
Question	Which of following activation function can't be used at output layer to classify an image
A	Tanh
B	If (x>5,1,0)
C	ReLU
D	None of the above
Answer	
Marks	1.5
Unit	III

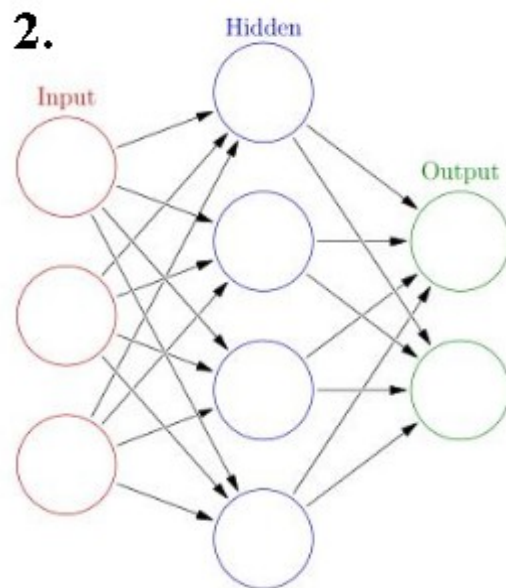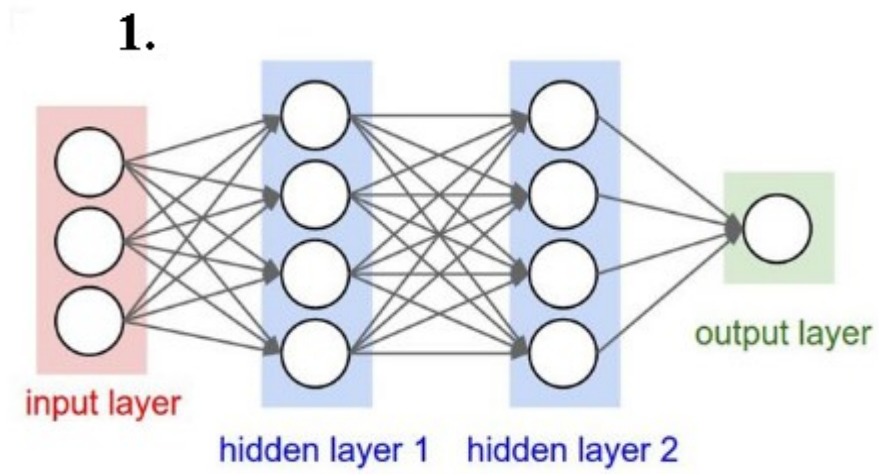| Id	42
Question	**Which of the following would have a constant input in each epoch of training a Deep Learning model?**
A	Weight between input and hidden layer
B	Weight between hidden and output layer
C	Biases of all hidden layer neurons
D	Activation function of output layer
Answer	
Marks	1.5
Unit	II

Id	43
Question	**What value would be in place of question mark?** INPUT FILTER CONVOLVED FEATURE Input matrix: 1 1 1 0 0 0 1 1 1 0 0 0 1 1 1 0 0 1 1 0 0 1 1 0 0 Filter: 1 0 1 0 1 0 1 0 1 Convolved Feature: ? Here we see a convolutional function being applied to input.
A	3
B	4
C	5
D	6
Answer	
Marks	1.5
Unit	III

Id	44
Question	**For a binary classification problem, which of the following architecture would you choose?**

1.

input layer

hidden layer 1   hidden layer 2

output layer

2.

Input

Hidden

Output

A	1
B	2
C	Any one of these
D	None of these
Answer	
Marks	1.5
Unit	I

Id	45
Question	Which of the following statement is true regrading dropout? 1: Dropout gives a way to approximate by combining many different architectures 2: Dropout demands high learning rates 3: Dropout can help preventing overfitting
A	Both 1 and 2
B	Both 1 and 3
C	Both 2 and 3
D	All 1,2 and 3
Answer	
Marks	1.5
Unit	III

Id	46
Question	**What steps can we take to prevent overfitting in a Neural Network?**
A	Weight Sharing
B	Early Stopping
C	Dropout
D	All of the above
Answer	
Marks	1.5
Unit	II

Id	47
Question	You are building a binary classifier for recognizing cucumbers (y=1) vs. watermelons (y=0). Which one of these activation functions would you recommend using for the output layer?
A	ReLU
B	LeakyReLU
C	sigmoid
D	Parametric ReLU
Answer	
Marks	1.5
Unit	II

Id	48
Question	Suppose you have built a neural network. You decide to initialize the weights and biases to be zero. Which of the following statements is true?
A	Each neuron in the first hidden layer will perform the same computation. So even after multiple iterations of gradient descent each neuron in the layer will be computing the same thing as other neurons.
B	Each neuron in the first hidden layer will perform the same computation in the first iteration. But after one iteration of gradient descent they will learn to compute different things because we have "broken symmetry".
C	Each neuron in the first hidden layer will compute the same thing, but neurons in different layers will compute different things, thus we have accomplished "symmetry breaking" as described in lecture.
D	The first hidden layer's neurons will perform different computations from each other even in the first iteration; their parameters will thus keep evolving in their own way.
Answer	
Marks	1.5
Unit	I

Id	49
Question	What are the issues on which biological networks proves to be superior than Al networks?
A	robustness and fault tolerance
B	flexibility
C	collective computation
D	all of the mentioned
Answer	
Marks	1.5
Unit	I

Id	50
Question	GPU stands for
A	Graphics Processing Unit
B	Gradient Processing Unit
C	General Processing Unit
D	Good Processing Unit.
Answer	
Marks	1.5
Unit	II

Id	51
Question	-------------is a Neural Nets way of classifying inputs.
A	Learning
B	Forward Propagation
C	Activation
D	Classification
Answer	
Marks	1.5
Unit	I

Id	52
Question	Name the component of a Neural Network where the true value of the input is not observed.
A	Hidden Layer
B	Gradient Descent
C	Activation Function
D	Output Layer
Answer	
Marks	1.5
Unit	I

Id	53
Question	_____ works best for Image Data.
A	Random Forest
B	Convolution Networks
C	single Layer Perceptrons
D	AutoEncoders
Answer	
Marks	1.5
Unit	II

Id	54
Question	_____ is a recommended Model for Pattern Recognition in Unlabeled Data.
A	Random Forest
B	Convolution Networks
C	single Layer Perceptrons
D	AutoEncoders
Answer	
Marks	1.5
Unit	II

Id	55
Question	Process of improving the accuracy of a Neural Network is called........................
A	Training
B	Random Walk
C	Cross Validation
D	Forward Propagation
Answer	
Marks	1.5
Unit	II

Id	56
Question	Support Vector Machines, Naive Bayes and Logistic Regression are used for solving problems.
A	Regression Time Series
B	Classification
C	Clustering
D	Image processing
Answer	
Marks	1.5
Unit	I

Id	57
Question	What does LSTM stand for?
A	Long Short Threshold Memory
B	Least Square Time Mean
C	Long Short Term Memory
D	Least Squares Term Memory
Answer	
Marks	1.5
Unit	III

Id	58
Question	What is the method to overcome the Decay of Information through time in RNN known as?
A	Gating
B	Back Propagation
C	Gradient Descent
D	Activation
Answer	
Marks	1.5
Unit	III

Id	59
Question	What is the best Neural Network Model for Temporal Data?
A	Multi Layer Perceptrons
B	Temporal Neural Networks
C	Convolution Neural Networks
D	Recurrent Neural Network
Answer	
Marks	1.5
Unit	III

Id	60
Question	ReLU stands for ….........
A	Rectified Linear Unit
B	Rectified Lagrangian Unit
C	Regressive Linear Unit
D	Regressive Lagrangian Unit
Answer	
Marks	1.5
Unit	II

Id	61
Question	Why is the Pooling Layer used in a Convolution Neural Network? They are of no use in CNN.
A	Image Sensing
B	Object Recognition
C	Dimension Reduction
D	Pattern Recognition
Answer	
Marks	1.5
Unit	II

Id	62
Question	What are the two layers of a Restricted Boltzmann Machine called?
A	Input and Output Layers
B	Recurrent and Convolution Layers
C	Activation and Threshold Layers
D	Hidden and Visible Layers
Answer	
Marks	1.5
Unit	II

Id	63
Question	The measure of Difference between two probability distributions is know as ……..
A	Probability Difference
B	Cost
C	KL Divergence
D	Error
Answer	
Marks	1.5
Unit	III

Id	64
Question	A_____matches or surpasses the output of an individual neuron to a visual stimull.
A	Convolution
B	Gradient
C	Cost
D	Max Pooling
Answer	
Marks	1.5
Unit	I

Id	65
Question	The rate at which cost changes with respect to weight or bias is called_____
A	Loss
B	Rate of Change
C	Gradient
D	Derivative
Answer	
Marks	1.5
Unit	I

Id	66
Question	Autoencoders are trained using...........
A	They do not require Training
B	Back Propagation
C	Reconstruction
D	Feed Forward
Answer	
Marks	1.5
Unit	II

Id	67
Question	De-noising and Contractive are examples of _____
A	Recurrent Neural Networks
B	Convolution Neural Networks
C	Autoencoders
D	Shallow Neural Networks
Answer	
Marks	1.5
Unit	II

Id	68
Question	What is the purpose of the Gradient Descent algorithm?
A	To normalize the inputs
B	To minimize the weights and bias
C	To minimize the loss function
D	To prevent model from overfitting
Answer	
Marks	1.5
Unit	III

Id	69
Question	Which of the following pre-trained models in Keras can be fine-tuned for image classification?
A	VGG16
B	MobileNet
C	InceptionResNetV2
D	All of the above
Answer	
Marks	1.5
Unit	III

Id	70
Question	Least Squares Estimation minimizes:
A	summation of squares of errors
B	summation of errors
C	summation of absolute values of errors
D	All
Answer	
Marks	1.5
Unit	III

Id	71
Question	Parameter Estimation problem is about:
A	Identifying Input Parameters
B	Identifying Output Parameters
C	Identifying Model Parameters
D	All
Answer	
Marks	1.5
Unit	III

Id	72
Question	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. The output will be:
A	238
B	76
C	119
D	123
Answer	
Marks	1.5
Unit	I

Id	73
Question	In the case of an algebraic model for a straight line, if a value for the x variable is specified, then …...............
A	the exact value of the response variable can be computed
B	the computed response to the independent value will always give a minimal residual
C	the computed value of y will always be the best estimate of the mean response
D	none of these alternatives is correct.
Answer	
Marks	1.5
Unit	I

Id	74
Question	How can states of units be updated in hopfield model?
A	synchronously
B	asynchronously
C	synchronously and asynchronously
D	none of the mentioned
Answer	
Marks	1.5
Unit	II

Id	75
Question	What is synchronous update in hopfield model?
A	all units are updated simultaneously
B	a unit is selected at random and its new state is computed
C	a predefined unit is selected and its new state is computed
D	none of the mentioned
Answer	
Marks	1.5
Unit	II

Id	76
Question	What is asynchronous update in hopfield model?
A	all units are updated simultaneously
B	a unit is selected at random and its new state is computed
C	a predefined unit is selected and its new state is computed
D	none of the mentioned
Answer	
Marks	1.5
Unit	II

Id	77
Question	What is gradient descent?
A	method to find the absolute minimum of a function
B	method to find the absolute maximum of a function
C	maximum or minimum, depends on the situation
D	none of the mentioned
Answer	
Marks	1.5
Unit	1.5
Id	III

Id	78
Question	method to find the absolute minimum of a function
A	all units are updated simultaneously
B	a unit is selected at random and its new state is computed
C	a predefined unit is selected and its new state is computed
D	none of the mentioned
Answer	
Marks	1.5
Unit	III

Id	79
Question	If pattern is to be stored, then what does stable state should have updated value of?
A	current sate
B	next state
C	both current and next state
D	none of the mentioned
Answer	
Marks	1.5
Unit	I

Id	80
Question	How can error in recall due to false minima be reduced?
A	deterministic update for states
B	stochastic update for states
C	not possible
D	none of the mentioned
Answer	
Marks	1.5
Unit	I

Id	81
Question	Pattern storage problem which cannot be represented by a feedback network of given size can be called as?
A	easy problems
B	hard problems
C	no such problem exist
D	none of the mentioned
Answer	
Marks	1.5
Unit	III

Id	82
Question	For what purpose Feedback neural networks are primarily used?
A	classification
B	feature mapping
C	pattern mapping
D	none of the mentioned
Answer	
Marks	1.5
Unit	II

Id	83
Question	Presence of false minima will have what effect on probability of error in recall?
A	directly
B	inversely
C	no effect
D	directly or inversely
Answer	
Marks	1.5
Unit	II

Id	84
Question	How is effect false minima reduced
A	deterministic update of weights
B	stochastic update of weights
C	deterministic or stochastic update of weights
D	none of the mentioned
Answer	
Marks	1.5
Unit	II

Id	85
Question	Match the following knowledge representation techniques with their applications: List – I (a) Frames (b) Conceptual dependencies (c) Associative networks (d) Scripts List – II (i) Pictorial representation of objects, their attributes and relationships (ii) To describe real world stereotype events (iii) Record like structures for grouping closely related knowledge (iv) Structures and primitives to represent sentences code: a b c d
A	(iii) (iv) (i) (ii)
B	(iii) (iv) (ii) (i)
C	(iv) (iii) (i) (ii)
D	(iv) (iii) (ii) (i)
Answer	
Marks	1.5
Unit	II

Id	86
Question	Slots and facets are used in
A	Semantic Networks
B	Frames
C	Rules
D	All of these
Answer	
Marks	1.5
Unit	III

Id	87
Question	Consider the following statements: (a) If primal (dual) problem has a finite optimal solution, then its dual (primal) problem has a finite optimal solution. (b) If primal (dual) problem has an unbounded optimum solution, then its dual (primal) has no feasible solution at all. (c) Both primal and dual problems may be infeasible. Which of the following is correct?
A	(a) and (b) only
B	(a) and (c) only
C	(b) and (c) only
D	(a), (b) and (c)
Answer	
Marks	1.5
Unit	III

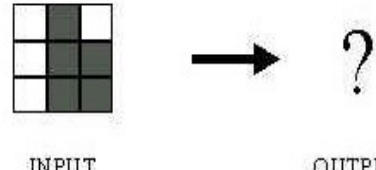| Id	88
Question	Consider the following statements : (a) Assignment problem can be used to minimize the cost. (b) Assignment problem is a special case of transportation problem. (c) Assignment problem requires that only one activity be assigned to each resource. Which of the following options is correct?
A	(a) and (b) only
B	(a) and (c) only
C	(b) and (c) only
D	(a), (b) and (c)
Answer	
Marks	1.5
Unit	III

Id	89
Question	What is purpose of Axon in Biological Neural network?
A	receptors
B	transmitter
C	transmission
D	none of the mentioned
Answer	
Marks	1.5
Unit	I

Id	90
Question	**A neural network model is said to be inspired from the human brain. The neural network consists of many neurons, each neuron takes an input, processes it and gives an output.**
A	A neuron has a single input and a single output only
B	A neuron has multiple inputs but a single output only
C	A neuron has a single input but multiple outputs
D	All of the above statements are valid
Answer	
Marks	1.5
Unit	I

Id	91
Question	Consider the following statements : (a) If you increase the number of hidden layers in a Multi Layer Perceptron, the classification error of test data always decreases. (b) A neuron has multiple inputs and multiple outputs . True or False
A	a)False b) True
B	a)False b) False
C	a)True b) True
D	a)True b) False
Answer	
Marks	1.5
Unit	II

Id	92
Question	Consider the following statements : (a) Convolutional Neural Networks can perform various types of transformation (rotations or scaling) in an input". (b) Bagging perform similar operations as dropout in a neural network . True or False
A	a)False b) True
B	a)False b) False
C	a)True b) True
D	a)True b) False
Answer	
Marks	1.5
Unit	II

Id	93
Question	Consider the following statements : (a) Convolutional Neural Networks can perform various types of transformation (rotations or scaling) in an input". (b) a neural network model the function (y=1/x) **True or False**
A	a)False b) True
B	a)False b) False
C	a)True b) True
D	a)True b) False
Answer	
Marks	1.5
Unit	II

Id	94
Question	**The network shown in Figure is trained to recognize the characters H and T as shown below:** INPUT → OUTPUT, INPUT → OUTPUT. If the following pattern was given INPUT → ? OUTPUT
A	
B	
C	
D	Could be A or B depending on the weights of neural network
Answer	
Marks	1.5
Unit	II

Id	95
Question	Consider the following statements : (a) **Suppose a convolutional neural network is trained on ImageNet dataset (Object recognition dataset). This trained model is then given a completely white image as an input.The output probabilities for this input would be equal for all classes. True or False?** (b) **When pooling layer is added in a convolutional neural network, translation invariance is preserved. True or False**
A	a)False b) True
B	a)False b) False
C	a)True b) True
D	a)True b) False
Answer	
Marks	1.5
Unit	II

Id	96
Question	**Which tool is best suited for solving Deep Learning problems**
A	R
B	Sk-learn
C	Excel
D	TensorFlow
Answer	
Marks	1.5
Unit	III

Id	97
Question	**Kears is a deep learning framework on which tool**
A	R
B	Sk-learn
C	Excel
D	TensorFlow
Answer	
Marks	1.5
Unit	II

Id	98
Question	**Which tool is a deep learning wrapper on TensorFlow**
A	Python
B	Keras
C	PyTourch
D	Azure
Answer	
Marks	1.5
Unit	II

Id	99
Question	**How deep learning models are built on Keras**
A	by using sequential models
B	by using feed_dict
C	by creating place holders and computational graphs
D	by creating data frames
Answer	
Marks	1.5
Unit	II

Id	100
Question	**Can we use GPU for faster computations in TensorFlow**
A	No, not possible
B	Possible only on cloud
C	Possible only with small datasets
D	Yes, possible
Answer	
Marks	1.5
Unit	II

Id	101
Question	Boltzman learning is a?
A	fast process
B	steady process
C	slow process
D	none of the mentioned
Answer	
Marks	1.5
Unit	I

Id	102
Question	In boltzman learning which algorithm can be used to arrive at equilibrium?
A	hopfield
B	mean field
C	hebb
D	none of the mentioned
Answer	
Marks	1.5
Unit	I

Id	103
Question	What happens when we use mean field approximation with boltzman learning?
A	it slows down
B	it get speeded up
C	nothing happens
D	may speedup or speed down
Answer	
Marks	1.5
Unit	I

Id	104
Question	For practical implementation what type of approximation is used on boltzman law?
A	max field approximation
B	min field approximation
C	hopfield approximation
D	none of the mentioned
Answer	
Marks	1.5
Unit	I

Id	105
Question	What should be the aim of training procedure in boltzman machine of feedback networks?
A	to capture inputs
B	to feedback the captured outputs
C	to capture the behaviour of system
D	none of the mentioned
Answer	
Marks	1.5
Unit	I

Id	106
Question	What consist of boltzman machine?
A	fully connected network with both hidden and visible units
B	asynchronous operation
C	stochastic update
D	all of the mentioned
Answer	
Marks	1.5
Unit	I

Id	107
Question	For what purpose is pattern environment useful?
A	determining structure
B	determining desired outputs
C	determining future inputs
D	none of the mentioned
Answer	
Marks	1.5
Unit	I

Id	108
Question	What is pattern environment?
A	probability of desired patterns
B	probability of given patterns
C	behaviour of system
D	none of the mentioned
Answer	
Marks	1.5
Unit	I

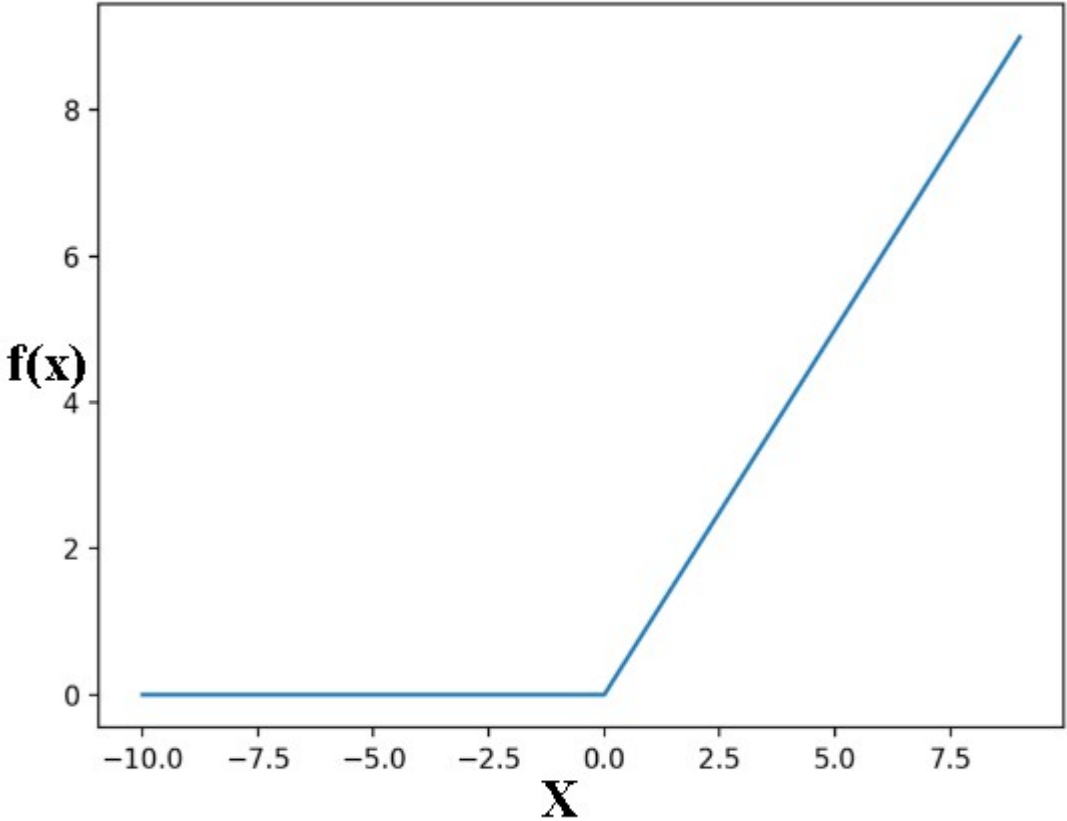| Id	109
Question	Probability of error in recall of stored patterns can be reduced if?
A	patterns are stored appropriately
B	inputs are captured appropriately
C	weights are chosen appropriately
D	none of the mentioned
Answer	
Marks	1.5
Unit	I

Id	110
Question	Why does change in temperature doesn't effect stochastic update?
A	shape landscape depends on the network and its weights which varies accordingly and compensates the effect
B	shape landscape depends on the network and its weights which is fixed
C	shape landscape depends on the network, its weights and the output function which varies accordingly and compensates the effect
D	shape landscape depends on the network, its weights and the output function which is fixed
Answer	
Marks	1.5
Unit	I

Id	111
Question	As temperature increase, what happens to stochastic update?
A	increase in update
B	decrease in update
C	no change
D	none of the mentioned
Answer	
Marks	1.5
Unit	II

Id	112
Question	In the case of an algebraic model for a straight line, if a value for the x variable is specified, then
A	the exact value of the response variable can be computed
B	the computed response to the independent value will always give a minimal residual
C	the computed value of y will always be the best estimate of the mean response
D	none of these alternatives is correct.
Answer	
Marks	1.5
Unit	I

Id	113
Question	Identify the activation function from the given plot

A	Binary Step Function
B	ReLU
C	Signmoid
D	Tanh
Answer	
Marks	1.5
Unit	II
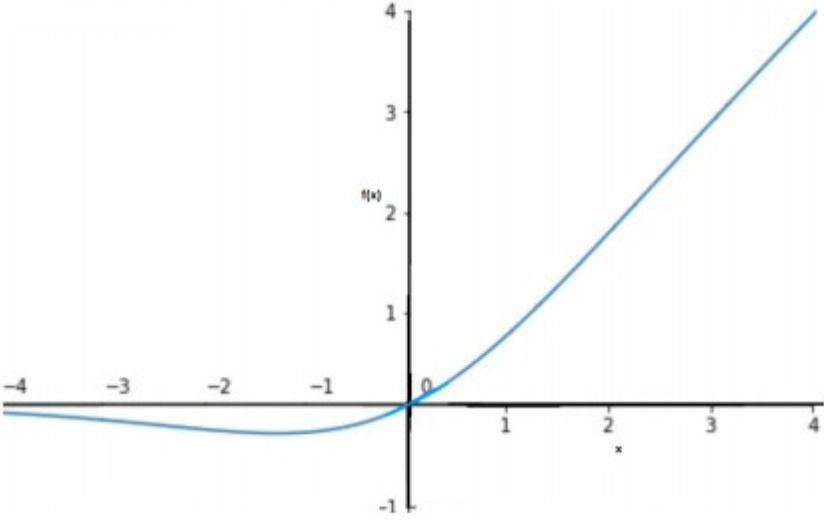
Id	114
Question	Identify the activation function from the given plot

A	Binary Step Function
B	ReLU
C	Signmoid
D	Tanh
Answer	
Marks	1.5
Unit	II

Id	115
Question	Identify the activation function f(x)=max(0,x)
A	Binary Step Function
B	ReLU
C	Signmoid
D	Tanh
Answer	
Marks	1.5
Unit	II
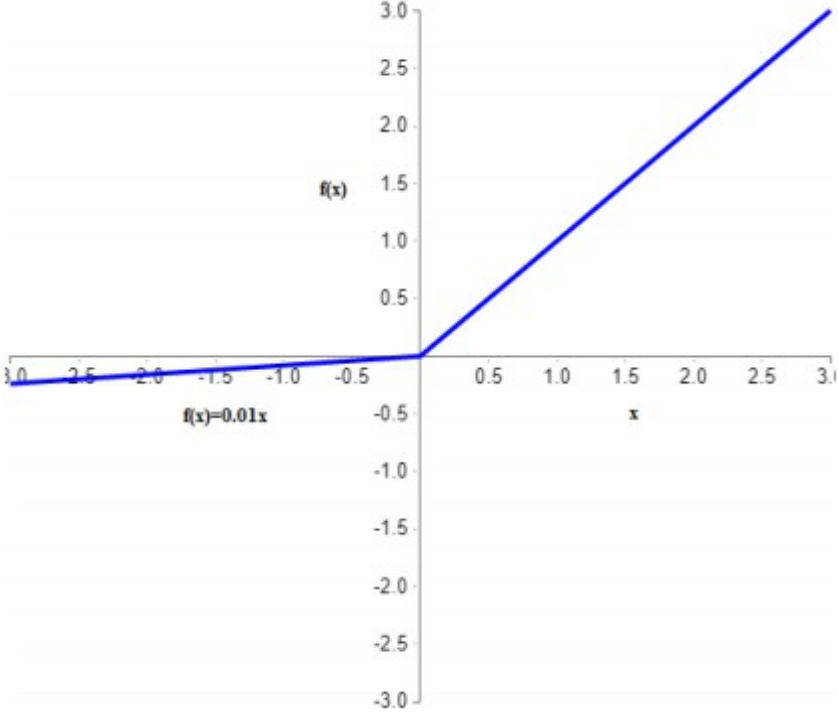
Id	116
Question	Identify the activation function $$f(x) = \frac{1}{1 + e^{-x}}$$
A	Binary Step Function
B	ReLU
C	Signmoid
D	Tanh
Answer	
Marks	1.5
Unit	II

Id	117
Question	Identify the activation function $$f(x) = \begin{cases} 0 & \text{for} \quad x < 0 \\ x & \text{for} \quad x \geq 0 \end{cases}$$
A	Binary Step Function
B	ReLU
C	Signmoid
D	Tanh
Answer	
Marks	1.5
Unit	II
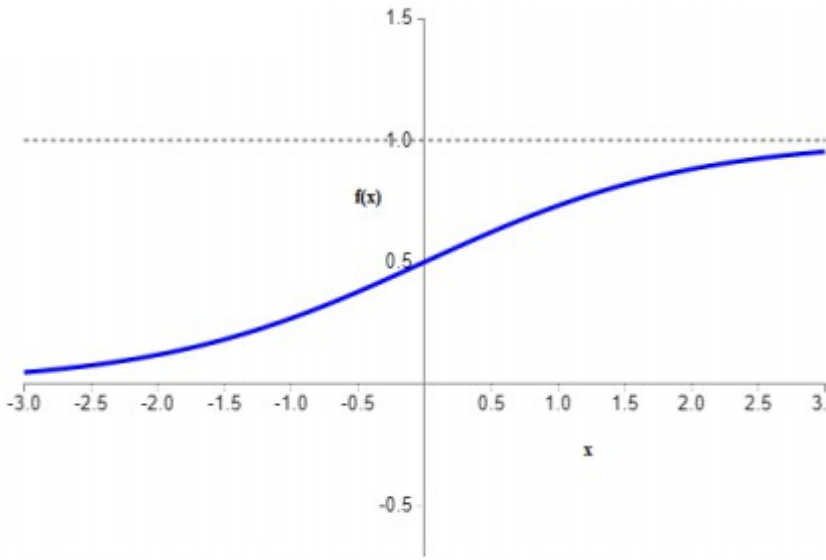
Id	118
Question	Identify the activation function $$f(x) = \begin{cases} x & \text{for} & x \geq 0 \\ \alpha x & \text{for} & x < 0 \end{cases}$$
A	Binary Step Function
B	ReLU
C	Leaky ReLU
D	Tanh
Answer	
Marks	1.5
Unit	II

Id	119
Question	Identify the activation function given in the plot blow.

A	Binary Step Function
B	ReLU
C	Swish
D	Tanh
Answer	
Marks	1.5
Unit	II

Id	120
Question	Identify the activation function given in the plot blow.

f(x)

f(x)=0.01x

x

A	Binary Step Function
B	LeakyReLU
C	Swish
D	Tanh
Answer	
Marks	1.5
Unit	II

Id	121
Question	Identify the activation function $$f(x) = \frac{x}{1 + e^{-x}}$$
A	Binary Step Function
B	ReLU
C	Leaky ReLU
D	Swish
Answer	
Marks	1.5
Unit	II

Id	122
Question	Identify the activation function from the given plot

A	Binary Step Function
B	ReLU
C	Signmoid
D	Tanh
Answer	
Marks	1.5
Unit	II

Id	123
Question	Statement 1: Recurrent Neural Networks are best suited for Text Processing. Statement 2: Recurrent Networks work best for Speech Recognition. Which of the statements given above is true?
A	Statement 1 is true while Statement 2 is false
B	Statement 2 is true while statement 1 is false
C	Both statements are true
D	Both statements are false
Answer	
Marks	1.5
Unit	I

Id	124
Question	Statement 1: Neural Networks Algorithms are inspired from the structure and functioning of the Human Biological Neuron. Statement 2: In a Neural Network, all the edges and nodes have the same Weight and Bias values. Which of the statements given above is true?
A	Statement 1 is true while Statement 2 is false
B	Statement 2 is true while statement 1 is false
C	Both statements are true
D	Both statements are false
Answer	
Marks	1.5
Unit	I

Id	125
Question	Data Collected from Survey results is an example of _____.
A	Data
B	Information
C	Structured Data
D	Unstructured Data
Answer	
Marks	1.5
Unit	I

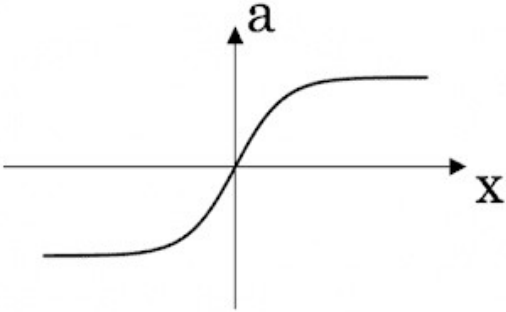| Id	126
Question	Statement 1: A Shallow Neural Network has only one hidden layer between Input and Output layers. Statement 2: All the Visible Layers in a Restricted Boltzmannn Machine are connected to each other. Which of the statements given above is true?
A	Statement 1 is true while Statement 2 is false
B	Statement 2 is true while statement 1 is false
C	Both statements are true
D	Both statements are false
Answer	
Marks	1.5
Unit	I
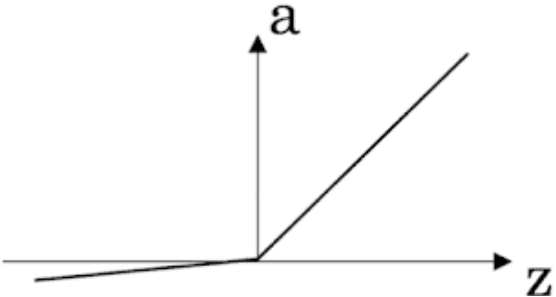
Id	127
Question	Statement 1: All the neurons in a convolution layer have different Weights and Biases. Statement 2: Recurrent Network can input Sequence of Data Points and Produce a Sequence of Output. Which of the statements given above is true?
A	Statement 1 is true while Statement 2 is false
B	Statement 2 is true while statement 1 is false
C	Both statements are true
D	Both statements are false
Answer	
Marks	1.5
Unit	II

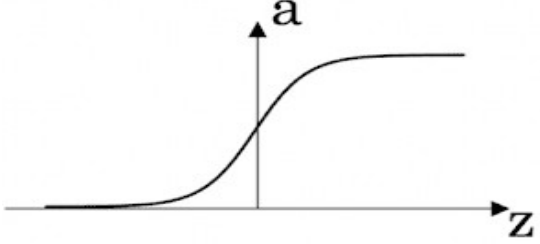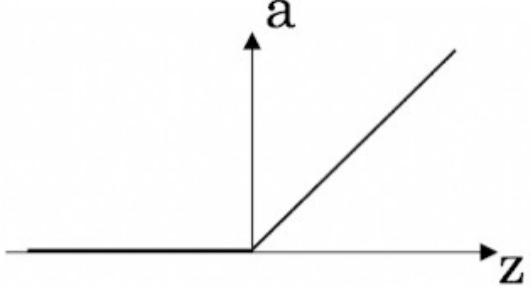| Id	128
Question	Statement 1: A Deep Belief Network is a stack of Restricted Boltzmann Machines. Statement 2: Restricted Boltzmann Machine expects the data to be labeled for Training. Which of the statements given above is true?
A	Statement 1 is true while Statement 2 is false
B	Statement 2 is true while statement 1 is false
C	Both statements are true
D	Both statements are false
Answer	
Marks	1.5
Unit	I

Id	129
Question	Statement 1: Autoencoders cannot be used for Dimensionality Reduction. Statement 2: Restricted Boltzmann Machine expects the data to be labeled for Training. Which of the statements given above is true?
A	Statement 1 is true while Statement 2 is false
B	Statement 2 is true while statement 1 is false
C	Both statements are true
D	Both statements are false
Answer	
Marks	1.5
Unit	II

Id	130
Question	What is hebb's rule of learning ?
A	the system learns from its past mistakes
B	the system recalls previous reference inputs & respective ideal outputs
C	the strength of neural connection get modified accordingly
D	none of the mentioned
Answer	
Marks	1.5
Unit	I

Id	131
Question	Which one of these plots represents a ReLU activation function?
A	
B	
C	
D	
Answer	
Marks	1.5
Unit	II

Id	132
Question	What does a neuron compute?
A	A neuron computes an activation function followed by a linear function (z = Wx + b)
B	A neuron computes the mean of all features before applying the output to an activation function
C	A neuron computes a function g that scales the input x linearly (Wx + b)
D	A neuron computes a linear function (z = Wx + b) followed by an activation function
Answer	
Marks	1.5
Unit	I

Id	133
Question	Statement 1: Boltzman law is practical for implementation. Statement 2: False minima can be reduced by deterministic updates Which of the statements given above is true?
A	Statement 1 is true while Statement 2 is false
B	Statement 2 is true while statement 1 is false
C	Both statements are true
D	Both statements are false
Answer	
Marks	1.5
Unit	I

Id	134
Question	Which of the following layers is NOT a part of CNN?
A	Convolutional Layer
B	Pooling Layer
C	Code Layer
D	Fully connected Layer
Answer	
Marks	1.5
Unit	II

Id	135
Question	In which of the following situations, you should NOT prefer Keras over TensorFlow?
A	When you want to quickly build prototype using neural networks
B	When you want to implement simple neural networks in your initial learning phase
C	When you are doing a critical and intensive research on any field
D	When you want to create simple tutorials for your students and subscribers
Answer	
Marks	1.5
Unit	II

Id	136
Question	Which of the following is an application of RNN?
A	Natural Language Processing
B	Audio and Video analysis
C	Stock market prediction
D	All of the above
Answer	
Marks	1.5
Unit	II

Id	137
Question	Which of the following types of RNN is best suited for machine translation?
A	One to One
B	One to Many
C	Many to One
D	Many to Many
Answer	
Marks	1.5
Unit	II

Id	138
Question	Which of the following is TRUE about TensorFlow?
A	Tensor with rank zero is also called scaler
B	Tensor with rank one is also called matrix
C	Tensor with rank two is also called vector
D	All of the above
Answer	
Marks	1.5
Unit	II

Id	139
Question	Gradient Descent computes derivative of loss function w.r.t:
A	input
B	activation value
C	weight
D	bias
Answer	
Marks	1.5
Unit	I

Id	140
Question	Which of the following is FALSE about step activation function?
A	It is a differentiable activation function
B	Derivative of a step function is always zero
C	It is non-linear in nature
D	None of the above
Answer	
Marks	1.5
Unit	I

Id	141
Question	Which of the following is NOT an application of Autoencoders?
A	Image coloring
B	Image captioning
C	Anomalies and outliers detection
D	Dimensionality reduction
Answer	
Marks	1.5
Unit	II

Id	142
Question	Which of the following is TRUE about Leaking ReLU?
A	Leaky ReLU is the most common method to alleviate a dying ReLU
B	Leaking ReLU adds a slight slope in the negative range to prevent dying ReLU issue
C	Parametric ReLU (PReLU) is a type of leaky ReLU that uses parameterized slope value
D	All of the above
Answer	
Marks	1.5
Unit	II

Id	143
Question	ReLU activation function outputs zero when:
A	Input is zero
B	Input is less than or equal to zero
C	Input is greater than or equal to zero
D	Input is zero or one
Answer	
Marks	1.5
Unit	II

Id	144
Question	Which of the following parameters is not required while compiling a model in Keras?
A	activation function
B	loss
C	metrics
D	optimizer
Answer	
Marks	1.5
Unit	II

Id	145
Question	Which of the following is NOT an application of Restricted Boltzmann Machines RBM ?
A	Dimensionality reduction
B	Image captioning
C	Feature learning and topic modelling
D	Collaborative filtering for recommender systems
Answer	
Marks	1.5
Unit	I

Id	146
Question	CNN is best suited for:
A	Image Classification
B	Natural Language Processing
C	Image Captioning
D	All of the above
Answer	
Marks	1.5
Unit	I

Id	147
Question	Which of the following MUST be initialized in TensorFlow?
A	Variables
B	Placeholders
C	Constants
D	Sessions
Answer	
Marks	1.5
Unit	I

Id	148
Question	Which of the following frameworks can be used as backend in Keras?
A	Swift
B	Gluon
C	TensorFlow
D	MATLAB
Answer	
Marks	1.5
Unit	I

Id	149
Question	Which of the following neural networks uses supervised learning? (A) Multilayer perceptron (B) Self organizing feature map (C) Hopfield network
A	(A) only
B	(B) only
C	(A) and (B) only
D	(A) and (C) only
Answer	
Marks	1.5
Unit	I

Id	150
Question	Which of the following statement is true regrading dropout? 1: Dropout gives a way to approximate by combining many different architectures 2: Dropout demands high learning rates 3: Dropout can help preventing overfitting
A	Both 1 and 2
B	Both 1 and 3
C	Both 2 and 3
D	All 1,2 and 3
Answer	
Marks	1.5
Unit	II