

AI Neural Networks MCQ

This section focuses on "Neural Networks" in Artificial Intelligence. These Multiple Choice Questions (mcq) should be practiced to improve the AI skills required for various interviews (campus interviews, walk-in interviews, company interviews), placements, entrance exams and other competitive examinations.

1. Who was the inventor of the first neurocomputer?

- A. Dr. John Hecht-Nielsen
- B. Dr. Robert Hecht-Nielsen
- C. Dr. Alex Hecht-Nielsen
- D. Dr. Steve Hecht-Nielsen

Ans : B

Explanation: The inventor of the first neurocomputer, Dr. Robert Hecht-Nielsen.

2. How many types of Artificial Neural Networks?

- A. 2
- B. 3
- C. 4
- D. 5

Ans : A

Explanation: There are two Artificial Neural Network topologies : FeedForward and Feedback.

3. In which ANN, loops are allowed?

- A. FeedForward ANN
- B. FeedBack ANN
- C. Both A and B
- D. None of the Above

Ans : B

Explanation: FeedBack ANN loops are allowed. They are used in content addressable memories.

4. What is the full form of BN in Neural Networks?

- A. Bayesian Networks
- B. Belief Networks
- C. Bayes Nets
- D. All of the above

Ans : D

Explanation: The full form BN is Bayesian networks and Bayesian networks are also called Belief Networks or Bayes Nets.

5. What is the name of node which take binary values TRUE (T) and FALSE (F)?

- A. Dual Node
- B. Binary Node
- C. Two-way Node
- D. Ordered Node

Ans : B

Explanation: Boolean nodes : They represent propositions, taking binary values TRUE (T) and FALSE (F).

6. What is an auto-associative network?

- A. a neural network that contains no loops
- B. a neural network that contains feedback
- C. a neural network that has only one loop
- D. a single layer feed-forward neural network with pre-processing

Ans : B

Explanation: An auto-associative network is equivalent to a neural network that contains feedback. The number of feedback paths(loops) does not have to be one.

7. What is Neuro software?

- A. A software used to analyze neurons
- B. It is powerful and easy neural network
- C. Designed to aid experts in real world
- D. It is software used by Neurosurgeon

Ans : B

Explanation: Neuro software is powerful and easy neural network.

8. Neural Networks are complex _____ with many parameters.

- A. Linear Functions
- B. Nonlinear Functions
- C. Discrete Functions
- D. Exponential Functions

Ans : A

Explanation: Neural networks are complex linear functions with many parameters.

9. Which of the following is not the promise of artificial neural network?

- A. It can explain result
- B. It can survive the failure of some nodes
- C. It has inherent parallelism
- D. It can handle noise

Ans : A

Explanation: The artificial Neural Network (ANN) cannot explain result.

10. The output at each node is called_____.

- A. node value
- B. Weight
- C. neurons
- D. axons

Ans : A

Explanation: The output at each node is called its activation or node value.

11. What is full form of ANNs?

- A. Artificial Neural Node
- B. AI Neural Networks
- C. Artificial Neural Networks
- D. Artificial Neural numbers

Ans : C

Explanation: Artificial Neural Networks is the full form of ANNs.

12. In FeedForward ANN, information flow is _____.

- A. unidirectional
- B. bidirectional
- C. multidirectional
- D. All of the above

Ans : A

Explanation: FeedForward ANN the information flow is unidirectional.

13. Which of the following is not an Machine Learning strategies in ANNs?

- A. Unsupervised Learning
- B. Reinforcement Learning
- C. Supreme Learning
- D. Supervised Learning

Ans : C

Explanation: Supreme Learning is not an Machine Learning strategies in ANNs.

14. Which of the following is an Applications of Neural Networks?

- A. Automotive
- B. Aerospace
- C. Electronics
- D. All of the above

Ans : D

Explanation: All above are application of Neural Networks.

15. What is perceptron?

- A. a single layer feed-forward neural network with pre-processing
- B. an auto-associative neural network
- C. a double layer auto-associative neural network
- D. a neural network that contains feedback

Ans : A

Explanation: The perceptron is a single layer feed-forward neural network.

16. 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, 3, 2 and 1 respectively. What will be the output?

- A. 30
- B. 40
- C. 50
- D. 60

Ans : B

Explanation: The output is found by multiplying the weights with their respective inputs, summing the results and multiplying with the transfer function. Therefore: $\text{Output} = 2 * (1*4 + 2*3 + 3*2 + 4*1) = 40$.

17. What is back propagation?

- A. It is another name given to the curvy function in the perceptron
- B. It is the transmission of error back through the network to adjust the inputs
- C. It is the transmission of error back through the network to allow weights to be adjusted so that the network can learn
- D. None of the Above

Ans : C

Explanation: Back propagation is the transmission of error back through the network to allow weights to be adjusted so that the network can learn.

18. The network that involves backward links from output to the input and hidden layers is called _____

- A. Self organizing map
- B. Perceptrons
- C. Recurrent neural network
- D. Multi layered perceptron

Ans : C

Explanation: RNN (Recurrent neural network) topology involves backward links from output to the input and hidden layers.

19. The first artificial neural network was invented in _____.

- A. 1957
- B. 1958
- C. 1959
- D. 1960

Ans : B

Explanation: The first artificial neural network was invented in 1958.

1: ANN is composed of large number of highly interconnected processing elements(neurons) working in unison to solve problems.

- a) True
- b) False

Ans:True

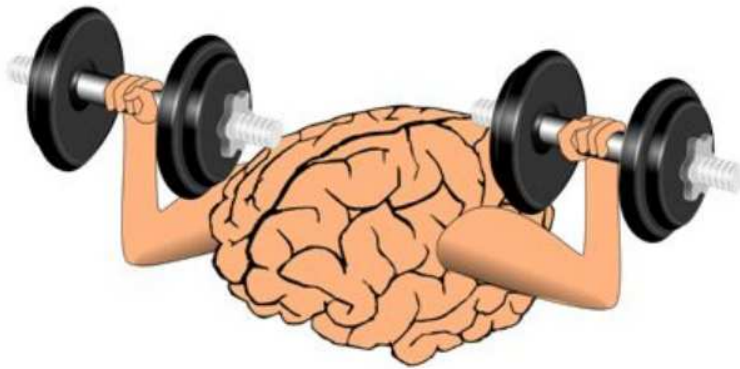
2: Artificial neural network used for

- a) Pattern Recognition**
- b) Classification**
- c) Clustering**
- d) All of these**

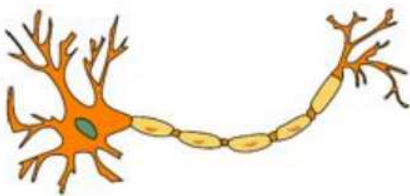
Ans:d

Questions and Answers

Q1. 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. Here's a diagrammatic representation of a real neuron.



Which of the following statement(s) correctly represents a real neuron?

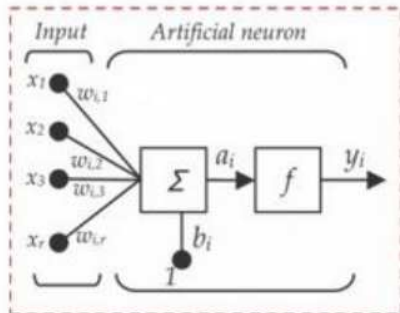
- A. A neuron has a single input and a single output only
- B. A neuron has multiple inputs but a single output only
- C. A neuron has a single input but multiple outputs
- D. A neuron has multiple inputs and multiple outputs

E. All of the above statements are valid

Solution: (E)

A neuron can have a single Input / Output or multiple Inputs / Outputs.

Q2. Below is a mathematical representation of a neuron.



The different components of the neuron are denoted as:

- x_1, x_2, \dots, x_N : These are inputs to the neuron. These can either be the actual observations from input layer or an intermediate value from one of the hidden layers.
- w_1, w_2, \dots, w_N : The Weight of each input.
- b_i : Is termed as Bias units. These are constant values added to the input of the activation function corresponding to each weight. It works similar to an intercept term.
- a : Is termed as the activation of the neuron which can be represented as
- and y : is the output of the neuron

$$a = f\left(\sum_{i=0}^N w_i x_i\right)$$

Considering the above notations, will a line equation ($y = mx + c$) fall into the category of a neuron?

A. Yes

B. No

Solution: (A)

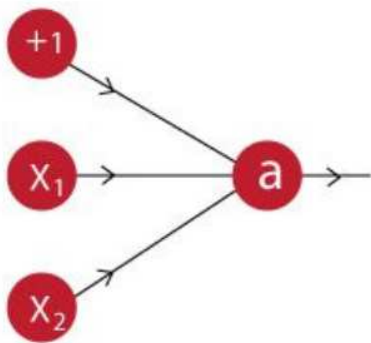
A single neuron with no non-linearity can be considered as a linear regression function.

Q3. Let us assume we implement an AND function to a single neuron. Below is a tabular representation of an AND function:

X1	X2	X1 AND X2
0	0	0
0	1	0
1	0	0
1	1	1

The activation function of our neuron is denoted as:

$$f(x) = \begin{cases} 0, & \text{for } x < 0 \\ 1, & \text{for } x \geq 0 \end{cases}$$



What would be the weights and bias?

(Hint: For which values of w1, w2 and b does our neuron implement an AND function?)

- A. Bias = -1.5, w1 = 1, w2 = 1
- B. Bias = 1.5, w1 = 2, w2 = 2
- C. Bias = 1, w1 = 1.5, w2 = 1.5
- D. None of these

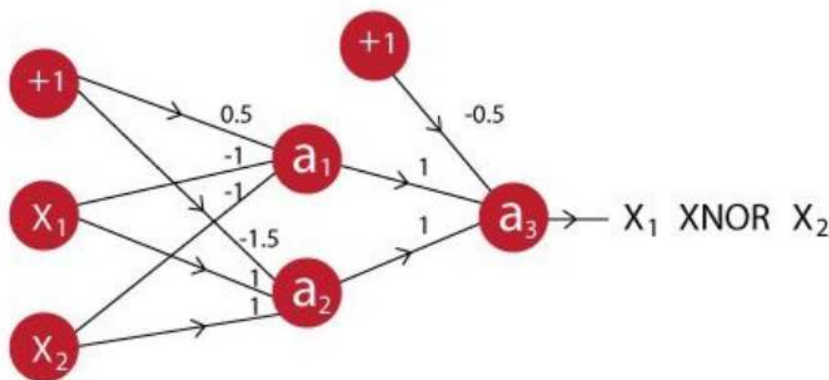
Solution: (A)

A.

1. $f(-1.5*1 + 1*0 + 1*0) = f(-1.5) = 0$
2. $f(-1.5*1 + 1*0 + 1*1) = f(-0.5) = 0$
3. $f(-1.5*1 + 1*1 + 1*0) = f(-0.5) = 0$
4. $f(-1.5*1 + 1*1 + 1*1) = f(0.5) = 1$

Therefore option A is correct

Q4. 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:

$$f(x) = \begin{cases} 0, & \text{for } x < 0 \\ 1, & \text{for } x \geq 0 \end{cases}$$

Suppose X1 is 0 and X2 is 1, what will be the output for the above neural network?

A. 0

B. 1

Solution: (A)

Output of a1: $f(0.5*1 + -1*0 + -1*1) = f(-0.5) = 0$

Output of a2: $f(-1.5*1 + 1*0 + 1*1) = f(-0.5) = 0$

Output of a3: $f(-0.5*1 + 1*0 + 1*0) = f(-0.5) = 0$

So the correct answer is A

Q5. 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 and pray to God they are correct
- 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 to make them better
- D. None of these

Solution: (C)

Option C is the description of gradient descent.

Q6. 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

Solution: (D)

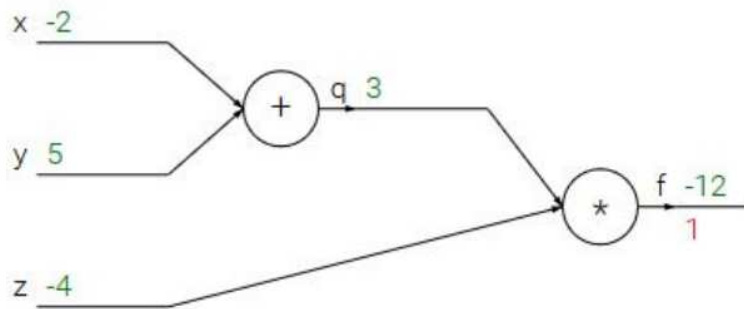
Option D is correct

Q7. 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?

(HINT: To calculate gradient, you must find (df/dx) , (df/dy) and (df/dz))

- A. (-3,4,4)
- B. (4,4,3)
- C. (-4,-4,3)
- D. (3,-4,-4)

Solution: (C)

Option C is correct.

Q8. Now let's revise the previous slides. We have learned that:

- A neural network is a (crude) mathematical representation of a brain, which consists of smaller components called neurons.
- Each neuron has an input, a processing function, and an output.
- These neurons are stacked together to form a network, which can be used to approximate any function.
- To get the best possible neural network, we can use techniques like gradient descent to update our neural network model.

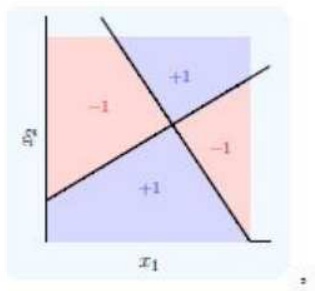
Given above is a description of a neural network. When does a neural network model become a deep learning model?

- A. When you add more hidden layers and increase depth of neural network
- B. When there is higher dimensionality of data
- C. When the problem is an image recognition problem
- D. None of these

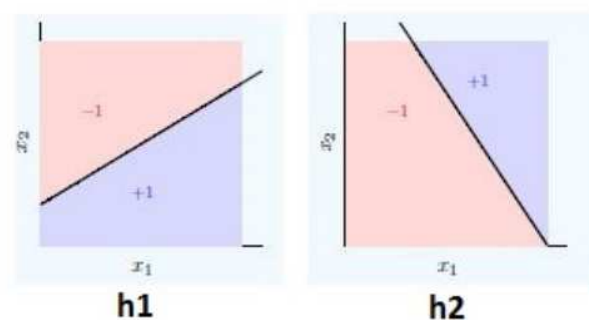
Solution: (A)

More depth means the network is deeper. There is no strict rule of how many layers are necessary to make a model deep, but still if there are more than 2 hidden layers, the model is said to be deep.

Q9. A neural network can be considered as multiple simple equations stacked together. Suppose we want to replicate the function for the below mentioned decision boundary.



Using two simple inputs h1 and h2



What will be the final equation?

- A. (h1 AND NOT h2) OR (NOT h1 AND h2)
- B. (h1 OR NOT h2) AND (NOT h1 OR h2)

C. (h1 AND h2) OR (h1 OR h2)

D. None of these

Solution: (A)

As you can see, combining h1 and h2 in an intelligent way can get you a complex equation easily. Refer Chapter 9 of [this book](#)

Q10. “Convolutional Neural Networks can perform various types of transformation (rotations or scaling) in an input”. Is the statement correct True or False?

A. True

B. False

Solution: (B)

Data Preprocessing steps (viz rotation, scaling) is necessary before you give the data to neural network because neural network cannot do it itself.

Q11. Which of the following techniques perform similar operations as dropout in a neural network?

A. Bagging

B. Boosting

C. Stacking

D. None of these

Solution: (A)

Dropout can be seen as an extreme form of bagging in which each model is trained on a single case and each parameter of the model is very strongly regularized by sharing it with the corresponding parameter in all the other models. Refer [here](#)

Q 12. Which of the following gives non-linearity to a neural network?

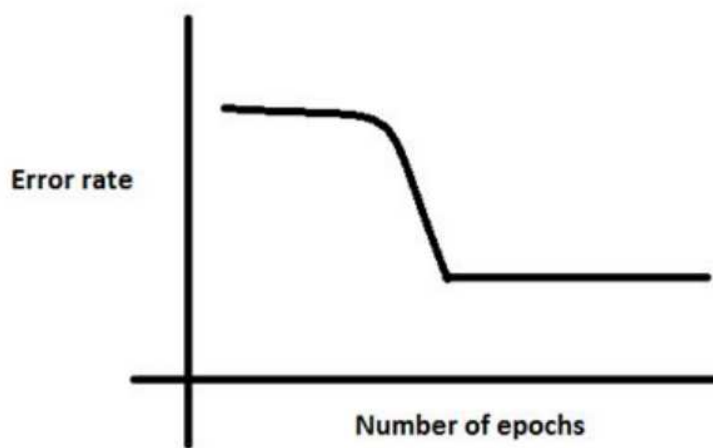
A. Stochastic Gradient Descent

- B. Rectified Linear Unit
- C. Convolution function
- D. None of the above

Solution: (B)

Rectified Linear unit is a non-linear activation function.

Q13. 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 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

Solution: (D)

The problem can occur due to any of the reasons mentioned.

Q14. Which of the following is true about model capacity (where model capacity means the ability of neural network to approximate complex functions) ?

- A. As number of hidden layers increase, model capacity increases
- B. As dropout ratio increases, model capacity increases
- C. As learning rate increases, model capacity increases
- D. None of these

Solution: (A)

Only option A is correct.

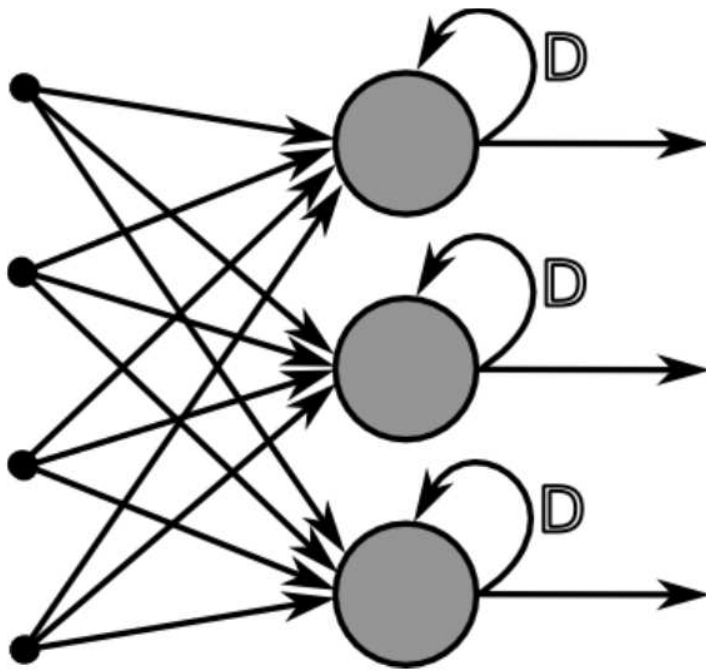
Q15. If you increase the number of hidden layers in a Multi Layer Perceptron, the classification error of test data always decreases. True or False?

- A. True
- B. False

Solution: (B)

This is not always true. Overfitting may cause the error to increase.

Q16. 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

Solution: (A)

Option A is correct.

Q17. What is the sequence of the following tasks in a perceptron?

1. Initialize weights of perceptron randomly
2. Go to the next batch of dataset
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

Solution: (D)

Sequence D is correct.

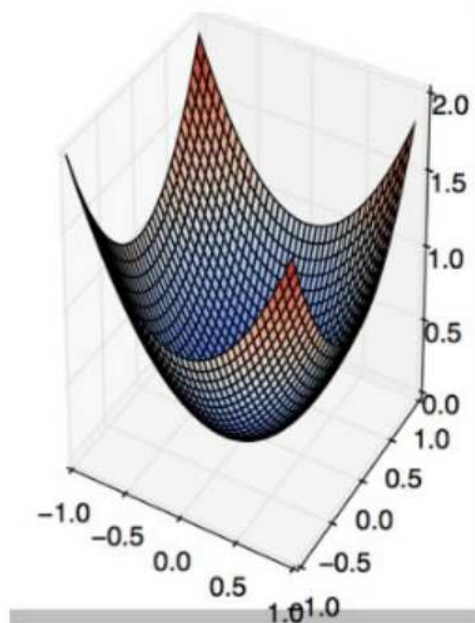
Q18. Suppose that you have to minimize the cost function by changing the parameters. Which of the following technique could be used for this?

- A. Exhaustive Search
- B. Random Search
- C. Bayesian Optimization
- D. Any of these

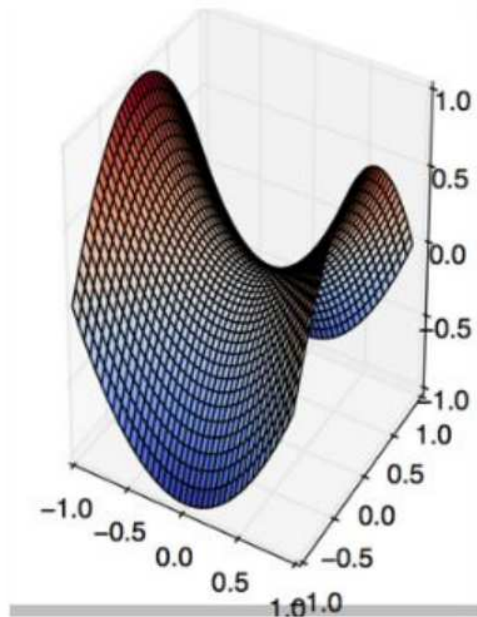
Solution: (D)

Any of the above mentioned technique can be used to change parameters.

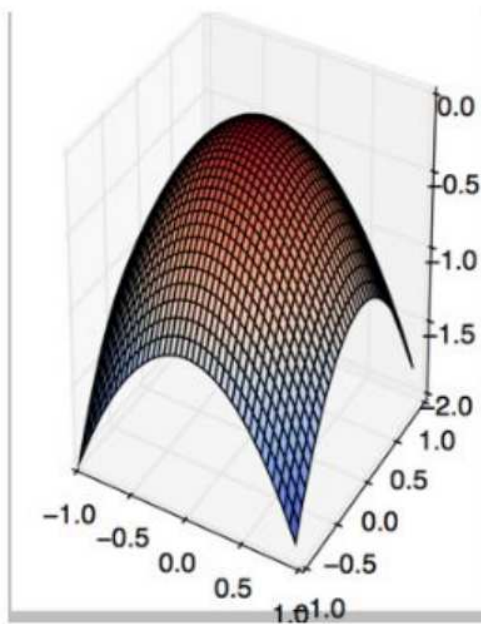
Q19. First Order Gradient descent would not work correctly (i.e. may get stuck) in which of the following graphs?



A.



B.



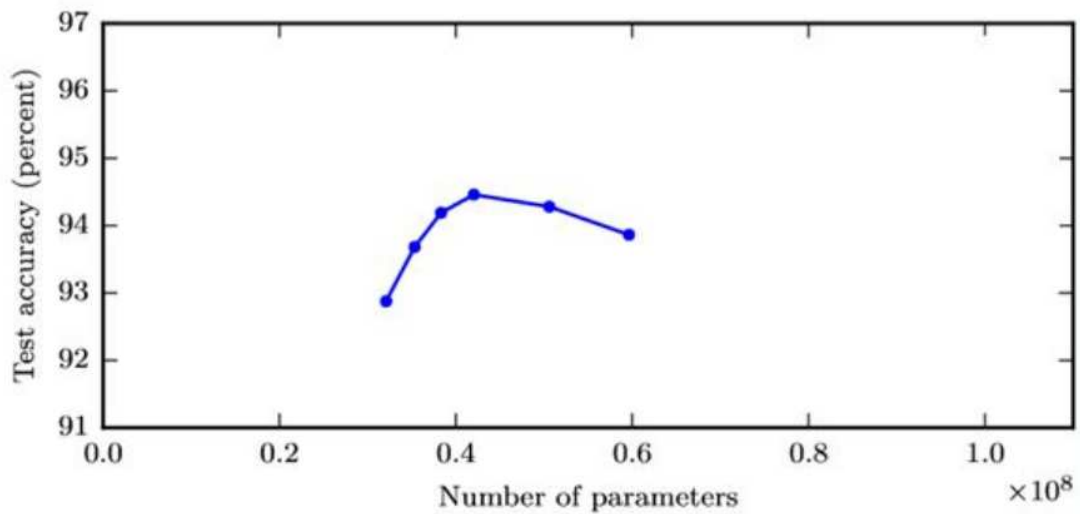
C.

D. None of these

Solution: (B)

This is a classic example of saddle point problem of gradient descent.

Q20. The below graph shows the accuracy of a trained 3-layer convolutional neural network vs the number of parameters (i.e. number of feature kernels).



The trend suggests that as you increase the width of a neural network, the accuracy increases till a certain threshold value, and then starts decreasing.

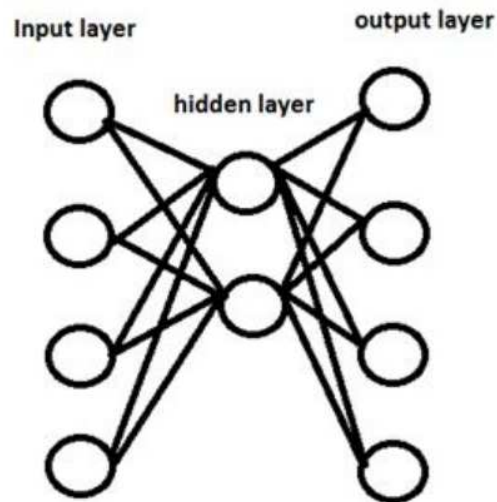
What could be the possible reason for this decrease?

- A. Even if number of kernels increase, only few of them are used for prediction
- B. As the number of kernels increase, the predictive power of neural network decrease
- C. As the number of kernels increase, they start to correlate with each other which in turn helps overfitting
- D. None of these

Solution: (C)

As mentioned in option C, the possible reason could be kernel correlation.

Q21. Suppose we have one hidden layer neural network as shown above. The hidden layer in this network works as a dimensionality reducer. Now instead of using this hidden layer, we replace it with a dimensionality reduction technique such as PCA.



Would the network that uses a dimensionality reduction technique always give same output as network with hidden layer?

A. Yes

B. No

Solution: (B)

Because PCA works on correlated features, whereas hidden layers work on predictive capacity of features.

Q22. Can a neural network model the function ($y=1/x$)?

A. Yes

B. No

Solution: (A)

Option A is true, because activation function can be reciprocal function.

Q23. 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

Solution: (D)

Option D is correct.

Q24. 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

Solution: (A)

To read more about batch normalization, see refer [this video](#)

Q25. 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

Solution: (D)

In reality achieving accurate prediction is a myth. So we should hope to achieve an “achievable result”.

Q26. The number of neurons in the output layer should match the number of classes (Where the number of classes is greater than 2) in a supervised learning task. True or False?

A. True

B. False

Solution: (B)

It depends on output encoding. If it is one-hot encoding, then its true. But you can have two outputs for four classes, and take the binary values as four classes(00,01,10,11).

Q27. 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

Solution: (D)

All of the techniques can be used to deal with overfitting.

Q28. $Y = ax^2 + bx + c$ (polynomial equation of degree 2)

Can this equation be represented by a neural network of single hidden layer with linear threshold?

A. Yes

B. No

Solution: (B)

The answer is no because having a linear threshold restricts your neural network and in simple terms, makes it a consequential linear transformation function.

Q29. 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

Solution: (A)

Option A is correct.

Q30. 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

Solution: (B)

Option B is correct.

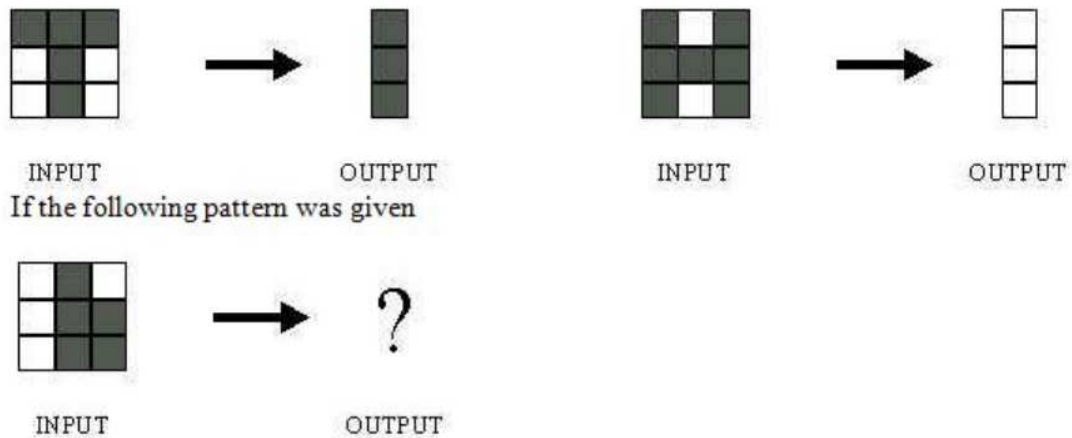
Q31. What if we use a learning rate that's too large?

- A. Network will converge
- B. Network will not converge
- C. Can't Say

Solution: B

Option B is correct because the error rate would become erratic and explode.

Q32. The network shown in Figure 1 is trained to recognize the characters H and T as shown below:



What would be the output of the network?

- A.
- B.
- C.
- D. Could be A or B depending on the weights of neural network

Solution: (D)

Without knowing what are the weights and biases of a neural network, we cannot comment on what output it would give.

Q33. 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?

- A. True
- B. False

Solution: (B)

There would be some neurons which are do not activate for white pixels as input. So the classes wont be equal.

Q34. When pooling layer is added in a convolutional neural network, translation invariance is preserved. True or False?

A. True

B. False

Solution: (A)

Translation invariance is induced when you use pooling.

Q35. Which gradient technique is more advantageous when the data is too big to handle in RAM simultaneously?

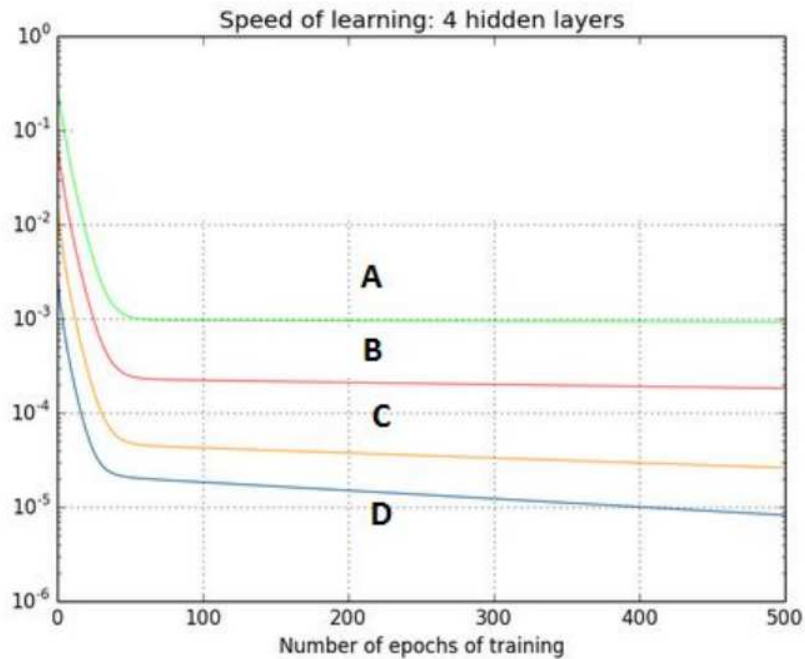
A. Full Batch Gradient Descent

B. Stochastic Gradient Descent

Solution: (B)

Option B is correct.

Q36. The graph represents gradient flow of a four-hidden layer neural network which is trained using sigmoid activation function per epoch of training. The neural network suffers with the vanishing gradient problem.



Which of the following statements is true?

- A. Hidden layer 1 corresponds to D, Hidden layer 2 corresponds to C, Hidden layer 3 corresponds to B and Hidden layer 4 corresponds to A
- B. Hidden layer 1 corresponds to A, Hidden layer 2 corresponds to B, Hidden layer 3 corresponds to C and Hidden layer 4 corresponds to D

Solution: (A)

This is a description of a vanishing gradient problem. As the backprop algorithm goes to starting layers, learning decreases.

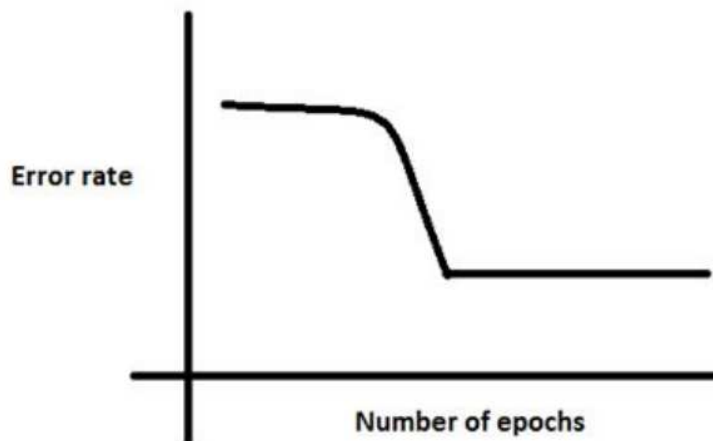
Q37. For a classification task, instead of random weight initializations in a neural network, we set all the weights to zero. Which of the following statements is true?

- 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

Solution: (B)

Option B is correct.

Q38. 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.



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

Solution: (C)

Option C can be used to take a neural network out of local minima in which it is stuck.

Q39. For an image recognition problem (recognizing a cat in a photo), which architecture of neural network would be better suited to solve the problem?

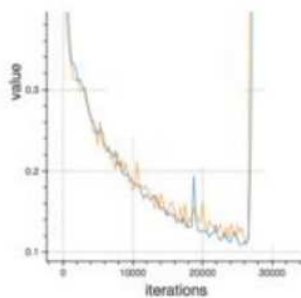
- A. Multi Layer Perceptron
- B. Convolutional Neural Network
- C. Recurrent Neural network

D. Perceptron

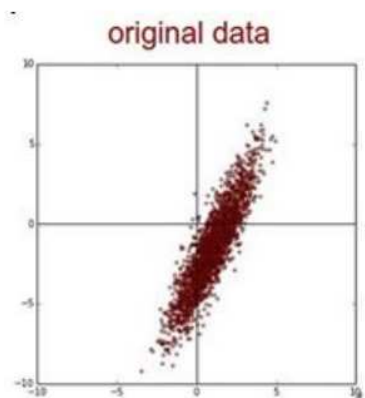
Solution: (B)

Convolutional Neural Network would be better suited for image related problems because of its inherent nature for taking into account changes in nearby locations of an image

Q40. Suppose while training, you encounter this issue. The error suddenly increases after a couple of iterations.



You determine that there must a problem with the data. You plot the data and find the insight that, original data is somewhat skewed and that may be causing the problem.



What will you do to deal with this challenge?

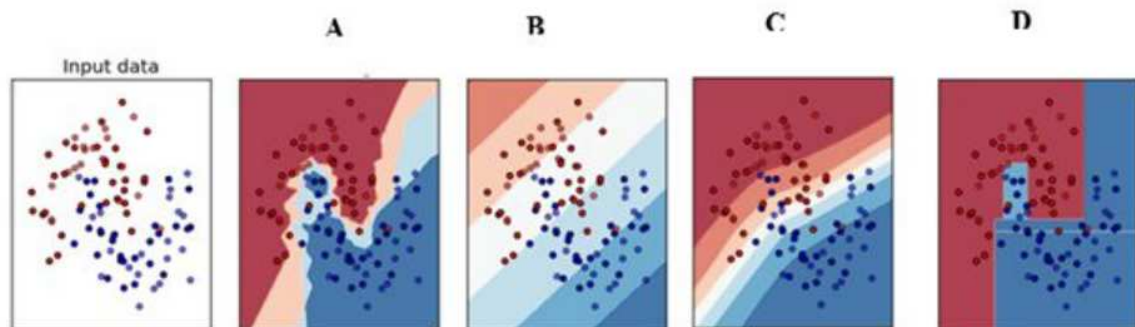
- A. Normalize
- B. Apply PCA and then Normalize
- C. Take Log Transform of the data

D. None of these

Solution: (B)

First you would remove the correlations of the data and then zero center it.

Q41. Which of the following is a decision boundary of Neural Network?



A) B

B) A

C) D

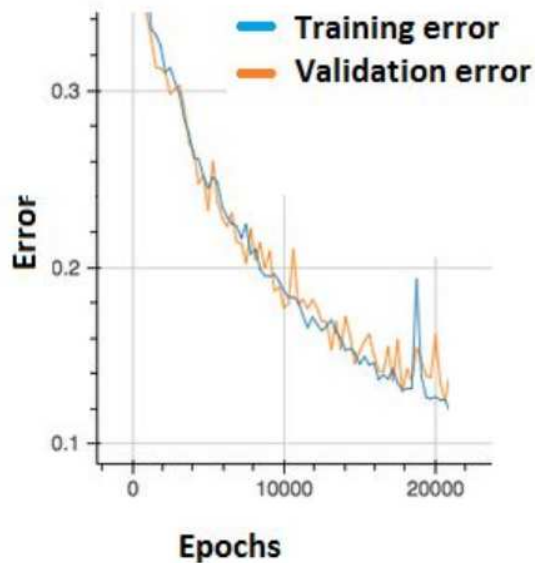
D) C

E) All of these

Solution: (E)

A neural network is said to be a universal function approximator, so it can theoretically represent any decision boundary.

Q42. In the graph below, we observe that the error has many “ups and downs”



Should we be worried?

- A. Yes, because this means there is a problem with the learning rate of neural network.
- B. No, as long as there is a cumulative decrease in both training and validation error, we don't need to worry.

Solution: (B)

Option B is correct. In order to decrease these “ups and downs” try to increase the batch size.

Q43. 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

Solution: (D)

All of the above factors are important to select the depth of neural network

Q44. 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

Solution: (D)

If the dataset is mostly similar, the best method would be to train only the last layer, as previous all layers work as feature extractors.

Q45. Increase in size of a convolutional kernel would necessarily increase the performance of a convolutional network.

- A. True
- B. False

Solution: (B)

Increasing kernel size would not necessarily increase performance. This depends heavily on the dataset.

1. A perceptron is

- A. a single layer feed-forward neural network with pre-processing
- B. an auto-associative neural network
- C. a double layer auto-associative neural network
- D. a neural network that contains feedback

Answer-A

2. An auto-associative network is

- A. a neural network that contains no loops
- B. a neural network that contains feedback
- C. a neural network that has only one loop
- D. a single layer feed-forward neural network with pre-processing

Answer-B

3. Which of the following is true

- (i) On average, neural networks have higher computational rates than conventional computers.
- (ii) Neural networks learn by example.
- (iii) Neural networks mimic the way the human brain works.

- A. All of these
- B. (ii) and (iii) are true
- C. (i), (ii) and (iii) are true
- D. None of these

Answer-A

4. Which of the following is true for neural networks

- (i) The training time depends on the size of the network.
- (ii) Neural networks can be simulated on a conventional computer.
- (iii) Artificial neurons are identical in operation to biological ones.

- A. All of these
- B. (ii) is true
- C. (i) and (ii) are true
- D. None of these

Answer-C

5. What are the advantages of neural networks over conventional computers

- (i) They have the ability to learn by example
- (ii) They are more fault tolerant
- (iii) They are more suited for real time operation due to their high 'computational' rates

- A. (i) and (ii) are true
- B. (i) and (iii) are true
- C. Only (i)
- D. All of these

Answer-D

6. Which of the following is true

Single layer associative neural networks do not have the ability to

- (i) perform pattern recognition
- (ii) find the parity of a picture
- (iii) determine whether two or more shapes in a picture are connected or not

- A. (ii) and (iii) are true
- B. (ii) is true
- C. All of these
- D. None of these

Answer-A

7. Which is true for neural networks

- A. It has set of nodes and connections
- B. Each node computes its weighted input
- C. Node could be in excited state or non-excited state
- D. All of these

Answer-D

8. Neuro software is

- A. A software used to analyze neurons
- B. It is powerful and easy neural network
- C. Designed to aid experts in real world
- D. It is software used by Neuro surgeon

Answer-B

9. Why is the XOR problem exceptionally interesting to neural network researchers

- A. Because it can be expressed in a way that allows you to use a neural network
- B. Because it is complex binary operation that cannot be solved using neural networks
- C. Because it can be solved by a single layer perceptron
- D. Because it is the simplest linearly inseparable problem that exists.

Answer-D

10. What is back propagation

- A. It is another name given to the curvy function in the perceptron
- B. It is the transmission of error back through the network to adjust the inputs
- C. It is the transmission of error back through the network to allow weights to be adjusted so that the network can learn.
- D. None of these

Answer-C

11. Why are linearly separable problems of interest of neural network researchers

- A. Because they are the only class of problem that network can solve successfully
- B. Because they are the only class of problem that Perceptron can solve successfully
- C. Because they are the only mathematical functions that are continue
- D. Because they are the only mathematical functions you can draw

Answer-D

12.Which of the following is not the promise of artificial neural network

- A. It can explain result
- B. It can survive the failure of some nodes
- C. It has inherent parallelism
- D. It can handle noise

Answer-A

13.Neural Networks are complex _____ with many parameters

- A. Linear Functions
- B. Nonlinear Functions
- C. Discrete Functions
- D. Exponential Functions

Answer-A

14.A perceptron adds up all the weighted inputs it receives, and if it exceeds a certain value, it outputs a 1, otherwise it just outputs a 0.

- A. True
- B. False
- C. Sometimes – it can also output intermediate values as well
- D. Can't say

Answer-A

15.The name for the function in question 16 is

- A. Step function
- B. Heaviside function
- C. Logistic function
- D. Perceptron function

Answer-B

16.The network that involves backward links from output to the input and hidden layers is called as_____.

- A. Self organizing maps
- B. Perceptrons

- C. Recurrent neural network
- D. Multi layered perceptron

Answer-C

17.Which of the following is an application of NN (Neural Network)

- A. Sales forecasting
- B. Data validation
- C. Risk management
- D. All of these

Answer-D

18.Different learning method does not include

- A. Memorization
- B. Analogy
- C. Deduction
- D. Introduction

Answer-D

19.Which of the following is the model used for learning

- A. Decision trees
- B. Neural networks
- C. Propositional and FOL rules
- D. All of these

Answer-D

20.Automated vehicle is an example of _____

- A. Supervised learning
- B. Unsupervised learning
- C. Active learning
- D. Reinforcement learning

1. What are the issues on which biological networks proves to be superior than AI networks?

- a) robustness & fault tolerance
- b) flexibility
- c) collective computation
- d) all of the mentioned

Answer:

d

Explanation: AI network should be all of the above mentioned.

2. The fundamental unit of network is
- a) brain
 - b) nucleus
 - c) neuron
 - d) axon

Answer:

c

Explanation: Neuron is the most basic & fundamental unit of a network .

3. What are dendrites?
- a) fibers of nerves
 - b) nuclear projections
 - c) other name for nucleus
 - d) none of the mentioned

Answer:

a

Explanation: Dendrites tree shaped fibers of nerves.

4. What is shape of dendrites like
- a) oval
 - b) round
 - c) tree
 - d) rectangular

Answer:

c

Explanation: Basic biological q&a.

5. Signal transmission at synapse is a?
- a) physical process
 - b) chemical process
 - c) physical & chemical both
 - d) none of the mentioned

Answer:

b

Explanation: Since chemicals are involved at synapse , so its an chemical process.

6. How does the transmission/pulse acknowledged ?
- a) by lowering electric potential of neuron body
 - b) by raising electric potential of neuron body
 - c) both by lowering & raising electric potential
 - d) none of the mentioned

Answer: c

Explanation: There is equal probability of both.

7. When the cell is said to be fired?
- a) if potential of body reaches a steady threshold values
 - b) if there is impulse reaction
 - c) during upbeat of heart
 - d) none of the mentioned

Answer: a

Explanation: Cell is said to be fired if & only if potential of body reaches a certain steady threshold values.

8. Where does the chemical reactions take place in neuron?
- a) dendrites
 - b) axon
 - c) synapses
 - d) nucleus

Answer: c

Explanation: It is a simple biological fact.

9. Function of dendrites is?
- a) receptors
 - b) transmitter
 - c) both receptor & transmitter
 - d) none of the mentioned

Answer: a

Explanation: Dendrites are tree like projections whose function is only to receive impulse.

10. What is purpose of Axon?
- a) receptors
 - b) transmitter
 - c) transmission
 - d) none of the mentioned

Answer: c

Explanation: Axon is the body of neuron & thus cant be at ends of it so cant receive & transmit signals.

SET-2 (Characteristics)

1. What is approx size of neuron body(in micrometer)?

- a) below 5
- b) 5-10
- c) 10-80
- d) above 100

Answer: c

Explanation: Average size of neuron body lies in the above limit.

2. What is the gap at synapses(in nanometer)?

- a) 50
- b) 100
- c) 150
- d) 200

Answer: d

Explanation: It is near to 200nm.

3. What is charge at protoplasm in state of inactivity?

- a) positive
- b) negative
- c) neutral
- d) may be positive or negative

Answer: b

Explanation: It is due to the presence of potassium ion on outer surface in neural fluid.

4. What is the main constituent of neural liquid?

- a) sodium
- b) potassium
- c) Iron
- d) none of the mentioned

Answer: a

Explanation: Potassium is the main constituent of neural liquid & responsible for potential on neuron body.

5. What is average potential of neural liquid in inactive state?

- a) +70mv

- b) +35mv
- c) -35mv
- d) -70mv

Answer: d

Explanation: It is a basic fact, founded out by series of experiments conducted by neural scientist.

6. At what potential does cell membrane loses its impermeability against Na^+ ions?

- a) -50mv
- b) -35mv
- c) -60mv
- d) -65mv

Answer: c

Explanation: Cell membrane loses its impermeability against Na^+ ions at -60mv.

7. What is the effect on a neuron as a whole when its potential gets raised to -60mv?

- a) it gets fired
- b) no effect
- c) it gets compressed
- d) it expands

Answer: a

Explanation: Cell membrane loses its impermeability against Na^+ ions at -60mv.

8. The membrane which allows neural liquid to flow will?

- a) never be imperturbable to neural liquid
- b) regenerate & retain its original capacity
- c) only the certain part gets affected, while rest becomes imperturbable again
- d) none of the mentioned

Answer: b

Explanation: Each cell of human body (internal) has regenerative capacity.

9. How fast is the propagation of discharge signal in cells of human brain?

- a) less than 0.1m/s
- b) 0.5-2m/s
- c) 2-5m/s
- d) 5-10m/s

Answer: b

Explanation: The process is very fast but comparable to the length of neuron.

10. What is the function of neurotransmitter ?

- a) they transmit data directly at synapse to other neuron
- b) they modify conductance of post synaptic membrane for certain ions
- c) cause polarisation or depolarisation
- d) both polarisation & modify conductance of membrane

Answer: d

Explanation: Excitatory & inhibitory activities are result of these two process.

SET-3 (Characteristics)

1. The cell body of neuron can be analogous to what mathematical operation?

- a) summing
- b) differentiator
- c) integrator
- d) none of the mentioned

Answer: a

Explanation: Because adding of potential(due to neural fluid) at different parts of neuron is the reason of its firing.

2. What is the critical threshold voltage value at which neuron get fired?

- a) 30mv
- b) 20mv
- c) 25mv
- d) 10mv

Answer: d

Explanation: This critical is founded by series of experiments conducted by neural scientist.

3. Does there is any effect on particular neuron which got repeatedly fired ?

- a) yes
- b) no

Answer: a

Explanation: The strength of neuron to fire in future increases.

4. What is name of above mechanism?

- a) hebb rule learning

- b) error correction learning
- c) memory based learning
- d) none of the mentioned

Answer: a

Explanation: It follows from basic definition of hebb rule learning.

5. 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: c

Explanation: The strength of neuron to fire in future increases, if it is fired repeatedly.

6. Are all neuron in brain are of same type?
- a) yes
 - b) no

Answer: b

Explanation: Follows from the fact no two body cells are exactly similar in human body, even if they belong to same class.

7. What is estimate number of neurons in human cortex?
- a) 108
 - b) 105
 - c) 1011
 - d) 1020

Answer: c

Explanation: It is a fact !

8. what is estimated density of neuron per mm^2 of cortex?
- a) $15 \times (10^2)$
 - b) $15 \times (10^4)$
 - c) $15 \times (10^3)$
 - d) $5 \times (10^4)$

Answer: b

Explanation: It is a biological fact !

9. Why can't we design a perfect neural network?
- a) full operation is still not known of biological neurons
 - b) number of neuron is itself not precisely known
 - c) number of interconnection is very large & is very complex
 - d) all of the mentioned

Answer: d

Explanation: These are all fundamental reasons, why can't we design a perfect neural network !

10. How many synaptic connection are there in human brain?

- a) 1010
- b) 1015
- c) 1020
- d) 105

Answer: b

Explanation: You can estimate this value from number of neurons in human cortex & their density.

SET-4(History)

1. Operations in the neural networks can perform what kind of operations?

- a) serial
- b) parallel
- c) serial or parallel
- d) none of the mentioned

Answer: c

Explanation: General characteristics of neural networks.

2. Does the argument information in brain is adaptable, whereas in the computer it is replaceable is valid?

- a) yes
- b) no

Answer: a

Explanation: Its a fact & related to basic knowledge of neural networks !

3. Does there exist central control for processing information in brain as in computer?

- a) yes
- b) no

Answer: b

Explanation: In human brain information is locally processed & analysed.

4. Which action is faster pattern classification or adjustment of weights in neural nets?

- a) pattern classification
- b) adjustment of weights
- c) equal
- d) either of them can be fast, depending on conditions

Answer: a

Explanation: Memory is addressable, so thus pattern can be easily classified.

5. What is the feature of ANNs due to which they can deal with noisy, fuzzy, inconsistent data?

- a) associative nature of networks
- b) distributive nature of networks
- c) both associative & distributive
- d) none of the mentioned

Answer: c

Explanation: General characteristics of ANNs.

6. What was the name of the first model which can perform weighted sum of inputs?

- a) McCulloch-pitts neuron model
- b) Marvin Minsky neuron model
- c) Hopfield model of neuron
- d) none of the mentioned

Answer: a

Explanation: McCulloch-pitts neuron model can perform weighted sum of inputs followed by threshold logic operation.

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

- a) McCulloch-pitts
- b) Marvin Minsky
- c) Hopfield
- d) none of the mentioned

Answer: b

Explanation: In 1954 Marvin Minsky developed the first learning machine in which connection strengths could be adapted automatically & efficiently.
advertisement

8. Who proposed the first perceptron model in 1958?

- a) McCulloch-pitts
- b) Marvin Minsky
- c) Hopfield
- d) Rosenblatt

Answer: d

Explanation: Rosenblatt proposed the first perceptron model in 1958 .

9. John Hopfield was credited for what important aspect of neuron?

- a) learning algorithms
- b) adaptive signal processing
- c) energy analysis
- d) none of the mentioned

Answer: c

Explanation: It was of major contribution of his works in 1982.

10. What is the contribution of Ackley, Hinton in neural?

- a) perceptron
- b) boltzman machine
- c) learning algorithms
- d) none of the mentioned

Answer: b

Explanation: Ackley, Hinton built the boltzman machine.

SET-5(Terminology)

1. What is ART in neural networks?

- a) automatic resonance theory
- b) artificial resonance theory
- c) adaptive resonance theory
- d) none of the mentioned

Answer: c

Explanation: It is full form of ART & is basic q&a.

2. What is an activation value?

- a) weighted sum of inputs
- b) threshold value
- c) main input to neuron
- d) none of the mentioned

Answer: a

Explanation: It is definition of activation value & is basic q&a.

3. Positive sign of weight indicates?

- a) excitatory input
- b) inhibitory input
- c) can be either excitatory or inhibitory as such
- d) none of the mentioned

Answer: a

Explanation: Sign convention of neuron.

4. Negative sign of weight indicates?

- a) excitatory input
- b) inhibitory input
- c) excitatory output
- d) inhibitory output

Answer: b

Explanation: Sign convention of neuron.

5. The amount of output of one unit received by another unit depends on what?

- a) output unit

- b) input unit
- c) activation value
- d) weight

Answer: d

Explanation: Activation is sum of weighted sum of inputs, which gives desired output..hence output depends on weights.

6. The process of adjusting the weight is known as?

- a) activation
- b) synchronisation
- c) learning
- d) none of the mentioned

Answer: c

Explanation: Basic definition of learning in neural nets .

7. The procedure to incrementally update each of weights in neural is referred to as?

- a) synchronisation
- b) learning law
- c) learning algorithm
- d) both learning algorithm & law

Answer: d

Explanation: Basic definition of learning law in neural.

8. In what ways can output be determined from activation value?

- a) deterministically
- b) stochastically
- c) both deterministically & stochastically
- d) none of the mentioned

Answer: c

Explanation: This is the most important trait of input processing & output determination in neural networks.

9. How can output be updated in neural network?

- a) synchronously
- b) asynchronously
- c) both synchronously & asynchronously
- d) none of the mentioned

Answer: c

Explanation: Output can be updated at same time or at different time in the networks.

10. What is asynchronous update in neural networks?

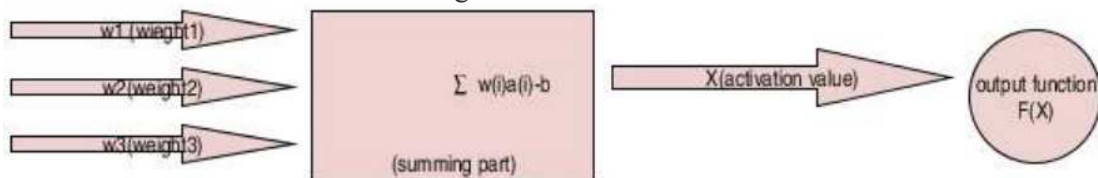
- a) output units are updated sequentially
- b) output units are updated in parallel fashion
- c) can be either sequentially or in parallel fashion
- d) none of the mentioned

Answer: a

Explanation: Output are updated at different time in the networks.

SET-6(Model 1)

1. What is the name of the model in figure below?



neural-networks-questions-answers-models-1-q1

- a) Rosenblatt perceptron model
- b) McCulloch-pitts model
- c) Widrow's Adaline model
- d) None of the mentioned

Answer: b

Explanation: It is a general block diagram of McCulloch-pitts model of neuron.

2. What is nature of function $F(x)$ in the figure?

- a) linear
- b) non-linear
- c) can be either linear or non-linear
- d) none of the mentioned

Answer: b

Explanation: In this function, the independent variable is an exponent in the equation hence non-linear.

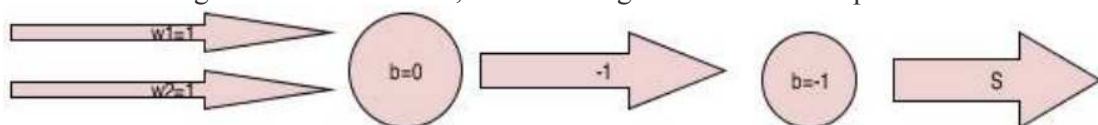
3. What does the character 'b' represents in the above diagram?

- a) bias
- b) any constant value
- c) a variable value
- d) none of the mentioned

Answer: a

Explanation: More appropriate choice since bias is a constant fixed value for any circuit model.

4. If 'b' in the figure below is the bias, then what logic circuit does it represents?



neural-networks-questions-answers-models-1-q4

- a) or gate
- b) and gate
- c) nor gate
- d) nand gate

Answer: c

Explanation: Form the truth table of above figure by taking inputs as 0 or 1.

5. When both inputs are 1, what will be the output of the above figure?

- a) 0
- b) 1
- c) either 0 or 1
- d) z

Answer: a

Explanation: Check the truth table of nor gate.

6. When both inputs are different, what will be the output of the above figure?

- a) 0
- b) 1
- c) either 0 or 1
- d) z

Answer: a

Explanation: Check the truth table of nor gate.

7. Which of the following model has ability to learn?

- a) pitts model
- b) rosenblatt perceptron model
- c) both rosenblatt and pitts model
- d) neither rosenblatt nor pitts

Answer: b

Explanation: Weights are fixed in pitts model but adjustable in rosenblatt.

8. When both inputs are 1, what will be the output of the pitts model nand gate ?

- a) 0
- b) 1
- c) either 0 or 1
- d) z

Answer: a

Explanation: Check the truth table of simply a nand gate.

9. When both inputs are different, what will be the logical output of the figure of question 4?

- a) 0
- b) 1
- c) either 0 or 1

d) z

Answer: a

Explanation: Check the truth table of nor gate.

10. Does McCulloch-pitts model have ability of learning?

a) yes

b) no

Answer: b

Explanation: Weights are fixed.

SET-7(Model 2)

1. Who invented perceptron neural networks?

a) McCulloch-pitts

b) Widrow

c) Minsky & papert

d) Rosenblatt

Answer: d

Explanation: The perceptron is one of the earliest neural networks. Invented at the Cornell Aeronautical Laboratory in 1957 by Frank Rosenblatt, the Perceptron was an attempt to understand human memory, learning, and cognitive processes.

2. What was the 2nd stage in perceptron model called?

a) sensory units

b) summing unit

c) association unit

d) output unit

Answer: c

Explanation: This was the very speciality of the perceptron model, that is performs association mapping on outputs of the sensory units.

3. What was the main deviation in perceptron model from that of MP model?

a) more inputs can be incorporated

b) learning enabled

c) all of the mentioned

d) none of the mentioned

Answer: b

Explanation: The weights in perceptron model are adjustable.

4. What is delta (error) in perceptron model of neuron?

a) error due to environmental condition

b) difference between desired & target output

- c) can be both due to difference in target output or environmental condition
- d) none of the mentioned

Answer: a

Explanation: All other parameters are assumed to be null while calculating the error in perceptron model & only difference between desired & target output is taken into account.

5. If $a(i)$ is the input, \hat{e} is the error, n is the learning parameter, then how can weight change in a perceptron model be represented?

- a) $n a(i)$
- b) $n \hat{e}$
- c) $\hat{e} a(i)$
- d) none of the mentioned

Answer: d

Explanation: The correct answer is $n \hat{e} a(i)$.

6. What is adaline in neural networks?

- a) adaptive linear element
- b) automatic linear element
- c) adaptive line element
- d) none of the mentioned

Answer: a

Explanation: adaptive linear element is the full form of adaline neural model.

7. who invented the adaline neural model?

- a) Rosenblatt
- b) Hopfield
- c) Werbos
- d) Widrow

Answer: d

Explanation: Widrow invented the adaline neural model.

8. What was the main point of difference between the adaline & perceptron model?

- a) weights are compared with output
- b) sensory units result is compared with output
- c) analog activation value is compared with output
- d) all of the mentioned

Answer: c

Explanation: Analog activation value comparison with output, instead of desired output as in perceptron model was the main point of difference between the adaline & perceptron model.

9. In adaline model what is the relation between output & activation value(x)?

- a) linear
- b) nonlinear
- c) can be either linear or non-linear
- d) none of the mentioned

Answer: a

Explanation: $s, \text{output} = f(x) = x$. Hence its a linear model.

10. what is the another name of weight update rule in adaline model based on its functionality?

- a) LMS error learning law
- b) gradient descent algorithm
- c) both LMS error & gradient descent learning law
- d) none of the mentioned

Answer: c

Explanation: weight update rule minimizes the mean squared error(delta square), averaged over all inputs & this laws is derived using negative gradient of error surface weight space, hence option a & b.

SET-8(Topology)

1. In neural how can connectons between different layers be achieved?

- a) interlayer
- b) intralayer
- c) both interlayer and intralayer
- d) either interlayer or intralayer

Answer: c

Explanation: Connections between layers can be made to one unit to another and within the units of a layer.

2. Connections across the layers in standard topologies & among the units within a layer can be organised?

- a) in feedforward manner
- b) in feedback manner
- c) both feedforward & feedback
- d) either feedforward & feedback

Answer: d

Explanation: Connections across the layers in standard topologies can be in feedforward manner or in feedback manner but not both.

3. What is an instar topology?

- a) when input is given to layer F1, the the jth(say) unit of other layer F2 will be activated to maximum extent
- b) when weight vector for connections from jth unit (say) in F2 approaches the activity pattern in F1(comprises of input vector)
- c) can be either way

d) none of the mentioned

Answer: a

Explanation: Restatement of basic definition of instar.

4. What is an outstar topology?

- a) when input is given to layer F1, the the jth(say) unit of other layer F2 will be activated to maximum extent
- b) when weight vector for connections from jth unit (say) in F2 approaches the activity pattern in F1 (comprises of input vector)
- c) can be either way
- d) none of the mentioned

Answer: b

Explanation: Restatement of basic definition of outstar.

5. The operation of instar can be viewed as?

- a) content addressing the memory
- b) memory addressing the content
- c) either content addressing or memory addressing
- d) both content & memory addressing

Answer: a

Explanation: Because in instar, when input is given to layer F1, the the jth(say) unit of other layer F2 will be activated to maximum extent.

6. The operation of outstar can be viewed as?

- a) content addressing the memory
- b) memory addressing the content
- c) either content addressing or memory addressing
- d) both content & memory addressing

Answer: b

Explanation: Because in outstar, when weight vector for connections from jth unit (say) in F2 approaches the activity pattern in F1 (comprises of input vector).

7. If two layers coincide & weights are symmetric($w_{ij}=w_{ji}$), then what is that structure called?

- a) instar
- b) outstar
- c) autoassociative memory
- d) heteroassociative memory

Answer: c

Explanation: In autoassociative memory each unit is connected to every other unit & to itself.

8. Heteroassociative memory can be an example of which type of network?

- a) group of instars
- b) group of outstar
- c) either group of instars or outstars

d) both group of instars or outstars

Answer: c

Explanation: Depending upon the flow, the memory can be of either of the type.

9. What is STM in neural network?

- a) short topology memory
- b) stimulated topology memory
- c) short term memory
- d) none of the mentioned

Answer: c

Explanation: Full form of STM.

10. What does STM corresponds to?

- a) activation state of network
- b) encoded pattern information pattern in synaptic weights
- c) either way
- d) both way

Answer: a

Explanation: Short-term memory (STM) refers to the capacity-limited retention of information over a brief period of time, hence the option.
advertisement

11. What LTM corresponds to?

- a) activation state of network
- b) encoded pattern information pattern in synaptic weights
- c) either way
- d) both way

Answer: b

Explanation: Long-term memory (LTM-the encoding and retention of an effectively unlimited amount of information for a much longer period of time) & hence the option.

SET-9(Learning 1)

1. On what parameters can change in weight vector depend?

- a) learning parameters
- b) input vector
- c) learning signal
- d) all of the mentioned

Answer: d

Explanation: Change in weight vector corresponding to jth input at time (t+1) depends on all of these parameters.

2. If the change in weight vector is represented by Δw_{ij} , what does it mean?

- a) describes the change in weight vector for ith processing unit, taking input vector jth into account
- b) describes the change in weight vector for jth processing unit, taking input vector ith into account
- c) describes the change in weight vector for jth & ith processing unit.
- d) none of the mentioned

Answer: a

Explanation: $\Delta w_{ij} = \mu f(w_i a)_j$, where a is the input vector.

3. What is learning signal in this equation $\Delta w_{ij} = \mu f(w_i a)_j$?

- a) μ
- b) $w_i a$
- c) a_j
- d) $f(w_i a)$

Answer: d

Explanation: This is the non linear representation of output of the network.

4. State whether Hebb's law is supervised learning or of unsupervised type?

- a) supervised
- b) unsupervised
- c) either supervised or unsupervised
- d) can be both supervised & unsupervised

Answer: b

Explanation: No desired output is required for its implementation.

5. Hebb's law can be represented by equation?

- a) $\Delta w_{ij} = \mu f(w_i a)_j$
- b) $\Delta w_{ij} = \mu (s_i) a_j$, where (s_i) is output signal of ith input
- c) both way
- d) none of the mentioned

Answer: c

Explanation: $(s_i) = f(w_i a)$, in Hebb's law.

6. State which of the following statements hold for perceptron learning law?

- a) it is supervised type of learning law
- b) it requires desired output for each input
- c) $\Delta w_{ij} = \mu (b_i - s_i) a_j$
- d) all of the mentioned

Answer: d

Explanation: all statements follow from $\Delta w_{ij} = \mu (b_i - s_i) a_j$, where b_i is the target output & hence supervised learning.

7. Delta learning is of unsupervised type?

- a) yes
- b) no

Answer: b

Explanation: Change in weight is based on the error between the desired & the actual output values for a given input.

8. Widrow & Hoff learning law is special case of?

- a) Hebb learning law
- b) perceptron learning law
- c) delta learning law
- d) none of the mentioned

Answer: c

Explanation: Output function in this law is assumed to be linear, all other things same.

9. What's the other name of Widrow & Hoff learning law?

- a) Hebb
- b) LMS
- c) MMS
- d) None of the mentioned

Answer: b

Explanation: LMS, least mean square. Change in weight is made proportional to negative gradient of error & due to linearity of output function.

10. Which of the following equations represent perceptron learning law?

- a) $\Delta w_{ij} = \mu(s_i) a_j$
- b) $\Delta w_{ij} = \mu(b_i - s_i) a_j$
- c) $\Delta w_{ij} = \mu(b_i - s_i) a_j \dot{A}(x_i)$, where $\dot{A}(x_i)$ is derivative of x_i
- d) $\Delta w_{ij} = \mu(b_i - (w_i a)) a_j$

Answer: b

Explanation: Perceptron learning law is supervised, nonlinear type of learning.

SET-10 (Learning 2)

1. Correlation learning law is special case of?

- a) Hebb learning law
- b) Perceptron learning law
- c) Delta learning law
- d) LMS learning law

Answer: a

Explanation: Since in Hebb is replaced by b_i (target output) in correlation.

2. Correlation learning law is what type of learning?

- a) supervised
- b) unsupervised
- c) either supervised or unsupervised
- d) both supervised or unsupervised

Answer: a

Explanation: Supervised, since depends on target output.

3. Correlation learning law can be represented by equation?

- a) $\Delta w_{ij} = \mu(s_i) a_j$
- b) $\Delta w_{ij} = \mu(b_i - s_i) a_j$
- c) $\Delta w_{ij} = \mu(b_i - s_i) a_j \dot{A}(x_i)$, where $\dot{A}(x_i)$ is derivative of x_i
- d) $\Delta w_{ij} = \mu b_i a_j$

Answer: d

Explanation: Correlation learning law depends on target output(b_i).

4. The other name for instar learning law?

- a) loser take it all
- b) winner take it all
- c) winner give it all
- d) loser give it all

Answer: b

Explanation: The unit which gives maximum output, weight is adjusted for that unit.

5. The instar learning law can be represented by equation?

- a) $\Delta w_{ij} = \mu(s_i) a_j$
- b) $\Delta w_{ij} = \mu(b_i - s_i) a_j$
- c) $\Delta w_{ij} = \mu(b_i - s_i) a_j \dot{A}(x_i)$, where $\dot{A}(x_i)$ is derivative of x_i
- d) $\Delta w_k = \mu (a - w_k)$, unit k with maximum output is identified

Answer: d

Explanation: Follows from basic definition of instar learning law.

6. Is instar a case of supervised learning?

- a) yes
- b) no

Answer: b

Explanation: Since weight adjustment don't depend on target output, it is unsupervised learning.

7. The instar learning law can be represented by equation?

- a) $\Delta w_{jk} = \mu(b_j - w_{jk})$, where the k th unit is the only active in the input layer
- b) $\Delta w_{ij} = \mu(b_i - s_i) a_j$
- c) $\Delta w_{ij} = \mu(b_i - s_i) a_j \dot{A}(x_i)$, where $\dot{A}(x_i)$ is derivative of x_i
- d) $\Delta w_{ij} = \mu(s_i) a_j$

Answer: a

Explanation: Follows from basic definition of outstar learning law.

8. Is outstar a case of supervised learning?

- a) yes
- b) no

Answer: a

Explanation: Since weight adjustment depend on target output, it is supervised learning.

9. Which of the following learning laws belongs to same category of learning?

- a) hebbian, perceptron
- b) perceptron, delta
- c) hebbian, widrow-hoff
- d) instar, outstar

Answer: b

Explanation: They both belongs to supervised type learning.

10. In hebbian learning intial weights are set?

- a) random
- b) near to zero
- c) near to target value
- d) near to target value

Answer: b

Explanation: Hebb law lead to sum of correlations between input & output, inorder to achieve this, the starting initial weight values must be small.

1. What strategies can help reduce overfitting in decision trees?

- (i) Enforce a maximum depth for the tree
 - (ii) Enforce a minimum number of samples in leaf nodes
 - (iii) Pruning
 - (iv) Make sure each leaf node is one pure class
- (A) All
(B) (i), (ii) and (iii)
(C) (i), (iii), (iv)
(D) None

Answer

Correct option is B

2. Which of the following is a widely used and effective machine learning algorithm based on the idea of bagging?

- (A) Decision Tree
- (B) Random Forest
- (C) Regression
- (D) Classification

Answer

Correct option is B

3. To find the minimum or the maximum of a function, we set the gradient to zero because which of the following

- (A) Depends on the type of problem
- (B) The value of the gradient at extrema of a function is always zero
- (C) Both (A) and (B)
- (D) None of these

Answer

Correct option is B

4. Which of the following is a disadvantage of decision trees?

- (A) Decision trees are prone to be overfit
- (B) Decision trees are robust to outliers
- (C) Factor analysis
- (D) None of the above

Answer

Correct option is A

5. What is perceptron?

- (A) A single layer feed-forward neural network with pre-processing
- (B) A neural network that contains feedback
- (C) A double layer auto-associative neural network
- (D) An auto-associative neural network

Answer

Correct option is A

6. Which of the following is true for neural networks?

- (i) The training time depends on the size of the network.
 - (ii) Neural networks can be simulated on a conventional computer.
 - (iii) Artificial neurons are identical in operation to biological ones.
- (A) All
 - (B) Only (ii)
 - (C) (i) and (ii)
 - (D) None

Answer

Correct option is C

7. What are the advantages of neural networks over conventional computers?

- (i) They have the ability to learn by example.
 - (ii) They are more fault tolerant.
 - (iii) They are more suited for real time operation due to their high 'computational' rates.
- (A) (i) and (ii)
(B) (i) and (iii)
(C) Only (i)
(D) All
(E) None

Answer

Correct option is D

8. What is Neuro software?

- (A) It is software used by Neurosurgeon
(B) Designed to aid experts in real world
(C) It is powerful and easy neural network
(D) A software used to analyze neurons

Answer

Correct option is C

9. Which is true for neural networks?

- (A) Each node computes its weighted input
(B) Node could be in excited state or non-excited state
(C) It has set of nodes and connections
(D) All of the above

Answer

Correct option is D

10. What is the objective of backpropagation algorithm?

- (A) To develop learning algorithm for multilayer feedforward neural network, so that network can be trained to capture the mapping implicitly
(B) To develop learning algorithm for multilayer feedforward neural network

- (C) To develop learning algorithm for single layer feedforward neural network
- (D) All of the above

Answer

Correct option is A

11. Which of the following is true?

Single layer associative neural networks do not have the ability to:-

- (i) Perform pattern recognition
 - (ii) Find the parity of a picture
 - (iii) Determine whether two or more shapes in a picture are connected or not
- (A) (ii) and (iii)
 - (B) Only (ii)
 - (C) All
 - (D) None

Answer

Correct option is A

12. The backpropagation law is also known as generalized delta rule.

- (A) True
- (B) False

Answer

Correct option is A

13. Which of the following is true?

- (i) On average, neural networks have higher computational rates than conventional computers.
 - (ii) Neural networks learn by example.
 - (iii) Neural networks mimic the way the human brain works.
- (A) All
 - (B) (ii) and (iii)
 - (C) (i), (ii) and (iii)
 - (D) None

Answer

Correct option is A

14. What is true regarding backpropagation rule?
- (A) Error in output is propagated backwards only to determine weight updates
 - (B) There is no feedback of signal at any stage
 - (C) It is also called generalized delta rule
 - (D) All of the above

Answer

Correct option is D

15. There is feedback in final stage of backpropagation algorithm.
- (A) True
 - (B) False

Answer

Correct option is B

16. An auto-associative network is
- (A) A neural network that has only one loop
 - (B) A neural network that contains feedback
 - (C) A single layer feed-forward neural network with pre-processing
 - (D) A neural network that contains no loops

Answer

Correct option is B

17. A 3-input neuron has weights 1, 4 and 3. The transfer function is linear with the constant of proportionality being equal to 3. The inputs are 4, 8 and 5 respectively. What will be the output?
- (A) 139
 - (B) 153
 - (C) 612
 - (D) 160

Answer

Correct option is B

18. What of the following is true regarding backpropagation rule?
- (A) Hidden layers output is not all important, they are only meant for supporting input and

output layers

- (B) Actual output is determined by computing the outputs of units for each hidden layer
- (C) It is a feedback neural network
- (D) None of the above

Answer

Correct option is B

19. What is back propagation?

- (A) It is another name given to the curvy function in the perceptron
- (B) It is the transmission of error back through the network to allow weights to be adjusted so that the network can learn
- (C) It is another name given to the curvy function in the perceptron
- (D) None of the above

Answer

Correct option is B

20. The general limitations of back propagation rule is/are

- (A) Scaling
- (B) Slow convergence
- (C) Local minima problem
- (D) All of the above

Answer

Correct option is D

21. What is the meaning of generalized in statement “backpropagation is a generalized delta rule” ?

- (A) Because delta is applied to only input and output layers, thus making it more simple and generalized
- (B) It has no significance
- (C) Because delta rule can be extended to hidden layer units
- (D) None of the above

Answer

Correct option is C

22. Neural Networks are complex _____ functions with many parameters.

- (A) Linear
- (B) Non linear
- (C) Discreate
- (D) Exponential

Answer

Correct option is A

23. The general tasks that are performed with backpropagation algorithm

- (A) Pattern mapping
- (B) Prediction
- (C) Function approximation
- (D) All of the above

Answer

Correct option is D

24. Backpropagaion learning is based on the gradient descent along error surface.

- (A) True
- (B) False

Answer

Correct option is A

25. In backpropagation rule, how to stop the learning process?

- (A) No heuristic criteria exist
- (B) On basis of average gradient value
- (C) There is convergence involved
- (D) None of these

Answer

Correct option is B

26. Applications of NN (Neural Network)

- (A) Risk management
- (B) Data validation

- (C) Sales forecasting
- (D) All of the above

Answer

Correct option is D

27. The network that involves backward links from output to the input and hidden layers is known as

- (A) Recurrent neural network
- (B) Self organizing maps
- (C) Perceptrons
- (D) Single layered perceptron

Answer

Correct option is A

28. Decision Tree is a display of an algorithm.

- (A) True
- (B) False

Answer

Correct option is A

29. Which of the following is/are the decision tree nodes?

- (A) End Nodes
- (B) Decision Nodes
- (C) Chance Nodes
- (D) All of the above

Answer

Correct option is D

30. End Nodes are represented by which of the following

- (A) Solar street light
- (B) Triangles
- (C) Circles
- (D) Squares

Answer
Correct option is B

31. Decision Nodes are represented by which of the following
- (A) Solar street light
 - (B) Triangles
 - (C) Circles
 - (D) Squares

Answer
Correct option is D

32. Chance Nodes are represented by which of the following
- (A) Solar street light
 - (B) Triangles
 - (C) Circles
 - (D) Squares

Answer
Correct option is C

33. Advantage of Decision Trees
- (A) Possible Scenarios can be added
 - (B) Use a white box model, if given result is provided by a model
 - (C) Worst, best and expected values can be determined for different scenarios
 - (D) All of the above

Answer
Correct option is D

1. Which of the following guidelines is applicable to initialization of the weight vector in a fully connected neural network.

- a) Should not set it to zero since otherwise it will cause overfitting
- b) Should not set it to zero since otherwise (stochastic) gradient descent will explore a very small space

- c) Should set it to zero since otherwise it causes a bias
- d) Should set it to zero in order to preserve symmetry across all neurons

[View Answer](#)

Answer: (b) should not set it to zero since otherwise gradient descent will explore a very small space

If we initialize all the weights to zero, the neural network will train but all the neurons will learn the same features during training. Setting all weights to zero makes your model equivalent to a linear model. When you set all weight to 0, the derivative with respect to loss function is the same for every w in weight matrix, thus, all the weights have the same values in the subsequent iteration. Hence, they must be initialized to random numbers.

2. Given two Boolean random variables, A and B , where $P(A) = 1/2$, $P(B) = 1/3$, and $P(A | \neg B) = 1/4$, what is $P(A | B)$?

- a) $1/6$
- b) $1/4$
- c) $3/4$
- d) 1

[View Answer](#)

Answer: (d) 1

$$P(A | B) = (P(B | A) P(A)) / P(B) = 3/2 P(B | A).$$

$$P(B | A) = 1 - P(\neg B | A), \text{ and } P(\neg B | A) = (P(A | \neg B) P(\neg B)) / P(A) = 1/3,$$

so $P(B | A) = 2/3$ and therefore

$$P(A | B) = (2/3)(3/2) = 1$$

3. For a neural network, which one of these structural assumptions is the one that most affects the trade-off between underfitting (i.e. a high bias model) and overfitting (i.e. a high variance model):

- a) The number of hidden nodes
- b) The learning rate
- c) The initial choice of weights
- d) The use of a constant-term unit input

[View Answer](#)

Answer: (a) The number of hidden nodes

The number of hidden nodes. 0 will result in a linear model, which many (with non-linear activation) significantly increases the variance of the model. A feed forward neural network without hidden nodes can only find linear decision boundaries.

The Hidden nodes perform computations and transfer information from the input nodes to the output nodes. A collection of hidden nodes forms a “Hidden Layer”. While a feedforward network will only have a single input layer and a single output layer, it can have zero or multiple Hidden Layers.

Overfitting: If there are so many neurons in the hidden layers it might cause Overfitting. Overfitting occurs when unnecessary more neurons are present in the network.

Underfitting: If the number of neurons are less as compared to the complexity of the problem data it takes towards the Underfitting. It occurs when there are few neurons in the hidden layers to detect the signal in complicated data set.

4. You've just finished training a decision tree for spam classification, and it is getting abnormally bad performance on both your training and test sets. You know that your implementation has no bugs, so what could be causing the problem?

- a) Your decision trees are too shallow.
- b) You need to increase the learning rate.
- c) You are overfitting.
- d) None of the above.

[View Answer](#)

Answer: (a) your decision trees are too shallow

Shallow decision trees - trees that are too shallow might lead to overly simple models that can't fit the data.

A model that is underfit will have high training and high testing error. Hence, bad performance on training and test sets indicates underfitting which means the set of hypotheses are not complex enough (decision trees that are shallow) to include the true but unknown prediction function.

The shallower the tree the less variance we have in our predictions; however, at some point we can start to inject too much bias as shallow trees (e.g., stumps) are not able to capture interactions and complex patterns in our data.

5. _____ refers to a model that can neither model the training data nor generalize to new data.

- a) good fitting
- b) overfitting
- c) underfitting
- d) all of the above

[View Answer](#)

Answer: (c) underfitting

An underfit machine learning model is not a suitable model and will be obvious as it will have poor performance on the training data. Usually, a model that is underfit will have high training and high testing error
