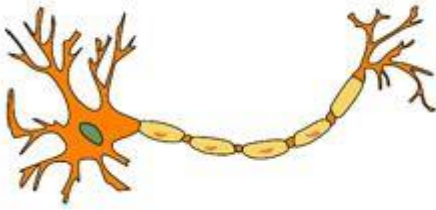**1)The neural network consists of many neurons, each neuron takes an input, processes it and gives an output. Here's a diagrammatic representation of a real neuron.**

 **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

2) 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:

$$f(x) = \begin{cases} 0, & for\ x < 0 \\ 1, & for\ x \geq 0 \end{cases}$$
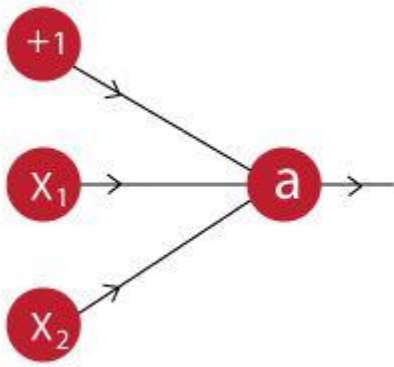
What would be the weights and bias?

(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

3) What are the steps for using a gradient descent algorithm?

3
5
2
1
4

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

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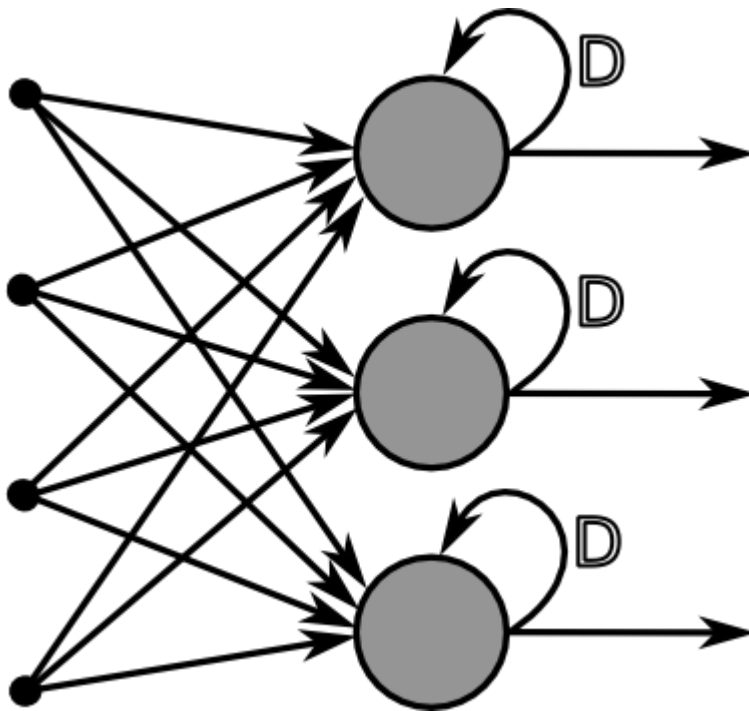
A. True

B. False

5) 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

6) 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

7) What is the sequence of the following tasks in a perceptron?

1
4
3
2

   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**
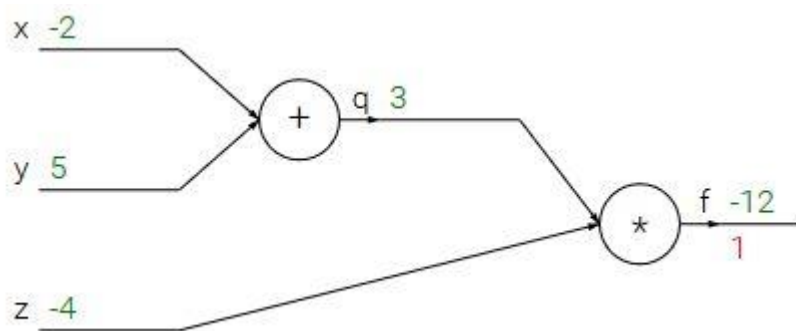
A. 1, 2, 3, 4

B. 4, 3, 2, 1

C. 3, 1, 2, 4

D. 1, 4, 3, 2

8) 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)

9)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

10)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

11) 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

12) 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

13) 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

14) 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

15) What if we use a learning rate that's too large?

A. Network will converge

B. Network will not converge

C. Can't Say

D. All of above

16) 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

C. Mini Batch Gradient Descent

D None Of Above

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

A. True

B. False

18) 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

19) 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

20) 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

21) What are the benefits of mini-batch gradient descent?

a) This is more efficient compared to stochastic gradient descent.
b) The generalization by finding the flat minima.
c) Mini-batches allow help to approximate the gradient of the entire training set which helps us to avoid local minima.
d) All of above

22)Which of the given point is true about auto encoder

a) It is an unsupervised ML algorithm similar to Principal Component Analysis
b) It minimizes the same objective function as Principal Component Analysis
c) The neural network's target output is its input
d) All of above

23) What are Hyperparameters are set before training.

a) Activation Function
b) Learning Rate
c) Momentum
d) All of above

24) In training a neural network, you notice that the loss does not decrease in the few starting epochs.
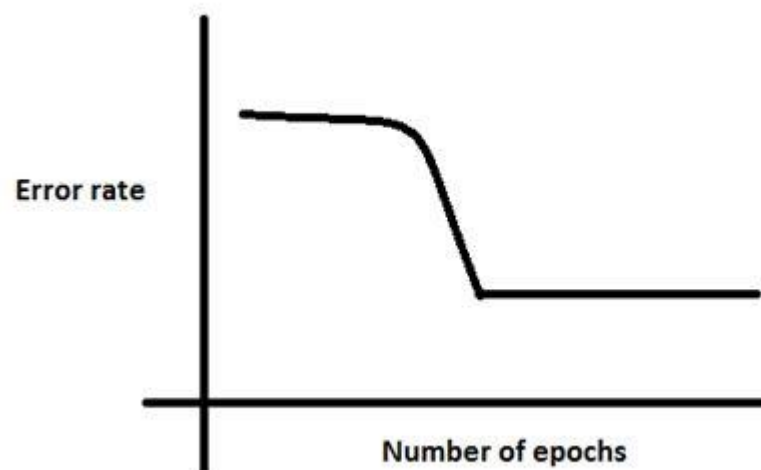
The reasons for this could be:

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

25) Which statement is true about NAG?

a) It calculates gradient from future position.

b) It calculates gradient from current position.

c)It calculates gradient from previous position.

d)None

26)Which of the following statement is FALSE about ADAGRAD?

a) There is change in learning rate for each update of weights
b) It will decrease learning rate if weights are being updated in short amount of time
c) It will increase learning rate if weights are not being updated
d) Learning rate remains same it doesn't modify

27) Gradient Descent with Momentum focus more on finding _____

a) Local minima
b) Global minima
c) Close to local minima

28) How many states are transferred from previous cell to next cell in LSTM

a) 1
b) 2
c) 3
d) 4

29)Sigmoid function of forget get in LSTM is responsible for __

a) Which values to keep

b) Which values to discard

c)Both

d)None

30) Which of the following is authentic approximately model capability (in which version capacity method the potential of the neural community to approximate complex capabilities) in TensorFlow?

   a) As range of hidden layers boom, model capability will increase
   b) As dropout ratio increases, version capacity increases
   c) As mastering charge will increase, model capacity will increase
   d) None of theseThrough map reduce tasks

31) Can a neural network model the characteristic (y=1/x) in TensorFlow?

   a) True
   b) False
   c) Can't Say it depends on dataset
   d) None

32) The memory blocks are responsible for remembering things and manipulations to this memory is done through _____ in LSTM

   a) Cell
   b) Gates
   c) Cell State
   d) Activation Functions

33) Alexnet have _____ conv layers and ___ fully connected layers.

   a) 3,5
   b) 5,3
   c) 5,2
   d) 4,3

34) Broadly how many kinds of stability can be defined in neural networks?
   a) 1
   b) 3
   c) 2
   d) 4

35) Why is the training of basis function is faster than MLFFNN?
   a) because they are developed specifically for pattern approximation
   b) because they are developed specifically for pattern classification
   c) because they are developed specifically for pattern approximation or classification
   d) none of the mentioned

36) What consist of a basic counterpropagation network?
     a) a feedforward network only
     b) a feedforward network with hidden layer
     c) two feedforward network with hidden layer
     d) none of the mentioned

37)How many gates are there in GRU?

  a) 1
  b) 2
  c) 3
  d) 4

38) How many types of GAN?

  a) 3
  b) 4
  c) 6
  d) 5

39)Which loss function we use if input values are in range [0,1] in auto encoders.

  a) Mean squared error
  b) Binary cross entropy
  c) Both a and b
  d) None

40) "Convolutional Neural Networks can carry out various forms of transformation (rotations or scaling) in an enter". Is the assertion correct true or false in TensorFlow?

  a) True
  b) False