Deep Learning Interview Questions

By Arya | Last updated on May 22, 2024 | 41243 Views

Previous   Next

Tutorial Playlist

ML Job Preparation

Machine Learning Interview Questions and Answers

Machine Learning Course Syllabus - 2024

Deep Learning Interview Questions

TensorFlow Interview Questions

Top 10 Machine Learning Project Ideas for Beginners

Machine Learning Engineer Salary in India

20K+
career-aspirant learners have read this article

Reviewed and fact-checked by

Akash Pushkar

Principal Data Scientist

Did you know?

- Deep Learning **Diagnoses Diseases and Develops Drugs:** From accurately identifying tumors in medical scans to designing personalized cancer treatments, deep learning is transforming healthcare. Its ability to analyze vast amounts of medical data leads to faster diagnoses, better treatments, and potentially cures for previously untreatable diseases.

- **Bias and Explainability in Deep Learning Systems:** As deep learning becomes more integrated into society, concerns about **bias and explainability arise**. Algorithms trained on biased data can perpetuate unfairness, and their decision-making processes can be opaque, r concerns about transparency and accountability.

- **Creative Revolution:** Deep learning **paints, writes, and composes**: Beyond traditional tasks, deep learning is now generating artistic masterpieces.

Hey There!!! Chat with Us?

From writing poetry and scripts to composing music and painting stunning landscapes, these models are blurring the lines between human and machine creativity.

Table of content

Show More

Deep Learning is a well-covered skill to possess in the 21st century. Working with it requires a lot of effort and this is seen in the interviews as well. The questions can sometimes get a bit tough.

This 'Top Deep Learning Interview Questions' blog is put together with questions sourced from experts in the field, which have the highest probability of occurrence in interviews. Studying these questions will help you ace your next Deep Learning interview.

Check out our Deep Learning Interview Questions And Answers on YouTube, designed especially for beginners:

Deep Learning Interview Questions And Answers | AI & Deep Lear...

Learn for free !
Subscribe to our youtube channel.
GET STARTED
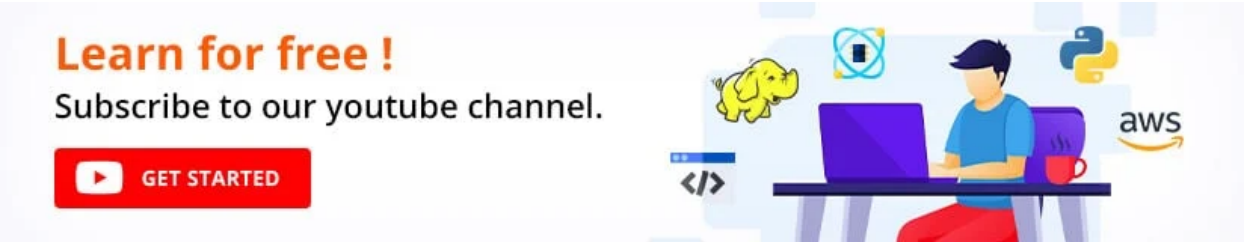
Basic Deep Learning Interview Questions for Freshers

1. What is the difference between Machine Learning and Deep Learning?

Machine Learning forms a subset of Artificial Intelligence, where we use statistics and algorithms to train machines with data, thereby, helping them improve with experience.

Deep Learning is a part of Machine Learning, which involves mimicking the human brain in terms of structures called neurons, thereby, forming **neural networks**.

Enroll in our [Machine Learning Training](#) in collaboration with IIT Madras and learn from Expert to be an Expert.

2. What is a perceptron?

A [perceptron](#) is similar to the actual neuron in the human brain. It receives inputs from various entities and applies functions to these inputs, which transform them to be the output.

A perceptron is mainly used to perform binary classification where it sees an input, computes functions based on the weights of the input, and outputs the required transformation.

3. How is Deep Learning better than Machine Learning?

Machine Learning is powerful in a way that it is sufficient to solve most of the problems. However, **Deep Learning** gets an upper hand when it comes to working with data that has a large number of dimensions. With data that is large in size, a Deep Learning model can easily work with it as it is built to handle this.

Learn more about Machine Learning. Enroll in [Machine Learning training in Bangalore](#).

4. What are some of the most used applications of Deep Learning?

Deep Learning is used in a variety of fields today. The most used ones are as follows:

- Sentiment Analysis
- Computer Vision
- Automatic Text Generation
- Object Detection
- Natural Language Processing
- Image Recognition

5. What is the meaning of overfitting?

Overfitting is a very common issue when working with Deep Learning. It is a scenario where the [Deep Learning algorithm](#) vigorously hunts through the data to obtain some valid information.

This makes the Deep Learning model pick up noise rather than useful data, causing very high variance and low bias. This makes the model less accurate, and this is an undesirable effect that can be prevented.

6. What are activation functions?

[Activation functions](#) are entities in Deep Learning that are used to translate inputs into a usable output parameter. It is a function that decides if a neuron needs activation or not by calculating the weighted sum on it with the bias.

Using an activation function makes the model output to be non-linear. There are many types of activation functions:

- ReLU
- Softmax
- Sigmoid
- Linear
- Tanh

7. Why is the Fourier transform used in Deep Learning?

Fourier transform is an effective package used for analyzing and managing large amounts of data present in a [database](#). It can take in real-time array data and process it quickly. This ensures that high efficiency is maintained and also makes the model more open to processing a variety of signals.

8. What are the steps involved in training a perception in Deep Learning?

There are five main steps that determine the learning of a perceptron:

1. Initialize thresholds and weights
2. Provide inputs
3. Calculate outputs
4. Update weights in each step
5. Repeat steps 2 to 4

9. What is the use of the loss function?

The [loss function](#) is used as a measure of accuracy to see if a neural network has learned accurately from the training data or not. This is done by comparing the training dataset to the testing dataset.

The loss function is a primary measure of the performance of the neural network. In Deep Learning, a good performing network will have a low loss function at all times when training.

10. What are some of the Deep Learning frameworks or tools that you have used?

This question is quite common in a Deep Learning interview. Make sure to answer based on the experience you have with the tools.

However, some of the top Deep Learning frameworks out there today are:

- TensorFlow
- Keras
- PyTorch
- Caffe2
- CNTK
- MXNet
- Theano

11. What is the use of the swish function?

The swish function is a self-gated activation function developed by Google. It is now a popular activation function used by many as Google claims that it outperforms all of the other activation functions in terms of computational efficiency.

12. What are autoencoders?

Autoencoders are artificial neural networks that learn without any supervision. Here, these networks have the ability to automatically learn by mapping the inputs to the corresponding outputs.

Autoencoders, as the name suggests, consist of two entities:

- Encoder: Used to fit the input into an internal computation state
- Decoder: Used to convert the computational state back into the output

13. What are the steps to be followed to use the gradient descent algorithm?

There are five main steps that are used to initialize and use the gradient descent algorithm:

- Initialize biases and weights for the network
- Send input data through the network (the input layer)
- Calculate the difference (the error) between expected and predicted values
- Change values in neurons to minimize the loss function
- Multiple iterations to determine the best weights for efficient working

14. Differentiate between a single-layer perceptron and a multi-layer perceptron.

Here is the differnece between single-layer perceptron and a multi-layer perceptron :

Single-layer Perceptron	Multi-layer Perceptron
Cannot classify non-linear data points	Can classify non-linear data
Takes in a limited amount of	Withstands a lot of parameters

parameters	
Less efficient with large data	Highly efficient with large datasets

15. What is data normalization in Deep Learning?

Data normalization is a preprocessing step that is used to refit the data into a specific range. This ensures that the network can learn effectively as it has better convergence when performing backpropagation.

16. What is forward propagation?

Forward propagation is the scenario where inputs are passed to the hidden layer with weights. In every single hidden layer, the output of the activation function is calculated until the next layer can be processed. It is called forward propagation as the process begins from the input layer and moves toward the final output layer.

17. What is backpropagation?

Backpropagation is used to minimize the cost function by first seeing how the value changes when weights and biases are tweaked in the neural network. This change is easily calculated by understanding the gradient at every hidden layer. It is called backpropagation as the process begins from the output layer, moving backward to the input layers.

18. What are hyperparameters in Deep Learning?

Hyperparameters are variables used to determine the structure of a neural network. They are also used to understand parameters, such as the learning rate and the number of hidden layers, and more, present in the neural network.

19. How can hyperparameters be trained in neural networks?

Hyperparameters can be trained using four components as shown below:

- Batch size: This is used to denote the size of the input chunk. Batch sizes can be varied and cut into sub-batches based on the requirement.
- Epochs: An epoch denotes the number of times the training data is visible to the neural network so that it can train. Since the process is iterative, the number of epochs will vary based on the data.
- Momentum: Momentum is used to understand the next consecutive steps that occur with the current data being executed at hand. It is used to avoid oscillations when training.
- Learning rate: Learning rate is used as a parameter to denote the time required for the network to update the parameters and learn.

20. What is Deep Learning?

Deep learning is a subset of machine learning that is completely based on Artificial Intelligence. It used to teach computers to process data in a way that was inspired by the human brain; it recognized the complex patterns in pictures, text, sounds, and so on.

21. What are Neural Networks?

Neural Network is also known as an Artificial Neural Network. It is a subset of machine learning that consists of interconnected nodes or neurons that process and learn from the data.

22. What are the advantages and disadvantages of neural networks?

**Advantages of Neural Networks:**

- Neural networks can learn complex models and non-linear relationships.
- It stores all the information on the entire network with the help of nodes.
- Neural networks also can work with unorganized data.
- Neural networks can perform more than one function at a time.
- If one or more than one cell is corrupted, even though the output doesn't have an impact.

Disadvantages of Neural Networks:

- Due to their quick adaptation to the changing requirements, neural networks require heavy machinery and hardware to work.
- Neural networks depend on a lot of training data, which leads to the problem of overfitting.
- Neural networks require lots of computational power because they act like a human brain and are composed of many interconnected nodes, and each node computes based on weights.
- Neural networks are much more complex and hard to explain than other models.
- Neural network models need careful attention in data preparation because it's a crucial step in machine learning and harms the input data.

23. What is the Learning Rate in the context of Neural Network Models?

The learning rate is a hyperparameter that controls the size of the updates that were created by the weights during data training. It also determines the size of each step in each training iteration. The default value of the learning rate is 0.1 or 0.01, and it's represented by the character 'a'.
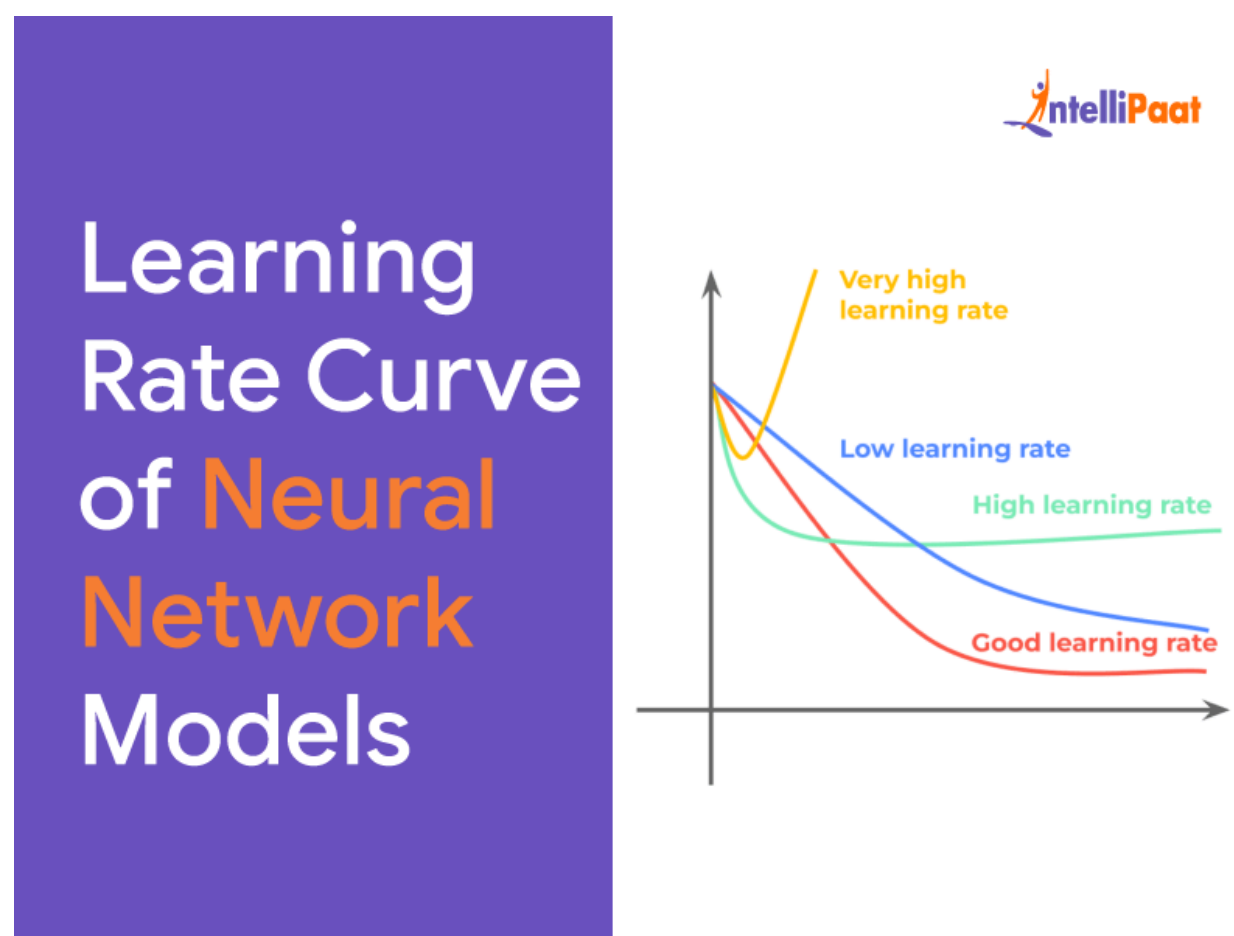
24. What is a Deep Neural Network?

A deep neural network is a machine learning algorithm that mimics the brain's information processing. It's made up of multiple layers of nodes known as neurons. DNN is used in complex mathematical modeling.

25. What are the different types of Deep Neural Networks?

There are 4 types of deep neural networks:

1. Feed Forward Neural Network: The Feed Forward Neural Network is the basic neural network, whose flow control starts from the input layer and moves forward to the output layer. The data will flow only in a single direction; there is no backpropagation mechanism.
2. Recurrent Neural Network: A recurrent neural network is another type of deep neural network in which the data will flow in a single direction. In this neural network, each neuron is present in the hidden layer, and they receive the input with a specific delay in time.
3. Convolutional Neural Network: A convolutional neural network is a special kind of neural network that we can use for image classification, clustering of images, and so on.
4. Restricted Boltzmann Machine: Restricted Boltzmann Machine is another type of Boltzmann Machine where the neurons present in the input layer and the hidden layer are surrounded by symmetric connections. This machine algorithm can be used in filtering, feature learning, and risk detection.

26. Explain Data Normalization. What is the need for it?

Data normalization helps us normalize the neural network nodes into different branches. It works by subtracting the mean and dividing it by the standard deviation.

Data normalization helps to make the data stable because whatever the features are in the dataset, they are not on the same scale, which makes the data difficult to learn.

Next up on this top Deep Learning interview questions and answers blog, let us take a look at the intermediate questions.

Intermediate Deep Learning Interview Questions

27. What is the meaning of dropout in Deep Learning?

Dropout is a technique that is used to avoid overfitting a model in Deep Learning. If the dropout value is too low, then it will have minimal effect on learning. If it is too high, then the model can under-learn, thereby, causing lower efficiency.

28. What are tensors?

Tensors are multidimensional arrays in Deep Learning that are used to represent data. They represent the data with higher dimensions. Due to the high-level nature of the programming languages, the syntax of tensors is easily understood and broadly used.

29. What is the meaning of model capacity in Deep Learning?

In Deep Learning, model capacity refers to the capacity of the model to take in a variety of mapping functions. Higher model capacity means a large amount of information can be stored in the network.

We will check out neural network interview questions alongside as it is also a vital part of Deep Learning.

30. What is a Boltzmann machine?

A Boltzmann machine is a type of recurrent neural network that uses binary decisions, alongside biases, to function. These neural networks can be hooked up together to create deep belief networks, which are very sophisticated and used to solve the most complex problems out there.

31. What are some of the advantages of using TensorFlow?

TensorFlow has numerous advantages, and some of them are as follows:

- High amount of flexibility and platform independence
- Trains using CPU and GPU
- Supports auto differentiation and its features
- Handles threads and asynchronous computation easily
- Open-source
- Has a large community

32. What is a computational graph in Deep Learning?

A computation graph is a series of operations that are performed to take inputs and arrange them as nodes in a graph structure. It can be considered as a way of implementing mathematical calculations into a graph. This helps in parallel processing and provides high performance in terms of computational capability.

If you are looking forward to becoming an expert in Deep Learning, make sure to check out Intellipaat's AI Engineer Course.

33. What is a CNN?

CNNs are convolutional neural networks that are used to perform analysis on image annotation and visuals. These classes of neural networks can input a multi-channel image and work on it easily.

These Deep Learning questions must be answered in a concise way. So make sure to understand them and revisit them if necessary.

34. What are the various layers present in a CNN?

There are four main layers that form a convolutional neural network:

- Convolution: These are layers consisting of entities called filters that are used as parameters to train the network.
- ReLu: It is used as the activation function and is always used with the convolution layer.
- Pooling: Pooling is the concept of shrinking the complex data entities that form after convolution and is primarily used to maintain the size of an image after shrinkage.
- Connectedness: This is used to ensure that all of the layers in the neural network are fully connected and activation can be computed using the bias easily.

35. What is an RNN in Deep Learning?

RNNs stand for [recurrent neural networks](#), which form to be a popular type of artificial neural network. They are used to process sequences of data, text, genomes, handwriting, and more. RNNs make use of backpropagation for the training requirements.

36. What is a vanishing gradient when using RNNs?

Vanishing gradient is a scenario that occurs when we use RNNs. Since RNNs make use of backpropagation, gradients at every step of the way will tend to get smaller as the network traverses through backward iterations. This equates to the model learning very slowly, thereby, causing efficiency problems in the network.

37. What is exploding gradient descent in Deep Learning?

Exploding [gradients](#) are an issue causing a scenario that clumps up the gradients. This creates a large number of updates of the weights in the model when training.

The working of gradient descent is based on the condition that the updates are small and controlled. Controlling the updates will directly affect the efficiency of the model.

38. What is the use of LSTM?

[LSTM](#) stands for long short-term memory. It is a type of RNN that is used to sequence a string of data. It consists of feedback chains that give it the ability to perform like a general-purpose computational entity.

39. Where are autoencoders used?

[Autoencoders](#) have a wide variety of usage in the real world. The following are some of the popular ones:

- Adding color to black–white images
- Removing noise from images
- Dimensionality reduction
- Feature removal and variation

40. What are the types of autoencoders?

There are four main types of autoencoders:

- Deep autoencoders
- Convolutional autoencoders
- Sparse autoencoders
- Contractive autoencoders

41. What is a Restricted Boltzmann Machine?

A Restricted Boltzmann Machine, or RBM for short, is an undirected graphical model that is popularly used in Deep Learning today. It is an algorithm that is used to perform:

- [Dimensionality reduction](#)
- [Regression](#)
- [Classification](#)
- Collaborative filtering

- Topic modeling

42. What do you mean by end-to-end learning?

In end-to-end learning, the model will learn all the steps between the input and the output result. The model learns all the useful features extracted from the data, which helps train the model for the complex dataset.

43. What is Forward and Back Propagation in Deep Learning?

Forward Propagation is the way data moves from left to right in the neural network, ie. from the input layer to the output layer.

Back Propagation is the way data moves from right to left, i.e., from the output layer to the input layer. Both ways help the data train properly; once the corrected weight is learned, it will be able to converge and generalize the data better.

44. What would happen if we set all the biases and weights to zero to train a neural network?

Yes, if all the biases are set to zero, then the neural network model has a chance of learning.
No, if the training model is set to zero, because the neural network will never learn the complete task. If the weights are set to zero then the derivatives for each weight remain constant, which leads the neurons to learn the same features in each iteration and generate poor results.

45. Explain the difference between a Shallow Network and a Deep Network.

Shallow Network: The shallow network has only one hidden layer; it will fit in any function, and it also requires a large number of input parameters. Shallow neural networks tell us exactly what is going on inside the deep neural network.

Deep Network: The deep network has numerous hidden layers, and it will also fit in any function. Deep neural networks are mostly used for data-driven modeling.

46. For the application of Face Detection, which deep learning algorithm would you use?

The best algorithm for face detection is Convolutional Neural Networks because CNN gives us better accuracy in object detection tasks, and it is a two-stage architecture with a region proposal network that improves localization.

47. What is an Activation Function?

The activation function in artificial neural networks helps the network learn the complex patterns in the data. The activation function is responsible for what data is to be fired to the next neurons at the end of the process.

48. What do you mean by an Epoch in the context of deep learning?

In deep learning, an [epoch](#) is a term that refers to the number of passes the machine has made across the fully trained dataset. The number of epochs is equal to the number of iterations if the batch size is the entire training dataset.

```
1  Formula,
2  d*e = i*b
```

where,
d → dataset size
e → number of epoch
i → number of iterations
b → batch size

Next up on this top Deep Learning interview questions and answers blog, let us take a look at the advanced questions.

Advanced Deep Learning Interview Questions for Experienced

49. What are some of the limitations of Deep Learning?

There are a few disadvantages of Deep Learning as mentioned below:

- Networks in Deep Learning require a huge amount of data to train well.
- Deep Learning concepts can be complex to implement sometimes.
- Achieving a high amount of model efficiency is difficult in many cases.

These are some of the vital advanced deep learning interview questions that you have to know about!

50. What are the variants of gradient descent?

There are three variants of gradient descent as shown below:

- Stochastic gradient descent: A single training example is used for the calculation of gradient and for updating parameters.
- Batch gradient descent: Gradient is calculated for the entire dataset, and parameters are updated at every iteration.
- Mini-batch gradient descent: Samples are broken down into smaller-sized batches and then worked on as in the case of stochastic gradient descent.

51. Why is mini-batch gradient descent so popular?

Mini-batch gradient descent is popular as:

- It is more efficient when compared to stochastic gradient descent.
- Generalization is done by finding the flat minima.
- It helps avoid the local minima by allowing the approximation of the gradient for the entire dataset.

52. What are deep autoencoders?

Deep autoencoders are an extension of the regular autoencoders. Here, the first layer is responsible for the first-order function execution of the input. The second layer will take care of the second-order functions, and it goes on.

Usually, a deep autoencoder is a combination of two or more symmetrical deep-belief networks where:

- The first five shallow layers consist of the encoding part
- The other layers take care of the decoding part

On the next set of Deep Learning questions, let us look further into the topic.

53. Why is the Leaky ReLU function used in Deep Learning?

Leaky ReLU, also called LReL, is used to manage a function to allow the passing of small-sized negative values if the input value to the network is less than zero.

54. What are some of the examples of supervised learning algorithms in Deep Learning?

There are three main supervised learning algorithms in Deep Learning:

- Artificial neural networks
- Convolutional neural networks
- Recurrent neural networks

55. What are some of the examples of unsupervised learning algorithms in Deep Learning?

There are three main unsupervised learning algorithms in Deep Learning:

- Autoencoders
- Boltzmann machines
- Self-organizing maps

Next up, let us look at more neural network interview questions that will help you ace the interviews.

56. Can we initialize the weights of a network to start from zero?

Yes, it is possible to begin with zero initialization. However, it is not recommended to use because setting up the weights to zero initially will cause all of the neurons to produce the same output and the same gradients when performing backpropagation. This means that the network will not have the ability to learn at all due to the absence of asymmetry between each of the neurons.

57. What is the meaning of valid padding and same padding in CNN?

- Valid padding: It is used when there is no requirement for padding. The output matrix will have the dimensions $(n – f + 1)$ X $(n – f + 1)$ after convolution.
- Same padding: Here, padding elements are added all around the output matrix. It will have the same dimensions as the input matrix.

58. What are some of the applications of transfer learning in Deep Learning?

Transfer learning is a scenario where a large model is trained on a dataset with a large amount of data and this model is used on simpler [datasets](#), thereby resulting in extremely efficient and accurate neural networks.

The popular examples of transfer learning are in the case of:

- BERT
- ResNet
- GPT-2
- VGG-16

59. How is the transformer architecture better than RNNs in Deep Learning?

With the use of sequential processing, programmers were up against:

- The usage of high processing power
- The difficulty of parallel execution

This caused the rise of the transformer architecture. Here, there is a mechanism called attention mechanism, which is used to map all of the dependencies between sentences, thereby making huge progress in the case of NLP models.

Work on your NLP skills by enrolling in [NLP Training in Chennai](#).
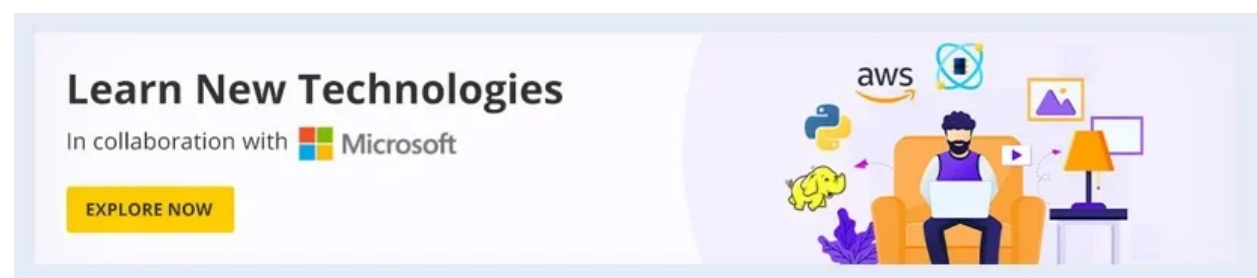
60. What are the steps involved in the working of an LSTM network?

There are three main steps involved in the working of an LSTM network:

- The network picks up the information that it has to remember and identifies what to forget.
- Cell state values are updated based on Step 1.
- The network calculates and analyzes which part of the current state should make it to the output.

61. What are the elements in TensorFlow that are programmable?

In TensorFlow, users can program three elements:

- [Constants](#)
- Variables
- Placeholders

62. What is the meaning of bagging and boosting in Deep Learning?

[Bagging](#) is the concept of splitting a dataset and randomly placing it into bags for training the model.

Boosting is the scenario where incorrect data points are used to force the model to produce the wrong output. This is used to retrain the model and increase accuracy.

63. What are generative adversarial networks (GANs)?

Generative adversarial networks are used to achieve generative modeling in Deep Learning. It is an unsupervised task that involves the discovery of patterns in the input data to generate the output.

The generator is used to generate new examples, while the discriminator is used to classify the examples generated by the generator.

Learn more about [Generative AI](#)

64. Why are generative adversarial networks (GANs) so popular?

Generative adversarial networks are used for a variety of purposes. In the case of working with images, they have a high amount of traction and efficient working.

- Creation of art: GANs are used to create artistic images, sketches, and paintings.
- Image enhancement: They are used to greatly enhance the resolution of the input images.
- Image translation: They are also used to change certain aspects, such as day to night and summer to winter, in images easily.

If you are looking forward to becoming an expert in Deep Learning, make sure to check out Intellipaat's [AI Course](#). With this program, you can become proficient in all of the concepts of Deep Learning and AI and earn a course certificate as well.

65. How does the choice of cost function impact the convergence properties of a deep neural network?

The convergence properties of a deep neural network are heavily influenced by the choice of cost function as it defines the gradient landscape. For example, cross-entropy loss creates a more aggressive gradient, providing more "signal" per update when the predictions are wrong, thus often leading to faster convergence, particularly in classification problems, where the output probabilities are being modeled.

66. Can you explain the advantages of using a cross-entropy loss over mean squared error in classification tasks?

Cross-entropy loss tends to work better for classification since it penalizes incorrect classifications more heavily than mean squared error (MSE), which can lead to quicker and more stable training. Cross-entropy aligns with the gradient updates of probabilistic outcomes, directly correlating to the likelihood of predicting true labels.

67. Describe a scenario where you would use a custom loss function and how you would go about implementing it.

If a problem has a unique cost structure (e.g., a different cost for different types of misclassifications), I would design a custom loss function. This requires ensuring the custom loss is differentiable for gradient-based optimization, and I would use

ML Job Preparation

automatic differentiation capabilities of deep learning frameworks like TensorFlow or PyTorch for implementation.

68. Describe the role of convolutional layers in CNNs and how they differ from fully connected layers regarding feature extraction.

Convolutional layers act as feature extractors that slide across input space and produce feature maps, highlighting features like edges or textures, whereas fully connected layers take those features to learn non-linear combinations that aid in classification.

69. Discuss the concept of bias-variance trade-off in the context of neural network weights and model complexity.

The bias-variance trade-off in the context of neural networks is about finding the right balance between a model that is too simple (high bias) and one that is too complex (high variance). If the weights are poorly chosen, the network can either fail to capture the underlying patterns (underfitting) or capture too much noise (overfitting).

Have a look at this blog to have an understanding of *Statistics for Data Science*

70. Explain the concept of receptive field in a CNN and how it relates to the architecture's ability to recognize patterns of different scales.

The receptive field in a CNN is the area of the input image that a neuron 'sees.' Initially, it captures basic features like edges, and in deeper layers, it represents more complex patterns due to larger receptive fields. The architecture is designed so that these fields grow progressively to recognize objects of various sizes, balancing the need to detect both fine details and broader patterns.

71. How does the concept of feature map concatenation in networks like DenseNet affect the performance and parameter efficiency of a model?

Feature map concatenation in DenseNet architectures allows each layer to access feature maps from all preceding layers, promoting feature reuse, which significantly improves the parameter efficiency. By concatenating, instead of summing, we provide subsequent layers with a rich, diverse set of features. This enhances the network's representational power and tends to improve model performance, particularly on complex tasks. Moreover, it leads to a reduction in the number of parameters compared to traditional CNNs, since each layer is thinner and only contributes a small number of feature maps.

72. How can we detect and prevent the vanishing or exploding gradients problem in deep neural networks?

To detect vanishing or exploding gradients, monitor the magnitude of gradients during backpropagation. If they are too small or too large, that's a sign of trouble. To prevent these issues, we typically use better weight initialization methods like Xavier or He initialization, employ batch normalization, use appropriate activation functions like ReLU, and potentially apply gradient clipping to cap the gradients during training.

73. Discuss how L1 and L2 regularization terms affect the distribution of weights in a neural network model.

L1 regularization, also known as Lasso regularization, tends to push the weights towards zero, creating a sparse solution where some weights can become exactly zero. This is useful for feature selection in high-dimensional datasets. On the other hand, L2 regularization, also known as Ridge regularization, encourages the weights to be small but not necessarily zero, leading to a more diffuse, small weight distribution. It helps in preventing overfitting by penalizing the magnitude of the weights without promoting sparsity.

74. How would you design a CNN to handle input images of varying sizes?

To design a CNN for handling input images of varying sizes, one approach is to incorporate global average pooling layers towards the end of the network. This allows the network to aggregate feature information efficiently, resulting in a fixed-length output regardless of the input image's dimensions, which is particularly useful when you're dealing with images of different resolutions.

75. What do you know about Dropout?

Dropout is a regularization approach that helps to avoid overfitting and improve the generalizability of the dataset. During the training, randomly selected neurons are ignored for each pass or update of the model; this means that during each iteration, a random subset of neurons is excluded and the model is trained on the remaining neurons.

76. What is the Vanishing Gradient Problem in Artificial Neural Networks?

The vanishing gradient problem is part of an artificial neural network with a gradient-based learning method. In this method, each of the neural networks receives the weights and updates them proportional to the partial derivative of the error function concerning the current weight in each iteration.

77. What exactly do you mean by Exploding and Vanishing Gradients?

Exploding Gradient: Exploding gradient is a problem that occurs during the training of deep neural networks, which leads to the gradients of the network losing weight. Vanishing Gradient: Vanishing gradient is a problem that occurs when gradients used to update the network become very small as they are back propagated from the output layer to the earlier layers.

78. What is the difference between Batch Normalization, Instance Normalization, and Layer Normalization?

Batch Normalization: In batch normalization, the mean and variance are calculated for each channel across all samples and their relative dimensions, i.e., the height of each activation map (H) and the width of each activation map(W).

Instance Normalization: In Instance normalization, the mean and variance are calculated for each channel for each sample across both the height of each activation map (H) and the width of each activation map (W).

Layer Normalization: In layer normalization, the mean and the variance are calculated for each sample across all channels and their relative dimensions i.e., the height of each activation map (H) and the width of each activation map (W).

79. What's the difference between GAN and Autoencoders?

GAN: Generative Adversarial Networks (GAN) is used as an adversarial feedback loop to learn how to generate some information that seems real.
Autoencoder: An autoencoder is used to learn some input information with high efficiency and, subsequently, how to reconstruct the input from its compressed form.

80. What's the difference between Recurrent Neural Networks and Recursive Neural Networks?

Recurrent Neural Network: It is used for sequential inputs where the time factor is the main differentiating factor between the elements of the sequence. Due to this, it's commonly used in time series, and the weights are shared with the length of the sequence.
Recursive Neural Network: It is more like a hierarchical network where there is no time aspect to the input sequence, but the input has to be processed hierarchically in a tree fashion, and the weights are shared at every node.

81. What is the importance of using the Non-linear Activation Function?

Neural networks with only linear activation do not gain from increasing the number of layers in them since all linear functions add up to a single linear function.

Non-linear activation functions allow us to stack different layers, and they will not be treated like a single layer as in the linear activation layer.

The derivation of a linear function has no relation to the input, so it is not possible to use backpropagation when it comes to linear functions. Non-linear functions allow backpropagation because they can be differentiated, and their derivative is related to the input.

Learn about [Linear Discriminant Analysis](#)

Deep Learning Engineer Salary on the Basis of Skills

Job Role	Average Salary in India	Average Salary in the USA
Deep Learning Engineer – Experience (0 – 9) years	[Minimum – 6 LPA](#)	[Minimum – 86,993 USD](#)
	Average – 10 LPA	Average – 157,361 USD
	Highest – 18 LPA	Highest – 284,649 USD

Deep Learning Job Trends in 2024

According to the Bureau of Labor Statistics US, the employment of Deep Learning Engineers will be projected to grow by [35%](#) from 2022 to 2032.

1. Global Demand: With more than [50,000 open jobs on LinkedIn](#) in the United States and more than [6,000 open jobs on LinkedIn](#) in India.
2. Growth Projections: The growth suggested by the [Bureau of Labor Statistics is 35%](#) in the field of Deep learning, which might surpass all other occupation fields' growth of 8%.

Job Opportunities in Deep Learning

Job Role	Description
Deep Learning Engineer/Scientist	The primary role is to design, implement, and optimize complex deep learning models.
Machine Learning Engineer	Developing and implementing machine learning models for data-driven solutions.
Data Scientist	Analyzing complex data, extracting insights, and informing strategic decision-making processes.
Computer Vision Engineer	Creating algorithms for image and video analysis using computer vision.
Natural Language Processing (NLP) Engineer	Developing NLP algorithms for text analysis, sentiment, and language processing.
Research Scientist (Deep Learning)	Advancing deep learning through research, publications, and academic contributions.
AI Consultant	Providing strategic guidance and solutions for implementing artificial intelligence projects.

Roles and Responsibilities in Deep Learning

According to the Job posted on Naukri.com by [Digiai Solutions](#)

Role: Machine Learning Engineer – Python/Deep Learning

1. **Responsibilities**
    1. Design and implement machine learning solutions to solve complex problems.
    2. Collaborate with cross-functional teams to understand project requirements and deliver scalable ML models.

3. Evaluate and choose appropriate models, algorithms, and tools for different tasks.
4. Optimize and fine-tune existing models for enhanced performance.

2. **Skill Required:**
   1. Strong programming skills in languages such as Python and proficiency in ML libraries.
   2. Experience with data preprocessing, feature engineering, and model evaluation.
   3. Solid understanding of deep learning, neural networks, and other ML algorithms.

Conclusion

I hope this set of Deep Learning Engineer Interview Questions will help you prepare for your interviews. Best of luck in your endeavors!

If you are looking to embark on a deep learning journey that will uplift your career in AI and Data Science, check out Intellipaat's [Deep Learning course](#), or enroll in Intellipaat's [Executive Post Graduate certification in AI and ML](#) for an enriching learning experience and career growth.

Got any questions regarding deep learning? Post your query in the [Intellipaat Community](#) space, and we will get back to you.

Course Schedule

Name	Date	Details
[Machine Learning Course](#)	**17 Aug 2024** (Sat-Sun) Weekend Batch	[View Details](#)
[Machine Learning Course](#)	**24 Aug 2024** (Sat-Sun) Weekend Batch	[View Details](#)
[Machine Learning Course](#)	**31 Aug 2024** (Sat-Sun) Weekend Batch	[View Details](#)

About the Author
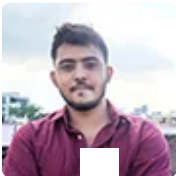
[Arya](#)
Senior Research Analyst

As a Senior Research Analyst, Arya Karn brings expertise in crafting compelling technical content in Data Science and Machine Learning. With extensive knowledge in AI/ML, NLP, DBMS, and Generative AI, his works get lakhs of views across social platforms that benefit both technical and business spheres.

Find Data Science Course Training in Other Regions

Australia | Melbourne | Delhi | Dublin | Hong Kong | Kolkata | Mumbai | Chennai | Dallas | Noida |

Pune | Singapore | Sydney | Bangalore | Chicago | Hyderabad | San Francisco | London | New York |

Toronto | India | Dubai | Houston | Jersey | Los Angeles | San Jose | Jaipur | Gurgaon | Indore |

Ahmedabad | Coimbatore | Kochi | Chandigarh | Bhubaneswar | United States | Germany | Texas |

Phoenix | Seattle | Atlanta | Washington | Kansas City | Boston | United Kingdom | Perth | Brisbane |

Canberra | Trichy | Nagpur | Vizag | Trivandrum | Mountain View | Ashburn | Charlotte | Austin |

San Diego | Irving | Columbus | Sunnyvale | Fremont | Denver | Philadelphia | Vijayawada | Thane |

Ireland | Jamshedpur | Kerala | Lucknow | Tamil Nadu | Vadodara | Ludhiana | Surat | Abu Dhabi |

Ernakulam | Kottayam