1. **. What is Deep Learning?**  
   ➔ Deep Learning is a part of Machine Learning that uses multiple layers of neural networks to learn from data.
2. **What is a neural network?**  
   ➔ A neural network is a collection of neurons (nodes) connected through weighted edges, inspired by the human brain.
3. **What is the difference between AI, ML, and DL?**  
   ➔ AI is the broad field, ML is a subset of AI, and DL is a subset of ML using neural networks.
4. **What is a neuron in deep learning?**  
   ➔ A neuron receives input, processes it (with weights and bias), and produces an output.
5. **What is weight in a neural network?**  
   ➔ Weights determine the importance of inputs to a neuron.
6. **What is bias in a neural network?**  
   ➔ Bias allows the model to shift the activation function curve to better fit the data.
7. **What happens in forward propagation?**  
   ➔ Data passes through the layers, and output is calculated.
8. **What happens in backpropagation?**  
   ➔ Errors are calculated and propagated backward to update weights.
9. **What is an optimizer?**  
   ➔ Optimizer adjusts weights to reduce the loss function.
10. **Name a few optimizers.**  
    ➔ SGD, Adam, RMSprop, Adagrad.
11. **What is SGD (Stochastic Gradient Descent)?**  
    ➔ It updates weights using a small random batch of data each time.
12. **What is Adam optimizer?**  
    ➔ Adam is an advanced optimizer that combines momentum and adaptive learning rates.
13. **What is loss function?**  
    ➔ It calculates the error between predicted and actual outputs.
14. **Name some loss functions.**  
    ➔ MSE (Mean Squared Error), Cross-Entropy Loss, Hinge Loss.
15. **What is the purpose of activation functions?**  
    ➔ They introduce non-linearity into the network.
16. **Why is ReLU preferred over sigmoid?**  
    ➔ ReLU avoids the vanishing gradient problem common in sigmoid.
17. **What is vanishing gradient problem?**  
    ➔ Gradients become too small, stopping the network from learning.
18. **What is exploding gradient problem?**  
    ➔ Gradients become too large, making the network unstable.
19. **What is dropout?**  
    ➔ Dropout randomly disables neurons during training to prevent overfitting.
20. **What is batch normalization?**  
    ➔ It normalizes layer inputs to stabilize and speed up training.
21. **What is a fully connected layer?**  
    ➔ Every neuron is connected to every neuron in the next layer.
22. **What is feature extraction?**  
    ➔ It is the process of finding useful information from raw input data.
23. **What is a CNN?**  
    ➔ CNN (Convolutional Neural Network) is a type of deep network for image data.
24. **What is the role of filters in CNN?**  
    ➔ Filters extract features like edges and patterns from images.
25. **What is pooling in CNN?**  
    ➔ Pooling reduces the size of feature maps while retaining important information.
26. **Types of pooling?**  
    ➔ Max pooling and Average pooling.
27. **What is a kernel (filter) in CNN?**  
    ➔ A small matrix that moves across input data to perform convolution.
28. **What is stride in CNN?**  
    ➔ Stride is the number of pixels by which the filter moves during convolution.
29. **What is padding in CNN?**  
    ➔ Padding adds extra pixels around the input to control the output size.
30. **What is an RNN?**  
    ➔ RNN (Recurrent Neural Network) handles sequential data by having loops in neurons.
31. **What is the problem with basic RNNs?**  
    ➔ They suffer from vanishing gradients during long sequence training.
32. **What are LSTM and GRU?**  
    ➔ LSTM and GRU are special types of RNNs that solve the vanishing gradient problem.
33. **What is a sequence-to-sequence model?**  
    ➔ A model that maps input sequences to output sequences, like language translation.
34. **What is transfer learning?**  
    ➔ Using a pre-trained model and fine-tuning it for a new task.
35. **What is fine-tuning in deep learning?**  
    ➔ Slightly training a pre-trained model on new data to adapt it to a new task.
36. **What is pre-training?**  
    ➔ Training a model on a large dataset before using it on a specific task.
37. **What is an autoencoder?**  
    ➔ A neural network that learns to compress and reconstruct input data.
38. **Applications of autoencoders?**  
    ➔ Data compression, denoising, anomaly detection.
39. **What is GAN (Generative Adversarial Network)?**  
    ➔ GANs have two networks (generator and discriminator) that compete to create realistic data.
40. **What is the role of generator in GAN?**  
    ➔ The generator creates fake data to fool the discriminator.
41. **What is reinforcement learning?**  
    ➔ Learning by interacting with the environment and receiving rewards or punishments.
42. **What is supervised learning?**  
    ➔ Learning from labeled data.
43. **What is unsupervised learning?**  
    ➔ Learning patterns from data without labels.
44. **What is semi-supervised learning?**  
    ➔ Using a small amount of labeled data and a large amount of unlabeled data.
45. **What is one-hot encoding?**  
    ➔ Representing categorical variables as binary vectors.
46. **What is the purpose of softmax?**  
    ➔ It turns raw outputs into probability distributions
47. **What is a perceptron?**  
    ➔ A simple model of a single neuron that performs binary classification using weights, bias, and activation.
48. **What is an activation function?**  
    ➔ It introduces non-linearity into the network so it can learn complex relationships.
49. **Name some activation functions.**  
    ➔ Sigmoid, ReLU, Tanh, and Softmax.
50. **What is forward propagation?**  
    ➔ It is the process where inputs pass through the network to produce an output.
51. **What is backpropagation?**  
    ➔ It is the method used to update model weights by calculating error gradients backward through the network.
52. **What is an epoch?**  
    ➔ One full pass of the entire training dataset through the model.
53. **What is batch size?**  
    ➔ The number of training samples processed before the model’s parameters are updated.
54. **What is learning rate?**  
    ➔ It controls how much the weights are adjusted with respect to the loss gradient.
55. **What is overfitting?**  
    ➔ When the model performs well on training data but poorly on unseen data because it memorized instead of learning.
56. **How can we avoid overfitting?**  
    ➔ By using techniques like regularization, dropout, adding more data, or early stopping.
57. **What is dropout?**  
    ➔ A regularization technique where random neurons are temporarily removed during training to prevent overfitting.
58. **What is convolution in CNNs?**  
    ➔ Convolution is applying a small filter over an image to detect features like edges and patterns.
59. **What is pooling in CNNs?**  
    ➔ Pooling reduces the size of feature maps by taking maximum or average values, helping in reducing computation.
60. **What is the difference between CNN and RNN?**  
    ➔ CNNs are used for spatial data like images, RNNs are used for sequential data like text or time series.
61. **What is ReLU activation?**  
    ➔ ReLU (Rectified Linear Unit) sets all negative inputs to zero and keeps positive values as they are.
62. **What is gradient descent?**  
    ➔ It's an optimization algorithm that updates weights to minimize the loss function.
63. **What is a loss function?**  
    ➔ It measures how far the model's predictions are from the actual values.
64. **What is the role of optimizer?**  
    ➔ The optimizer updates the weights of the network to reduce the loss during training.
65. **What is softmax function?**  
    ➔ Softmax converts raw outputs into probabilities that sum to one, used for multi-class classification.
66. **Why do we normalize data before training?**  
    ➔ Normalization scales data to a standard range, helping the model to converge faster.