Neural Networks Viva Questions with Answers (100)

1. 1. What is a neural network?

→ A neural network is a computing system inspired by the biological brain, consisting of interconnected units (neurons) that process information using dynamic state responses to inputs.

2. 2. What are the basic components of a neural network?

→ Input layer, hidden layers, output layer, weights, biases, and activation functions.

3. 3. What is a neuron?

→ A neuron is the basic unit of a neural network that receives input, applies a function, and passes the output to the next layer.

4. 4. What is the role of weights in neural networks?

→ Weights determine the strength of the connection between neurons and influence how input data is transformed.

5. 5. What is an activation function?

→ An activation function introduces non-linearity into the network, allowing it to learn complex patterns.

6. 6. Name some common activation functions.

→ ReLU, Sigmoid, Tanh, Softmax.

7. 7. What is forward propagation?

→ The process of passing input data through the network layers to get the output prediction.

8. 8. What is backpropagation?

→ An algorithm for training neural networks by calculating the gradient of the loss function and updating the weights.

9. 9. What is gradient descent?

→ An optimization algorithm used to minimize the loss function by updating weights in the opposite direction of the gradient.

10. 10. What is the loss function?

→ A function that measures the error between predicted and actual output values.

11. 11. What is the role of the hidden layer?

→ To learn complex patterns by applying transformations to the input data.

12. 12. Why do we use activation functions?

→ To introduce non-linearity into the model, enabling the network to solve complex tasks.

13. 13. What is ReLU?

→ ReLU (Rectified Linear Unit) is an activation function defined as f(x) = max(0, x).

14. 14. What is the vanishing gradient problem?

→ A problem where gradients become too small, slowing or stopping learning in deep networks.

15. 15. How does ReLU help with vanishing gradient?

→ ReLU avoids vanishing gradients by allowing gradients to flow for positive inputs.

16. 16. What is softmax?

→ An activation function used in the output layer for multi-class classification that converts logits into probabilities.

17. 17. What is one-hot encoding?

→ A method of representing categorical variables as binary vectors.

18. 18. What is a model's accuracy?

→ The percentage of correct predictions out of all predictions made by the model.

19. 19. What is training error and testing error?

→ Training error is the error on the training data; testing error is on unseen data.

20. 20. What is early stopping?

→ A technique to stop training when the model performance on validation data stops improving.

21. 21. What is transfer learning?

→ Using a pre-trained model on a new task to leverage learned features.

22. 22. What is a validation set?

→ A subset of data used to tune hyperparameters and evaluate model performance during training.

23. 23. What is bias-variance tradeoff?

→ A tradeoff between a model's ability to generalize (bias) and sensitivity to data (variance).

24. 24. What is a confusion matrix?

→ A table used to evaluate classification models by showing true vs. predicted values.

25. 25. What is precision and recall?

→ Precision is the ratio of true positives to predicted positives; recall is the ratio of true positives to actual positives.

26. 26. What is F1 score?

→ The harmonic mean of precision and recall.

27. 27. What is a learning curve?

→ A plot that shows model performance over time or number of epochs.

28. 28. What is a neural network optimizer?

→ An algorithm that adjusts the weights to minimize the loss function.

29. 29. What is the Adam optimizer?

→ An optimization algorithm that combines momentum and RMSProp techniques.

30. 30. What is momentum in optimization?

→ A technique to accelerate gradient descent by considering past gradients.

31. 31. What is a bias node in a neural network?

→ A node that provides a constant value to help shift the activation function.

32. 32. Why are weights initialized randomly?

→ To break symmetry and allow neurons to learn different features.

33. 33. What happens if all weights are initialized to zero?

→ All neurons learn the same features and the network fails to train properly.

34. 34. What is Xavier initialization?

→ A weight initialization method that keeps the variance of outputs across layers balanced.

35. 35. What is batch normalization?

→ A technique to normalize inputs of each layer to improve training speed and stability.

36. 36. What is a receptive field in CNN?

→ The region of the input image that affects a particular output neuron.

37. 37. What is a pooling layer in CNN?

→ A layer that reduces spatial dimensions using operations like max pooling.

38. 38. What is stride in CNN?

→ The step size with which the filter moves across the input image.

39. 39. What is padding in CNN?

→ Adding extra pixels to the input image to control output size.

40. 40. What is a fully connected layer?

→ A layer where each neuron is connected to all neurons in the previous layer.

41. 41. What is the main use of CNNs?

→ Image processing and computer vision tasks.

42. 42. What is the main use of RNNs?

→ Sequence modeling, such as language modeling and time series prediction.

43. 43. What is an LSTM?

→ A type of RNN that can remember long-term dependencies.

44. 44. Why do RNNs face vanishing gradient problems?

→ Due to repeated multiplication of small gradients over time steps.

45. 45. What is BPTT (Backpropagation Through Time)?

→ An extension of backpropagation used in training RNNs.

46. 46. What are autoencoders?

→ Neural networks used to learn compressed representations of data.

47. 47. What is a bottleneck in autoencoders?

→ The central compressed representation of input data.

48. 48. What is a generative model?

→ A model that can generate new data similar to the training data.

49. 49. What is a GAN (Generative Adversarial Network)?

→ A generative model made of two networks: generator and discriminator.

50. 50. What are some applications of neural networks?

→ Image classification, speech recognition, self-driving cars, etc.

51. 51. What is model generalization?

→ It is the model's ability to perform well on new, unseen data.

52. 52. What is the difference between classification and regression?

→ Classification predicts categories, regression predicts continuous values.

53. 53. What is multi-class classification?

→ A classification task where each input belongs to one of three or more classes.

54. 54. What is multi-label classification?

→ A classification task where each input can belong to multiple classes simultaneously.

55. 55. What is the difference between AI, ML, and neural networks?

→ AI is a broad field, ML is a subset of AI, and neural networks are a method within ML.

56. 56. What are the types of neural networks?

→ Feedforward NN, CNN, RNN, LSTM, GANs, Autoencoders, etc.

57. 57. What is the difference between shallow and deep learning?

→ Shallow learning uses 1–2 layers, deep learning uses multiple layers to extract features.

58. 58. What is hyperparameter tuning?

→ The process of selecting the best model parameters like learning rate, batch size, etc.

59. 59. What is a cost function?

→ A function that measures how far off a model's predictions are from actual outcomes.

60. 60. What is the difference between cost and loss function?

→ Loss is for one sample; cost is the average of losses over the dataset.

61. 61. What is learning rate decay?

→ Reducing the learning rate during training to fine-tune the model.

62. 62. What are epochs, iterations, and batch size?

→ Epoch = one full pass of dataset, iteration = one update, batch size = number of samples per iteration.

63. 63. What is model convergence?

→ When the model stops improving significantly during training.

64. 64. What is model evaluation?

→ Assessing a model's performance using metrics like accuracy, precision, etc.

65. 65. Why do we use dropout?

→ To prevent overfitting by randomly dropping units during training.

66. 66. What is fine-tuning in neural networks?

→ Taking a pre-trained model and adjusting it slightly on a new dataset.

67. 67. What is a deep belief network (DBN)?

→ A generative neural network with multiple layers of hidden units.

68. 68. What is the vanishing/exploding gradient problem?

→ Gradients become too small or too large, hindering learning in deep networks.

69. 69. What is the difference between supervised and unsupervised learning?

→ Supervised uses labeled data, unsupervised does not.

70. 70. What is model interpretability?

→ How easily a human can understand why a model made a certain prediction.

71. 71. What is the difference between AI and Deep Learning?

→ AI is the broad field; deep learning is a specific subset that uses neural networks.

72. 72. What is a transformer model?

→ A deep learning model using self-attention, used in NLP tasks.

73. 73. What is attention mechanism?

→ A way to focus on relevant parts of input while processing sequences.

74. 74. What is self-attention?

→ A mechanism where each input element relates to all others, used in transformers.

75. 75. What are the steps in training a neural network?

→ Data preprocessing, model design, training, evaluation, and testing.

76. 76. What is model capacity?

→ The ability of a model to fit a wide variety of functions.

77. 77. What is a dead neuron?

→ A neuron that never activates across any inputs.

78. 78. What is the dying ReLU problem?

→ When ReLU neurons output zero for all inputs and stop learning.

79. 79. What is a skip connection?

→ A shortcut path in neural networks that skips one or more layers.

80. 80. What is residual learning?

→ Learning the difference between input and output, used in ResNets.

81. 81. What is ensemble learning?

→ Combining multiple models to improve overall performance.

82. 82. What is bagging and boosting?

→ Bagging reduces variance; boosting reduces bias by combining weak learners.

83. 83. What is dropout rate?

→ The proportion of nodes dropped during training to prevent overfitting.

84. 84. What is the difference between ReLU and Leaky ReLU?

→ Leaky ReLU allows a small gradient when input is negative.

85. 85. What is the purpose of using multiple hidden layers?

→ To learn more complex patterns and representations.

86. 86. What is feature extraction in neural networks?

→ Identifying important patterns or representations in input data.

87. 87. What is transfer function in a neuron?

→ The function used to convert input signals into output (activation function).

88. 88. What is online learning?

→ Model updates continuously as new data arrives.

89. 89. What is a pre-trained model?

→ A model trained on a large dataset that can be fine-tuned for another task.

90. 90. Give some real-world applications of neural networks.

→ Face recognition, medical diagnosis, fraud detection, etc.

91. 91. What is data augmentation?

→ Creating new training samples by modifying existing ones.

92. 92. What is a confusion matrix used for?

→ Evaluating classification performance by showing TP, FP, TN, FN.

93. 93. What is overfitting?

→ When a model performs well on training data but poorly on unseen data.

94. 94. What is underfitting?

→ When a model cannot capture patterns even in training data.

95. 95. What is regularization?

→ A technique to reduce overfitting by adding a penalty to the loss function.

96. 96. What is L1 and L2 regularization?

→ L1 adds absolute values of weights, L2 adds squared values to the loss.

97. 97. What is a perceptron?

→ The simplest type of neural network with a single layer of output.

98. 98. What is mini-batch gradient descent?

→ Gradient descent performed on small random subsets of data.

99. 99. What is the sigmoid function?

→ An activation function that outputs values between 0 and 1.

100. 100. What is the tanh function?

→ An activation function that outputs values between -1 and 1.