

DA-5 Deep Learning

1. What is the main role of the activation function in a neural network?

- ☐ (A) To initialize weights
- ☐ (B) To introduce non-linearity
- ☐ (C) To reduce gradient vanishing
- ☐ (D) To normalize input features

2. In CNNs, what do filters (kernels) primarily help with?

- ☐ (A) Feature extraction like edges and patterns
- ☐ (B) Reducing training data size
- ☐ (C) Performing backpropagation faster
- ☐ (D) Increasing fully connected layers

3. The vanishing gradient problem is most commonly associated with which activation function?

- ☐ (A) ReLU
- ☐ (B) Tanh
- ☐ (C) Sigmoid
- ☐ (D) Softmax

4. Suppose you are analyzing stock market data, where each day's value depends on past values. Which model would best capture this temporal dependency?

- ☐ (A) CNN
- ☐ (B) RNN / LSTM
- ☐ (C) Autoencoder
- ☐ (D) GAN

5. You train a neural network and observe that training loss keeps decreasing, but validation loss starts increasing after a point. What should you try first?

- ☐ (A) Increase learning rate
- ☐ (B) Add dropout or regularization
- ☐ (C) Increase number of epochs
- ☐ (D) Remove activation functions

6. Which application typically uses Generative Adversarial Networks (GANs)?

- ☐ (A) Object detection in autonomous vehicles
- ☐ (B) Fake image and video generation (deepfakes)
- ☐ (C) Predicting stock prices
- ☐ (D) Machine translation

7. Which of the following is not a deep learning framework?

- ☐ (A) TensorFlow
- ☐ (B) PyTorch
- ☐ (C) Scikit-learn
- ☐ (D) Keras

8. The basic computational unit of a neural network is called a:

- ☐ (A) Filter
- ☐ (B) Neuron
- ☐ (C) Weight
- ☐ (D) Gradient

9. Which type of NN is most effective for sequential data like speech and text?

- ☐ (A) CNN
- ☐ (B) RNN
- ☐ (C) Autoencoder
- ☐ (D) GAN

10. In a fully connected NN, the connection between neurons is represented by:

- ☐ (A) Activation function
- ☐ (B) Bias
- ☐ (C) Weights
- ☐ (D) Gradient descent

11. What is the main role of the Softmax function in the output layer?

- ☐ (A) Normalizes inputs
- ☐ (B) Introduces non-linearity
- ☐ (C) Converts logits to probabilities
- ☐ (D) Prevents vanishing gradient

12. Dropout in neural networks is used to:

- ☐ (A) Speed up backpropagation
- ☐ (B) Reduce overfitting
- ☐ (C) Normalize input data
- ☐ (D) Increase training accuracy

13. In Keras, which function is used to define a sequential model?

- ☐ (A) `keras.Sequential()`
- ☐ (B) `keras.Model()`
- ☐ (C) `keras.Network()`
- ☐ (D) `keras.Layer()`

14. In Keras, `model.compile()` requires:

- ☐ (A) Only optimizer
- ☐ (B) Optimizer, loss, and metrics
- ☐ (C) Only metrics
- ☐ (D) Dataset path

15. In Keras, which method is used to start training the model?

- ☐ (A) `model.train()`
- ☐ (B) `model.fit()`
- ☐ (C) `model.run()`
- ☐ (D) `model.start()`

16. In a feed-forward neural network, information flows:

- ☐ (A) From input to output, possibly looping back
- ☐ (B) From input to output in one direction only
- ☐ (C) From output to input repeatedly
- ☐ (D) Randomly between layers

17. If you want to stop training a Keras model once validation accuracy stops improving, which callback is most suitable?

- ☐ (A) TensorBoard
- ☐ (B) EarlyStopping
- ☐ (C) ReduceLROnPlateau
- ☐ (D) ModelCheckpoint

18. Backpropagation is primarily used for:

- ☐ (A) Initializing weights
- ☐ (B) Calculating and propagating errors backward to update weights
- ☐ (C) Converting outputs into probabilities
- ☐ (D) Reducing the dataset size

19. Backpropagation uses which rule to compute gradients?

- ☐ (A) Bayes' theorem
- ☐ (B) Chain rule of calculus
- ☐ (C) Central limit theorem
- ☐ (D) Linear regression formula

20. The purpose of gradient descent in training neural networks is:

- ☐ (A) To increase loss
- ☐ (B) To minimize the loss function
- ☐ (C) To maximize weights
- ☐ (D) To normalize features

21. The perceptron is best described as:

- ☐ (A) A multi-layer network with backpropagation
- ☐ (B) A single-layer binary linear classifier
- ☐ (C) A clustering algorithm
- ☐ (D) A type of CNN

22. A single perceptron cannot solve which type of problem?

- ☐ (A) AND gate
- ☐ (B) OR gate
- ☐ (C) XOR gate
- ☐ (D) NAND gate

23. In a perceptron, the output is obtained by applying:

- ☐ (A) Activation function (usually step function)
- ☐ (B) Softmax function
- ☐ (C) Gradient descent
- ☐ (D) Chain rule

24. The perceptron learning algorithm converges if:

- ☐ (A) Data is non-linear
- ☐ (B) Data is linearly separable
- ☐ (C) Learning rate is zero
- ☐ (D) Epochs are infinite

25. A shallow neural network typically has:

- ☐ (A) Only an input and output layer
- ☐ (B) One hidden layer
- ☐ (C) More than 5 hidden layers
- ☐ (D) No activation functions

26. A deep neural network usually refers to a network with:

- ☐ (A) Exactly 2 layers
- ☐ (B) At least 3 layers (including input/output)
- ☐ (C) More than one hidden layer
- ☐ (D) No constraints on number of layers

27. One major challenge in training very deep neural networks is:

- ☐ (A) Lack of enough neurons
- ☐ (B) Vanishing or exploding gradients
- ☐ (C) Too few layers to model complexity
- ☐ (D) Inability to use dropout

28. Which Keras layer is used for fully connected layers?

- ☐ (A) Dense
- ☐ (B) Conv2D
- ☐ (C) Dropout
- ☐ (D) Flatten

29. In CNNs, which Keras layer performs feature extraction using filters?

- ☐ (A) Dense
- ☐ (B) Conv2D
- ☐ (C) MaxPooling2D
- ☐ (D) Embedding

30. The Keras layer used to convert multi-dimensional feature maps into a 1D vector before feeding into dense layers is:

- ☐ (A) Dropout
- ☐ (B) Flatten
- ☐ (C) Embedding
- ☐ (D) Conv1D

31. To prevent overfitting, which Keras layer randomly drops some neurons during training?

- ☐ (A) Dense
- ☐ (B) Conv2D
- ☐ (C) Dropout
- ☐ (D) ReLU

32. In Keras, the input shape is usually specified in the

- ☐ (A) First hidden layer
- ☐ (B) Output layer
- ☐ (C) Optimizer
- ☐ (D) Loss function

33. You are classifying handwritten digits (MNIST). The input is a 28×28 image. Which Keras layer should you add before the Dense layer?

- ☐ (A) Dropout
- ☐ (B) Flatten
- ☐ (C) Embedding
- ☐ (D) Conv2D

34. You are comparing two regression models. Model A has RMSE = 3, Model B has RMSE = 5. Which is better?

- ☐ (A) Model A
- ☐ (B) Model B
- ☐ (C) Both are equally good
- ☐ (D) Cannot say without accuracy

35. In Keras, a callback is used to

- ☐ (A) Define the model architecture
- ☐ (B) Monitor or modify the training process
- ☐ (C) Preprocess the dataset
- ☐ (D) Initialize weights

36. EarlyStopping callback is used to

- ☐ (A) Stop training after fixed epochs
- ☐ (B) Stop training if validation metric stops improving
- ☐ (C) Reduce learning rate
- ☐ (D) Save the model

37. TensorBoard is used to

- ☐ (A) Train models faster
- ☐ (B) Visualize training metrics, losses, and graphs
- ☐ (C) Initialize neural network layers
- ☐ (D) Preprocess data

38. Logs for TensorBoard are saved in:

- ☐ (A) Optimizer object
- ☐ (B) Directory specified in log_dir
- ☐ (C) GPU memory
- ☐ (D) Model weights only

39. What is the primary purpose of the Keras Flatten layer?

- ☐ (A) Reduce overfitting
- ☐ (B) Convert multi-dimensional input into 1D vector
- ☐ (C) Apply activation functions
- ☐ (D) Reduce learning rate

40. If the input to Flatten is (batch_size, 28, 28, 3), the output shape is:

- ☐ (A) (batch_size, 28, 28)
- ☐ (B) (batch_size, 28*28*3)
- ☐ (C) (batch_size, 3)
- ☐ (D) (28, 28, 3)