Status	Finished
Started	Monday, 16 June 2025, 4:59 PM
Completed	Monday, 16 June 2025, 5:06 PM
Duration	6 mins 49 secs
Grade	**4.75** out of 5.00 (**95**%)

Question 1

Correct

Mark 0.25 out of 0.25

Which type of neural network architecture is primarily used for image recognition tasks?

A. Long Short-Term Memory Networks (LSTMs)

B. Recurrent Neural Networks (RNNs)

C. Convolutional Neural Networks (CNNs)

D. Multilayer Perceptrons (MLPs)

Your answer is correct.

The correct answer is: Convolutional Neural Networks (CNNs)

Question 2

Correct

Mark 0.25 out of 0.25

What does the width of a Deep Neural Network refer to?

A. The number of layers in the network

B. The number of parameters in the model

C. The number of input features

D. The number of neurons in each hidden layer ✓

Your answer is correct.

The correct answer is: The number of neurons in each hidden layer

Question 3

Incorrect

Mark 0.00 out of 0.25

The complexity of the forward pass in a deep neural network is typically considered to be:

A. O(N), where N is the number of neurons ✗

B. O(1)

C. O(L), where L is the number of layers

D. O(M), where M is the number of training examples

Your answer is incorrect.

The correct answer is: O(L), where L is the number of layers

Question 4

Correct

Mark 0.25 out of 0.25

Which activation function is commonly used in the output layer of an MLP for multiclass classification?

A. Tanh

B. ReLU

C. Sigmoid

D. Softmax ✓

Your answer is correct.

The correct answer is: Softmax

Question 5

Correct

Mark 0.25 out of 0.25

During the backpropagation process, the error is propagated from the output layer to the input layer using:

A. The chain rule of calculus ✓

B. The sigmoid activation function

C. A gradient descent algorithm

D. Random initialization of weights

Your answer is correct.

The correct answer is: The chain rule of calculus

Question 6

Correct

Mark 0.25 out of 0.25

If the target label for a binary classification problem is 0 or 1, the loss function typically used to compare the predicted output with the true label is:

A. Mean Squared Error

B. Cross-Entropy Loss ✓

C. L2 Loss

D. Hinge Loss

Your answer is correct.

The correct answer is: Cross-Entropy Loss

Question 7

Correct

Mark 0.25 out of 0.25

What makes MLPs universal approximators?

A. Their capacity to minimize error without a loss function

B. Their use of multiple layers and non-linear activation functions ✓

C. Their ability to handle linearly separable data efficiently

D. Their reliance on Boolean gates for decision-making

Your answer is correct.

The correct answer is: Their use of multiple layers and non-linear activation functions

Question 8

Correct

Mark 0.25 out of 0.25

If the learning rate in gradient descent is too large, what might happen?

- A. The weight updates may overshoot, causing the model to diverge and fail to minimize the loss. ✓
- B. There will be no impact on the training process.
- C. The model will converge faster to the minimum of the loss function.
- D. The weight updates will be too small, slowing down the learning process.

Your answer is correct.

The correct answer is: The weight updates may overshoot, causing the model to diverge and fail to minimize the loss.

Question 9

Correct

Mark 0.25 out of 0.25

In multiclass classification, the Cross-Entropy Loss compares:

- A. The Euclidean distances between feature vectors
- B. The true label and the predicted label directly
- C. The actual probabilities of two different models
- D. The one-hot encoded true label vector and the predicted probability vector ✓

Your answer is correct.

The correct answer is: The one-hot encoded true label vector and the predicted probability vector

Question 10

Correct

Mark 0.25 out of 0.25

Which deep learning architecture is frequently used for sequence prediction tasks, such as language modeling or machine translation?

A. Convolutional Neural Networks (CNNs)

B. Recurrent Neural Networks (RNNs) ✓

C. Generative Adversarial Networks (GANs)

D. Fully connected neural networks

Your answer is correct.

The correct answer is: Recurrent Neural Networks (RNNs)

Question 11

Correct

Mark 0.25 out of 0.25

What is the output of the XOR function for inputs (1, 0) and (0, 0)?

A. 0, 0

B. 1, 0 ✓

C. 1, 1

D. 0, 1

Your answer is correct.

The correct answer is: 1, 0

Question 12

Correct

Mark 0.25 out of 0.25

Which of the following is a common application of deep learning in image recognition?

- A. Object detection ✓
- B. Machine translation
- C. Time series forecasting
- D. Sentiment analysis

Your answer is correct.

The correct answer is: Object detection

Question 13

Correct

Mark 0.25 out of 0.25

Cross-entropy loss is commonly used for:

- A. Linear regression
- B. Multiclass classification problems ✓
- C. Unsupervised learning
- D. Dimensionality reduction

Your answer is correct.

The correct answer is: Multiclass classification problems

Question 14

Correct

Mark 0.25 out of 0.25

In a feedforward neural network, increasing the width of the network (i.e., adding more neurons to hidden layers) will typically result in:

A. A lower risk of overfitting

B. A faster training time

C. An increase in model capacity to learn complex patterns

D. No change in the models performance

Your answer is correct.

The correct answer is: An increase in model capacity to learn complex patterns

Question 15

Correct

Mark 0.25 out of 0.25

What does the forward pass in a neural network primarily consist of?

A. Calculating the loss function

B. Matrix multiplications followed by activation functions

C. Backpropagation of gradients

D. Weight updates

Your answer is correct.

The correct answer is: Matrix multiplications followed by activation functions

Question 16

Correct

Mark 0.25 out of 0.25

In gradient descent, the weight vector is updated in the direction of:

- A. Maximum ascent along the error surface
- B. Maximum descent along the error surface ✓
- C. Perpendicular to the error surface
- D. Any direction that decreases the error

Your answer is correct.

The correct answer is: Maximum descent along the error surface

Question 17

Correct

Mark 0.25 out of 0.25

Which dataset is an example of non-linearly separable data that the perceptron cannot classify correctly?

- A. OR gate
- B. NOR gate
- C. XOR gate ✓
- D. AND gate

Your answer is correct.

The correct answer is: XOR gate

Question 18

Correct

Mark 0.25 out of 0.25

If you double the number of neurons in a hidden layer of a DNN, the network will have:

A. Faster convergence during training

B. A lower risk of overfitting

C. More parameters to optimize ✓

D. A simpler optimization process

Your answer is correct.

The correct answer is: More parameters to optimize

Question 19

Correct

Mark 0.25 out of 0.25

What is the range of output values produced by the Softmax activation function?

A. [0, 1] and sums to 1 ✓

B. $(-\infty, \infty)$

C. (0, 1)

D. (−1, 1)

Your answer is correct.

The correct answer is: [0, 1] and sums to 1

Question 20

Correct

Mark 0.25 out of 0.25

Which of the following is the output of the computation graph for binary classification when using a sigmoid activation function?

A. A single scalar representing the predicted class (0 or 1)

B. A vector of weights

C. A vector of features

D. A scalar value between 0 and 1 representing the probability of belonging to the positive class

Your answer is correct.

The correct answer is: A scalar value between 0 and 1 representing the probability of belonging to the positive class

← Previous
Group Self Selection

Next →
Assignment 1