TopicQuiz5

Started: 6 May at 21:09

Quiz instructions

Question 1	1 pts

Despite theoretical concerns about ANNs getting stuck in local optima during training, why has this limitation been found to be benign in practice, especially for larger neural networks?

○ Training algorithms have become more efficient at escaping local optima

○ Local optima often perform almost as well as the global optimum

○ The global optimum is not important for model performance

○ Larger neural networks have fewer local optima

Question 2	2 pts

Given the perceptron's threshold logic unit (TLU) architecture, what would be the output of a TLU with the following input weights, input values, and bias: w = [0.5, -0.2], x = [2, 3], and b = 0.1?

○ -0.1

○ 0

○ 1

○ 0.6

Question 3	1 pts

In the context of neural networks, what does the term "fully connected layer" or "dense layer" mean?

○ A layer that is solely responsible for calculating the loss function

○ A layer where each neuron is connected to a single input

○ A layer in which every input is connected to every output

○ A layer that contains only a single threshold logic unit (TLU)

Question 4   1 pts

Given a neural network with an input matrix X of shape (m, n) and a weight matrix W of shape (n, p), what are the dimensions of X and W, respectively?

○ n instances with m features for X; p inputs with n neurons for W

○ m instances with n features for X; n inputs with p neurons for W

○ m instances with p features for X; n inputs with m neurons for W

○ p instances with n features for X; m inputs with n neurons for W

Question 5   1 pts

In the context of neural networks and data science, what does "broadcasting" refer to when adding a vector to a matrix?

○ Dividing the matrix by the vector

○ Adding the vector to every column in the matrix

○ Multiplying the vector by each element in the matrix

○ Adding the vector to every row in the matrix

Question 6

1 pts

What is the main principle behind Hebbian learning in the context of neural networks?

○ Neurons that fire independently, lose their connections

○ Neurons that fire together, wire together

○ Neurons that fire together, lose their connections

○ Neurons that fire independently, wire together

Question 7

1 pts

Given the perceptron learning rule, calculate the new weight $w_{i,j}$ after a single update for the following scenario: $w_{i,j} = 0.5$, $x_i = 2$, $y_j = 1$, $\hat{y}_j = 0.8$, and $\eta = 0.1$.

○ 0.6

○ 0.54

○ 0.56

○ 0.48

Question 8

1 pts

What does the Perceptron Convergence Theorem state about the learning process of a perceptron?

○ The perceptron learning process converges only for certain activation functions

○ The perceptron learning process is guaranteed to converge if the data is linearly separable

○ The perceptron learning process always converges, regardless of the data distribution

○ The perceptron learning process never converges for non-linearly separable data

Question 9   1 pts

What kind of problems can a multilayer perceptron (MLP) solve that a simple perceptron cannot?

○ Problems with non-linearly separable data

○ Problems with linearly separable data only

○ Problems that involve unsupervised learning

○ Problems that require reinforcement learning

Question 10   1 pts

Which of the following is a reason why perceptrons may not generalize as well as logistic regression or linear SVM classifiers?

○ Logistic regression and linear SVM classifiers cannot handle non-linearly separable data

○ Perceptrons stop training as soon as there are no more prediction errors on the training set

○ Perceptrons require more computational resources

○ Perceptrons use regularization by default

Question 11   1 pts

In the context of neural networks, what does the term "gradients of the neural network's error" refer to?

○ The change in error with respect to the number of layers in the network

○ The change in error with respect to the change in model parameters, such as weights and biases

○ The change in error with respect to the choice of activation functions

○ The change in error with respect to the change in input values

Question 12   1 pts

Why is it important to initialize the weights of a neural network's hidden layers randomly?

○ Random initialization breaks symmetry, allowing backpropagation to train a diverse team of neurons

○ Random initialization ensures faster convergence

○ Random initialization ensures all neurons in a layer are identical

○ Random initialization reduces the need for regularization techniques

Question 13   1 pts

Why does the backpropagation algorithm not work well with step activation functions in neural networks?

○ Step activation functions have no gradient, making it impossible to update weights using gradient-based methods

○ Step activation functions require more computational resources, making backpropagation inefficient

○ Step activation functions have a very large gradient, causing instability during training

○ Step activation functions have a constant gradient, leading to slow convergence

Question 14

1 pts

why are activation functions necessary in neural networks?

○ To simplify the network architecture by removing unnecessary connections

○ To introduce non-linearity between layers, enabling the network to learn complex patterns and effectively utilize multiple layers

○ To reduce the need for regularization techniques

○ To ensure faster convergence during training

Question 15

1 pts

What is the main difference between the Softplus and ReLU activation functions?

○ Softplus is a smooth approximation of ReLU

○ Softplus is linear, while ReLU is non-linear

○ ReLU is linear, while Softplus is non-linear

○ ReLU is a smooth approximation of Softplus

Question 16

2 pts

Given a training set with 405 images and a default batch size of 19, how many batches would the model go through per epoch?

○ 21 batches of size 19 and 1 batch of size 16

○ 20 batches of size 19 and 1 batch of size 5

○ 21 batches of size 19 and 1 batch of size 6

○ 20 batches of size 20 and 1 batch of size 5

Question 17   1 pts

What is transfer learning in the context of deep neural networks?

○ Initializing the weights and biases of a new network with random values

○ Reusing the lower layers of a pre-trained network and fine-tuning the higher layers for a new, related task

○ Training a network from scratch for each new task

○ Reusing the entire pre-trained network for a new, unrelated task

Question 18   1 pts

What is the main concern when a layer in a neural network has too few neurons?

○ The layer will not have enough representational power to preserve all the useful information from the inputs, potentially leading to unrecoverable information loss

○ The layer will require more training time to converge

○ The layer will be more susceptible to vanishing gradients

○ The layer will have an increased risk of overfitting the training data

Question 19   1 pts

Given the following activation function equation, identify its name:

f(x) = 1 / (1 + exp(-x))

○ Linear function

○ Sigmoid function

○ Hyperbolic Tangent (tanh)

○ Rectified Linear Unit (ReLU)

Question 20

1 pts

Which activation function is commonly used in the output layer of a neural network for binary classification problems?

○ Tanh

○ ReLU

○ Sigmoid

○ Softmax

Not saved   Submit quiz