

# Reading



This Quiz Will Be Submitted In Thirty Minutes



Started: Feb 1 at 4:10pm

## Quiz Instructions



The following questions are based on the assigned reading (Alexnet paper: [Readings](https://canvas.asu.edu/courses/212341/discussion_topics/6267399)) ([https://canvas.asu.edu/courses/212341/discussion\\_topics/6267399](https://canvas.asu.edu/courses/212341/discussion_topics/6267399))



Question 1 1 pts

Alexnet (the network described in the assigned reading/paper) has seven convolutional layers.



True



False



Question 2 1 pts

The authors used a technique called \_\_\_\_\_ to prevent overfitting by randomly setting neuron outputs to zero during training.



Question 3 1 pts

The model was trained using \_\_\_\_\_ descent with a batch size of 128 examples.



Question 4 1 pts

What is the main advantage of using ReLU activation over traditional activation functions like sigmoid or tanh?



It provides probabilistic output



It allows the network to be trained without dropout

○ It speeds up training

 This Quiz Will Be Submitted In Thirty Minutes



Question 5 1 pts

What methods of data augmentation are used by the authors?

(Two are correct)

☐

Adding extra images to the dataset

☐

Adding Gaussian noise to the images

☐

Generating image translations and horizontal reflections

☐

Altering the intensities of the RGB channels in training images



The following questions are based on the video recording published on Friday: [Rest of the ML/DL review lecture \(https://canvas.asu.edu/courses/212341/discussion\\_topics/6266679\)](https://canvas.asu.edu/courses/212341/discussion_topics/6266679)



Question 6 1 pts

Which layer is used as the last layer in classification networks to produce a probability distribution (all entries in the output vector sum to 1).

☐

Batch Normalization

☐

Softmax

☐

Pooling



Question 7 1 pts

In a CNN, fully connected (FC) or linear layers are typically added after the convolutional layers. If the output of the last convolution layer has height = 32, width = 32 and number of channels = 64, then what is the number of inputs going into the FC layer?

☐

65536

☐

1024

☐

64



This Quiz Will Be Submitted In Thirty Minutes

Question 8 1 pts

What are the common names for the gradient calculations during back propagation?

(Two are correct)



fprop



dgrad



weight update



wgrad



Question 9 1 pts

Which component of the Transformer network is responsible for capturing the relationships between words in a sentence?



Positional encoding



Feed forward layers



Attention layers



Add &amp; Norm layers



Question 10 1 pts

What is the original application for which the Transformer neural network was proposed?



Computer vision



Linear regression



Natural language processing



Recommendations



This Quiz Will Be Submitted In Thirty Minutes

Submit Quiz