1. Given the syntax below, select the option that will best improve a CNN model that you are trying to fit.

model.fit(x\_train, y\_train, batch\_size=batch\_size, epochs=100, validation\_data=(x\_test, y\_test))

* 1. Add shuffling, by adding “, shuffle=True” at the end.

1. Which of the following statements is **TRUE** about a kernel in a Convolutional Layer applied to an image?
   1. Kernels detect local features in an image such as lines, corners, and edges.
2. The main idea of transfer learning of a neural network is:
   1. To keep the early layers of a pre-trained network and re-train the later layers for a specific application.
3. In the context of transfer learning, which is a guiding principle of fine tuning?
   1. Using data that is similar to the pre-trained network
4. In the context of transfer learning, what do we call the process in which you only train the last or a few layers instead of all layers of a neural network?
   1. Frozen layers
5. This concept came as a solution to CNNs in which each layer is turned into branches of convolutions:
   1. Inception
6. Which CNN Architecture is considered the flash point for modern Deep Learning?
   1. AlexNet
7. Which CNN Architecture can be described as a "simplified, deeper LeNet" in which the more layers, the better?
   1. VGG
8. Which CNN Architecture is the precursor of using convolutions to obtain better features and was first used to solve the MNIST data set?
   1. LeNet
9. The motivation behind this CNN Architecture was to solve the inability of deep neural networks to fit or overfit the training data better when adding layers.
   1. ResNet
10. This CNN Architecture keeps passing both the initial unchanged information and the transformed information to the next layer.
    1. ResNet
11. This concept came as a solution to CNNs in which each layer is turned into branches of convolutions:
    1. Inception
12. What is the main function of backpropagation when training a Neural Network?
    1. Make adjustments to the weights
13. (True/False) The “vanishing gradient” problem can be solved using a different activation function.
    1. True
14. (True/False) Every node in a neural network has an activation function.
    1. True
15. These are all activation functions except:
    1. Leaky hyperbolic tangent
16. Deep Learning uses deep Neural Networks for all these uses, except:
    1. Cases in which explainability is the main objective
17. These are all activation functions for CNN, except:
    1. Pruning
18. (True/False) Optimizer approaches for Deep Learning Regularization use gradient descent:
    1. False
19. Stochastic gradient descent is this type of batching method:
    1. online learning
20. The main purpose of data shuffling during the training of a Neural Network is to aid convergence and use the data in a different order each epoch.
    1. True
21. Which of the following IS NOT a benefit of Transfer Learning?
    1. Improving the speed at which large models can be trained from scratch
22. Which of the following statements about using a Pooling Layer is TRUE?
    1. Pooling can reduce both computational complexity and overfitting.