Keras and Spark - Quiz

Warm-Up:

1. Describe one potential advantage and disadvantage of adding an additional layer to a convolutional neural network.
2. In Keras, what is the name of the fully-connected layer? What are its parameters?
3. What parameters does the Conv2D layer take? Which are optional and which are required?
4. How does Spark provide a clean interface to interact, process, and clean data before using a machine learning model?
5. Give examples of potential datasets where MapReduce can be used as a data processing tool in Spark.
6. True or False: It is possible to never have all the data present in a Numpy array or Pandas DataFrame during training.
7. True or False: The most common activation function used in convolutional neural networks is the rectified linear unit.
8. True or False: Spark only provides data processing tools such as the inbuilt Dataset and various map/reduce tools.
9. True or False: Spark is often more efficient than Pandas given a larger dataset with more parallelizable tasks.

Coding Question:

10. Fill in the blanks to properly evaluate the Keras model given that we wish to center the data outside of the model. Assume that we trained the model on centered data and that we have a loss function named `loss`.

```
def evaluate_mean_cent(model_loc, X_val, y_val):
      model = tf.keras.models.load_model(model_loc)
      loss_val = 0
      ____________________________________________________________
      ____________________________________________________________
      ____________________________________________________________
      ____________________________________________________________
      ____________________________________________________________
      ____________________________________________________________

      return loss_val
```