Please provide complete and well-written solutions to the following exercises.

Due November 21, 4PM PST, to be uploaded as a single PDF or Jupyter Notebook document to brightspace.

Homework 8

Exercise 1. Use a convolutional neural network to classify the CIFAR 10 and CIFAR 100 datasets found here:

https://www.cs.toronto.edu/~kriz/cifar.html

https://keras.io/api/datasets/cifar10/

https://keras.io/api/datasets/cifar100/

What error rate can you obtain? For CIFAR-10, I would consider an error rate below 20% to be quite good. Does ResNet50 offer better performance?

Do the same for the Fashion MNIST dataset found here:

https://keras.io/api/datasets/fashion mnist/

Is your performance with Fashion MNIST similar to what we did in class with MNIST?

(Optional: make heatmaps for these tasks, as we did in class.)

In all cases, use a test set whose size is 10% of the size of the whole dataset.

Exercise 2. Fill in any missing details of the 6.7 billion parameter Llama-3 definition, by examining the code at

https://github.com/meta-llama/llama3/blob/main/llama/model.py.

(You don't have describe every single detail, just try to fill in more details than we did in class.)

Exercise 3. Describe in roughly a page what you did on your project this week.