STAT 6340 (Statistical and Machine Learning, Fall 2024)

Bonus Project

Instructions:

- Due date: Dec 8, 2024 (Use Submission Link on eLearning to upload by 11:59 pm)
- Total points = 30
- This project is optional. No partial credit will be given.
- Submit a typed report.
- It is OK to consult with others regarding this project or use any publicly available resources, but you must write code and answers on your own. Violation of this policy will be considered academic dishonesty, and you will referred to appropriate university authorities. If you don't do the work by yourself, how will you learn?
- Do a good job.
- You must use the following template for your report:

  Mini Project #
  Name
  Section 1. Answers to the specific questions asked
  Section 2: Code. Your code must be annotated. No points may be given if a brief look at the code does not tell us what it is doing.
- Section 1 of the report must be limited to *five* pages. Also, only those software outputs should be provided in this section that are referred to in the report.
- You can code in `R` or `Python`.
- You are encouraged to use `R Markdown` or `Jupyter Notebook` to write your report.

---

1. (15 points) Consider the `MNIST` dataset from the previous project. It contains a training set of 60,000 28×28 grayscale images of 10 handwritten digits (from 0 to 9), along with a test set of 10,000 images. We would like to build a convolutional neural network model to identify the digit on the image. This is a multiclass classification problem with 10 output classes. For fitting the model below, use ReLU activation for the hidden layers, softmax activation for the output layer, and minibatches of size 64.

   (a) Consider a convnet with the following architecture (this is also mentioned on a lecture notes slide):
      i. One convolution layer with thirty-two 3×3 filters, followed by max-pooling over 2×2 blocks.
      ii. One convolution layer with sixty-four 3×3 filters, followed by max-pooling over 2×2 blocks.
      iii. One convolution layer with sixty-four 3×3 filters. The shape of the output of this layer should be (3, 3, 64). Flatten this array before giving it as input for the next layer.
      iv. One dense layer of size 64.
      v. An output layer.

   (b) Specify the number of parameters in each layer. What is the total number of parameters?

   (c) Fit the model in (a) with 5 epochs and report its accuracy on the test data.

2. (15 points) Do Exercise #13 from Chapter 10. Textbook Python
   pg 474
   pg 462