San José State University
Computer Science Department
CS156, Introduction to Artificial Intelligence, Spring 2021

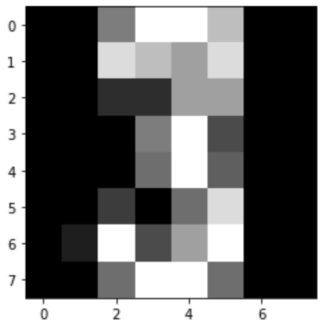
Homework #7

Objective:

This homework's objective is to implement several multi-layer perceptron models to classify handwritten digits (MNIST dataset) and compare these models' performance.

Details:

For this assignment you will be using the handwritten digits dataset MNIST. You will implement digit classification from image data using multi-layer perceptron (MLP) classifier. Remember that MNIST images provided by scikit learn libraries are 8x8 pixel images. For example, the image below shows one of the observations labeled "3" in this dataset:



Use *load_digits()* function in scikitlearn to load the data. Once the data is loaded, each image will have to be reshaped/flattened into a single array. Also, note that MNIST data will need to be normalized. Here's an example of an easy way to do so (here X contains the flattened image data):

X = X.astype("float32") / 255

Split the data into training and test sets with 80/20 proportion. Make sure to stratify your split:

train_test_split(X, Y, test_size=0.2, random_state=0, stratify=Y)

For this assignment you will define and train 7 different MLP models:

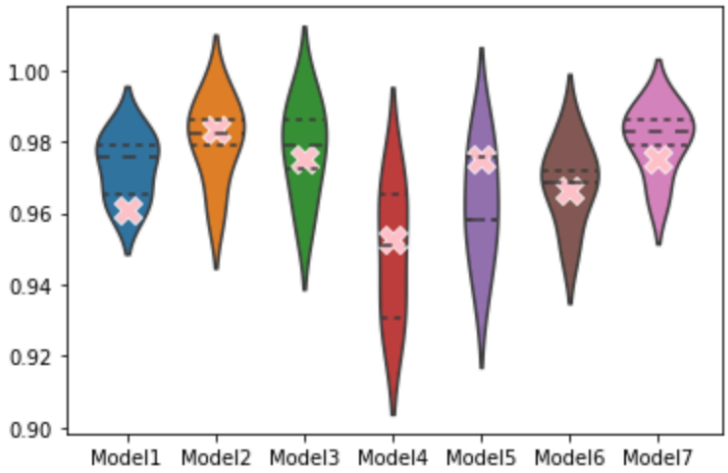
1. A model with default parameters for MLPClassifier other than r*andom_state=1* and *max_iter=max_iter*

   model1 = MLPClassifier(random_state=1, max_iter=max_iter)
2. Three hidden layer sizes: 400,150,50; with Relu activation function
3. Three hidden layer sizes: 400,150,50; with logistic activation function

4. Three hidden layer sizes: 64,32,8; with Relu activation function
5. Two hidden layer sizes: 32,16; with Relu activation function
6. Three hidden layer sizes: 120,64,16; with Relu activation function
7. Three hidden layer sizes: 320,120,32; with Relu activation function

For each of the seven models perform stratified 5-fold cross-validation and collect prediction accuracies for each fold. In addition, compute prediction accuracies on the held-out test set.

In a single plot show the results from both cross-validation and the test set. For each model, on the x-axis, plot results from 5-fold cross-validation predictions as a violin or a box plot (or a similar visualization technique). For each model also indicate the accuracies obtained on the test set. Clearly identify each model on the x-axis with a name descriptor. You can use any of the plotting libraries you prefer, for example *seaborn*. Your plot might look similar to the example below (here violin plots show cross-validation accuracies and the pink X marks indicate test set accuracies):



Submission:

Email your assignment submission to me at Yulia.Newton@sjsu.edu and the grader (Akshay Kajale) at akshay.kajale@sjsu.edu. Make sure to email this submission by 11:59pm on the due date listed in Canvas. Your sent email is the proof of submission. The subject of the email should say "CS156 Assignment 7". In the body of the email list your name as it appears on the class roster and your student ID. Attach to this email both the pdf of your Jupyter notebook, which contains the solution for this homework assignment, as well as the notebook itself (the notebook file with .ipynb extension). Make sure to submit both files, otherwise the submission will not be considered complete.

Grading:

I will return the grades as fast as we can grade this homework. Normally it should not take more than a few weeks.

A total of 10 points are possible for this homework assignment.