ECE5984 SP22 - Prof. Jones – HW5

Due Tuesday, April 19, 2022 – 11:59 PM via Canvas

In this assignment you are to investigate the effect of different neural network architectures on the performance of a classification problem. Specifically, which activation functions, and number and sizes of hidden layers, give the best performance for a given dataset.

Here is what you are to do.
1. Load the "ccpp.xlsx" dataset. Use the data on the sheet named "allBin". This data has an ID field called "ID" and a binary target called "TG".
2. Normalize the data appropriately (in scikit.preprocessing terminology, "scale" the data).
3. Divide the data into training and test sets. Use a specified random seed so the split is done the same way each time.
4. For a variety of architectures (see below):
   a. Train an MLPClassifier on the training data. Use early stopping; you may need to experiment a bit to find suitable learning rate and tolerance values.
   b. Measure the performance on the test set using two different measures: AUROC and misclassification rate.
5. Build two tables: the ten best <u>and</u> ten worst model architectures by AUROC, and the ten best and ten worst model architectures by misclassification rate. Each model should be identified by the hidden layer sizes and the activation function.
6. Identify the best model using each of the two measures of performance: do they agree or do they indicate different models?
7. Summarize your results and your findings.

Note that I want to see all combinations of number of hidden layers (from 1 to 3), numbers of nodes in hidden layers (from 1 to 10 – and <u>note</u>: do not assume that the layers have the same number of nodes), and all choices for internal activation functions (relu, logistic, identity and tanh).
You will go crazy unless you write code to automatically cycle through all possible combinations of these parameters and save or print out the parameters and the results. Be aware that this program may take a long time to run: mine took almost 40 minutes to try all possible combinations!

Your submission should be a Word or PDF document including all of your code (pasted in as unformatted text, not formatted on a dark background or as a screen shot!), your two "top-ten" tables, your choice(s) for the best model architecture, and your discussion of the results and your observations. Also attach your Python code file(s) as .py files (not as .ipnyb files).
Submit your work in the usual way via Canvas.