CSCI 315: Artificial Intelligence

Assignment 2: Feed-Forward Neural Network with Scikit-Learn

**Due Date: February 13th @ 11:59 PM**

Introduction

For this assignment you will be implementing a feed-forward neural network with 2 hidden layers. You will be training your FFNN to analyze the MNIST dataset. MNIST is a dataset of hand-written digits which you can predict using a SoftMax layer. Many of the scikit-learn API calls will be useful in terms of creating and running your network.

You will be responsible for creating this network without much prewritten code. You will need to figure out how to import, transform, and model your data.

Assignment

There are three major parts to this assignment. The first is the ability to handle the MNIST data points. You will need to find this dataset (available via Keras), but you will need to process these images so that your FFNN can handle them.

Feed forward neural networks can only deal with standardized images. That means you will need to ensure that the images are in a format that a neural network can work with. You will need to reshape and normalize these images.

The second part is that you will need to create your network. You will likely need to read up and refer to the neural network class in scikit-learn. You may want to visit this website (<https://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPClassifier.html>) which should give you all the necessary options to create a good network. We will be using the relu activation function with stochastic gradient descent as our optimization algorithm. You will also be using a learning rate of 0.01, define the max iterations to be 50, define the verbose parameter to be 10, define the random state to be 1 and define alpha to be 1e-5.

The final part of your assignment will be to evaluate your model. This will include reserving a part of your dataset purely for evaluation. For this assignment, you only need to worry about the measurement of accuracy.

You should also display a graph of the overall error of the model, this should be a general trend downward. There will undoubtably be components of your code that will be different, and not every line has been described in this assignment. Please feel free to use scikit-learns website to explore the options to create this network. Implementation details can vary from what is described here, but the final model should yield above 95% accuracy for your test data. From the metrics library of scikit learn, please generate a confusion matrix of the test data.

Turn-in

As with every assignment this semester, you need to submit an extensive readme that describes the functions you use and why you use them as well as citing all your sources. The README is about 40% of your grade so make sure you spend a good bit of time on it. Please zip up your assignment in a file named assignment2.zip. Within that zip file you should include your model source code, which you should name neuralnetwork.py, and your README file, which should be named analysis.txt, then submit the zip file to canvas.