

Problem Set 10

Data Science 602: Data Analysis and Machine Learning

Spring 2022

For this assignment, you will develop and train a sequential neural network in Keras to classify the MNIST digits dataset.

1. **Generate the model.** Develop a function

```
build_network(nslayers, n_neurons_per_layer, activation_fn)
```

The function should return a compiled model with the following structure:

- An Input node accepting an image of dimensions 28×28
- A Flatten node
- n hidden layer nodes, each containing `n_neurons_per_layer` neurons and using the activation function `activation_fn`.
- An output layer (Dense layer) of 10 neurons that uses the softmax activation function.

The model should be compiled as such:

- Optimizer: `sgd`
- metrics: `["accuracy"]`
- loss: `sparse_categorical_crossentropy` (because the target variable is represented as a single value, as opposed to being one-hot encoded)

2. **Load MNIST.** Load the keras MNIST dataset. (call `keras.datasets.mnist.load_data("mnist.npz")`, which returns `(X_train, y_train)`, `(X_test, y_test)`). Split the training dataset into a training and validation set.
3. **Train the model.** Call `build_network` with parameters of your choice (4-8 layers, 50-150 neurons per layer, and ReLU activation (`relu`) is a reasonable starting point). Train the model against the training dataset. To reduce training time, an early stopping callback is advised. Evaluate the model using the validation dataset. What is the prediction accuracy of the neural net?
4. **Optimize the model.** Use one of the hyperparameter optimization frameworks discussed in class, such as `scikit-optimize`, to find optimal values of the number of layers, activation function, and neurons per layer for this neural network. Use a budget of about 20 runs. Use the below tables as rough guideline for parameter space. Note that depending on your configuration and optimization framework, this optimization may require a few hours to run. What combination of parameters produces the highest accuracy, and what is that accuracy?

Parameter	Space
Activation function	<code>["relu", "sigmoid"]</code>
Number of layers	2-20 (integer)
Neurons per hidden layer	10-300 (integer)