Machine Learning and Computational Physics
Fall 2020

Assignment 1

Due: September 9^{th} 2020, 11:59:59 PM PDT

Expressive power of deep networks

In this assignment, you will design simple feedforward networks. The primary objective is to understand how the expressive power of the network varies with depth and width.

1. Write a function that takes as input:

   - `input_dim`: dimension of the input vector
   - `output_dim`: dimension of the output vector/prediction
   - `width`: width of each hidden layer (number of neurons)
   - `depth`: depth of the neural network (number of hidden layers + output layer)

   First layer is layer0

   The function should construct a network based on these input arguments. The weights and biases of each layer should be initialized using uniform distribution between (-1,1). The `tanh` activation function should be used in all layers. The constructed and initialized network should be returned by a function. The following TensorFlow tools might be useful

   - `keras.Sequential()`
   - `keras.layers.Dense()`
   - `keras.initializers.RandomUniform()`

2. Using the above function, construct different instances of the network by fixing `input_dim = 1`, `output_dim = 1` and varying `depth` $= 10, 15, 20, 25$ and `width = 10,15,20,25`.

   (a) For each configuration, calculate and print the total number of network parameters (weights and biases).

   (b) Create an array of 10,000 uniformly chosen points between -1 and 1. Using this as the input to the network, evaluate the network prediction for each of the above configurations. Note that there is no training involved here. Plot the outputs from each networks in a grid of 4 X 4 subplots (corresponding to 16 different configurations of depth and width), where the width varies along each column and the depth varies along each row.

   evaluate what kind of prediction

   Do we need to compile the model

   (c) Plot the FFT of each of the network output (produced in previous step) in another figure in similar 4 X 4 grid format (`numpy.fft.fft` might be useful for this) such that (i, j) subplot of this figure is a FFT of the network output shown in (i, j) subplot in previous figure.

   What is FFT

   (d) What trends do you observe as a function of depth and width?

   (e) How much do these trends vary if you re-run the script? What is the cause of this variation, if any?

(f) What is your conclusion about the expressivity of neural networks as a function of depth and width? Keep in mind the number of parameters used in each configuration.

Instructions:

- You need to submit your work as a single notebook saved as `A1_FirstName_LastName.ipnyb` (for example `A1_Tommy_Trojan.ipnyb`). You can create this notebook locally (on your computer using Jupyter notebook) or on cloud using Google Colab (which we recommend). If you are using Google Colab, then please make sure that you are signed in to your USC Google account before starting. This will make sharing your saved work little easier.
- At the very beginning of your notebook insert a text cell and write your name and USC id.
- For questions requiring descriptive output (such as question `2d, 2e, 2f` in this assignment) use individual text cell.
- Make sure that your entire notebook runs successfully on Google Colab before submitting it. It is your responsibility to ensure this.
- Once you finish the assignment save it and share it with `dhruvvpa@usc.edu`. (If you are using Google Colab, then the notebook will automatically be saved to your Google Drive. Once you locate it in your Google Drive, right click on it and share it with `dhruvvpa@usc.edu`). While sharing make sure that you enable "editor" option, so that we can run your notebook on our end while grading it.