Machine Learning and Computational Physics
Fall 2020

Assignment 3

Due: 10/07/2020, 11:59:59 PM PDT

Convolutional Neural Network (CNN)

In this assignment, you will train a deep convolutional network to classify images of hand-written digits. The primary objective of this assignment is to learn how to build and train a CNN for image processing tasks and visualize the learned filters.

1. For this assignment we will be using MNIST dataset, which is a dataset of 60,000 gray-scale images of hand-written digits of size 28×28.
2. Download these images and re-scale them to have values between -1 and 1. For downloading and splitting these into a training and testing set use the command: `datasets.mnist.load_data`
3. Plot 25 training images in a 5×5 grid.
4. Then construct a CNN based classifier with the following architecture (in all convolution and fully connected layers use `ReLU` activation except in the last fully-connected layer):
   - Convolution layer with 32 (3×3) filters and zero padding.
   - A 2×2 Max. pooling layer.
   - Convolution layer with 64 (3×3) filters with zero padding.
   - A 2×2 Max. pooling layer.
   - Convolution layer with 64 (3×3) filters with no padding.
   - Flatten the output of the previous layer and connect it to a fully connected layer with width = 64.
   - A final fully connected layer with width = number of classes = 10 with softmax activation.
5. Train this network with the following parameters:
   - learning rate = 0.01
   - optimizer = Adam
   - number of epochs = 5
   - loss function = sparse categorical cross-entropy.
   - batch size = 64
6. Create a plot of accuracy as a function of number of epoch for the training and test set.

7. For the trained network extract the learned filters from the first convolutional layer. (Hint: you might want to use combination of `model.layers` and `Layer.get_weights()`). For each of the 32 filters of the first layer, subtract their respective means and then plot all 32 filters of size (3×3). For each subplot show the colorbar, and choose its limits symetrically about zero. Note the point of creating these plots is to see if you can identify these filters (see the next task).

8. Identify at least two filters that approximate a first derivative and one filer that approximates a second derivative. Justify your assertion. Also say along which direction these derivatives are computed.

Instructions:

- You need to submit your work as a single notebook saved as `A3_FirstName_LastName.ipnyb` (for example `A3_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 **email address**.
- For Question 8, use 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.
- Submit you notebook only once. Resubmissions will not be considered.