Assignment 5 Solutions

December 17, 2020

1 CS 335: Lab problems

1.1 ReLU and Softmax

Autograded

1.2 Blocks of neural network

Autograded

1.3 Applications of neural nets

In case the autograder fails, try with the seeds mentioned by students to verify their reported stats.

1. **XOR dataset**: minimal topology is FullyConnectedLayer(2,4,'relu'), FullyConnectedLayer(4,2,'softmax'). For this topology reported, give marks according to the autograder. In case the student has a larger topology than this in the report while reporting accuracies for topologies, i.e. they have demonstrated that their topology is minimum, deduct 25% marks given by the autograder. In case such an effort is not made, deduct 50% of marks given by autograder. (2 marks)
2. **Circle dataset**: minimal topology is FullyConnectedLayer(2,2,'relu'), FullyConnectedLayer(2,2,'softmax'). For this topology reported, give marks according to the autograder. In case the student has a larger topology than this in the report while reporting accuracies for smaller topologies as well, i.e. they have demonstrated that their topology is minimum, deduct 25% marks given by the autograder. In case such an effort is not made, deduct 50% of marks given by autograder. (2 marks)
3. **MNIST**: Autograded (2 marks)
4. **CIFAR10**: Autograded (4 marks)

2 CS 337: Theory Problem

2.1 Task 1

Expect depiction of the Kernel/filter on the image and description in pictures/words of how (i) local connections help detect, vertical and horizontal edges, parts of cars such as wheels/body/chasis

etc.. (i) weight sharing helps detect the same vehicle/car or its parts in different locations (iii) equivariance ensures that moving an object followed by convolution is the same as convolution followed by moving the object and finally (iv) max pooling does down sampling so that large contiguous sections such as the body of the car can have a more compact representation. These are the broad outlines but the pictorial depiction may not be unique and TAs can give the benefit of doubt to the student. (5 marks)

2.2 Task 2

A naive solution is to change the output layer (like in RNN) to predict not just one class, but one class for each pixel. Thus, change the output layer to a 1024×1024 matrix of softmax units. Possible problem is that we are making prediction at every pixel - which can be expensive. Optional: A better alternative (not expected as answer) is to predict bounding boxes and restrict the prediction of objects at the level of bounding boxes. So smaller middle layer, upsamping, Bounding box etc is what RCNN and Faster RCNN do. Give marks for this or any reasonable answer. (4 marks)

2.3 Task 3

The solution to Task 2 will detect exactly one vehicle at every pixel. So a simple modification can be to replace the 1024×1024 matrix of softmax units with a 1024×1024 matrix, with each cell now being an N dimensional vector of sigmoids so that we can deal with multi-label classification. Extra information: This is technically also called a Mask. Thus, each convolutional neuron processes data only for its receptive field. Although fully connected feedforward neural networks can be used to learn features as well as classify data, it is not practical to apply this architecture to images. A very high number of neurons would be necessary, even in a shallow (opposite of deep) architecture, due to the very large input sizes associated with images, where each pixel is a relevant variable. For instance, a fully connected layer for a (small) image of size 100 x 100 has 10000 weights for each neuron in the second layer. The convolution operation brings a solution to this problem as it reduces the number of free parameters, allowing the network to be deeper with fewer parameters. For instance, regardless of image size, tiling regions of size 5 x 5, each with the same shared weights, requires only 25 learnable parameters. (3 marks)