COMP9444 Neural Networks and Deep Learning
Term 3, 2020

Exercises 5: Hidden Units and Convolution

This page was last updated: 10/26/2020 11:19:48

---

1. **Hidden Unit Dynamics**

   Consider a fully connected feedforward neural network with 6 inputs, 2 hidden units and 3 outputs, using tanh activation at the hidden units and sigmoid at the outputs. Suppose this network is trained on the following data, and that the training is successful.

Item	Inputs	Outputs
	123456	123
1.	100000	000
2.	010000	001
3.	001000	010
4.	000100	100
5.	000010	101
6.	000001	110

   Draw a diagram showing:
   a. for each input, a point in hidden unit space corresponding to that input, and
   b. for each output, a line dividing the hidden unit space into regions for which the value of that output is greater/less than one half.

2. **Softmax**

   Recall that the formula for Softmax is

   $\text{Prob}(i) = \exp(z_i) / \Sigma_j \exp(z_j)$

   Consider a classification task with three classes 1, 2, 3. Suppose a particular input is presented, producing outputs

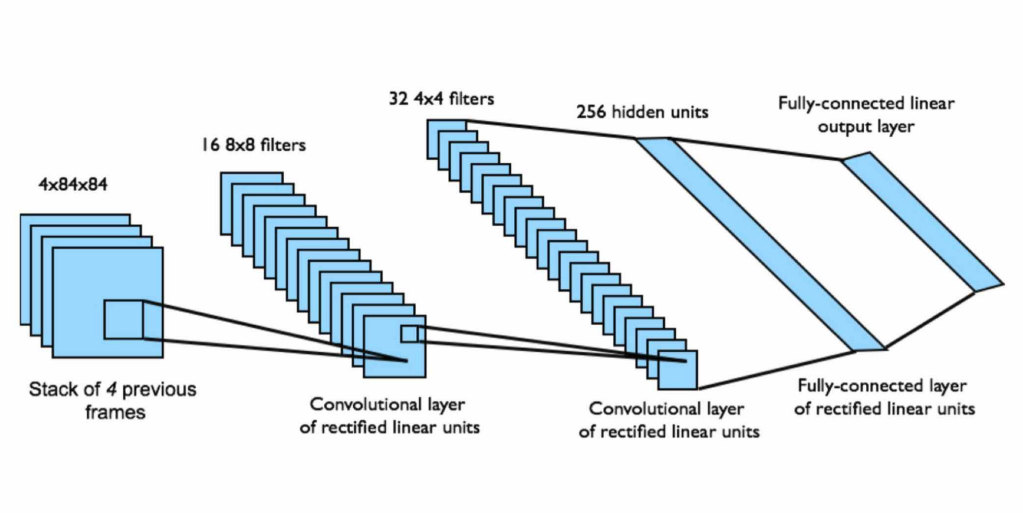
   $z_1=1.0, z_2=2.0, z_3=3.0$

   and that the correct class for this input is Class 2. Compute the following, to two decimal places:

   a. Prob(i), for i = 1, 2, 3
   b. $d(\log \text{Prob}(2))/dz_j$, for j = 1, 2, 3

3. **Convolutional Network Architecture**

   One of the early papers on Deep Q-Learning for Atari games (Mnih et al, 2013) contains this description of its Convolutional Neural Network:



> "The input to the neural network consists of an 84 × 84 × 4 image. The first hidden layer convolves 16 8 × 8 filters with stride 4 with the input image and applies a rectifier nonlinearity. The second hidden layer convolves 32 4 × 4 filters with stride 2, again followed by a rectifier nonlinearity. The final hidden layer is fully-connected and consists of 256 rectifier units. The output layer is a fully-connected linear layer with a single output for each valid action. The number of valid actions varied between 4 and 18 on the games we considered."

For each layer in this network, compute the number of

a. weights per neuron in this layer (including bias)
b. neurons in this layer
c. connections into the neurons in this layer
d. independent parameters in this layer

You should assume the input images are gray-scale, there is no padding, and there are 18 valid actions (outputs).

Make sure you attempt the questions yourself, before looking at the Sample Solutions.