Deep Learning

Assignment 2

Niranjan Ravi (3507597, M.Sc. Information Technology)

*University of Stuttgart*

1. Determine the size of all feature maps of the first model created in the

code skeleton, which is named MyModel.

**Ans: Output = 10**

2. Determine the ratio between the number of parameters used for convolutional and dense layers in MyModel.

Convolution Layer:

**Formula:**

Kl 2 (Dl -1) \* Dl = 900

**Dense Layer:**

There are three dense layers

Number of parameters for each layer is calculated as ML (ML-1 +1) ; 1<= L<=3

128(258 + 1) + 64(128 + 1) + 10( 10 + 1) = 41518

**Ratio: 450 / 20759**

3. Consider a weight matrix W of a dense layer and a random vector x~ N(\_; I) as its input, i.e. the activation is a = Wx. How does a dropout layer applied before the dense layer affect the expected activation ua = E[a]? What is a simple way of ensuring that the expected activation is not affected if such a dropout layer is applied?

When we train a single network with a large capacity, it often happens that the weights of some neurons are sensitive to the weights of other neurons. They are co-adapted. By dropout, the weights of some neurons become more independent to other weights. This makes the model more robust. If a hidden neuron has to work well in different combinations with other hidden neurons, it’s more likely that this hidden neuron does something individually useful.

Even when such a dropout layer is applied, the final output of DNN is combination of all inputs in a random manner and the activation is not affected by this. Hence Dropout is simple to implement and does not boost the computational complexity.

4. How could you quantify how similar two domains are?

In DNN all inputs are taken and processed, i.e each neuron is connected to all inputs. They are used to make global decisions. Whereas in CNN each neuron is connected to a particular receptive field (small part of input), this greatly helps in learning local features. CNN is widely used for feature learning; it is very efficient and also has low complexity.

In CNN, ideally the output neuron is mapped to all input neurons but in layer level it has only small receptive field, this brings complex input output relationship and hence better learning. (I.e a neuron in a deep layer can still be indirectly connected to all input neurons.

These two domains can be quantified based on the Number of Parameters (memory storage) and number of multiplications (computational complexity).