Worksheet: Neural Networks

1163150 - Ausgewählte Kapitel sozialer Webtechnologien

2. Mai 2022

1. So far, you should be able to describe the following terms. Give a short and crisp explanation of:

    (a) Backprop

    (b) Artifical Neuron

    (c) (Hidden) Layer

    (d) Activation function

    (e) Fully connected layer (dense layer)

    (f) Hypothesis (score function)

    (g) Cost function (loss function)

    (h) Optimzier, i.e., Gradient Descent

2. Given is three layer neual network. Layer L_1 has 500 neurons, Layer L_2 has 250 neurons and Layer L_3 is the output layer. The network classifies the CIFAR dataset. A 10-class problem consisting of images (32x32x3) showing common objects (https://www.cs.toronto.edu/ kriz/cifar.html).

    (a) Assuming you are using the *sigmoid* function as activation function. What does the equation that computes the hypothesis (scores) look like?

    (b) What are the dimension of each matrix/vector in that equation?

    (c) Draw the network as a graph organized in a layer-structure.

    (d) Calculate the number of learnable parametes in the neural network.

3. Following information are given about a neural network N_{SIMPLE}:

    Weight matrices W_{HIDDEN} and W_{OUTPUT}, where the first column corresponds to a bias value:

$$W_{HIDDEN} = \begin{pmatrix} 10 & -20 & 20 & -40 \\ 20 & -40 & 0 & 0 \end{pmatrix}$$

$$W_{OUTPUT} = \begin{pmatrix} 20 & 40 & -40 \end{pmatrix}$$

Activation function $g(z)$, which applies to all neurons in the network:

$$g(z) = \begin{cases} 0 & z \leq -10 \\ 1 & z \geq 10 \\ 0.5 & else \end{cases}$$

(a) Draw a graph of the network N_{SIMPLE} including all neurons and their connections. Note all weight and bias values on the corresponding nodes and edges of the graph.

(b) Use the given vectors x_1, x_2, x_3 to create a mini-batch matrix as input for the network and calculate its output. Only use matrix operations for the calculation and note all intermediate results.

$$\vec{x}_{(1)} = [0,1,1], \ \vec{x}_{(2)} = [1,1,0], \ \vec{x}_{(3)} = [1,0,1]$$

4. From the teaching materials you know that the *ReLU* (rectified linear unit) function is more advantageous at least in the context of image classification.

   (a) Recap the equation for the *ReLU* and *sigmoid* function and sketch their plots.

   (b) Discuss why the *ReLU* function is more beneficial compared with the sigmoid function.

   (c) Which layer of a neural network does not have an activation function?

5. What does it mean that a neural network is a /textituniversal function approximator? Discuss why that could be important?