Assignment 1 - Forward Propagation
Winter 2024

Atieh Armin
Student ID : 14658308

1 Theory

1. Acknowledge teach of the following by **typing** your name after each of the following statements:

   (a) While you can use online resources, you may not copy them. **Atieh Armin** (1pt)

   (b) You cannot place any photos in your report. **Atieh Armin** (1pt)

   (c) You cannot place any code snippets in your report. **Atieh Armin** (1pt)

   (d) You cannot use any ML frameworks in your code, accept for some purposes explicitly mentioned. **Atieh Armin** (1pt)

   (e) While you may work in Jupyter notebook, or the like, you **must**, export your code as a python file for submission (and check that it works). **Atieh Armin** (1pt)

2. Given a single input observation $x = \begin{bmatrix}1 & 2 & 3\end{bmatrix}$ and a fully connected layer with weights of $W = \begin{bmatrix}1 & 2\\ 3 & 4\\ 5 & 6\end{bmatrix}$ as biases $b = \begin{bmatrix}-1 & 2\end{bmatrix}$, what are the output of the fully connected layer given x as its input (5pts)?

   Solution:

   $$\text{y} = \text{xW} + \text{b} = \begin{bmatrix}1 & 2 & 3\end{bmatrix}\begin{bmatrix}1 & 2\\ 3 & 4\\ 5 & 6\end{bmatrix} + \begin{bmatrix}-1 & 2\end{bmatrix} = \begin{bmatrix}22 & 28\end{bmatrix} + \begin{bmatrix}-1 & 2\end{bmatrix} = \begin{bmatrix}21 & 30\end{bmatrix}$$

3. Given an input, $h = [10, -1]$, what would be the output(s) if this data was processed by the following activation functions/layers (10pts)?

   (a) Linear

   (b) ReLU

   (c) Sigmoid

   (d) Hyperbolic Tangent

   (e) Softmax

Solution:

(a) Linear : $g(z) = z \longrightarrow g([10, -1]) = [10, -1]$

(b) ReLU : $g(z) = max(0, z) \longrightarrow g([10, -1]) = [10, 0]$

(c) Sigmoid : $g(z) = \frac{1}{1+e^{-z}} \longrightarrow g([10, -1]) = [\frac{1}{1+e^{-10}}, \frac{1}{1+e}] \approx [0.99, 0.27]$

(d) Hyperbolic Tangent : $g(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}} \longrightarrow g([10, -1]) = [\frac{e^{10} - e^{-10}}{e^{10} + e^{-10}}, \frac{e^{-1} - e}{e^{-1} + e}] \approx [1, -0.76]$

(e) Softmax : $g(z) = \frac{e^z}{\sum_i e^{z_i}} \longrightarrow g([10, -1]) = [\frac{e^{10}}{e^{10} + e^{-1}}, \frac{e^{-1}}{e^{10} + e^{-1}}] \approx [1, 0]$

2 Start Implementing Your Layers

You can find the answers of this section in the codes.

3 Connecting Layers and Forward Propagate

Now let's assemble a simple network and forward propagate data through it! With real data!

Our architecture will be:

Input→FC (1 output)→Logistic Sigmoid

We'll use the dataset mentioned earlier in the assignment. Read in the dataset as your input data X and pass it through the architecture mentioned above. From an implementation standpoint, you'll likely want to create instances of your classes and organize them in some sort of an ordered structure such that the output of one layer is the input of the next. See the example code at the end of the *Building Blocks* slides.

In your report, just provide the output of the last layer pertaining to the **first observation**.

Answer: The output of the last layer pertaining to the first observation is [0.50000607].

Submission

For your submission, upload to Blackboard a single zip file containing:

1. PDF Writeup
2. Source Code
3. readme.txt file

The readme.txt file should contain information on how to run your code to reproduce results for each part of the assignment.

The PDF document should contain the following:

1. Part 1: Your solutions to the theory question. **DO NOT put screen shots or code in this. You should be typesetting your numeric results.**
2. Parts2: Nothing
3. Part 3: The output pertaining to the first observation from the final layer, when given the augemented medical cost dataset as its input.