MATH *007*A: Calculus for Life Sciences I

Networking with Neural Networks Bonus

The goal of the problem set below is to provide an initial glimpse into how basic neural network machinery operates. Different types of neural network algorithms are now ubiquitous in machine learning and, therefore, in many aspects of daily life. For example, even the operation and safety systems of electric vehicles rely heavily on such algorithms.

At a high level, a machine *learns* from examples. The user supplies input data together with the correct (outcome) labels. The learning algorithm then adjusts its internal parameters to minimize an *error function*, typically the sum of squared differences between its predicted outputs and the actual, correct outputs.

In the simple example below, we illustrate one basic training procedure. The task is to "teach the machine" to distinguish between photos containing an elephant and those containing a rabbit. For simplicity, we assume:

- each photo contains exactly one animal;

- the animal is either an elephant or a rabbit;

- all photos were taken from approximately the same distance.

The only input feature (parameter) we will use is the proportion of the photo occupied by the animal, measured on a scale from 0 to 10.

We will search for a linear function of the form

$$g(x) = 0.5x + b,$$

and use it as our classification model. Here the input value x is the proportion of the picture occupied by the animal. If the output satisfies $g(x) > 0$, we classify the image as an elephant; if $g(x) < 0$, we classify it as a rabbit. Our goal will be to determine the optimal value of the parameter b that minimizes the discrepancy between the model's outputs and the correct labels in our training set.

Below is the data collected from 11 observed photos. The first row lists the x-values (the proportion of the picture occupied by the image of the animal, on a scale from 0 to 10). The second row specifies which animal appears in each photo.

3	8	7	6	4	5	2	3	1	9	2

Table 1: Observed animals

For example, the first photo contains a rabbit whose image occupies $3/10 = 30\%$ of the picture, while the third photo shows an elephant occupying $7/10 = 70\%$ of the picture (see the figures on the next page).

Figure 1: The first photo contains a rabbit; its image occupies $3/10 = 30\%$ of the picture.

Figure 2: The third photo contains an elephant; its image occupies $7/10 = 70\%$ of the picture.

Definition. To distinguish between the two types of animals in the dataset, we introduce the function f(**photo**). It assigns the value 1 to an elephant and the value -1 to a rabbit. For instance,

$$\text{f}\left(\text{\includegraphics{elephant}}\right) = 1, \qquad \text{f}\left(\text{\includegraphics{rabbit}}\right) = -1.$$

Problems

1. (**1 pt**) Fill in the missing row of the table with the values of f (1 for elephants, -1 for rabbits) according to the definition above.

x-value	3	8	7	6	4	5	2	3	1	9	2
animal											
f-value											

Table 2: Observed values

Our goal is to find the value of the parameter b that minimizes the error function:

$$\mathcal{L}(b) = \sum_{i=1}^{n} (g(x_i) - f(\text{photo}_i))^2 = \sum_{i=1}^{n} (0.5x_i + b - f(\text{photo}_i))^2$$
$$= (0.5x_1 + b - f(\text{photo}_1))^2 + (0.5x_2 + b - f(\text{photo}_2))^2 + \ldots + (0.5x_n + b - f(\text{photo}_n))^2,$$

where n is the total number of photos.

2. (**1 pt**) Consider the first 5 photos.Using the table above, write an explicit (full) formula for $\mathcal{L}(b)$ and simplify:
$\mathcal{L}(b) = (0.5x_1 + b - f(\text{photo}_1))^2 + (0.5x_2 + b - f(\text{photo}_2))^2 + \ldots + (0.5x_5 + b - f(\text{photo}_5))^2 = (0.5 \cdot 3 + b - (-1))^2 + \ldots =$

3. (**1 pt**) Compute the derivative $\mathcal{L}'(b)$.

4. (**1 pt**) Find the critical value b_0, i.e. the value for which $\mathcal{L}'(b_0) = 0$.

5. (**1 pt**) Verify that $\mathcal{L}(b)$ attains a minimum at the value you found in 4.
   Hint. use the second derivative test.

6. (**2 pts**) Use the function $g(x) = 0.5x + b_0$, with the value of b_0 obtained in part 3, to determine whether the machine correctly identifies Bugs Bunny in the photo below as a rabbit:

Assume that the occupancy value of the animal in the image is $x = 3.17$.

7. (**3 pts**) Let's *teach* the machine a bit more. Extend the procedure in $1 - 6$ to all $n = 11$ photos in the database.
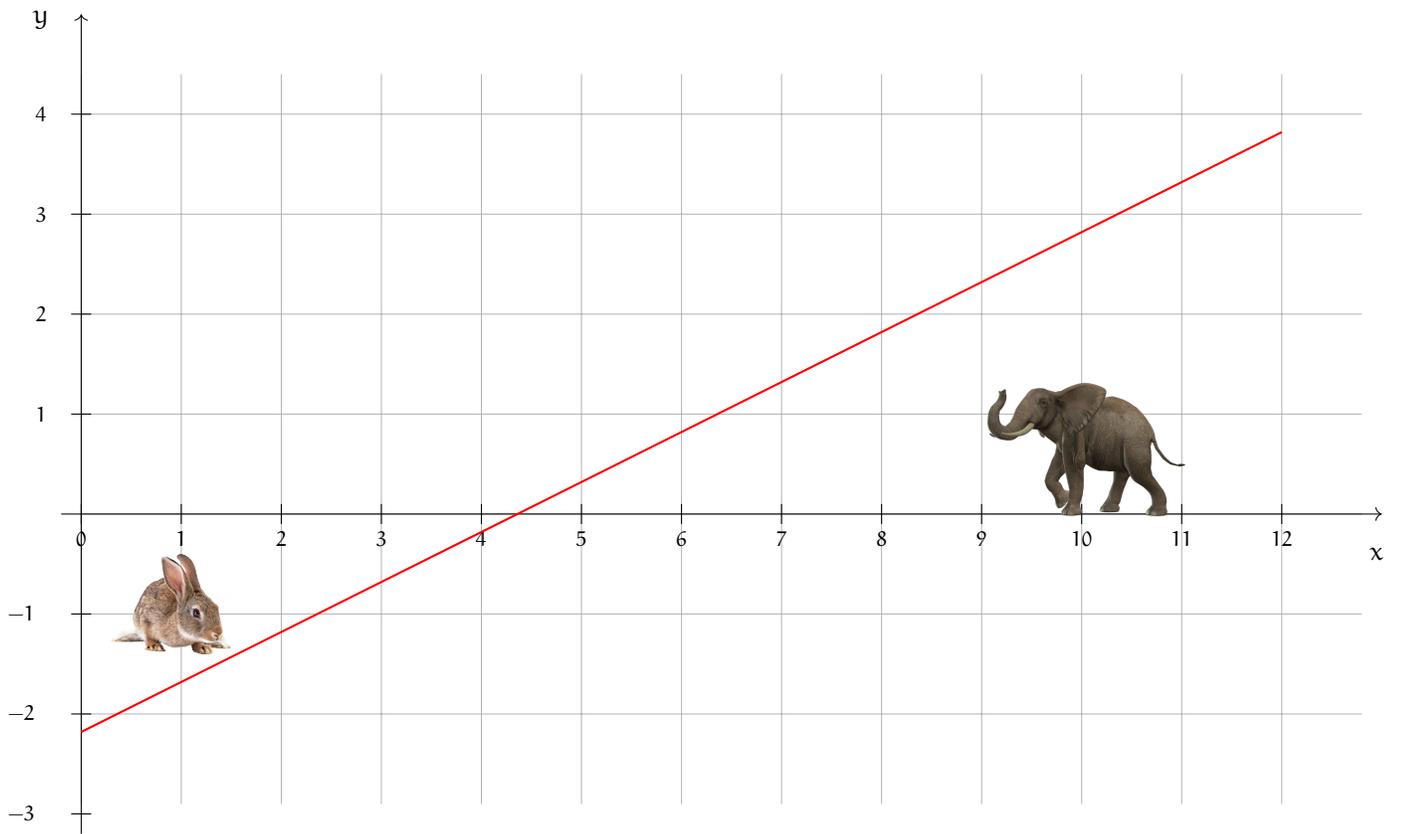
Figure 3: 🐰 Vs 🐘