HW4 - ML in HealthCare

Submission Guidelines: Your GitHub must contain at least the following files:

- Dry_part.pdf (that you create)- Containing your answers to the dry questions of the first part.
- ICA.ipynb - Implement your code according to the instructions in the notebook itself. Notice that your algorithm will be tested on the same unseen dataset.
- CNN_in_2D.ipynb - Implement your code according to the instructions within the notebook itself but before please follow the instructions below in order to work with GoogleColab.

In theoretical questions in all of the assignments' parts, it is better if you typed the calculations and\or your verbal explanations but you may also write them by hand (in English or Hebrew). Either way, make sure that your pdf file for submission can be easily read. An unreadable answer will not be marked (0 pt).

1 Part I: Theory

1. Let X,Y be two random variables.
   (i) Prove that if X and Y are statistically independent then they are not correlated as well.
   (ii) If X and Y are not correlated, can we deduce that X and Y are independent? If so prove it, otherwise give a counter example.
2. Let X be a random variable. Say which of the following sentences are true or false. If a sentence is false, give a counter example otherwise, prove that it is true.
   (i) X and X^2 are always independent.
   (ii) X and X^2 are never correlated.
   (iii) X and X^2 are always correlated.
   (iv) X and X^2 are never independent.
3. Let X be a 4×4 matrix and $C = X^T X$. Show that C is symmetric semi-positive definite.
4. Given the following matrix X = $\begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_6 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 2 & 0 \\ 3 & 0 \\ 5 & 6 \\ 6 & 6 \\ 7 & 6 \end{pmatrix}$ (i.e. $x_1 = (1\ 0)$, $x_2 = (2\ 0)$ etc.).

   The full answer should be detailed without any use of code.
   (i) Compute the covariance matrix. Notice that here, every row is an observation and every column is a feature.
   (ii) According to the spectral decomposition theorem, every real symmetric matrix has a spectrum, i.e. it can be diagonalized by an orthonormal matrix. Moreover, it can be shown that the diagonalizing matrix is its' eigen matrix. Use this to diagonalize the matrix C and elaborate the calculations.
   (iii) What is the main component?
   (iv) What is the main eigenvalue?

(v) What is the projection of the new data point $x_7 = \begin{pmatrix} 2 & 1 \end{pmatrix}$ on the main component?

5. In this part we would use a neural network for regression. In regression, we mostly estimate a single real value and thus we would have a single neuron in the output layer. Mostly, we will not apply a nonlinear activation in this output neuron. The nonlinear function that we would like to estimate is $y(x)$ shown in Fig. 1 colored in green.

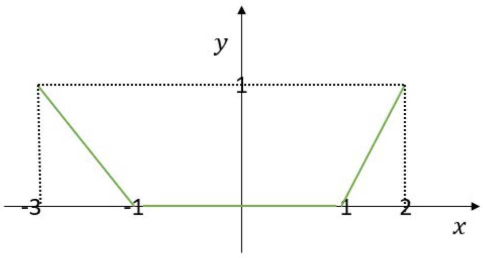

Figure 1: $y(x)$.

You are now also given with the following simple neural network:

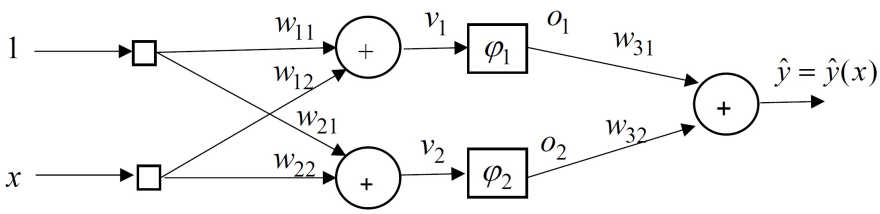

Figure 2: A basic MLP.

The activation functions φ_1, φ_2 are ReLU, i.e. $\varphi(z) = max\{0, z\}$.

(a) Write the matrices $W^{(1)}, W^{(2)}$ as seen in the tutorial according to the weights shown in Fig. 2.

(b) Write explicitly what is $\hat{y}(x)$ using $W^{(1)}, W^{(2)}$.

(c) Write $y(x)$ (given in Fig. 1) explicitly using $\varphi(x)$ and find a possible set of the 6 weights that will make the neural network output to be equal to the given function.

(d) Assume the square loss $L = (y - \hat{y}(x))^2$. Calculate $\vec{\nabla} L$ as a function of the weights x, y, and φ.
Before you start, think how many elements should $\vec{\nabla} L$ have. In addition, you may use the sum representation of $\hat{y}$ rather than the matrix form. In your derivation, you should encounter a term that will repeat itself in all of the elements, so it is best to simply define it as a variable and then to use it when needed.

(e) Using the previous section, write the update policy of every weight using some given learning rate η.

(f) The initialized weights were:

$$(w_{11}, w_{12}) = (w_{31}, w_{32}) = (1, 1) \text{ } and \text{ } (w_{21}, w_{22}) = (-1, -1)$$

The first example was $(x, y) = (0, 0)$. What was the first value of the loss?

2 Part II: Practice

Complete the two given notebooks according to the instructions written there.
For the CNN assignment you will first need to upload the folder named "data" to the root directory of your google drive. Then, open https://colab.research.google.com and connect (top right corner) to your google account. If it is the first time you are working with this notebook, please click on "file" (top left), import the notebook and select the notebook you want to load (i.e. your .ipynb file). If it is not your first time, simply click on "file", open the notebook and then select the notebook you want to open among the list. If you want to use a gpu (for the CNN assignment), go on runtime, change runtime execution and select GPU. If you do not want a gpu, please do the same and select None (even if Colab provide you a free gpu it is for a limited amount of hour per day so don't use it if not working on a task that requires a gpu). You can now start to run the cells like a classical notebook and save your work regularly with Ctrl+s or cmd+s if you work on Mac. It will be saved on the google Colab for the next time you open it including the folders etc.
PS: For the CNN task you will see a cell with the following code :
from google.colab import drive
drive.mount('/content/drive')
Once it is executed it will ask you to connect again to your google account to authorize Colab to have an access to your google drive and thus to the data folder.