Assignment 03

To be solved INDIVIDUALLY or in GROUP of at most two elements

Submit by November 23, 2019, 23h59 by email to jaime.cardoso@fe.up.pt

1. Use the logistic regression implemented with the newton method as made available in the class, logRegF.

Load the height/weight data from the file heightWeightData.txt. The first column is the class label (1=male, 2=female), the second column is height, the third weight. **Remember**: the math in the class (and my code) needs 0's and 1's for the labels.

a. Implement a logistic regression function using gradient descent, **logRegG, having the same inputs and output as logRegF**. Note: when doing the gradient update, use the average of the gradients over the data, not the sum of the gradients (it helps a bit setting the learning rate).
b. Initializing the model to zero (all coefficients equal to zero) find a suitable learning rate and total number of iterations to make the model converge to the optimal model (use the output of **logRegF** as reference). Write the values you choose.
c. Repeat b) but now but first pre-processing the data, centering the data in zero (=removing the mean of each attribute). The convergence should be facilitated. Write the values you choose.
d. Repeat b) but now but first pre-processing the data, centering the data in zero **and** making the variance unitary for each attribute. The convergence should be facilitated. Write the values you choose.

2. Several phenomena and concepts in real life applications are represented by angular data or, as is referred in the literature, directional data. Assume the directional variables are encoded as a periodic value in the range $[0, 2\pi]$.
Assume a two-class (y_0 and y_1), one dimensional classification task over a directional variable x, with equal a priori class probabilities.

a) If the class-conditional densities are defined as $p(x|y_0)= e^{2\cos(x-1)}/(2\,\pi\,2.2796)$ and $p(x|y_1)= e^{3\cos(x+0.9)}/(2\,\pi\,4.8808)$, what's the decision at x=0?
b) If the class-conditional densities are defined as $p(x|y_0)= e^{2\cos(x-1)}/(2\,\pi\,2.2796)$ and $p(x|y_1)= e^{3\cos(x-1)}/(2\,\pi\,4.8808)$, for what values of x is the prediction equal to y_0?
c) Assume the more generic class-conditional densities defined as $p(x|y_0)= e^{k_0\cos(x-\mu_0)}/(2\,\pi\,I(k_0))$ and $p(x|y_1)= e^{k_1\cos(x-\mu_1)}/(2\,\pi\,I(k_1))$. In these expressions, k_i and μ_i are constants and $I(k_i)$ is a constant that depends on k_i. Show that the posterior probability $p(y_0|x)$ can be written as $p(y_0|x) = 1/(1+e^{w_0+ w_1\sin(x-\Theta)})$, where w_0, w_1 and Θ are parameters of the model (and depend on k_i, μ_i and $I(k_i)$).