

Assignment Set 4 - Machine Learning

Instructor: Dr. Enrique Hortal

November 26, 2020

Group assignment

These assignments should be handed in and defended individually.

Handing in

Upload your documented code for Linear Regression to the Canvas' assignment in a single zip archive. In addition, upload a single report in the form of a PDF. We really mean PDF - no other format. The report should **NOT** be part of the zip archive so that we could provide you with feedback on that report (if needed). Submissions by email or other types of archives are not accepted. Thank you for your understanding.

1 Graded Assignment: Linear Regression

In this lab, you will implement the cost function and the gradient descent algorithms for Linear Regression. The implementation is expected to be made in Python. To do so, a start code is provided. See the file "LinearRegression.py" for a basis to start with. Code that you have to update is marked with "HERE YOU...".

After you implemented the required modifications, perform and document experiments to answer the following questions as part of your report:

1. What happens if the learning rate is too low?
2. What happens if the learning rate is too high?
3. Can Linear Regression really find the absolute global minimum? Explain why/why not.
4. What effect does it have if you change the initial guess for θ_0 and θ_1 for the gradient descent to something completely off?
5. What happens if you are not updating θ_0 and θ_1 "simultaneously" as you should but you are updating both parameters in separate for-loops (see code)?
6. How many iterations of the gradient descent algorithm do you have to calculate to reach the correct exact values of θ_0 and θ_1 ?

We really mean perform **and document** experiments. E.g. modify the learning rate, run your code, and see what happens if this parameter is very small. Just providing an answer will only give you very few points. Also provide a plot (just make a screenshot) to document the result of your experiment. Do not forget to provide a short explanation in form of one or two sentences so we know what is the cause for what you see in the plot. You can also try to generate linear data with some noise and check how it affects the training process and the results.

2 Extra implementation

If you manage to implement the linear regression solution satisfactory, we highly encourage you to modify your code in order to perform classification using the diabetes dataset used in Lab 1 by applying **Logistic regression**.