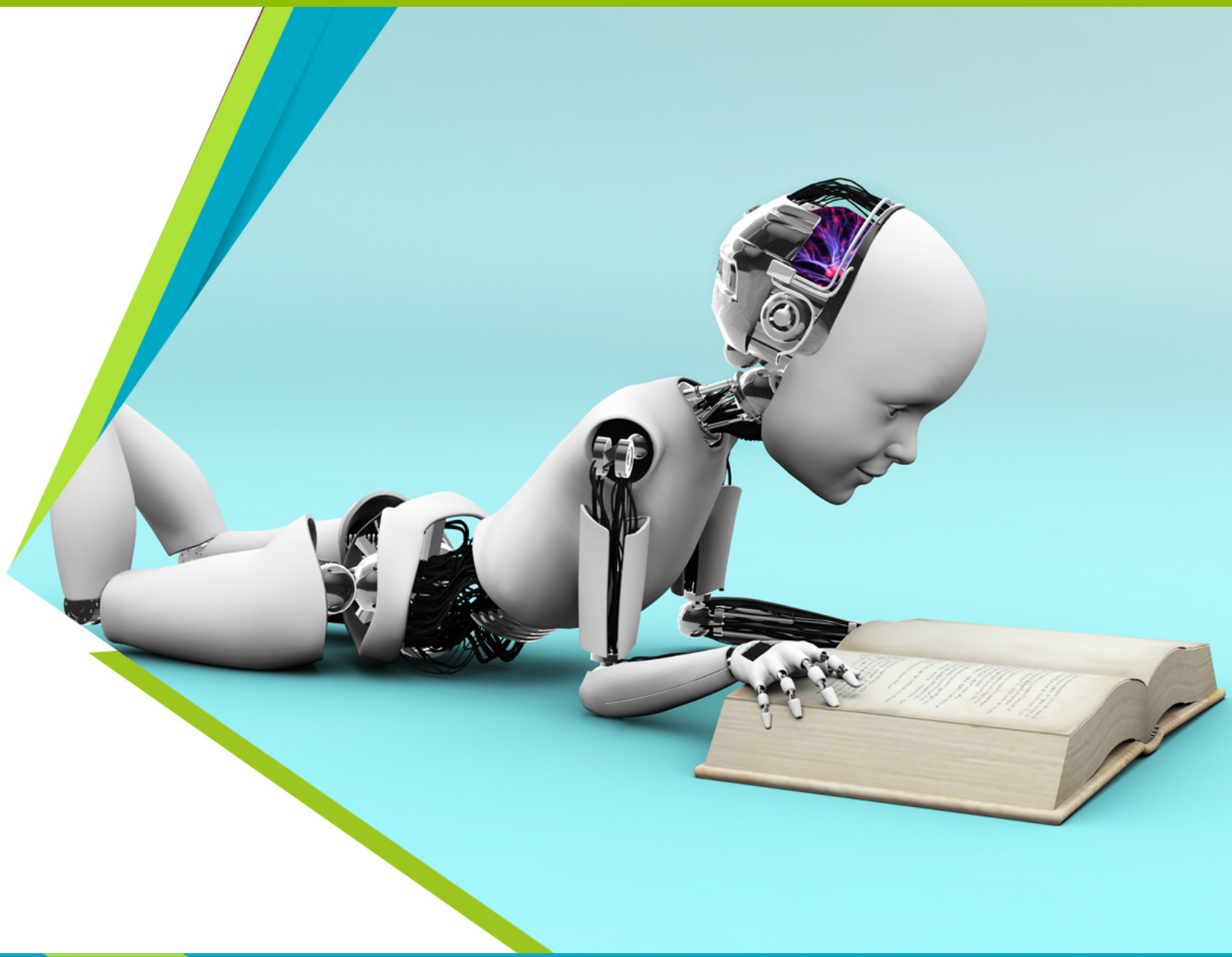
CS4013 Machine Learning

Assignment 3. Univariable Regression and Cost Functions



Tecnológico de Monterrey

Instructions:

From the Andrew Ng's course of Machine learning at Coursera, on the example "ex1", follow the experiments described in the document "ex1.pdf" and apply the following steps:

1. ***Graph the data (Plot).*** *Modify the color and shape of the examples (i.e. blue circles). Add three different variations and comment on the script section.*
2. ***Test the Gradient Descent algorithm.*** *Document how the cost function is constituted analytically, as well as the code (see section 2.2.1 and 2.2.2 of the document "ex1.pdf"), and comment on the result for:*
    a. *The solution: compute the resulting root mean square error and include the figure.*
    b. *The space of parameters: optimal value of the parameters, the profit and its graph.*
3. ***Add regularization criteria of the Gradient Descent algorithm****. Document how the resulting cost function is constituted, as well as the code and comment the result for:*
    a. *The solution: calcula el error cuadrático medio resultante e incluye la gráfica.*
    b. *The space of parameters: optimal value of the parameters, the profit and its graph.*
4. ***Comment and compare the results****.*
    a. *What happens when adding the regularization and when varying the value of lambda?*
    b. *Discuss and compare the solution and the parameter space with 5 lambda values (including lambda=0).*

Apply it in the following data:

1. The example database Ex1, ex1data1.txt
2. From the database "deaths" of Assignment 2, for the variables n vs deviation (figure 2 (a) and (b), pag 17 y 18), comment the result for both cases (lineal and logaritmic).

Realize a document using the LNCS format, available for download on the following page http://www.springer.com/gp/computer-science/lncs/conference-proceedings-guidelines and upload it on the google classroom.

Tecnológico de Monterrey