MACHINE LEARNING
Lecturer: Esteban Abelardo Hernandez Vargas

The project test consists of 2 specific problems and a third one which is open.

A brief written report with pictures and explanations of the results should be provided.

Remark: Provide a brief reasoning in the report why you choose certain algorithm.

Please provide in your report full name and matriculate number.

Problem 1. Using the data set called „problem1.csv (x_training, y_training)" :
   a) Find the polynomial that fits the best the training data
   b) Using the AIC criteria, find the best polynomial that can fit the data.
   c) Cross validate the polynomial with the data set called "problem1.csv (x_test, y_test)"

Problem 2. From a clinical trial, we have 12 patients with HIV infection. After treatment, the disease progressed in 6 patients (1) and in 6 patients the infection did not progress (0). Four measurements are taken in the 12 patients (Age, sugar levels, T cell levels and Cholesterol). Which measurement can be used as a marker to describe progression of the disease? Which will be the criteria to predict the progression? The data can be found in „problem2.csv (x_age, x_sugar, x_Tcell, x_cholesterol, outcome). Arrange the data and briefly explain your results. The variable "y" (target) is a vector of 0 and 1 to represent the progression.

Problem 3. The third problem is flexible.

The student can either find any data set online and try to apply any of the tools we learn in the course

or

Make an assay of 500 words of one of the following papers

- LeCun_2015_Deep learning
- Prosperi _Causal inference and counterfactual prediction in machine learning for actionable healthcare
- Backpropagation Algorithm
- Or any other attached in the platform

Please explain in your own words what is the paper about, provide key aspects you consider relevant about.