Lab Assignment #5

(due April 29, 2020)

<u>Healthcare Application</u>: COVID'19 prediction based on machine learning.

This lab assignment is all about using machine learning; however, you are not expected to know machine learning *per se*—only how to use it in the context of Amazon services.

To work on the assignment, you will be using the Amazon *SageMaker* (https://aws.amazon.com/sagemaker/), which is a fully-managed machine -earning service for building and training learning models, and then deploying them directly into a production-ready hosted environment. It provides an integrated *Jupyter* (sic!) notebook instance for easy access to your data sources for exploration and analysis, so you don't have to manage servers. It also provides common machine-learning algorithms that are optimized to run efficiently on very large data in a distributed environment.

Please perform the steps outlined below.

A. Download the data and code provided at
    https://drive.google.com/drive/folders/1thu0jOyjMAJ33gdtcC-25OHm-N78PYHc. The five files there are
    1) 04-18-2020.csv (data)
    2) time_series_covid19_confirmed_global.csv (data)
    3) time_series_covid19_confirmed_global.csv (data)
    4) time_series_covid19_recovered_global.csv (data)
    5) **coronavirus-covid-19-visualization-prediction.ipynb (code)**

B. Create an Amazon S3 bucket and load the above data into this bucket.

C. Open Amazon SageMaker

D. Launch a Notebook instance (and at the same time create the new IAM permissions for the Amazon S3 bucket you had created)

E. Load the notebook

F. Edit the bucket name and run the code on each line of the notebook

G. Check the results in the *output folde*r in the Amazon S3 bucket you had created.

H. Prepare a pdf file with your report, which must

    1) describe all the steps you have taken (further supported by the screenshots and all the graphics generated in the process).

    2) Have a one-paragraph conclusion describing the results of the study