Deep Learning Group Project
Master's Degree in Data Science 23/24

This document contains all the instructions for successfully developing the project for this course. You have two options for developing the project and your **group can only choose ONE.**

Free Choice

The main idea of this option is to apply some of the concepts covered in the course to address a problem that is of interest to you. The problem considered could be a regression or classification task. However, there are several points which need to be addressed during this choice:

- Task definition: What does the proposed system do (what is its input and output)? What real-world problem does this system try to solve? The scope of the project must not be too narrow or broad. The definition of the task is one of the most relevant parts of the project.
- Evaluation measure(s). An important part of defining the task is the evaluation. In other words, how is it possible to measure the success of the proposed system? For instance, a natural evaluation metric for a classification problem is accuracy, but memory or running time could also be considered. To evaluate the system, it is necessary to obtain a reasonably sized dataset of example input-output pairs, either from existing sources or collecting one from scratch. If the project uses existing datasets that are already pre-processed (e.g., Kaggle), students will be expected to do more with the project. Additionally, in this case, existing solutions must be compared against the solution you're proposing.
- Approach. Identify the challenges of building the deep learning model and the phenomena in the data you are trying to capture. How should you model the task? What deep learning architectures are appropriate for handling the models you came up with, and what are the trade-offs between accuracy and efficiency? Are there any implementation choices specific to the considered problem?

Dermatology Problem

In this topic, you will develop a deep-learning model to address a dermatology classification problem. The dataset used in this project contains several images, each corresponding to only one specific disease. The folder contains only the metadata of each image with some disease information such as malignant/benign, angle scores, and the *URL* of the images of skin lesions. The column "label" from the *CSV* file is your target column. The annotated images represent 114 skin conditions, with at least 53 images and a maximum of 653 images per skin condition.

The dataset of images was taken from the "*FITZPATRICK*17" repository. The Fitzpatrick scale is a numerical classification schema for human skin color. It was developed in 1975 by American dermatologist *Thomas B. Fitzpatrick* to estimate the response of different types of skin to ultraviolet light. The dataset is a combination of two dermatology datasets,

"*DermaAmin*" and "*ATLAS Dermatologico*". Your task is to build a model that performs well on unseen images. For that, you have to split the dataset into three splits: train, validation, and test.

General Guidelines

Like almost every Machine Learning task, there is no "right solution" for this problem. You can create your model however you wish, **as long as you do not look at the test set.** To develop your model, feel free to apply different pre-processing techniques, test distinct combinations of hyperparameters, and make use of all the concepts covered in the course. Naturally, you can (and should) use available libraries such as Pandas, Keras, etc. While the approaches you try are likely to have already been tried by someone else, **you should not plagiarise code**.

Report

In addition to your code, you should deliver a short scientific report (maximum 5 pages + bibliography, if applicable) with the following information:
Group identification, with names and student ID.
A clear description of your best approach (why it is the best, etc..) and the steps you followed for arriving at that model, including (when applicable):

- The applied pre-processing.
- The model/approach implemented.

-Your experimental setup:

- The list of Python libraries needed to run your code and any other implementation details that can be relevant to reproduce your experiment.

-How you evaluated your approach:

- A brief description of your intermediate models (what have you tested and why it has not worked compared to your best approach).
- An evaluation of your best approach. Feel free to include results for other intermediate approaches you tried.

-A summary of the results reported:

- A brief error analysis: what are the most common errors made by your approach? Feel free to illustrate it with a confusion matrix and/or specific examples. (Maybe images of wrong classifications with a corresponding discussion)

-Future work:

- If you had more time, what would you have tried?

The report must be written in English.

Deliverables

By the 28th of April, you must deliver your source code and a report that discusses the previous points and where you explain your choices. You must deliver your work through Moodle (a Turnitin assignment will be available soon). The required files must be uploaded in a unique zip/rar file named with the group number (e.g., GROUP_1.rar). For this purpose, we have created an Excel file you must fill in with group information. Each group *MUST* consist of 5 students. There are no exceptions to this rule.

The grade of the exam will be obtained by evaluating the following features:

- Ability to design and implement Deep Learning models.
- Ability to analyze the results obtained and to use Deep Learning concepts correctly.
- Quality of the report and language adequacy.
- Comparative evaluations against other groups.

The collaboration between groups and the use of AI tools to generate code are strictly forbidden. Not respecting these rules will result in the immediate reprobation of the curricular unit in both epochs.

The fact that you are working in a group does not imply that all the components of the group will receive the same grade.