Welcome to the readme file. Follow the instructions written here, and we hope that the project will start on your machine.

Install the sys, os, numpy, pandas, keras, sklearn, xgboost, matplotlib, seaborn packages. If you have a tensorflow package, then you need to make sure that it is version 1.5. Or just execute two commands “pip uninstall tensorflow” and “pip install tensorflow == 1.5”

To execute the script, create the following folder structure in your folder where you want to run the project: data, doc, lib, output. Copy the main.py file to your working folder.

Arrange the files in folders as they are in the distribution.

Place four files in the ‘data’ folder: data.csv, data_Test.csv, labels.csv, labels_Test.csv. data.csv and labels.csv should be if you want to use real data. data_Test.csv and labels_Test.csv should be if you want to use test data.

Place five files in ‘lib’ folder: Examples_Keras.py, Examples_scikit_learn.py, Examples_XGBoost.py, heat_map_clus.py, open_Data.py.

The program is launched using from your working folder the terminal command “python main.py” or “python3 main.py” depending on your settings.

During the running of the program, you will be asked two basic questions about the calculation mode:

1. What kind of data to use? Test or real.

2. How many instances to use?

All other questions relate to the use of a program module.

After starting from the terminal and the appearance of pictures, they must be closed to continue the work of the program.

In addition, in the modules there are several parameters that are available from a text editor. The description of the parameters is located near the function name. For example, # file_of_data - data with gene expression, # file_of_labels - data with gene, # n_folds - number of folds to make the cross-validation.

The program uses third-party packages, which also contain a bunch of parameters. For example, train_test_split (file_of_data, file_of_labels, test_size = 0.3, random_state = 42). The sense of these parameters is recommended to read on the developer's website.

All files can be found here https://github.com/LosevAMU/M2_DLAD_Math