Report: Measure fairness in healtcare-related database using the **Aequitas** toolkit

Simone Callegarin

April 2023

The practical project on Data Ethics consisted in the creation and presentation of a python notebook.
The assignment required the use of the Aequitas toolkit to measure fairness in an healthcare-related dataset that in my case was related to diabetes.
The content of the notebook can be divided into three main parts:

1. Setup of the notebook

2. Data management

3. Fairness auditing

Each of these required the use of different tools imported in the first phase.

1 Notebook Setup

In the first part are indeed present all the various imports of the libraries that are used further on in the notebook. The notebook setup includes also a cell dedicated to store the API token of the GitHub repository (passed through a `json` file) that is later used to upload the processed Data file encoded in `base64` directly into the GitHub repository through the use of the `requests` library.

2 Data management

The Data went through a process that consisted in five different operations:

1. Data download (from the GitHub repository)

2. Data description

3. Application of a prediction algorithm

4. Predicted Data analysis

5. Data upload (into the GitHub repository)

Here tools as `GitHub` and `Pandas` have been involved in order to respectively download the `csv` file and turn it into a dataframe.

The operation on the dataframe were managed by `Pandas` and the analysis and description phases involved libraries such as `Seaborn` and `MatPlotLib` in particular for graphs plotting.

The Predictive algorithm used is the `Random Forest Classifier` that is imported from `sklearn` library that also offers tools to handle Data training and test splits.

3 Fairness auditing

The main tool used to audit the fairness of the dataset is `Aequitas`, that offers 4 main classes:

1. `Group`

2. `Bias`

3. `Fairness`

4. `Plotting/Plot`

Here `Aequitas` is used to first of all define groups from the attributes contained in the dataset, to calculate disparities on these groups and to assert the fairness of the some famous metrics over these groups.

For completeness the analysis has been completed first on the entire dataset and then only on the measures related to the sensitive attribute.

At the end the `Plotting` library has been used to support the disparities evaluation through the use of many different graph focused on each metric of interest.

4 Bibliography

A short bibliography have been provided with all the resources that helped the writing of the notebook.