

## Instructions:

- A training dataset ([train\\_dataset.csv](#)) consisting of around 51,018 rows and 22 columns contains features and their corresponding target labels.
- To import the 'train\_dataset.csv' file, which is located in the root path of your project, you should use the following path: './train\_dataset.csv'.
- A test dataset ([holdout.csv](#)) comprising 10,000 rows and 21 columns without target labels
- To import the 'holdout.csv' file, which is located in the root path of your project, you should use the following path: './holdout.csv'.
- You can find more information about the dataset [here](#).

## Tasks:

- Work on the provided training dataset to develop the best predictive model for diabetes risk prediction.
- Utilize techniques such as data preprocessing, feature selection, handling class imbalance, and selecting appropriate machine learning models.
- Apply the developed model to the test dataset to predict the target variable.
- Convert the predictions into a CSV file ([predictions.csv](#)) format for evaluation, and the column name should be 'predictions'.
- You can download the predictions.csv after it has been exported in the final task.
- Evaluation will be based on the F1-score calculated from the predicted results compared to the actual target labels in the test dataset.