Machine learning – 2nd software project

**Subject choose: Classification of diabetes using a decision tree.**

The whole project is written in **python 3.7** (in an anaconda environment)

**Libraries** used (only for preprocessing and visualization):

* For data manipulation and loading from archive files: **pandas, numpy and Scipy**
* For fast data visualization: **matplotlib**

**Dataset: Pima Indians Diabetes Dataset**

<https://www.kaggle.com/uciml/pima-indians-diabetes-database>