# Design Patterns Detection

The impact of design patterns on quality attributes has been extensively evaluated in studies with different perspectives, objectives, metrics, and quality attributes, leading to contradictive and hard to compare results. Knowing that a particular module implements a design pattern is a shortcut to design comprehension. Manually detecting design patterns is a time consuming and challenging task; therefore, researchers have proposed automatic design patterns detection techniques which include machine learning based approaches.

In this task you will be implementing a machine learning based approach for detecting design patterns category from opensource projects. The dataset is a subset of source code metrics for each an every java project. Each project in the dataset belongs to one of 3 categories. The categories are described **here** The design of the design pattern detection approach is up to you. The implementation should be in python programming language. Good programming practices will be rewarded and bad practices will be penalized. The list below outlines the main requirements for this task:

- Data extraction, preprocessing and feature engineering procedures should be clearly implemented and justified

- Use Ensemble Learning

- Select two more models to train and do hyper-parameters turning (grid or random search)

- Tune the following hyper-parameters of the estimators in all ensemble models using grid search:

    - n _estimators
    - max features −→ for the base estimators
    - min impurity decrease −→ for the base estimators

- Compare models and make conclusion

Dataset for the task can be downloaded from here: **LINK**