

# EDAMI laboratory

Lab – Clustering task

# Task

- **Aim:**

Determine the best grouping according to the given evaluation method.

- **Data Description:**

- <http://archive.ics.uci.edu/ml/datasets/Wine+Quality> – for the Thursday "10.15" group
- <http://archive.ics.uci.edu/ml/datasets/Cardiotocography> – for the Thursday "16.15" group

# Additional requirements - cardiotocography

- In the analysis the fields: Class and NSP should not be taken into consideration
- Maximum number of groups 15 (the references grouping has 10 groups).
- The reference grouping is defined by the Class attribute.
- Minimum 10 tests are required.
- The obtained result should be better than the result obtained by application k-means algorithm for 10 groups with default values of parameters and without pre-processing of data.

# Additional requirements - wines

- In the analysis (clustering) the quality field should not be taken into consideration
- Maximum number of groups 10 (the references grouping has 6 groups).
- The reference grouping is defined by the quality attribute.
- Minimum 10 tests are required.
- The obtained result should be better than the result obtained by application k-means algorithm for 6 groups with default values of parameters and without pre-processing of data.