Introduction to Artificial Intelligence

Summer 2022

EXERCISE 4: Regression and classification

Teacher: Tomasz Lehmann, tomasz.lehmann.dokt@pw.edu.pl

This exercise is done in pairs.

1. **Exercise details**

   Variant 2

Write a program that predicts red wine quality based on the provided dataset (**variant2**).

Use at least two different methods (e.g. Linear Regression, Logistic Regression, SVM, …) and compere their performance. You should also deliver report that includes the comparison of metrics for each used method and explanation of which method consider the best.

The solutions you deliver will be evaluated based on the results of the chosen models and the correctness of your research methods.

1) Properly divide, investigate, and preprocess the data.
2) Correctly implement the regression and classification algorithms.
3) Correctly implement the regression and classification algorithms.
4) Implement validation and testing and perform research on their results.

2. **Technical details**
   a. The solution must be implemented in Python
   b. Please ensure that your code adheres to basic standards of lean coding in accordance to PEP 8. Additionally, it should contain comments in the crucial parts to help with readability and understanding
   c. The clear instruction how to run and test the code should be included
   d. The submission of the final report is mandatory, and the task will not be accepted without it

3. **Handing-in guidelines**
   a. You should submit the source code of your solution and final report to tomasz.lehmann.dokt@pw.edu.pl not later than:
      i. **2023.04.17 9:59:59 (Monday) GMT+1 for the Monday group**
      ii. **2023.04.19 23:59:59 (Tuesday) GMT+1 for the Wednesday group**
      iii. **2023.04.21 23:59:59 (Thursday) GMT+1 for the Friday group**

      The late project submissions will result in a 20% reduction in points per week, up to a maximum of 2 weeks late. Please ensure that you submit your projects on time to avoid any point deductions

b. Please include "[EARIN] Exercise 4: Name Surname 1, Name Surname 2" in the title, and do not forget about adding emails of both team members in the email content.
c. You may get 0-5 pts for this assignment.
d. In case of questions, please contact me only via MS Teams (on the laboratory channel).
e. Online assessments are scheduled for the Monday, Wednesday, and Friday groups on **17.04.23**, **19.04.23**, and **21.04.23**, respectively. Further information regarding the details will be provided at a later time.

4. Assessment Criteria

The following criteria will be used to evaluate your work:

a. Proper data preparation: **1 point**
b. Proper algorithms implementation: **1 point**
c. Proper training procedures and results analysis: **1.5 points**
d. Final report including results, clear explanation of the programmed solutions and reflections on what was done well and what could be improved: **1.5 points**