

Artificial Intelligence

Practical Sheet 6: Classification



1. **“WINE” dataset problem.** Your task is to adapt the *“linear_regression.py”* script (available at the course web page), in order to perform classification, by means of logistic regression.

Start by downloading the data from the **“wine”** Dataset, available at the course web page. In this set, the goal is to use chemical analysis to determine the origin of different wines. As you can see, there are 13 attributes, all numeric (either integer or real numbers):

- 1) *Alcohol*
- 2) *Malic acid*
- 3) *Ash*
- 4) *Alcalinity of ash*
- 5) *Magnesium*
- 6) *Total phenols*
- 7) *Flavanoids*
- 8) *Nonflavanoid phenols*
- 9) *Proanthocyanins*
- 10) *Color intensity*
- 11) *Hue*
- 12) *OD280/OD315 of diluted wines*
- 13) *Proline*

The dependent variable is provided in the first column.



hugomcp@di.ubi.pt, 2022/23

- a. At first, implement only the “two-classes” version;
- b. Subsequently, generalize your script, to handle multi-class problems;

You should also consider different feature normalization strategies (Min-max, Z-score)