



**DM**  
DEPARTAMENTO  
DE MATEMÁTICA  
TÉCNICO LISBOA

## **Statistical Methods for Data Mining**

MEGIE, MECD AND MMA, 1<sup>st</sup> SEMESTER, 2022/2023

### **Project - Part II**

Handed out on December 5, 2022.

**To be handed back by January 8, 2023**

Using the same dataset from Project 1:

1. Apply clustering methods to your data. Interpret the results. Compare the obtained partitions with the true class each object belongs to.
2. Solve your classification problem using for classes the partition clusters obtained in step (1) and your best classifier obtained in Project 1. Compare the results with the ones obtained in Project 1. Discuss potential advantages and drawbacks of each strategy.
3. Discuss limitations of the analysis you have done and provide suggestions for future work.

#### **About the report:**

- The report should not exceed 20 pages.
- Do not forget topics such as objectives; estimation methods; decisions; conclusions of the study; and references.
- The R code (or other code), as well as the report, must be send by email to: `conceicao.amado@tecnico.ulisboa.pt`