

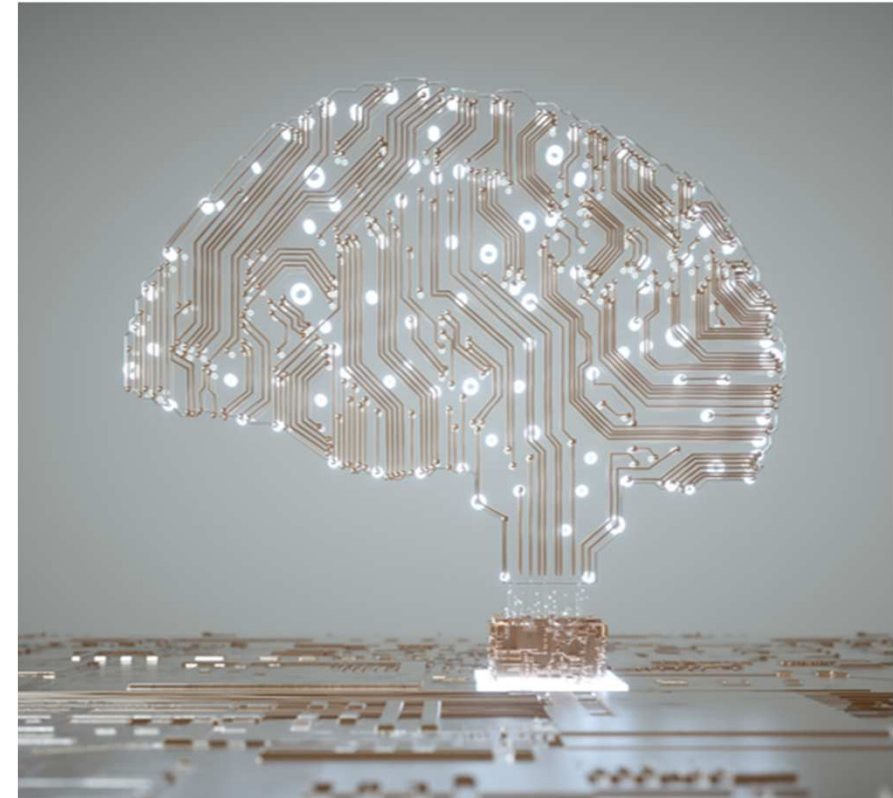
# Project Guide 02

## Fundamentals of Artificial Intelligence

MSc in Applied Artificial Intelligence, 2023-24

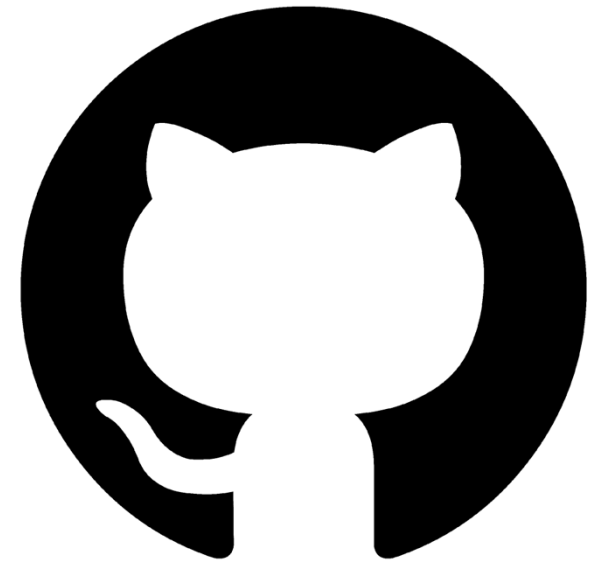
# Introduction

- This document presents the goal and the requirements for the development of project 02
- The goal of the project is to implement distinct **Machine Learning** (ML) approaches and methods to address a specific problem using a public dataset
- The work groups should be composed of **two students**
  - Different tasks must be assigned to each student
  - The project **score is individual**



# GitHub Platform

- The project must be available in a **GitHub repository**
- Each student must create an account on the GitHub platform
  - One of the members of the group must create a private repository
  - Use the nomenclature "**FAI23\_G##**", where G## corresponds to the group
  - Add the other members of the group as collaborators
- The lecturer/teacher should also be added to the repository and the project



# Project submission

- The **Github repository** must contain all files required to run the solution
- The code and the documentation must be merged into a **Jupyter Notebook**
- A **ZIP file** of the repository contents must be submitted on Moodle before the deadline
- The **link** for the repository must be included in the first markdown block of the Jupyter Notebook
- The project will be later presented and defended by the work group



# Notebook structure

# Introduction

- Identify the work group: student name and number
- Establish here the context and the purpose of project 02
- Provide the individual contribution of each student
- Each student should focus in one main task
  - **Supervised learning:** apply two automatic classification methods
  - **Unsupervised learning:** apply k-means and Apriori algorithms
- You can find many public datasets in <https://www.kaggle.com/datasets>
- The structure of the notebook can be adapted according to each project characteristics

# Supervised learning

- Dataset for automatic classification
  - Provide the source of the used dataset(s)
  - Present a brief description
  - Include the dataset metadata
  
- Automatic classification process
  - Describe the process and tasks for automatic classification
  - Define the business goals to be achieved
  - Select two algorithms and the parameters to be used
  - Present the data selection criteria

# Supervised learning (2)

- For each classification method (2 sections)
  - Explain how the data was prepared
  - Apply the ML algorithms and evaluate the generate models
  - Optimize the selected algorithm
  - Document the intermediate and final results
- Results Analysis
  - Present an analysis of the results based on performance metrics



# Unsupervised learning

- Dataset for unsupervised learning
  - Provide the source of the used dataset(s)
  - Present a brief description
  - Include the dataset metadata
  
- Clustering
  - Define the business goal to be achieved
  - Present the data selection criteria
  - Explain how the data was prepared
  - Apply and evaluate the K-Means algorithm
  - Optimize the algorithm parameters
  - Document the intermediate and final results

# Unsupervised learning (2)

- Association rules
  - Define the business goal to be achieved
  - Present the data selection criteria
  - Describe the data preparation steps
  - Apply and evaluate the Apriori algorithm, adjusting the algorithm parameters
  - Document the intermediate and final results
- Results Analysis
  - Present an analysis of the results based on performance metrics

Thank you!