

Introduction in Machine Learning - WBAI056-2024.2025

Assignment 1

Total points: **40**
Starting date: 18 November 2024
Submission deadline: **23:59, 28 November 2024**

Lecturers Dr. Matias Valdenegro, Dr. Andreea Sburlea.

General guidelines:

- The tasks are targeted at groups of three students. Please make sure that the load is well divided: every student should contribute.
- Please take advantage of the practical sessions to ask your questions about the tasks.
- Provide a (short but comprehensive) explanation of what you are doing for each task.
- A reviewer should be able to understand plots independently; be sure to label axes, a legend for colors, use an easily readable font size, etc.
- Refer to all plots, tables, code blocks, etc. in your report.
- For the report: you can use a jupyter notebook or write a PDF in a word processor of your preference. Please include code as `.py` or `.ipynb` files as attachments in Brightspace.
- In your report, please include a statement from all group members, indicating what was their contribution to this assignment.

Part I - Feature Type Examples

Maximum obtainable points: **10**

Consider features that can be continuous, discrete, and categorical. Give five examples of continuous features, three examples of discrete features, and two examples of categorical features (10 in total). For each example, give a brief description/argument on why that feature is of the specified type. Examples can be given from real life, applications, but should not be examples from the lectures.

Part II - Classifiers vs Regressors

Maximum obtainable points: **15**

Consider that you have a task that can be modelled as regression or classification, that is, the target variable can be continuous or discrete. Provide arguments on why you would prefer to model this task as regression or classification.

Here there is no right answer of choosing classification or regression, but we focus on what arguments you provide for your case. Argue for only one of the possible options.

Part III - Model Selection and Differences.

Maximum obtainable points: **15**

In the lecture we have seen several models for different tasks, including classification and regression models, but often students do not have clarity on how these models compare to each other.

First choose one of classification or regression, then consider all models we covered in lectures that can do such task, and choose 3-4 of them, and then describe what makes them different, you might consider factors like:

- Linearity.
- Shape of decision boundary.
- Model equations.
- Training process.
- Prediction/inference process.
- Advantages/disadvantages.
- Your own comparison criteria.

Each comparison should have a good quality description between your selected models, and consider all the models you selected in the comparison.