Challenge

Aprendizagem Computacional/Machine Learning, 2024
DEI, FCT, University of Coimbra

Context ang Goals

In this work, groups of up to two students are challenged to identify a Machine Learning case study and evaluate the capability of simple models to deal with it:

1. **Objective**: The objective of the challenge is to encourage students to identify a real-world problem that can be solved using machine learning and then construct simple models to tackle it. The challenge should focus on promoting creativity and knowledge in the field of machine learning.

2. **Case study**: A clear description of the proposed case study should be provided, including a brief introduction to the problem, a dataset description, and a detailed explanation of the used methodology.

3. **Data Sources**: Any publicly available dataset can be used. As an alternative it is possible to collect new data as long as it is made publicly available. Please ensure that the data is representative of the problem to be solved. If any preprocessing is performed it should be explained.

4. **Model Construction**: To tackle the problem outlined in their proposal, students must construct at least two simple machine learning models, Decision Trees and K-Nearest Neighbors are suggested, but clustering approaches or Logistic Regression can also be used. The models should be constructed using Python programming language and should be able to make predictions on new unobserved data.

5. **Evaluation Metrics**: The models should be evaluated using appropriate evaluation metrics, which may vary depending on the problem being solved. Students should choose evaluation metrics that are appropriate for their problem and explain their choice. Visual aids can also be used.

6. **Documentation**: Students should document the entire process, from data collection to model construction, and evaluation. They should include details of the methodology used and the results obtained, along with any challenges faced and how they overcame them.

7. **Submission**: The final submission should include the proposal, a graphical abstract, the code used to construct the model, the documentation, and any other relevant information.

8. **Criteria**: The submissions will be evaluated based on the relevance of the problem, the quality of the data, the performance of the model, the choice of evaluation metrics, the clarity and soundness of the documentation, and the discussion.

9. **Prize**: The best work will be chosen as the topic for the project of machine learning to be developed later in the semester. If no work is considered adequate, the professors will propose a different topic for the project.

Deliverables and deadlines

1. The problem should be informally approved by professors in classes. Discussion of themes must be fulfilled by the February 16th.

2. Challenge submission deadline: March 16th

    a. Submit a .pdf file into *Inforestudante* with your report including all details described above.
    b. Submit a zip file with all your code including setup information (data, software, etc.)

3. Discussion is mandatory and will take place in classes the following week.

FeedNPlay

Projects developed in the context of this course may be submitted to be displayed in FeedNPlay, a multimedia device with a total width of more than 6 meters, which is installed on the 1st Floor of the Department of Informatics Engineering. Students should follow the instructions available at the FeedNPlay website (https://feednplay.dei.uc.pt).