
Exercise: Supervised learning II

Machine Learning
2025/26

Contents

| | | |
|-----|---------------------|---|
| 1 | Presentation | 2 |
| 1.1 | Dataset | 2 |
| 2 | What to do | 2 |
| 3 | Deliverables | 3 |
| 4 | Evaluation criteria | 4 |

1 Presentation

In this last assignment of the subject, you'll need to apply your knowledge about machine learning. You'll learn the best possible model for the proposed dataset using an appropriate methodology, perform an overfitting study and analyze the problem after transforming it into a semi-supervised learning problem.

1.1 Dataset

We propose using a dataset generated to simulate registration of high-energy gamma particles in a ground-based gamma telescope, aimed to learn a classifier that discriminates the so-called showers caused by primary gammas (signal) from those initiated by cosmic rays in the upper atmosphere (background).

The **MagicTelescope** dataset (find it here) has 11 numerical features and a binary class variable. The number of samples is above 19k and 64% of them are positive.

If you'd like using a different dataset, please, check it with your professor beforehand.

2 What to do

1. Learn a model from the MagicTelescope data and measure its performance. Choose a type of model among the ones explained in the last part of the subject. Preprocess the data if required. Be careful with the methodology chosen for *hyperparameter tuning* and *performance evaluation*.

Motivate your decisions, explain your results. Discuss the performance reached and potential limitations.

2. Identify one relevant hyperparameter of the learning technique you use. Keep the rest constant while you perform a complexity analysis to assess overfitting. You can expect results similar (or not) to those in Figure 1.



Figure 1: Complexity study: how performance (in train and test data) changes as hyperparameter's values are modified. Sources: <https://medium.com/@brandon93.w/understanding-overfitting-and-underfitting-in-machine-learning-b699e0ed5b28> and <https://developers.google.com/machine-learning/crash-course/overfitting/overfitting>.

Motivate your decisions, explain your results. Why do you choose *that* parameters? What have you observed?

3. Take the hyperparameters of the best model you found. Perform a sample size analysis. You can expect results similar (or not) to those in Figure 2. You'll have to use training data subsets of increasing size.

Motivate your decisions, explain your results. What have you observed?

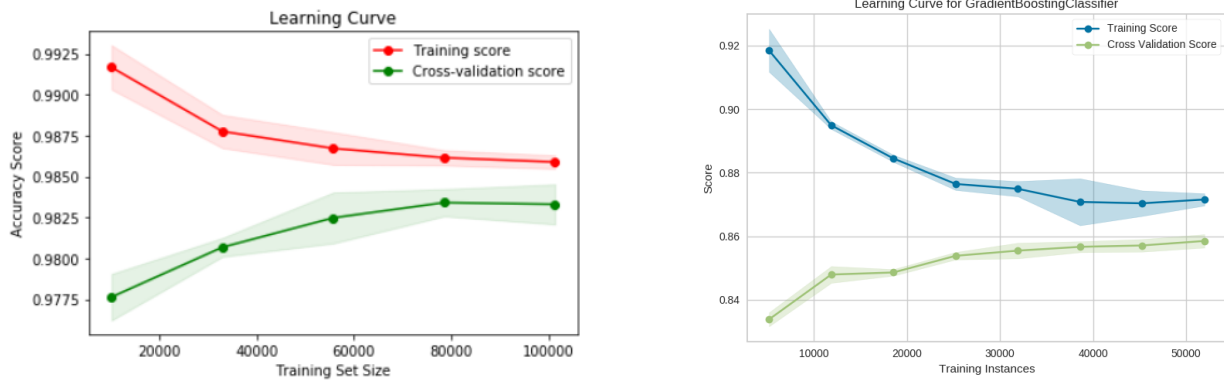


Figure 2: Sample size study: how performance (in train and test data) changes as the same size increases. Sources: <https://stats.stackexchange.com/questions/438632/assessing-overfitting-via-learning-curves> and <https://stackoverflow.com/questions/58221470/is-this-classification-model-overfitting>.

4. Transform the MagicTelescope data into an SSL problem. Design a learning technique to learn with SSL data. Perform an analysis about the amount of labeled data required to learn a model. You can expect results similar (or not) to those in Figure 3. You might want to compare it with the performance of the model learned with the fully labeled data.

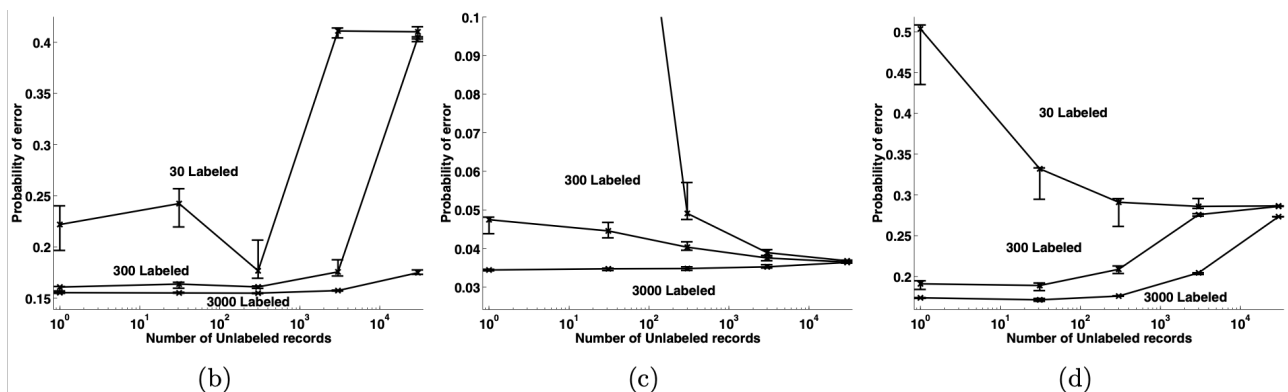


Figure 3: Semi-supervision study: how performance changes as the relative and absolute amounts of labeled data change. Sources: <https://doi.org/10.7551/mitpress/6173.003.0007>.

Motivate your decisions, explain your results. What have you observed?

You'll find in Moodle a notebook with 3 different ways to transform labeled data into SSL data.

3 Deliverables

- a) Submit a Jupyter notebook in PDF format¹, via Moodle, with the implementation and description of your analysis.

- Use Markdown cells to explain/justify all your decisions and motivate your analysis.
- Include a statement, and the corresponding links, about the use of external sources and tools. E.g.,
 - “I have based my implementation on StackOverflow’s code in this discussion”
 - “I asked ChatGPT to check my code for flaws in this conversation”

Properly cite any text you rely on to justify your decisions.

Failing to declare the use of any tool/source or to include the corresponding citations may be considered fraud or plagiarism, and we will proceed as stipulated by UdG (eBOU-1751, January 16th, 2019, Art. 21).

Check Moodle’s task for the submission’s deadline.

¹Run>“Restart kernel and run all cells” + File>“Save and export notebook as”>PDF
or Kernel>“Restart and run all” + File>Download as>PDF

b) Explain and defend your work to the professor in a face-to-face interview.

Follow Moodle's update for the instructions to schedule your interview.

4 Evaluation criteria

The following rubric will be used to calculate the final grade:

| Criteria | Degrees of accomplishment | | | |
|---|---|---|--|---|
| | Not fulfilled | Below expected | Good | Excellent |
| Methodology for hyperparameter tuning and model evaluation (up to 2 points) | Invalid approach (0%) | Heavy flaws or flawed limited analysis (33%) | Slight flaws or limited analysis (67%) | Correct methodology and analysis (100%) |
| Overfitting study as a function of sample size (up to 2 points) | No study or incorrect design (0%) | Heavy flaws or flawed limited analysis (33%) | Slight flaws or limited analysis (67%) | Well designed and insightful analysis (100%) |
| Overfitting study as a function of a key hyper-parameter's value (up to 2 points) | No study or incorrect design (0%) | Heavy flaws or flawed limited analysis (33%) | Slight flaws or limited analysis (67%) | Well designed and insightful analysis (100%) |
| Semisupervised-learning study for hyperparameter tuning and model evaluation (up to 2 points) | No study or incorrect design (0%) | Heavy flaws or flawed limited analysis (33%) | Slight flaws or limited analysis (67%) | Well designed and insightful analysis (100%) |
| Documentation, reasoning, motivation (both in the notebook and during the interview) (up to 2 points) | No documentation or minimal, not able to answer professor's questions. (0%) | Basic documentation that lacks detail, depth or formality. Only answer to few basic questions (33%) | Correctly documented, with some issues in completeness, depth or formality. Cannot answer some questions (67%) | Comprehensive, formal and well-organized documentation. Correct answers to professor's questions (100%) |