

2024 Semester 1

Home

Announcements

Echo 360

Course Readings

Modules

Discussions

Assignments

Grades

2

Assignment Help

LMS Analytics


Search

SELT

# Assignment 2

**Due** 26 Apr by 23:59    **Points** 60    **Submitting** an external tool    **Available** 1 Apr at 0:00 - 28 Apr at 23:59

In this assignment, you will apply several types of classifiers to the problem of breast cancer detection from images.

[This zip file](#)  contains a Jupyter notebook template and the dataset.

Only submit your notebook (.ipynb) and nothing else. You don't need to submit the data file, but your code should assume it is in the same directory, and its name has not changed.

The autograde of this assignment is worth **10%** of your overall marks for the course.

1	step01	1.0
2	step02	3.0
3	step03	3.0
4	step04	4.0
5	step05	2.0
6	step06	3.0
7	step07	3.0
8	step08	3.0
9	step09	2.0
10	step10	3.0
11	step11	2.0
12	step12	2.0
13	step13	2.0
14	step14	2.0
15	step15	0.0
16	step16	3.0
17	step17	4.0
18	step18	3.0
19	step19	2.0
20	step20	2.0
21	step21	3.0
22	step22	3.0
23	step23	2.0
24	step24	3.0

Some considerations:

- In Step5, the input *y\_pred* of the function **calculate metrics** should be the probability of belonging to the positive class (i.e values between 0 and 1). In this way, you can use this value directly (*y\_pred*) in the calculation of specific metrics. On the other hand, for the metrics where you need to use the predicted class as input (i.e. 0 or 1), you can take the probabilities given by *y\_pred* and use a threshold of 0.5 to determine the classes of your predictions inside your **calculate metrics function** before passing them to the metrics where you need to use the predicted class.
- Remember that if you need to get the probabilities of your predictions, you can use the function **predict\_proba** to obtain the probabilities. However, in the case of **SGDClassifier(log = “perceptron”)**, instead of using the function **predict\_proba**, use the function **decision\_function**, which gives a score of your prediction.
- SGDClassifier can use the parameter *log\_loss* (e.g., **SGDClassifier(loss = log\_loss)**); **however**, if you are using older Sklearn versions (which is better not to do so) try to use (e.g., **SGDClassifier(loss = log)**).

The submission portal will open on the 6th of April

This tool needs to be loaded in a new browser window

Load Assignment 2 in a new window