Project 2: Digit recognizer

Start oppgave

Forfall Fredag av 23.59 **Poeng** 25 **Må leveres** en filopplasting **Filtyper** pdf og zip
Tilgjengelig til 31. okt. i 23.59

In this project, we train a classifier that recognizes hand-written digits. The description of the task can be found here: **project2.pdf** (https://mitt.uib.no/courses/30808/files/3686005/download?download_frd=1)

Data files can be found here: **handwritten_digits_images.csv** (https://mitt.uib.no/courses/30808/files/3662002/download?download_frd=1) and **handwritten_digits_labels.csv** (https://mitt.uib.no/courses/30808/files/3662004/download?download_frd=1)

Return exactly two files: a pdf containing your report and a zip containing your code

If you have done the work with a pair, then do the following:

- First, both persons should sign-up into the same "Project 2" group.
- Then one person returns the project on behalf of the whole group.

Project 2

Kriterier	Vurderinger					Poeng
Has the student tried at least three classifiers?	**1 poeng** **Yes**		**0,5 poeng** **1-2 classifiers**		**0 poeng** **None**	1 poeng
Design choices Do the approach and design choices make sense? Are the candidate classifiers reasonable for this task? Has the student tried different hyperparameter values to choose the best model? Are the hyperparameter values reasonable?	**3 poeng** **Excellent**		**2 poeng** **Good**	**1 poeng** **Decent**	**0 poeng** **Bad**	3 poeng
Performance tuning Has the student made an effort to optimize performance? does the model achieve good performance?	**4 poeng** **Excellent** The student has clearly tried to and succeeded in getting good performance	**3 poeng** **Very good** Results are reasonable and the student has made a clear effort to optimize performance.	**2 poeng** **Decent** Results are reasonable but there is only a limited effort to optimize them.	**1 poeng** **Room for improvement** Results are reasonable but there is no visible effort to get good performance	**0 poeng** **No** There is clearly something wrong in the learning process	4 poeng
Model selection Has the student used a proper model selection procedure (e.g., holdout validation or cross-validation)? Has the model selection procedure been used properly?	**6 poeng** **Excellent** Basic principles and all details are correct.	**4,5 poeng** **Good** Basic principles are correct but there are some minor errors	**3 poeng** **Decent** Basic principles are correct but there are major errors.	**1,5 poeng** **Room for improvement** Basic principles are wrong but there are some correct elements	**0 poeng** **Bad** Totally wrong procedure or no model selection at all	6 poeng
Model evaluation Has the student properly evaluated the performance of the final classifier (that is, using unseen test data)?	**3 poeng** **Excellent** The final classifier is evaluated using unseen test data		**2 poeng** **Good**	**1 poeng** **Decent**	**0 poeng** **Bad** No separate evaluation done (for example, validation error is used instead of test error)	3 poeng

Kriterier	Vurderinger				Poeng
Quality of the report Are choices and decision well-justified? Is the report clearly written? Are the result described well (quality of plots, ...)? Are all required contents there?	**3 poeng** **Excellent** Report is clear and well-written. Results are illustrated with informative plots.	**2 poeng** **Good** Small unclarities or missing details. Plenty of plots, most of them clear.	**1 poeng** **Decent** Lots of unclarities and missing details. Few or very messy plots.	**0 poeng** **Bad** The report makes no sense.	3 poeng
Code quality Is the code clear? Has the student documented the code properly?	**3 poeng** **Excellent**	**2 poeng** **Good**	**1 poeng** **Decent**	**0 poeng** **Bad**	3 poeng
Reproducibility Are the results reproducible? Note: due to the limited time, graders have not always run the code but inspected the code.	**2 poeng** **Excellent** Results can be reproduced exactly.	**1 poeng** **Good** Results can be reproduced to a large degree but not exactly. For example, seeds are not fixed so every run gives slightly different results.	**0 poeng** **Bad** The results cannot be reproduced For example, a part of the code is missing and the test pipeline cannot be run		2 poeng
					Totalt antall poeng: 25