

## IST Austria: Statistical Machine Learning 2020/21

Christoph Lampert <chl@ist.ac.at>

TAs: Alexandra Peste <alexandra.peste@ist.ac.at>,

Bernd Prach <bernd.prach@ist.ac.at>

Final Project Description (project end 16/11/2020 10:15, reports due 23/11/2020 10:15)

Please send your reports via email to the TAs

# 1 Final Project

The final project should be solved in teams of exactly two participants each. If you cannot find a partner, please let us know and we'll team you up.

**Task 1)** Participate in the Kaggle competition *IST Austria - Statistical Machine Learning 2020* at <https://www.kaggle.com/c/ist-austria-sml-2020/>. 16.7% of your final grade will depend on the results you achieve here. It is completely up to you, which technique(s) you use. You can use external programs or your own implementation, etc. You can inspect the training data manually, subsample it if it's too large for your purposes, etc.

Do not look at the test data yourself, it should only be used to produce a submissions file that you can upload. Note: the challenge shows a public leaderboard. This score is computed from 20% of the test data you submit. Don't trust the value too much, though. *The actual score for your submission will be computed after the challenge is finished from the remaining 80% data.*

**Task 2)** Write a short scientific report (at most 2 pages, continuous text) about how you achieved the submitted results and send it to the TAs. 16.65% of your final grade will depend on this report.

The description must make the process reproducible by which you achieved the results you submitted (submitting code is not a replacement for this requirement). In particular, include the algorithms you used as well as potential preprocessing (manual or automatic) and model selection steps. If you choose any hyperparameters manually, please give the values and explain why these are the ones you chose. *All of these aspects will contribute to scoring of your report.*

**Task 3)** Present your results in a short presentation (approx. 10 minute plus 5 minutes of questions and answers), –tentatively– on November 16th 16.65% of your final grade will depend on this presentation.

As far as possible, the presentation should convey the same information as the report. *During the Q&A, both team members must be able to explain all aspects of the project.*