Data Analysis and Model Classification
Guidesheet I-5: Statistical significance and Final Classifier

Fumiaki Iwane Ping-Keng Jao Bastien Orset Julien Rechenmann
Ricardo Chavarriaga José del R. Millán

In this (very short) guidesheet, you will have to first evaluate the statistical significance of the overall performance measured by test error. Then, based on the previous guidesheets you will have to build your final classifier with the optimal hyperparameters that you found.

Statistical significance

Imagine that the mean test error across *outer* folds is 11%. This suggests that the accuracy is pretty good compared to the random level of 50% (2-class problem). Now imagine that the mean test error across *outer* folds is 46%. It is still lower than the random level, but the difference might not be statistically significant.

Hands on

- What is the mean test error across *outer* folds? what is the standard deviation of the test error across *outer* folds?
- How do the test error values compare to the random level?
- Do a t-test of the hypothesis that the test error values across *outer* folds come from a distribution with mean 50% (`[h, p] = ttest(error, 0.5)`)
- Can you reject the null hypothesis? Why?

Final Classifier

In the previous guidesheets, you have been introduced to feature selection methods *(feature forward selection, fisher score)* and how to optimize your number of features you need to train a classifier. We also introduced to you different type of classifiers *(diaglinear,diagquadratic,linear,quadratic)* as well as dimension reduction algorithm such as *Principal Component Analysis*. Now you have all the tools in hands to make a final "best" classifier (final model). How would you design it? Which hyperparameters would you use?

Include this final guidesheet of mini-project 1 in the report and submit it on Moodle before **November 12th, 11:59pm**. **Good luck for the mini-project!**