Machine Learning | Prof. Dr. Bernt Schiele
Dr. Paul Swoboda & Dr. Gerard Pons-Moll
Exercise Sheet 8 - 10.12.2018 | due: 17.12.2018, 14:15

Exercise 15 - ROC and AUC

You are reviewer for an international conference on machine learning. Among the submissions are two independent papers, each proposing a solution to the important problem of automatic classification of hand-crafted contemporary art pottery. As you are an expert in this emerging new field, you are asked to write a review of the two papers. In both papers, the quality of the proposed classifier is evaluated by plotting the ROC curves. Author 1 reports an AUC of 0.73, Author 2 reports an AUC of 0.70.

Both authors made the results of their classifier available as supplementary material. However, the authors used different datasets to evaluate their methods. You can find the dataset Y_1 (the true labels) and the prediction f_1 of the first author as well as the dataset Y_2 and the prediction f_2 of the second author in the file `auc.mat`.

a. **(4 Points)** Inspect the datasets and the predictions made by the two methods. Reproduce the values of the AUC reported by the two authors by means of the function `PlotROC`. Perform 1000 runs of a random classifier on each dataset and compare the ROC curves and the AUC values obtained by the two methods with the results of the random classifiers. Report your results. Which classifier do you think works better? Justify your decision.

Exercise 16 - Statistical Tests

We are in the setting of binary classification $\mathcal{Y} = \{-1, 1\}$. We want to develop statistical tests to evaluate if there is any dependency between the features X and the labels Y.

a. **(1 Point)** Prove the following statement:

$$X \text{ and } Y \text{ are independent if and only if } p(x|y = 1) = p(x|y = -1)$$

Thus we can reduce the question if X and Y are independent to the question if the class conditional distributions are equal.

b. **(3 Points)** We only have one feature, that is $x \in \mathbb{R}$. Suppose that $p(x|y = 1) \sim \mathcal{N}(\mu_+, \sigma^2)$ and $p(x|y = -1) \sim \mathcal{N}(\mu_-, \sigma^2)$ with known variance σ^2. Develop a modified version of the Gauss-test to test if $p(x|y = 1) = p(x|y = -1)$ (null-hypothesis). That requires to develop a test-statistic, derivation of the distribution of the test-statistic under the null-hypothesis, definition of a rejection region for a given significance level α, computation of the p-value. The number of samples from positive and negative class can be different.

c. **(2 Points)** Write a Matlab Function

```
pval = DoPermutationTest(X,Y),
```

which given the data $(x_i, y_i)_{i=1}^n$ does a permutation test for testing whether $p(x|y = 1)$ is equal to $p(x|y = -1)$ using the Fisher-score as the test-statistic (1000 permutations drawn uniform at random) and returns the p-value.

d. **(3 Points)** Compare the two developed tests, where

- first with data sampled from $p(x|y = 1) \sim \mathcal{N}(\mu, 1)$ and $p(x|y = -1) \sim \mathcal{N}(0, 1)$ (100 points from $Y = 1$ and 200 points from $Y = -1$) and vary $\mu =$ `0:0.1:0.5` and for each μ repeat the experiment 100 times. How often do you reject H_0 for the significance level 0.05 for both tests for each μ.

- second with data sampled from $p(x|y=1) \sim \texttt{Uni}[\mu, 1+\mu]$ and $p(x|y=-1) \sim \texttt{Uni}[0,1]$, where `Uni` denotes the uniform distribution (100 points from $Y=1$ and 200 points from $Y=-1$) and vary $\mu =$ `0:0.03:0.15` and for each μ repeat the experiment 100 times. How often do you reject H_0 for the significance level 0.05 for both tests for each μ.

Discuss the result of the two tests. Plot the number of rejections of H_0 as a function of μ.

Send the code for c) and the plots of d).

Hints:

- In order to get random permutations use the function `randperm(n)` which returns a permutation of the number $1, \dots, n$.
- For b) you may use that if $X \sim \mathcal{N}(\mu, \Sigma)$ where $X \in \mathbb{R}^n$, then for $A \in \mathbb{R}^{m \times n}$ the variable AX is distributed as $\mathcal{N}(A\mu, A\Sigma A^T)$.

Submission instructions

- We accept both handwritten and electronic submissions. So you can choose what is more convenient for you. In any case, you should specify full names and immatriculation IDs of all team members. Obviously, programming tasks you can submit only electronically.
- Handwritten submissions should be submitted in the lecture hall of Monday's lecture (before the lecture starts).
- Electronic submissions should be zipped, containing the m-files (`Basis` etc.), your plots (png files) and the matlab data files (.mat) and emailed to the corresponding tutor:

  a. Apratim Bhattacharyya (Wednesday 8-10): abhattac@mpi-inf.mpg.de
  b. Maksym Andriushchenko (Thursday 8-10): s8mmandr@stud.uni-saarland.de
  c. Max Losch (Friday 16-18): mlosch@mpi-inf.mpg.de

  If not all 3 students belong to the same tutorial group, then you should email your submission to **only** one tutor (e.g. to the tutor of the first author of your homework), so please do not put other tutors in copy of the email.

  The email subject must have the following form: "[ML18/19 Exercise] Sheet X", where X is the number of the current exercise sheet. Then please specify in the email full names and immatriculation IDs of all team members. Then please attach all your files as a single zip archive, which consists of your immatriculation IDs, e.g. "2561234_2561235_2561236.zip".
- Reminder: you should submit in groups of 3. Otherwise, we will later on merge the groups smaller than 3 students.