Technische Hochschule Ostwestfalen-Lippe
Fachbereich Elektrotechnik und Technische Informatik
Prof. Dr. Helene Dörksen

Exercises and Labs 7
for Lecture "Authentication „ (M.Sc.)

Lab 7.1 Consider datasets *xyz.dat* provided previously.
Corresponding scripts are *xyz_load.m*

Consider also the new dataset: *planning_relax.dat* and the corresponding script *planning_relax_load.dat*.

More details about datasets are found on
https://archive.ics.uci.edu/ml/index.php

Additional Matlab functions are provided.

For given datasets (start with *planning_relax.dat*)

a) Classify the datasets within *LDA* and *SVM* approaches in original form and within normalisation (i.e. preprocess features to have values $\mu = 0$, $\sigma^2 = 1$). Compare the corresponding accuracies.

b) Compare classification accuracies for data in original form and within softmatrix scaling (i.e. preprocess features with respect to the formula for softmatrix scaling).

c) Perform estimation of classification accuracy for *LDA* and *SVM* (if calculations possible) based on *Holdout Method* for train / test splits by 50 / 50%, 66.6 / 33.3% and 33.3 / 66.6%. Compare results for data in original form as well as within random permutation of objects.

d) Perform estimation of classification accuracy (i.e. averaged mean and averaged standard deviation) for *LDA* and *SVM* (if calculations possible) based on $k-$*fold cross-validation* for $k = 2, 5, 10$ and compare the results.

e) For *hepatitis.dat* and *seeds.dat* perform estimation of classification accuracy (i.e. averaged mean and averaged standard deviation) for *LDA* and *SVM* (if calculations possible) based on *leave-one-out* method.

f) Compare performance of *LDA* and *SVM* within *McNemar's test*. Here, values: e_{00}, e_{10}, e_{10} and e_{11} are based on the accuracy results for test sets for methods, e.g. *Holdout Method* or averaged accuracies for $k-$*fold cross-validation* or *leave-one-out*.

Lab 7.2 advanced self-study For given datasets compare performance of *LDA* and *SVM* (if calculations possible) within:

a) 5×2 *cv paired t test*;

b) 5×2 *cv paired F test*.