BIL443/553 Pattern Recognition
Homework 3

Exercise 1 (20%): Assume you have a set V of m real vectors, each with n dimensions. Let u be the mean vector of V. Prove that u is the vector with the least cumulative (Euclidian) distance to the vectors of V.

Exercise 2 (20%): Prove that XOR is not a linearly separable classification problem.

Exercise 3 (60%): select a dataset of any type of data available online, and train/test a SVM classifier with it using 5-fold cross validation. Conduct grid search to optimize your kernel parameters, and compare the performance of your classifier with linear, polynomial and RBF kernels (you can use software libraries/toolkits, e.g. libsvm).

Optional: watch this video on why/how the RBF kernel carries data to infinite dimensions:
https://www.youtube.com/watch?v=XUj5JbQihlU
(~ 25th minute)

Submit your report with your answers and findings.

+ 20 % points if the report is prepared with LateX.

Good luck.