Problem 1 (5 pt) The aim of this problem is to study how to generate Gaussian random variables.

- Generate 10000 samples from a standard normal distribution (zero mean and unit variance) using at least three methods we discussed in the class. Compare and comment on the merits and demerits of the various methods.
- Select your favorite method and extend it to draw 3 dimensional (multivariate) normal samples with mean $\mu = (0, 0, 0)$ and covariance $\Sigma = I$.
- Extend the above procedure to draw samples from an arbitrary 3 dimensional normal distribution with mean $\mu$ and covariance $\Sigma$. Describe clearly the procedure you employed. (Hint: You need to use the eigenvalue decomposition to solve this problem. We did not discuss this problem in the class).

Problem 2 (5 pt) Recall the regularized risk minimizing framework for deriving new classifiers. This time we will minimize the following objective function

$$J(w) = \frac{1}{2} ||w||^2 + \sum_{i=1}^{m} \max(0, 1 - y_i(w, \phi(x_i)))^2.$$ 

Clearly, the objective function is convex and smooth (check for yourself!). The aim here is to use coordinate descent to find the optimal $w$ by using coordinate descent. Unlike gradient descent and other procedures which
update all the coordinates of \( w \) at every iteration, the coordinate descent procedure updates only one randomly chosen coordinate per iteration (see details in the textbook). Your task is to

- Derive the one dimensional objective function in order to apply coordinate descent.
- Think about how you will perform a line search along the chosen coordinate.
- Implement your optimizer and test it on the ECML/PAKDD discovery challenge dataset that you used in the previous HW.
- Compare the coordinate descent procedure with any other optimizer of your choice (BFGS, Gradient Descent etc).
- Write a report to document your findings.