Name:
ID:
Date:
ITU, Computer Engineering Dept.
BLG527E, Machine Learning HW1
Due: October 30, 2016 23:00 through Ninova. NO LATE SUBMISSION WILL BE ACCEPTED. DO NOT SUBMIT THROUGH E-MAIL.

Grading: You must complete the table below according to what you expect to get out of each question.

		Q1	Q2	Q3	Q4	Q5	Total
Grade	Max	1	1	1	1	1	5 pts
	Expected						

Policy:
Please do your homeworks on your own. You are encouraged to discuss the questions with your class mates, but the code and the hw you submitted must be your own work. Cheating is highly discouraged for it could mean a zero or negative grade from the homework.
If a question is not clear, please let me know (via email or in class). Unless we indicate otherwise, do not use libraries for machine learning methods. When in doubt, email me.

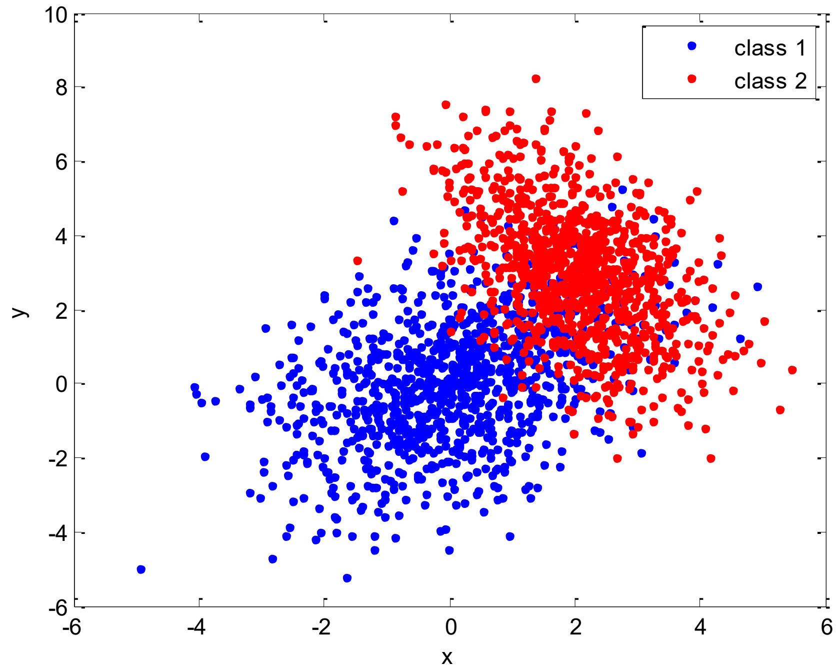
There will be 5 homeworks this term. Each hw is worth 5 points and each question will be evaluated on a 0/1 basis.

In order to be able to take the final exam for BLG527E you have to have a **weighted average score of 30 (over 100) for midterm and homeworks.** Otherwise you will get a VF from the course.

NO LATE SUBMISSIONS WILL BE ACCEPTED. PLEASE DO NOT EMAIL YOUR HWs.

QUESTIONS

You will use the following dataset given for this hw. The last column of the file represents the label (class 0 or class 1)



Q1) Examine the dataset. The number of features, the number of classes and their probabilities (P(Ci)). Plot the dataset. Compute the mean vectors and covariance matrices of the classes. Partition the dataset into 10 training and validation sets, preserving the class distributions in each fold.
http://www.csie.ntu.edu.tw/~b92109/course/Machine%20Learning/Cross-Validation.pdf

Design a quadratic discriminant classifier using different covariance matrices. Obtain the training and validation errors.

Draw the decision boundary on one fold.

Q2) Design a linear discriminant classifier using a shared covariance matrix. Obtain the training and validation errors.

Draw the decision boundary on one fold.

Q3) Generate a 2 dimensional 2 class dataset with 1000 examples for each class where

$\mu_1 = [1\ 2], \mu_2 = [3\ 5]$ and

$$\Sigma_1 = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}, \Sigma_2 = \begin{bmatrix} 1 & -0.8 \\ -0.8 & 3 \end{bmatrix}$$

plot the dataset.

In order to generate Gaussian dataset you can use the following steps:
(http://people.sc.fsu.edu/~jburkardt/f_src/normal_dataset/normal_dataset.html)

The multivariate normal distribution for the M dimensional vector X has the form:

pdf(X) = (2*pi*det(A))**(-M/2) * exp(-0.5*(X-MU)'*inverse(A)*(X-MU))

where MU is the mean vector, and A is a positive definite symmetric matrix called the variance-covariance matrix.
To create X, an MxN matrix containing N samples from this distribution, it is only necessary to

1- create an MxN vector Y, each of whose elements is a sample of the 1-dimensional normal distribution with mean 0 and variance 1;
2- determine the upper triangular Cholesky factor R of the matrix A, so that A = R' * R;
3- compute X = MU + R' * Y.

Find the Cov(X) in terms of MU + R' * Y and find the relation between A.

Plot the generated dataset as shown in Q1)

Q4) Design a quadratic discriminant classifier using the whole dataset obtained in Q3.

Draw the decision boundary.

Q5) Let X be random variable from the Binomial $B(m,\Theta)$ distribution where $P(X=x)=\binom{m}{x}\Theta^x(1-\Theta)^{m-x}$, and assume that you have a $Beta(\alpha,\beta)$ prior distribution for Θ. Given $x_1,x_2,\ldots x_N$ observations find the most probable estimate for Θ. (Hint: write down the posterior distribution using Bayesian Estimation and take the derivative with respect to Θ)