Istanbul Technical University- Spring 2017
BLG527E Machine Learning
Homework 1

Purpose: Better understanding of PR/ML basics.
Total worth: 6% of your grade.
Handed out: Tuesday, Feb 15, 2017.
Due: Monday, Mar 7, 2017 10:00pm. (through ninova!)
Instructor: Zehra Cataltepe (cataltepe@itu.edu.tr),
Assistants:
Mahiye Uluyağmur (muluyagmur@itu.edu.tr),
Hakan Gündüz (hakangunduz@itu.edu.tr)
Policy: Collaboration in the form of discussions is acceptable, but you should write your own answer/code by yourself. Cheating is highly discouraged for it could mean a zero or negative grade from the homework.
If a question is not clear, please let us know (via email, during office hour or in class).
Submission Instructions: Please submit through the class ninova site.
Please zip and upload all your files using filename studentID_HW1.zip. You must provide all functions you wrote with your zipped file. Functions you do not submit may cause you lose a portion of your grade. You must also include a .doc or pdf file with answers to the questions and how to call your matlab functions for each question so that we can run and check the results.

QUESTIONS:

Q1) [5 points][Central Limit Theorem]
Q1a) [2 points] Consider a univariate random variable x distributed according to

$$U\left(10-\frac{\sqrt{3}}{2}, 10+\frac{\sqrt{3}}{2}\right)$$

(i.e. minimum and maximum values x can take are $10-\frac{\sqrt{3}}{2}$ and $10+\frac{\sqrt{3}}{2}$ respectively.)

aa) Using Matlab or Python, draw **N=10** samples of x from the distribution, compute the **mean** and plot the **histogram of the means** for **500** different experiments.
ab) Repeat aa for **N=100** samples.

Q1b) [2 points] Repeat Q1a), but instead of **uniform density**, use **N(10,1)** (i.e. x is normally distributed with **mean 10** and **variance 1**.)

Q1c) [1 points] What are the differences and similarities between the plots you see in 1a) and 1b)? Why?

Hint: Scale x and y axis of plots in Q1a) and Q1b) to the same range in order to be able to compare them.

Q2: [7 points] [Bayesian Decision Theory]

Assume **a discriminant function** of the form:

$g_i(x) = \ln\left(P(x|w_i)\right) + \ln\left(P(w_i)\right)$ which achieves the minimum error classification.

Assume that $x \in R$, for **class 1:** $x \sim N(\mu_1, \sigma)$ and for **class 2:** $x \sim N(\mu_2, \sigma)$.
Also assume that $P(w_1) = P(w_2)$.

a) **[1 points]** Derive the discriminant functions $g_1(x)$ and $g_2(x)$.
b) **[1 points]** What is equation for the separating surface?
c) **[2 points]** For $\mu_1 = 5$, $\mu_2 = 15$ and $\sigma = 5$, plot the pdf's of the two classes inputs and also the separating surface.
d) **[1 points]** If $P(w_1)$ increases to **0.8**, where would the separating surface be?
e) **[2 points]** Generate random datasets for **c)** and **d)** and show histograms and the separating points on the histograms.