Name:
ID:
Date:
ITU, Computer Engineering Dept.
BLG527E, Machine Learning HW1
Due: September 30, 2014, 22:00 through Ninova.
Instructors: Zehra Cataltepe (cataltepe@itu.edu.tr), Yusuf Yaslan (yyaslan@itu.edu.tr)
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 us know (via email or in class). Unless we indicate otherwise, do not use libraries for machine learning methods. When in doubt, email us.

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.

Q1)
Make sure that you read Appendix A of the textbook and the resources on matrices, linear algebra and probability and statistics on ninova.

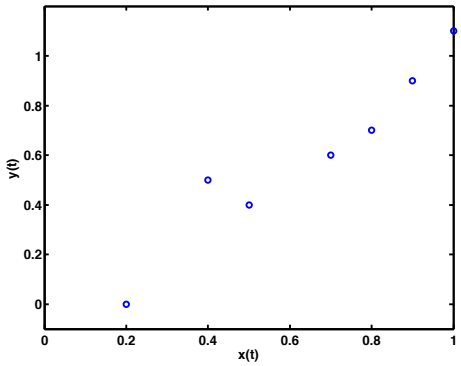
Given the table of joint probabilities between two discrete random variables X and Y evaluate:
Q1a) P(X=1|Y=-1)
Q1b) Are X and Y independent random variables, why or why not?

	X=-1	X=0	X=1
Y=-1	0.2	0.1	0.05
Y=0	0.05	0.1	0.12
Y=1	0.1	0.15	0.13

Q1c) What is the expected value of 5*X + 4*Y*Y?

Q2) Do this exercise in matlab/python/C/java. Submit your code in ninova with instructions on how to run it.
Q2a) Given the following data points, compute least squares regression line that passes through them.
Q2b) What are your predictions for x=-5, x=4?
Q2c) Compute the least squares model for a polynomial of degree 4 and compare the variances of the linear and degree 4 polynomial model by leaving one data point out at a time.

t	**1**	**2**	**3**	**4**	**5**	**6**	**7**
x(t)	**0.2**	**0.5**	**0.4**	**0.7**	**0.8**	**0.9**	**1.0**
y(t)	**0.0**	**0.4**	**0.5**	**0.6**	**0.7**	**0.9**	**1.1**



Q3) If **x** and **w** are d dimensional vectors and y is a real number, what is the derivative of E(**w**) with respect to **w**?

$g(\mathbf{x},\mathbf{w}) = x_1*w_1 + x_2*w_2 + \ldots + x_d*w_d$ (Model)
$X = \{(\mathbf{x}^1, y^1), \ldots, (\mathbf{x}^N, y^N)\}$ (Training Data)

$$E(\mathbf{w}) = \frac{1}{2N}\sum_{t=1}^{N}(g(\mathbf{x}^t,\mathbf{w}) - y^t)^2 \qquad \frac{dE}{d\mathbf{w}} = ?$$

Q4) Given a random sample $X = \{x^1, \ldots, x^N\}$ where each x^i is a nonnegative real number and are i.i.d. distributed according to the probability density function (pdf):

$$f_X(x) = k\frac{x}{a^2}e^{-\frac{x^2}{a^2}}$$

Q4a) What is the value of k to make $f_X(x)$ a valid pdf (a is a constant).
Q4b) Write the likelihood and log-likelihood functions for X in terms of the parameter a
Q4c) Find the maximum likelihood estimate for the parameter a

Q5) Given that çinekop and sarıkanat lengths (two different types of bluefish) are distributed according to Gaussians with means 11 and 18 respectively and both Gaussian have a variance of 3 and assuming that both fish have the same prior probability 0.5 of being caught, what is the best (maximizing the posterior probability of class given length) value of the threshold length that discriminates these two classes?