2413, Machine Learning, Assignment 1
Universität Bern
Due date: 2/11/2016

In this assignment your job is to solve a supervised classification problem. The objective is to learn how linear regression and gaussian discriminant analysis work in practice. You will need to implement linear classifiers based on these methods in MATLAB. You will also need to evaluate the performance of these methods from different aspects.

There are two datasets. The first one consists of face/non-face images and the second one consists of pedestrian/non-pedestrian images. In each datastes face/pedestrian has been indicated by ($y = 1$) and non-face/non-pedestrian has been indicated by ($y = 0$) images. To make the assignment more focused, we provide the script 'assignment1.m'. It handles the loading of the dataset and it shows some of the samples. It also divides the data into training and test sets, displays graphs with different evaluation metrics and shows the classifier output for some test samples. **In this assignment you need to upload a zip file to ILIAS which includes three .m files(gda.m, logistic_sgd.m, logistic_newton.m) and a pdf of your answers to the asked questions. The zip file's name must be FirstName_LastName.zip. Note that you need to use exact the same function signatures as they are provided. If your implementation requires auxiliary functions, you must implement that function inside the corresponding .m file.** The three .m files consists of a dummy implementation by default. The pdf file have to contain all your answers to the questions, tables and figures. We prefer pdf format (over docx) because it is universally supported on all platforms. Please also indicate your name in top of the first page in your pdf submission. Please do not print or display anything in the code (comment it out before submission).

1. In this exercise we consider the problem of face classification. There is a dataset attached to your homework containing face and non-face images, and their labels (1 for face and 0 for non-face). You are required to complete the script "assignment1.m". Implement a logistic regression classifier using the stochastic gradient descent algorithm. The specifications of the functions is provided below.

```
function theta = logistic_sgd_Your_Name( data, labels )
% logistic regression using stochastic gradient descent
% input: data: dxn matrix, where d is the dimensionality and
%                  n is the number of training samples. The column
%                  data(:,i) represents the training sample x^i
%           labels: 1xn row vector, labels(i) contains y^i
% output: theta: dx1 column vector, the parameters of the
%                       classifier
```

2. Plot a graph, where the horizontal axis shows the number of iterations and the vertical shows the objective function. Try different learning rates.

   - At what learning rate does the method converge?
   - What is the optimal learning rate?

3. Solve the logistic regression problem above by using Newton's method. The specifications are the same.

```
function theta = logistic_newton_Your_Name( data, labels )
% logistic regression using Newton's method
% input: data: dxn matrix, where d is the dimensionality and
%                  n is the number of training samples. The column
%                  data(:,i) represents the training sample x^i
%           labels: 1xn row vector, labels(i) contains y^i
% output: theta: dx1 column vector, the parameters of the
%                       classifier
```

4. Compare the results obtained by the two methods.

   - What do you expect to see (stochastic gradient vs. Newton) before running the experiment?
   - Which method takes less time to compute one step?
   - Which method requires fewer steps to converge?
   - Which method achieves better accuracy?

5. Solve the face classification problem with Gaussian Discriminant Analysis (GDA). Implement a function that learns the parameters of a multivariate normal distribution. The specification is the following.

```
function [ phi, mu0, mu1, Sigma ] = gda_Your_Name( data, labels )
% gaussian discriminant analysis
% input: data: dxn matrix, where d is the dimensionality and
%                  n is the number of training samples. The column
%                  data(:,i) represents the training sample x^i
%           labels: 1xn row vector, labels(i) contains y^i
% output: phi, mu0, mu1, Sigma: parameters of multivariate
%             gaussian distributions for positive and negative
%             samples
```

   Note that here the labels are different, 1 means the sample is a face and -1 means it is not.

6. Compare logistic regression and gaussian discriminant analysis.

   - Which method is more computationally efficient?
   - Which method achieves the best accuracy?