RPZ **January 21, 2016**

Name:

1. Formulate the logistic regression problem **(1 point)**. Derive the learning method as gradient descent of the log of the aposteriori probability **(2.5 points)**. Given the training set $T = \{(-1;-1), (-2;-1), (0;1), (2;1)\}$, starting point $w = (0, 0)$, and an initial step size of 1, do the first iteration of gradient descent to minimize the objective and give the step size that would be used in the 2nd iteration (Note: you should not perform the 2nd iteration). **(1.5 points)**.

2. Objects from two classes $k \in \{1, 2\}$ are classified given a single observation, a non-negative real number x. Conditional probability densities $p(x|k)$ are: $p(x|1) = 2(1-x)$, if $x \in [0, 1)$ and 0 elsewhere, $p(x|2) = \exp(-x)$. Find the decision strategy optimal for the
   (a) **(2 points)** the minmax formulation of the recognition task.
   (b) **(2 points)** the Neyman-Pearson formulation of the recognition task. The false negative rate (probability of misclassifying a dangerous object) is required to be below 0.1 (10%). Class 2 is considered dangerous.
   (c) **(1 point)** the Wald formulation of the recognition problem with the same threshold (0.1) on the classification error as in the previous task.

3. Compare the properties of Support Vector Machines and Decision trees. Consider: the type of observations it is suitable for, the size of the trainings set, generalisation properties, memory requirements, the speed of learning, the influence of outliers during training and in testing, and any other property not listed. **(Each considered property is for 1 point up to total of 5 points)**

4. Describe the AdaBoost learning algorithm **(2 points)**. Derive the formula for α_t **(1.5 points)**. A training set $T = \{(\mathbf{x}_i; k_i)\}$, where $\mathbf{x}_i \in \mathbb{R}^2$ and $k \in \{-1, 1\}$ is given as $T = \{(0, 0; 1), (1, 1; 1), (2, 2; 1), (0, 1; -1), (0, 3; -1), (2, 3; -1)\}$ and the following weak classifiers are available: $h_{1m} = \delta(x_1 > m/2)$, $h_{2m} = \delta(x_1 \leq m/2)$, $h_{3m} = \delta(x_2 > m/2)$, and $h_{4m} = \delta(x_2 \leq m/2)$, where $m \in \mathbb{Z}$ is a parameter of the classifiers and $\delta(C)$ equals 1 if the condition C is satisfied and -1 otherwise. Do two steps of learning, i.e. find a strong classifier of the length two. **(1.5 points)**: