Deadline: Th. Nov. 15, 10:00 am

Exercise 1 (Logistic Regression)**.** Consider the following data sets:

(a) "X2"

x_1	x_2	y
-1	-1	0
-1	1	1
1	-1	0
1	1	1

(b) "XOR"

x_1	x_2	y
-1	-1	0
-1	1	1
1	-1	1
1	1	0

(c) "AND"

x_1	x_2	y
-1	-1	0
-1	1	0
1	-1	0
1	1	1

1. (3p) Describe in your own words the difference between classification and regression problems.
2. (3p) Make plots of the datasets, including the class label. For which of these datasets can 100% correct classification be achieved using logistic regression?
3. (3p) Train a logistic regression classifier for each dataset by performing one iteration of gradient descent. Initialize with bias 0, $\beta_1 = 1$, $\beta_2 = 0$ and use step length $\alpha = 1$. Report the updated parameters in each case.
4. (1p) For each dataset draw the decision boundary of the classifier before and after performing the gradient descent step.

Exercise 2 (Multiclass classification)**.** In this exercise we want to perfom Linear Discrimant Analysis (LDA) on the famous Iris data set. Obtain the file 'iris.data' from the learnweb or from the UCI-repository[1]

1. (2p) Make scatter plots of sepal-length/sepal-width and petal-length/petal-width
2. (4p) For each class, compute the mean μ_k and covariance matrix Σ_k.
3. (4p) Perform the prediction. How many data points are predicted correctly? Highlight the false predictions in your scatter plots.

[1] https://archive.ics.uci.edu/ml/datasets/iris