Machine Learning – WS'12
Exercise-5

Prof. Dr. Dr. Lars Schmidt-Thieme, Umer Khan
Information Systems and Machine Learning Lab (ISMLL),
University of Hildesheim

Problem-1:

Given the following data:

y	x	y	x
0	9.5	1	11.1
0	9.6	1	11.1
0	9.7	1	11.1
0	9.8	1	11.5
0	9.9	1	11.8
0	10.5	1	11.9
0	11.0	1	12.1
0	11.2	1	12.2
0	11.5	1	12.5
0	11.7	1	12.6
0	12.1	1	12.6

a) Compute linear regression $y = \beta_0 + \beta_1 x$ for target variable y. Also compute the cost function (MSE)

b) Compute logistic regression $logit(y) = \beta_0 + \beta_1 x$ for target variable y. To compute this, use *Iteratively Reweighted Least Squares* (IRLS) algorithm. Stop after 2nd iteration and compare the parameters of 1st and 2nd iteration. Finally compare them with final estimates of $\hat{\beta}_0 = -23.35,\ \hat{\beta}_1 = 2.064$.

c) Plot the decision boundary from a) and b). Discuss the results. Why the decision curve of logistic regression is not linear´?

Hint: For IRLS, see Newton's Method slides of lecture.

Problem 2:

Maximum likelihood principle for estimation of parameters for linear and logistic regression models is considered the best in and for sufficiently big data sets. Pre-requisite for this is that model is probabilistic and it has finite number of parameters.

a) Consider the probabilistic and parametric model assumption of linear and logistic regression. Describe the difference. How many parameters they have?

b) Consider the following model: The observed data $y_i \in R \ \forall_i = i, \ldots, n$, are n independently sampled random instances of a normal distribution $y_i \sim N(\mu, \sigma^2)$, where σ^2 is known, but the second parameter μ (mean) is unknown. Use Maximum likelihood estimation to derive the mean of model.

Hint: Write the joint probability and consider the logarithm of this probability. And derive this expression with respect to μ, set it equal to zero and solve for μ.

Problem-3:

Find the first and second derivatives of the log-likelihood for logistic regression with one predictor variable. Explicitly write out the formula for doing one step of Newton's method. Explain how this relates to re-weighted least squares.