Machine Learning 2
Exercise Sheet 10

Prof. Dr. Dr. Lars Schmidt-Thieme, Nicolas Schilling
Information Systems and Machine Learning Lab
University of Hildesheim

June 21st, 2016

Submission until June 28th, 18.00 to schilling@ismll.de

Exercise 16: Laplace Priors (7 Points)

a) Given is a linear Regression model with Laplace prior:

$$p(y \mid x, \theta, \sigma_y^2) = \mathcal{N}(y \mid \theta^\top x, \sigma_y^2)$$
$$p(\theta) = \mathrm{Lap}(\theta \mid 0, 1/\gamma)$$

What is the purpose of the prior? Why do we use it and which other priors are commonly used?

b) Compute the MAP estimate

$$\theta^\star = \max_\theta \ \log(p(\theta \mid y))$$

of a linear Regression with Laplace Prior. Show that it basically optimizes the same objective function as L1 regression.

Exercise 17: Automatic Relevance Determination (7 Points)

a) What is the difference of ARD compared to classic regularized Regression?

b) Implement the EM version of ARD (inference for the β, update σ_y and all σ_{β_m}) for a data set of your choice.