N. Bertschinger
M. Kaschube
V. Ramesh

Machine Learning I
Exercise Sheet 9
Due on Mon, January 16, 12:15

If you turn in your solutions via email, please send to:
kamyshanska@fias.uni-frankfurt.de
Include [ML1-1617] in the subject line, turn in a single PDF file and include your name in the file name.

Exercise 1. *Verify the relation*

$$\frac{d\sigma}{da} = \sigma(1-\sigma)$$

for the derivative of the logistic sigmoid function defined by

$$\sigma(a) = \frac{1}{1+\exp(-a)}.$$

2 points

Exercise 2. *We consider the problem of two-class classification. The posterior probability of class C_1 can be written as a logistic sigmoid acting on a linear function of the feature vector $\boldsymbol{\phi}$ so that $p(C_1|\boldsymbol{\phi}) = y(\boldsymbol{\phi}) = \sigma(\mathbf{w}^T\boldsymbol{\phi})$ with $p(C_2|\boldsymbol{\phi}) = 1 - p(C_1|\boldsymbol{\phi})$ and weights $\mathbf{w}$.*

For a data set $\{\boldsymbol{\phi}_n, t_n\}$, where $t_n \in \{0,1\}$ and $\boldsymbol{\phi}_n = \boldsymbol{\phi}(\mathbf{x}_n)$ with $n = 1, \cdots, N$, the likelihood function can be written

$$p(\mathbf{t}|\mathbf{w}) = \prod_{n=1}^{N} y_n^{t_n}\{1-y_n\}^{1-t_n}$$

where $\mathbf{t} = (t_1, \cdots, t_N)^T$ and $y_n = p(C_1|\boldsymbol{\phi}_n)$. By taking the negative logarithm of the likelihood, the error function can be defined as

$$E(\mathbf{w}) = -\ln p(\mathbf{t}|\mathbf{w}) = -\sum_{n=1}^{N}\{t_n \ln y_n + (1-t_n)\ln(1-y_n)\}.$$

By making use of the expression for the derivative of the logistic sigmoid from exercise 1, show that the derivative of the error function for the logistic regression model is given by

$$\nabla E(\mathbf{w}) = \sum_{n=1}^{N}(y_n - t_n)\boldsymbol{\phi}_n.$$

4 points

Exercise 3. *Show that for a linearly separable data set, the maximum likelihood solution for the logistic regression model is obtained by finding a vector* $\mathbf{w}$ *whose decision boundary* $\mathbf{w}^{\mathbf{T}}\boldsymbol{\phi}(\mathbf{x}) = \mathbf{0}$ *separates the classes and then taking the magnitude of* $\mathbf{w}$ *to infinity.*

4 points