Machine Learning I WS 2016/17

Lecturer: Prof. Dr. Bethge

4. Logistic Regression, Linear Regression

Task 1: Derive Logistic Regression from Gaussian Bayes classifier

Derive Logistic Regression from a Gaussian Bayes classifier under the assumption that $P(\vec{X}|Y = y_i) = N(\vec{\mu}_i, \Sigma)$, i.e. the joint distribution of $(\vec{X}, Y)$ is a mixture of Gaussians, where both mixture components have the same covariance matrix, Σ, but different means $\vec{\mu}_i$. In other words: show that the conditional likelihood can be written as

$$P(Y|\vec{X} = \vec{x}) = \frac{1}{1 + \exp(w_0 + \sum_i w_i x_i)}$$

Task 2: Fit Logistic Regression

We will continue exploring the problem of topic classification using the Reuters-21578 dataset from last week. Specifically, we will use Logistic Regression to classify documents as belonging to the selected topic or not. Implement a Logistic Regression classifier using gradient ascent on the conditional log-likelihood.

Task 3: Compare Naïve Bayes and Logistic Regression

Compare the performance of your Logistic Regression classifier to the Naïve Bayes classifier. Explore how the performance of both depends on the size of the training set. To do so, select a random subset of examples from the training set to create smaller training sets of size 8, 16, 32, 64, …, 4096. For each training set size, run a few independent runs with new random draws of training examples from then full training set. For both classifiers, plot the average performance (percent correct) on the (full) test set as a function of the size of your training set.

Task 4: Linear regression

Derive the closed-form solution for the weights in linear regression, starting from the assumption that the conditional likelihood is

$$P(y|x; w) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\frac{\left(y - f(x; w)\right)^2}{\sigma^2}\right)$$