Übungen zur Vorlesung
„Maschinelles Lernen und Künstliche Intelligenz aus Sicht der Stochastik“

Sommersemester 2017, Blatt 7

Abgabetermin: 27.06.2017, spätestens zu Beginn der Vorlesung

Aufgabe 1 (4 Punkte)

Betrachten Sie ein Netzwerk mit mehrdimensionalem Input $\mathbf{x}$ und Output $a \in R$ definiert durch

$$y_j := \sigma\left(\sum_i u_{ji} x_i\right)$$

$$z_k := \sigma\left(\sum_j v_{kj} y_j\right)$$

$$a := \sum_k w_k z_k$$

1. Berechnen Sie $\frac{\partial a}{\partial w_k}, \frac{\partial a}{\partial v_{kj}}$ und $\frac{\partial a}{\partial u_{ji}}$
2. Angenommen es gibt M Input Trainingsdaten $\mathbf{x}^1, \ldots, \mathbf{x}^M$ mit Zielvariablen $s^1, \ldots, s_M$. Sei $J = \frac{1}{M} \sum\limits_m \frac{1}{2}(s^m - a^m)^2$.

   Berechnen Sie $\frac{\partial J}{\partial u_{ji}}$

Aufgabe 2 (4 Punkte)

Zeigen Sie das für ein Model in dem $Y = f(X) + \epsilon$ mit $\mathbb{E}[\epsilon] = 0$ und $\mathbb{V}[\epsilon] = \sigma^2$ für die Vorhersage $\hat{f}(X)$ an der Stelle $X = x_0$ gilt:

$$\mathbb{E}[(Y - \hat{f}(x_0))^2 | X = x_0] = \sigma^2 + \text{Bias}^2(\hat{f}(x_0)) + \mathbb{V}[\hat{f}(x_0)]$$

Berechnen Sie obige Terme für die Methode $\hat{f}(x_0) = \frac{1}{k} \sum\limits_{l=1}^{k} x_{(l)}$, wobei $x_{(i)}$ der i-tnächste Punkt zu x_0 in den Trainingsdaten (x_i, y_i), $i = 1, \ldots N$ ist. Sie können dafür annehmen das die x_i fixiert sind und die Erwartung nur über y_i bilden.

Für welches k werden die Terme minimiert.

Aufgabe 3 (4 Punkte)

Wenden Sie Regularisieren durch Bagging auf Ihre Klassifizierung der Trainingsdaten aus Aufgabe 3 von Blatt 6 an. Versuchen Sie Ihre Vorhersagegenauigkeit auf den Testdaten dadurch zu verbessern. Probieren Sie dabei verschieden Größen für die Baggingsamples aus.