

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

Sommersemester 2017, Blatt 2

Abgabetermin: 09.05.2017, spätestens zu Beginn der Vorlesung

(Bitte geben Sie auf jedem Lösungsblatt Ihren Namen an)

Bitte maximal zu zweit abgeben!

Bearbeiten Sie R-Aufgaben grundsätzlich mit „R Markdown“ und senden Sie sie an [maschinelleslernen17@gmail.com](mailto:maschinelleslernen17@gmail.com).

## Aufgabe 1 (4 Punkte)

Seien  $X_i \sim \mathcal{N}(\mu, \sigma^2)$  unabhängig identisch verteilt für  $i = 1, \dots, N$ . Berechnen Sie den maximum likelihood Schätzer für  $\sigma$ .

## Aufgabe 2 (4 Punkte)

Erstellen Sie in R eine Klassifizierung der Challenger O-ring data

1. mit logistischer Regression
2. mit Support Vector Machines
3. begrenzen Sie die Summe der slack variablen in unterschiedlicher Stärke und entscheiden Sie sich mit Kreuzvalidierung für den finalen Schätzer.
4. visualisieren Sie Ihr Ergebnis ansprechend in R

## Aufgabe 3 (4 Punkte)

Erstellen Sie mit `set.seed(25)` zehn zweidimensional standardnormalverteilte Punkte  $p_1, \dots, p_{10}$  mit Mittelwert  $(0, 1)$  in Orange und zehn zweidimensional standardnormalverteilte Punkte  $q_1, \dots, q_{10}$  mit Mittelwert  $(1, 0)$  in Blau.

```
library(MASS)
```

```
set.seed(25)
```

```
m1 = mvrnorm(10, c(1, 0), diag(2))
```

```
m2 = mvrnorm(10, c(0, 1), diag(2))
```

Erstellen Sie außerdem eine Stichprobe von je hundert Punkten in Orange und Blau, indem Sie aus  $\{p_1, \dots, p_{10}\}$  und  $\{q_1, \dots, q_{10}\}$  mit Zurücklegen Punkte ziehen und zu diesen iid normalverteilte Fehler  $\epsilon_i \sim \mathcal{N}((0, 0), \frac{1}{5}E_2)$  addieren.

1. Plotten Sie die simulierten Punkte.
2. Bestimmen und markieren Sie die optimale Bayes-Entscheidungsgrenze.
3. Zeichnen Sie die Klassifizierungsgrenze und die Margins einer geeigneten Support Vector Machine ein.