

## Übungsblatt 2

Regression und Klassifikation mit linearen Modellen

Abgabe online auf ILIAS bis 20. November 2025, 12:00 Uhr

Ladet für die Abgabe bitte ein PDF mit der eingescannten Lösung für Aufgabe 1 sowie das bearbeitete Notebook für Aufgabe 2 hoch.

### Aufgabe 1 *Lineare Regression*

(4 Punkte)

Gegeben sind die folgenden 4 Datenpunkte  $(x_i, y_i)$

$$(1, 1), (2, 3), (4, 3), (5, 5)$$

welche jeweils aus einer Eingangsgröße  $x_i$  und einem Funktionswert  $y_i$  bestehen. Ihre Aufgabe ist es den Gewichtsvektor  $\mathbf{w} = (w_1 \ w_0)$  eines linearen Regressionsmodells zu bestimmen welches auf diesen Datenpunkten trainiert wird:

$$y_{pred} = w_1 \cdot x + w_0$$

Geben sie den finalen Gewichtsvektor  $\mathbf{w}$  sowie ihren vollständigen Rechenweg an.

## Aufgabe 2 Klassifikation

(4 Punkte)

Es soll ein Modell trainiert werden welche eine seltene Krankheit vorhersagen soll. Die Krankheit tritt nur etwa bei 2 von 100 Personen auf. Das Modell soll auf einem Datensatz mit 10.000 Personen trainiert werden, die zufällig aus der Bevölkerung ausgewählt wurden. Durch initiale Tests wurde herausgefunden dass die Krankheit anhand von nur 2 Merkmalen  $x_1, x_2$  vorhergesagt werden kann.

- (a) Erklären Sie warum es irreführend sein kann, die Accuracy als Metrik für die Bewertung des Modells zu verwenden und nennen sie eine alternative Metrik welche hier eine bessere Aussagekraft hat. (2 Punkte)
- (b) Die folgende Abbildung zeigt die Verteilung der Datenpunkte im Merkmalsraum. Zeichnen Sie skizzenhaft die Entscheidungsgrenze eines logistischen Regressionsmodells ein, welches auf diesen Daten trainiert wurde. (2 Punkte)

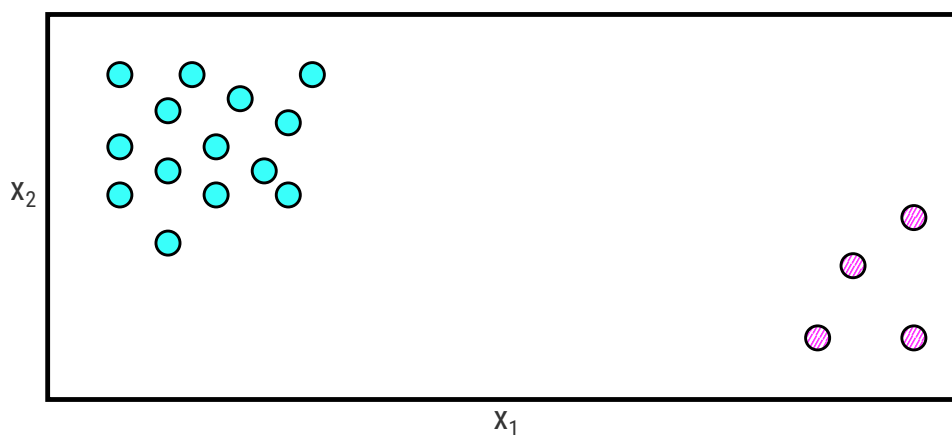


Figure 1: Verteilung der Datenpunkte im Merkmalsraum

### **Aufgabe 3** *Notebook*

(12 Punkte)

Diese Aufgabe ist eine Programmieraufgabe. Bearbeitet das Notebook wie unter Übungsbetrieb im Vorlesungsinhalt Verzeichnis auf ILIAS beschrieben und folgt den Anweisungen dort.