Arbeitsblatt: Logistische Regression

Binäre Klassifikation mit Wahrscheinlichkeiten

Informatik

1. Lernziel

Ich kann

- den Zweck der logistischen Regression erklären (Ja/Nein-Vorhersage)
- die Sigmoid-Funktion anwenden, um eine Wahrscheinlichkeit zu berechnen
- Kategorien vorhersagen und einfache Metriken (Accuracy) interpretieren

2. Was macht die logistische Regression?

Die Logistische Regression wird für **binäre Klassifikationsaufgaben** genutzt, also wenn die Antwort nur zwei Werte annehmen kann, z.B.: Ja/Nein, 0/1, Wahr/Falsch.

Analog zur linearen Regression wird auch hier eine Gerade berechnet, aber das Ergebnis wird **in eine Wahrscheinlichkeit umgewandelt** (z.B. die Wahrscheinlichkeit, ob ein:e Schüler:in den Test besteht oder nicht). Somit sagt die logistische Regression **keine direkten Kategorien** vorher, sondern **Wahrscheinlichkeiten** für die Kategorien.

Bedeutung der Wahrscheinlichkeit p:

- Beispiel: $p = 0.8$ bedeutet, dass das Modell zu 80% sicher ist, dass die Kategorie 1 (z.B. „bestanden“) zutrifft.
- Beispiel: $p = 0.3$ bedeutet, dass das Modell zu 70% sicher ist, dass die Kategorie 0 (z.B. „nicht bestanden“) zutrifft.

Nach Berechnung der Wahrscheinlichkeit wird eine **Schwelle** (z.B. 0.5) genutzt, um die Wahrscheinlichkeit in eine konkrete Kategorie umzuwandeln.

2.1 Ablauf in drei Schritten

1) Rechne eine **Gerade**: $z = a\,x + b$.
2) Setze z in die **Sigmoid-Funktion** $\sigma(z)$ ein und erhalte eine Wahrscheinlichkeit $p \in (0, 1)$.
3) Setze eine **Schwelle** (oft 0.5), um aus p eine Kategorie zu machen.

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

Anmerkungen:

- **Entscheidungsgrenze:** $p = 0.5$ entspricht $z = 0$.
- **Typische Regel:** $\hat{y} = 1$ falls $p \geq 0.5$, sonst $\hat{y} = 0$.
- **Schwelle anpassbar:** Höhere Schwelle = strenger, niedrigere Schwelle = grosszügiger.

3. Mini-Übung: Sigmoid skizzieren

- Zeichne die Kurve von $\sigma(z)$ für $z \in [-6, 6]$ in das Gitter unten.
- Berechne selbst die gerundeten Werte $\sigma(z)$ für die ganzzahligen Punkte $z = -6, \ldots, 6$ und trage sie in die Tabelle ein.

z	$\sigma(z)$
-6	
-5	
-4	
-3	
-2	
-1	
0	
1	
2	
3	
4	
5	
6	

- Zeige, dass die Kurve bei 0.5 „umkippt“ und sich bei 0 bzw. 1 asymptotisch annähert.

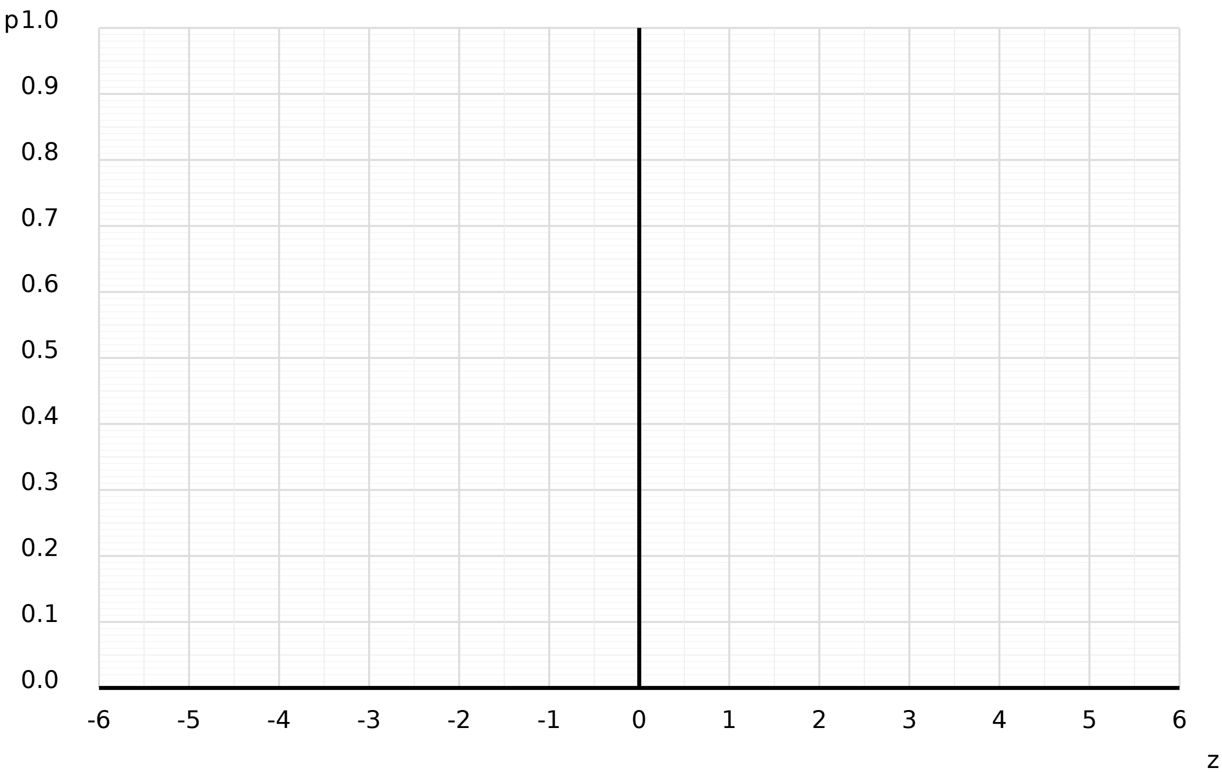

Abbildung 1: Koordinatengitter (nutzen für Sigmoid)

Beispiel 1: Besteht man den Test?

Gegebenes Modell (aus Training gelernt):

$$p = \sigma(1.2\,x - 2.0)$$

Interpretation: x = Lernzeit in Stunden, p = Wahrscheinlichkeit, den Test zu bestehen (Kategorie 1).

Datenpunkte

Lernzeit x (h)	Tatsächliches Label y (0 = nicht bestanden, 1 = bestanden)
0.5	0
1.0	0
1.5	0
2.0	0
2.5	1
3.0	1

Aufgabe 1a: Wahrscheinlichkeiten berechnen

Fülle die Tabelle aus (Zwischenschritte runden auf 3 Nachkommastellen):

x	$z = 1.2x - 2.0$	$p = \sigma(z)$	$\hat{y}$ (Schwelle 0.5)
0.5			
1.0			
1.5			
2.0			
2.5			
3.0			

Aufgabe 1b: Accuracy bestimmen

- Trage $\hat{y}$ aus der Tabelle ein und vergleiche mit y.
- Zähle die **korrekt** klassifizierten Beispiele.
- **Accuracy = (korrekte Vorhersagen) / (Anzahl Beispiele)**.

Accuracy =

Beispiel 2: Zwei Modelle vergleichen

Gleiche Datenpunkte wie oben.

Modelle

- **Modell A:** $p_A = \sigma(1.0\,x - 1.8)$
- **Modell B:** $p_B = \sigma(1.4\,x - 2.9)$

Aufgabe 2a: Wahrscheinlichkeiten & Kategorien

Berechne für beide Modelle p und $\hat{y}$ (Schwelle 0.5) und fülle die Tabellen.

Modell A

x	y	$z_A = 1.0x - 1.8$	p_A	$\hat{y}_A$
0.5	0			
1.0	0			
1.5	0			
2.0	0			
2.5	1			
3.0	1			

Modell B

x	y	$z_B = 1.4x - 2.9$	p_B	$\hat{y}_B$
0.5	0			
1.0	0			
1.5	0			
2.0	0			
2.5	1			
3.0	1			

Aufgabe 2b: Accuracy vergleichen

- Berechne die Accuracy von Modell A und B.
- Welches Modell ist besser? Gibt es Punkte, die ein Modell richtig und das andere falsch hat?

Accuracy A =

Accuracy B =

Aufgabe 2c: Kurze Interpretation (2–4 Sätze)

1. Welches Modell würdest du wählen und warum?
2. Wie könnte man die Schwelle 0.5 verändern, um mehr/ weniger falsch kategorisierte Zuordnungen zu bekommen?

Reflexion

1. In welchen Situationen ist die logistische Regression besser geeignet als die lineare Regression?
2. Warum sind Wahrscheinlichkeiten hilfreicher als nur harte 0/1-Entscheidungen?