

Klassifikation mit Wahrscheinlichkeiten

Wir lernen, wie man mit W schätzt, daraus Entscheidungen abzuleitet, Schwellenwerte wählt, Modelle fair bewertet, Klassengleichgewichte behandelt, und nicht lineare Entscheidungsgrenzen vergleicht.

A1. Aufgabe & Datenbild.

- Wir betrachten eine linäre Klassifikation $\{0, 1\}$.
- Eingaben sind Merkmale in einer Tabelle; jedes Beispiel gehört genau zu einer Klasse.
- Ziel ist eine Funktion die für jedes Beispiel eine W für Klasse „1“ liefert.

A2. Logistische Regression als W -Modell.

- Wir modellieren die s.g. logit-Funktion als lineare Funktion der Merkmale.
- Sei $\pi(x)$ die W für Klasse 1:

$$\log \frac{\pi(x)}{1-\pi(x)} = \beta_0 + \beta_1 x_1 + \dots + \beta_n x_n$$

Das Logarithmus der Quotient $\frac{p(1)}{p(0)} = \frac{\pi(x)}{1-\pi(x)}$ wird als lineare Kombination der Merkmale dargestellt.

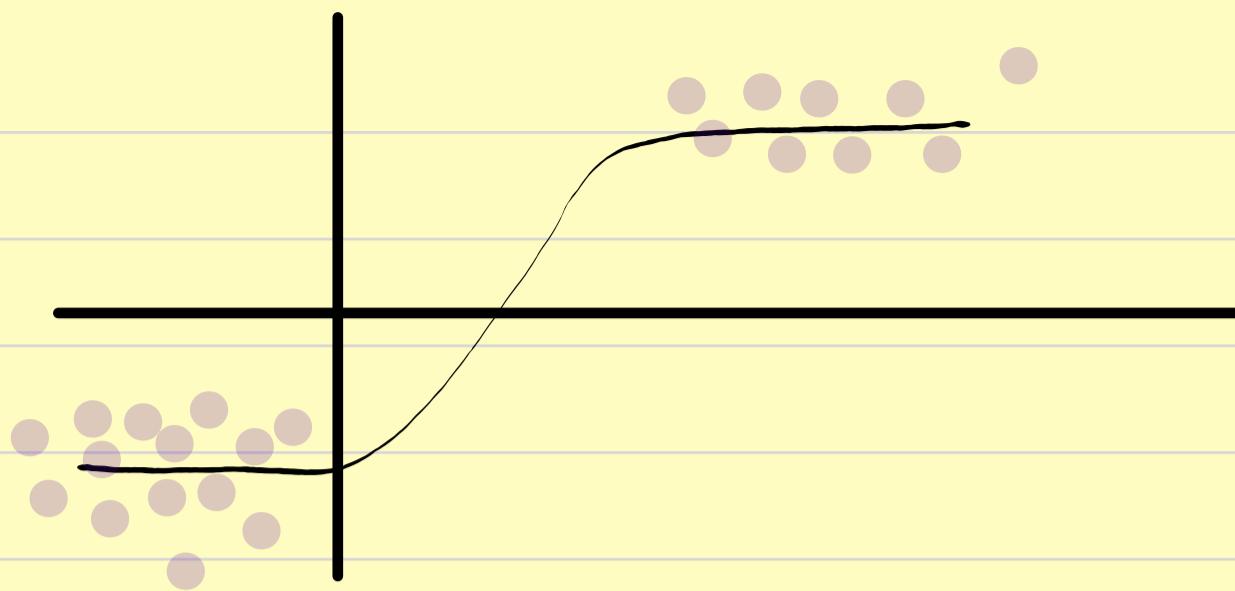
Daraus folgt die s.g. Sigmoid-Funktion

$$\pi(x) = \frac{1}{1 + \exp^{-(\beta_0 + \sum \beta_j x_j)}}$$

Schätzung der Gewichte erfolgt durch Maximierung der Likelihood: L.

$$L(\beta_0, \beta_1, \dots, \beta_n) = \prod_{i=1}^n \pi(x_i)^{y_i} (1 - \pi(x_i))^{1-y_i}$$

Interpretation: bei normierten Merkmalen, steht $\exp(\beta_j)$ für den Faktor, um den sich die Chancen auf Klasse 1 ändern, wenn x_j um eine Std. Varierung steigt, während andere Merkmale konstant bleiben.



A3. Wahl des Schwellenwertes

- Eine Vorhersage ist zunächst eine W.
Eine Klassenzuordnung entsteht erst durch einen Schwellenwert.
Standard ist 0'5.

A4. Verwechslungsmatrix & Grund-Metriken

- V-Matrix. 4-Felder. Richtig Positiv TP
Falsch Positiv FP

	0	430	30
	1	25	217

Wirklichkeit

- Richtig Negativ TN
Falsch Negativ FN

Schwelle 0'5

$$\sum = 430 + 30 + 25 + 217 = 702$$

$$TP = \sum \bar{1} \{ \hat{y}=1, y=1 \}$$

$$FP = \sum \bar{1} \{ \hat{y}=1, y=0 \}$$

$$TN = \sum \bar{1} \{ \hat{y}=0, y=0 \}$$

$$FN = \sum \bar{1} \{ \hat{y}=0, y=1 \}$$

$$TP = \frac{217}{702} = 0'309 \quad FP = \frac{25}{702} = 0'035$$

$$TN = \frac{430}{702} = 0'613 \quad FN = \frac{30}{702} = 0'043$$

	0	1
0	0'613	0'043
1	0'035	0'309

• **Treffergenauigkeit:** Anteil der korrekten Vorphersagen: Kann bei unausgewogenen Klassen trügen.

$$TS = \frac{TP+TN}{TP+TN+FP+FN}$$

• **Präzision:** Anteil wahrer positive Fälle unter allen als positiv vorhergesagten Fällen.

$$P = \frac{TP}{TP+FP}$$

• **Trefferquote:** Anteil wahrer positive Fälle, die korrekt als positiv erkannt werden.

$$Tq = \frac{TP}{TP+FN}$$

• F1-Wert: Mittel aus Präzision & Treffergüte.

$$F1 = 2 \cdot \frac{\text{Präz. Tq}}{\text{Präz} + \text{Tq}}$$

• Spezifität:

$$\mathcal{S}_p = \frac{TN}{TN+FP}$$

• Fehleralarmquote

$$\mathcal{F}_A = \frac{FP}{FP+TN}$$

• Brier-Wert: mittlere quadratische Abweichung zw. vorhergesagten W. & den beobachteten Klassen. Misst die Güte der W. selbst.

$$\text{Brier} = \frac{1}{m} \sum_{i=1}^m (\hat{\pi}_i - y_i)^2$$

