Machine Learning

6-10-25
(1.te Einheit)

① Lineare Modelle – Lineare Diskriminanz Analyse (LDA)

Wir betrachten Modelle der Form

$$(\ast)\quad \underbrace{p(\overset{\text{Klassen ID}}{y=c} \mid x, \theta)}_{\text{A-posteri Wahrscheinlichkeit der Klasse } c\text{, gegeben Datenpunkt } x \text{ und fixen Modellparametern } \theta} = \frac{p(x \mid y=c, \theta) \cdot p(y=c)}{\sum_k p(x \mid y=k, \theta) \cdot p(y=k)}$$

Label/Target, Daten, Modellparameter

A-posteri Wahrscheinlichkeit der Klasse c, gegeben Datenpunkt x und fixen Modell-parametern θ

$p(x \mid y=c, \theta)$ .... Klassenbedingte Wahrscheinlichkeitsdichte, gegeben Klasse c und fixen Modellparametern

$p(y=c)$ .... A-priori Wahrscheinlichkeit der Klasse c

Zu gegebenem x (Datenpunkt) und fixem θ, erhalten wir als Entscheidungsregel:

$$\boxed{k^* = \arg\max_c \, p(y=c \mid x, \theta)}$$

<u>Beispiel</u>: Wir betrachten den Fall $\underbrace{x \in \mathbb{R}^d}_{\text{Datenpunkt}}$ und $\underbrace{y \in \{0,1\}}_{\text{Label}}$.

<u>Annahme 1</u>: $p(x \mid y=c, \theta)$ lässt sich mittels einer multivariaten Normalverteilung beschreiben.

$$p(x \mid y=c, \theta_c) = (2\pi)^{-d/2} \cdot \sqrt{\det(\Sigma_c)} \cdot \exp\left(-\frac{1}{2} \cdot (x-\mu_c)^T \Sigma_c^{-1} (x-\mu_c)\right)$$

Multivariate Normalverteilung mit Kovarianz Matrix $\underbrace{\Sigma_c}_{d \times d \text{ Matrix}}$ und Lokationsparameter $\mu_c \in \mathbb{R}^d$

mit $\theta_c = \{\Sigma_c, \mu_c\}$.

Annahme 2: $\underbrace{\Sigma_1 = \Sigma_0 = \Sigma}_{\text{Kovarianz Matrizen der Klasse 1 und 0 sind gleich!}}$

Wir sehen aus (*), dass

$$p(y=c \mid x, \theta) \propto p(x \mid y=c, \theta) \cdot p(y=c)$$

proportional zu

(da Nenner in (*) immer gleich)

Wir setzen $p(y=0 \mid x, \theta) = p(y=1 \mid x, \theta)$ um zu sehen wie die Entscheidungsgrenze aussieht. Definieren $\pi_0 = p(y=0)$ und $\pi_1 = p(y=1)$.

$$\pi_0 \cdot \cancel{\frac{1}{\sqrt{(2\pi)^d \cdot \det(\Sigma_0)}}} \cdot \exp\left(-\frac{1}{2} \cdot (x-\mu_0)^T \Sigma^{-1} (x-\mu_0)\right) = \cancel{\frac{1}{\sqrt{(2\pi)^d \cdot \det(\Sigma_1)}}} \cdot \exp(\cdots\cdots) \cdot \pi_1$$

$$\Sigma_0 = \Sigma, \quad \Sigma_1 = \Sigma$$

$$\overset{\log}{\Leftrightarrow} \log(\pi_0) - \frac{1}{2} \cdot (x-\mu_0)^T \Sigma^{-1} (x-\mu_0) = \log(\pi_1) - \frac{1}{2} \cdot (x-\mu_1)^T \Sigma^{-1} (x-\mu_1)$$

$$\Leftrightarrow \log(\pi_0) - \frac{1}{2} \cdot x^T \Sigma^{-1} x + \frac{1}{2} x^T \Sigma^{-1} \mu_0 + \frac{1}{2} \mu_0^T \Sigma^{-1} x - \frac{1}{2} \mu_0^T \Sigma^{-1} \mu_0 =$$

$$\log(\pi_1) - \frac{1}{2} x^T \Sigma^{-1} x + \frac{1}{2} \cdot x^T \Sigma^{-1} \mu_1 + \frac{1}{2} \mu_1^T \Sigma^{-1} x - \frac{1}{2} \cdot \mu_1^T \Sigma^{-1} \mu_1$$

$$\Leftrightarrow \cancel{-\frac{1}{2}x^T\Sigma^{-1}x} + \mu_0^T\Sigma^{-1}x - \frac{1}{2}\mu_0^T\Sigma^{-1}\mu_0 + \log(\pi_0) =$$
$$\cancel{-\frac{1}{2}x^T\Sigma^{-1}x} + \mu_1^T\Sigma^{-1}x - \frac{1}{2}\mu_1^T\Sigma^{-1}\mu_1 + \log(\pi_1)$$

$$\Leftrightarrow \left(\mu_1^T\Sigma^{-1}x - \mu_0^T\Sigma^{-1}x\right) - \frac{1}{2}\mu_1^T\Sigma^{-1}\mu_1 + \frac{1}{2}\mu_0^T\Sigma^{-1}\mu_0 + \log\left(\frac{\pi_1}{\pi_0}\right) = 0$$

$$\Leftrightarrow \left(\Sigma^{-1}(\mu_1-\mu_0)\right)^T x + \frac{1}{2}(\mu_0-\mu_1)^T\Sigma^{-1}(\mu_0+\mu_1) + \log\left(\frac{\pi_1}{\pi_0}\right) = 0$$

$$\overset{\cdot 2}{\Leftrightarrow} \underbrace{2\cdot\left(\Sigma^{-1}(\mu_1-\mu_0)\right)^T x}_{\text{abhängig von } x!!!} + \underbrace{(\mu_0-\mu_1)^T\Sigma^{-1}(\mu_0+\mu_1)}_{\text{Skalar}} + \underbrace{2\cdot\log\left(\frac{\pi_1}{\pi_0}\right)}_{\text{skalar}} = 0$$

check!!!

$\Rightarrow$ Wir erhalten einen Ausdruck der Form $\boxed{a^T x + b = 0}$ mit

$$a = \Sigma^{-1}(\mu_1-\mu_0)\cdot 2$$
$$b = (\mu_0-\mu_1)^T\Sigma^{-1}(\mu_0+\mu_1) + 2\cdot\log\left(\frac{\pi_1}{\pi_0}\right)$$

Lineare Entscheidungsgrenze!

(daher der Name LDA)

$\downarrow$

Linear

Wichtig: Wir haben $\Sigma_1 = \Sigma_0 = \Sigma$ angenommen. (Annahme 2).