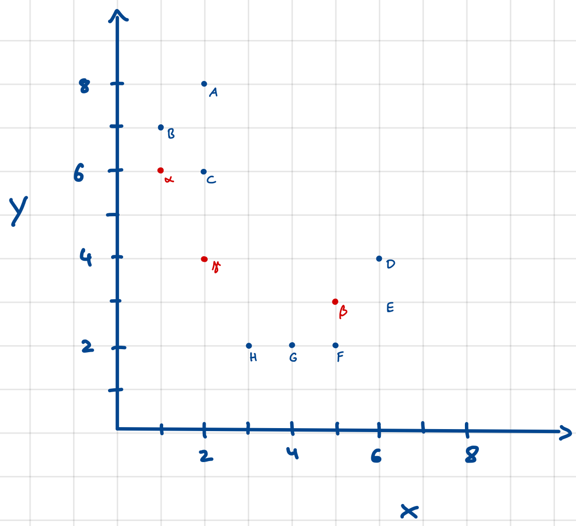
Übung 9

9.1)



Euklidische Distanzen:

$$d(\alpha, A) = \sqrt{(x_\alpha - x_A)^2 + (y_\alpha - y_A)^2}$$

	A	B	C	D	E	F	G	H
α	2,2	1	1	5,4	5,8	5,7	5	4,5
β	5,8	5,7	4,2	1,4	1	1	1,4	2,2
γ	4	3,2	2	4	4,1	3,6	2,8	2,2

immer z = 5 ⤳ z-Achse nicht relevant

k = 1 :
- α → gesundes Gewebe
- β → Tumorgewebe
- γ → gesundes Gewebe

k = 5 :
- α → gesundes Gewebe
- β → Tumorgewebe
- γ → Tumorgewebe

k = 3 :
- α → gesundes Gewebe
- β → Tumorgewebe
- γ → Tumorgewebe

k = 7 :
- α → Tumorgewebe
- β → Tumorgewebe
- γ → Tumorgewebe

→ je größer k gewählt wird, im Vergleich zur Probenanzahl, desto „falscher" wird das Ergebnis

→ Problem: keine gleiche Anzahl an jew. verglichenen Gewebeproben (3x gesundes Gewebe und 5x Tumorgewebe) ⤳ wenn dann k > 3 wird, muss der zu klassifizierende Datenpunkt zum Tumorgewebe zählen (wenn zuvor komplett zu gesundem Gewebe zählte)

→ außerdem: wenn relativ großes k gewählt wird (im Bezug auf Probengröße), werden Übereinstimmungen mit Gewebegruppen künstlich geschaffen

9.2) $p(x) = \dfrac{1}{1 + \exp(-\beta_0 + \beta_1 x)}$ $\quad \beta_0 = -5$, $\beta_1 = 1$ $\quad p(x) = 0{,}5$ → dann ist auch p(gesund) = 0,5

$$0{,}5 = \frac{1}{1 + e^{(5+x)}} \qquad | \cdot 1 + e^{(5+x)}$$

$$1 = 0{,}5 \cdot (1 + e^{(5+x)})$$
$$= 0{,}5 + 0{,}5 e^{(5+x)} \qquad | -0{,}5 \quad | \ln$$
$$\ln(0{,}5) = \ln(0{,}5) + 5 + x \qquad | -\ln(0{,}5)$$

$$0 = 5 + x$$
$$x = -5$$ → decision boundary

$x = 2$ $\quad p(x) = \dfrac{1}{1 + e^{(5+2)}} = 9 \cdot 10^{-3}$

$x = 6$ $\quad p(x) = 2 \cdot 10^{-5}$

$x = 10$ $\quad p(x) = 3 \cdot 10^{-7}$