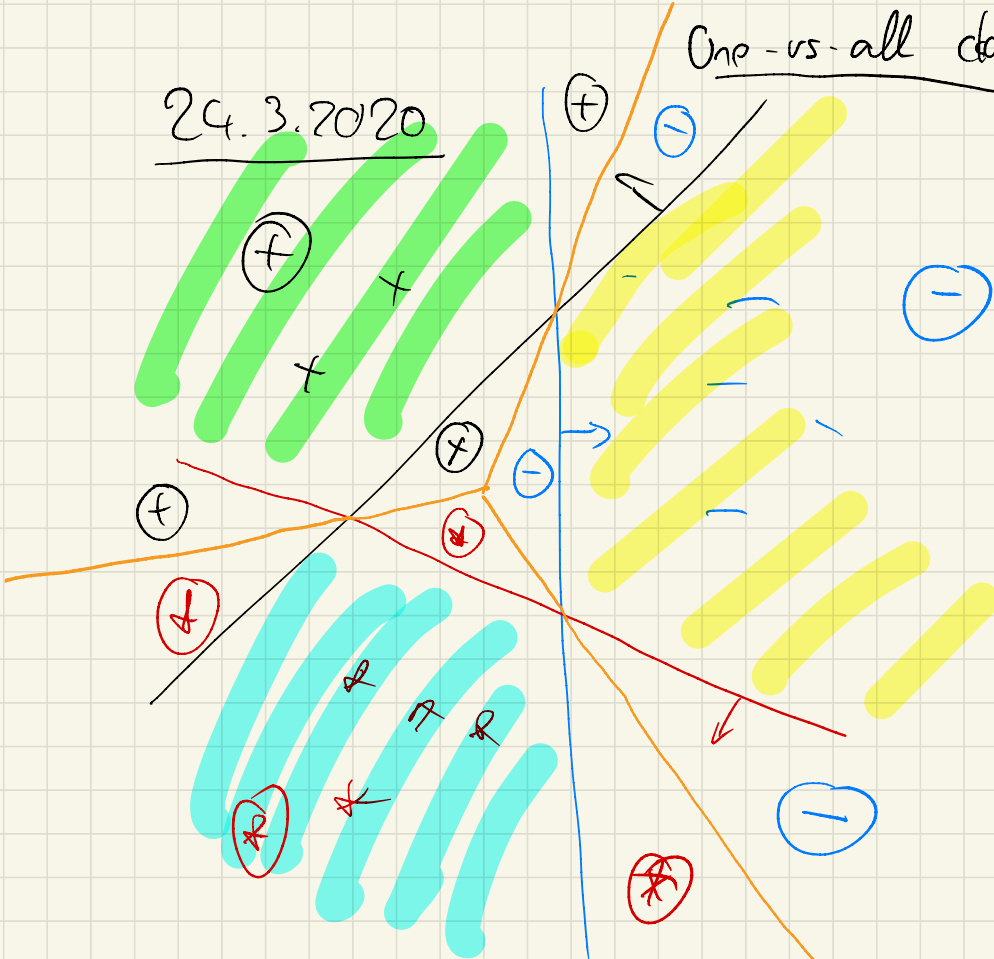


24.3.2020

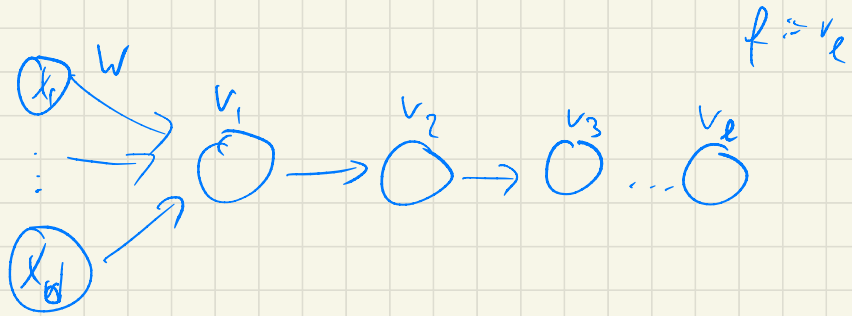
# One-vs-all classification



Predict:  $y(x)$  as  $\operatorname{argmax}_i w^{(i)T} x$

Sps.:  $\|w^{(i)}\|_2 = 1$

Neural nets with one unit per hidden layer  
yield linear classifiers



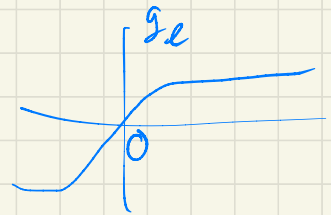
$$v_1 = \varphi_1(w^T x)$$

$$v_2 = \varphi_2(\alpha_2 v_1) =: \bar{\varphi}_2(v_1)$$

:

$$v_i = \varphi_i(\alpha_i v_{i-1}) =: \bar{\varphi}_i(v_{i-1}) = \bar{\varphi}_i(\underbrace{\bar{\varphi}_{i-1}(\dots(\bar{\varphi}_2(v_1))\dots)}_{g_i: \mathbb{R} \rightarrow \mathbb{R}})$$

$$f(x) = v_l = g_l(w^T x)$$



Classify:  $\text{sign } f(x)$

$$= \text{sign } g_l(w^T x)$$



scalar function

$\Rightarrow$  all levelsets of  $f$  are hyperplanes