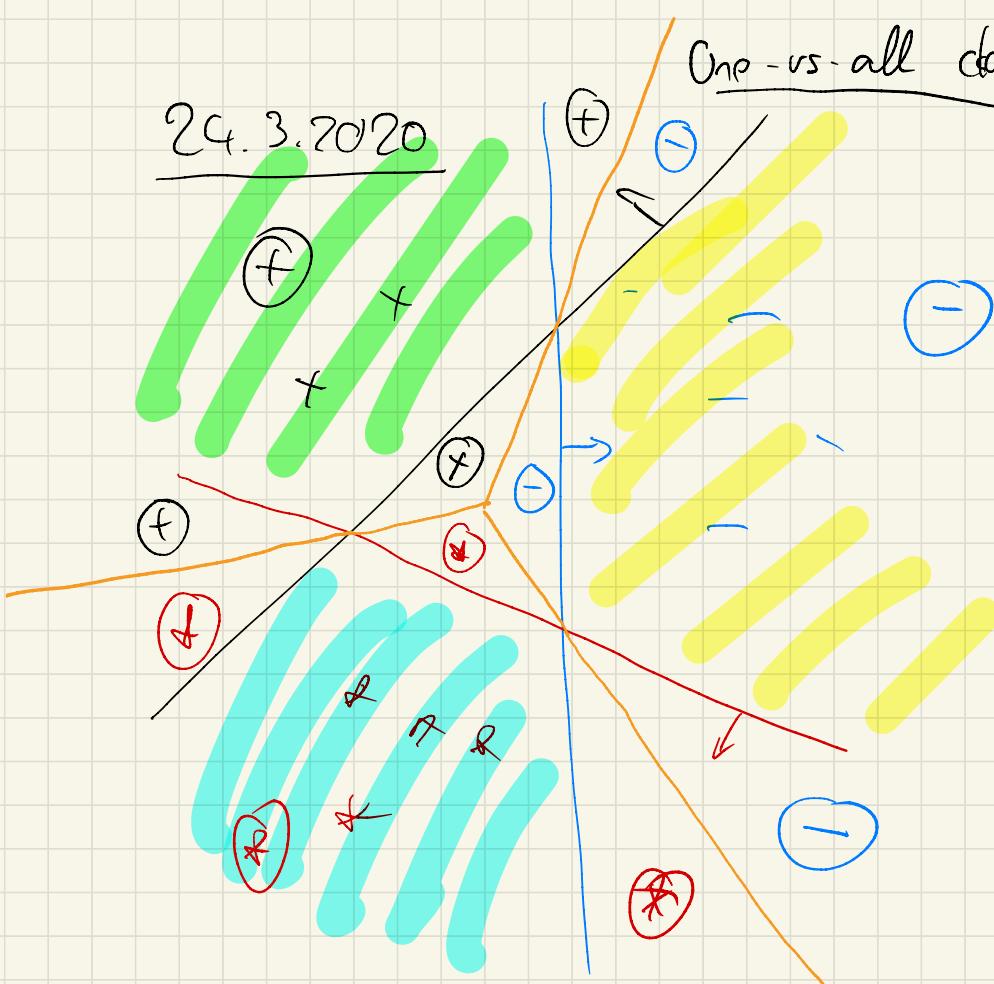


One-vs-all classification

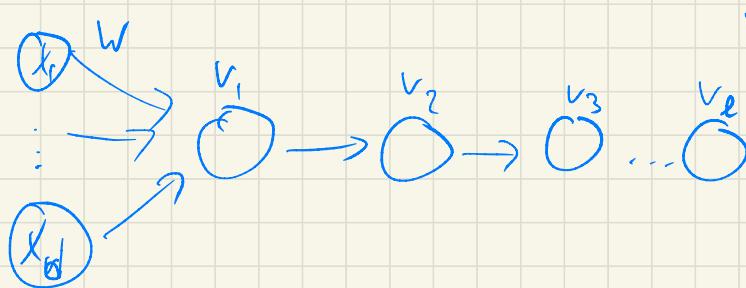
24. 3. 2020



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

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

Neural nets with one unit per hidden layer
yield linear classifiers



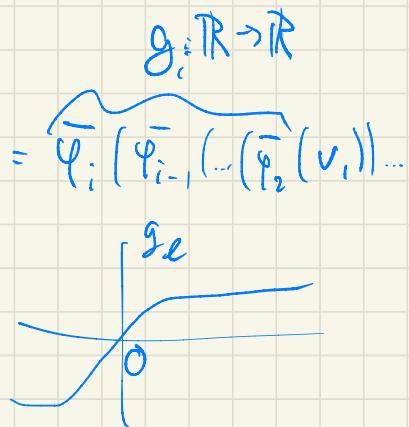
$$v_1 = \varphi_1(w^T x)$$

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

:

$$v_i = \varphi_i(\alpha_i v_{i-1}) =: \bar{\varphi}_i(v_{i-1}) = \overbrace{\bar{\varphi}_i}^g(\underbrace{\bar{\varphi}_{i-1}(\dots(\bar{\varphi}_2(v_1))\dots)}_e)$$

$$f(x) = v_e = g_e(w^T x)$$



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

$$= \text{sign } g_e(v^T x)$$



scalar function

\Rightarrow all levelsets of f are hyperplanes