Introduction to ML

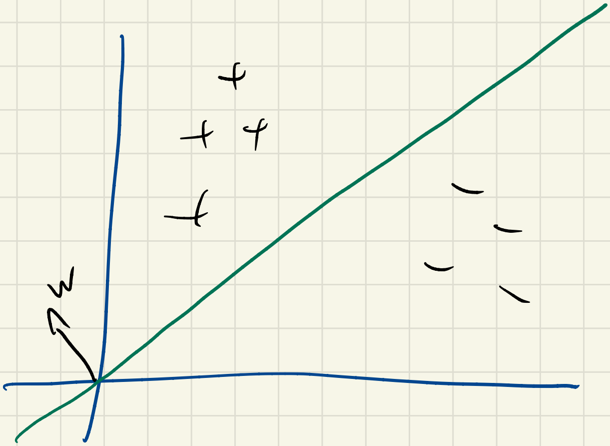
Flipped Classroom Notes

Introduction to Machine Learning

Flipped Classroom session 17.3.2020

Updates

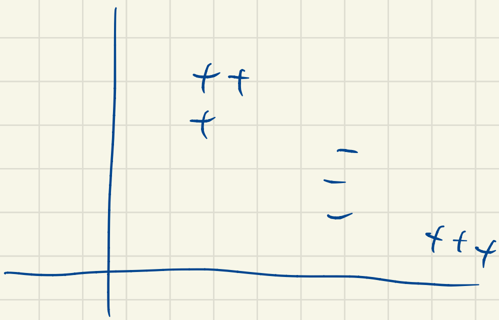
- Lecture: flipped classroom
- Exercises: Online only, no Q&A
- Project: As planned



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

$x_i \in \mathbb{R}^d$, $y_i \in \{+1, -1\}$

$D = \{(x_1, y_1), \dots (x_n, y_n)\}$ is called <u>linearly separable</u> if $\exists\, w \in \mathbb{R}^d$ s.t. $\forall (x_i, y_i) \in D$: $\text{sign}(w^T x_i) = y_i$



is <u>not</u> linearly separable.

D is linearly separable after transformation $\phi : \mathbb{R}^d \to \mathbb{R}^D$ iff $\exists\, w \in \mathbb{R}^D$ s.t. $\forall (x_i, y_i) \in D : \operatorname{sign}(\phi(x_i)^T w) = y_i$

Does for <u>every</u> data set D there exist a transform ϕ which linearly separates it?

In general: No.

Example: $D = \{(x, +1), (x, -1)\}$

If we disallow this special case, i.e., $\nexists\, (x, +1) \in D \wedge (x, -1) \in D$ for some x, then we can construct a mapping ϕ s.t. D becomes lin. separable!

Pick $\phi : \mathbb{R}^d \to \mathbb{R}^n$, s.t.

$\phi(x_i) := e_i$ ← ith position

$\uparrow [0, \ldots, 0\ 1\ 0 \ldots 0]$

and $\phi(x) = 0$ if $x \in D_x$

Then: $w_i := y_i$ separates the data

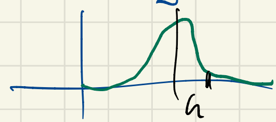
$$\text{sign}(w^T \phi(x_i)) = \text{sign}(w^T e_i)$$
$$= \text{sign}(w_i) = y_i$$

Consider Perceptron / SVM with Gaussian kernel.

Predict: $$y(x) = \text{sign}\Big(\underbrace{\sum_{i=1}^{n} \alpha_i y_i k(x_i, x)}_{\alpha^T \phi(x)}\Big)$$

$$\phi(x) = [y_1 k(x_1, x),\ y_2 k(x_2, x) \dots y_n k(x_n, x)]^T$$

$$k(x_i, x) = \exp\left(-\frac{\|x_i - x\|_2^2}{h^2}\right)$$



as $h \to 0$: $k(x_i, x) \to \begin{cases} 1 & \text{if } x_i = x \\ 0 & \text{otw} \end{cases}$