Hans U. Simon
Christoph Ries

Bochum, den 26. Juni 2018
Abgabe bis zum 03.07.2018

Übungen zur Vorlesung

Theorie des maschinellen Lernens

Sommer 18

Übungsblatt 09

Aufgabe 9.1 (4 Punkte)
Seien K_1 und K_2 zwei Kernels über $\mathcal{X} \times \mathcal{X}$ mit $\mathcal{X} \subseteq \mathbb{R}^n$, $a \in \mathbb{R}^+$ und f eine reell-wertige Funktion auf $\mathcal{X}$. Zeige, dass folgende Funktionen dann allesamt Kernels sind.

a) $K(x, z) = K_1(x, z) + K_2(x, z)$

b) $K(x, z) = aK_1(x, z)$

c) $K(x, z) = f(x)f(z)$

d) Seien $x, z \in \mathbb{R}^2$. Ist $K(x, z) = \sqrt{2}[(x_1z_1)^2 - 5(x_2z_1^2 + x_1^2z_2) + 25x_2z_2]$ ein Kernel?

Aufgabe 9.2 (4 Punkte)
Für alle $A \subseteq \{1, ..., n\}$ sei $\psi_A(x) = \prod_{i \in A} x_i$. Des Weiteren sei $\psi(x) = (\psi_A(x))_{A \subseteq \{1,...,n\}}$ und es gelte $K(x, z) = \langle \psi(x), \psi(z) \rangle$.
Finde eine effizient auswertbare Formel für $K(x, z)$ (poly(n) arithmetische Operationen).

Aufgabe 9.3 (4 Punkte)

a) Sei $K(x, y) = \langle \psi(x), \psi(y) \rangle$ ein Kernel für ein $\psi : \mathbb{R}^n \to \mathbb{R}^m$. Zeige, dass die Berechnung von Distanzen zwischen Trainingspunkten im Merkmalsraum indirekt über die Verwendung von K erfolgen kann.

b) Für $x, x' \in [N]$ definiere
$$K(x, x') = \min\{x, x'\}.$$
Finde eine Abbildung $\psi : [N] \to H$, wobei H ein geeigneter Hilbertraum ist, sodass für alle $x, x' \in [N]$ gilt
$$K(x, x') = \langle \psi(x), \psi(x') \rangle.$$

— Bitte wenden! —

Aufgabe 9.4 (4 Punkte)
Sei die Trainingsmenge $S = \{(x_i, y_i)\}_{i=1}^m$ über $\mathbb{R}^n$ mit Markierungen aus $\{-1, 1\}$ gegeben.

Erinnerung: Der Perzeptron-Algorithmus angesetzt auf S initialisiert w mit $0 \in \mathbb{R}^n$. Solange es noch einen Trainingspunkt (x_i, y_i) mit $y_i \langle w, x_i \rangle \leq 0$ gibt, wird w aktualisiert mittels

$$w \leftarrow w + y_i x_i \ . \tag{1}$$

Es gebe nun eine Merkmalsfunktion ψ, die die Daten in den $\mathbb{R}^d$ mit $d \geq n$ abbildet. Dies liefert die Trainingsmenge S'. Zeige, dass sich der Perzeptron-Algorithmus für S' so verändern lässt, dass er auf die Trainingsmenge S' nur implizit durch Aufrufe eines Kernels, $K(x_i, x_j) = \langle \psi(x_i), \psi(x_j) \rangle$, zu greift.

Hinweis: Es ist nicht mehr nötig w explizit wie in (1) zu bestimmen.