MACHINE LEARNING, 8. SEMINAR – SUPPORT VECTOR MACHINES

Aufgabe 1. Sei $x \in \mathbb{R}$ und $y \in \mathbb{R}$ zwei reellwertige Variablen. Lösen Sie

$$x^2 + y^2 \to \min_{xy}$$
$$\text{s.t. } 2x + y \geq 1.$$

a) Geben Sie zunächst eine geometrische Interpretation der Aufgabe an und leiten Sie daraus die Lösung ab.

b) Lösen Sie die Aufgabe mittels der Methode der Lagrange-Koeffizienten.

Aufgabe 2. Für Muster $x \in \mathscr{X} = \mathbb{R}$ wird die folgende Abbildung $\Phi : \mathscr{X} \to \mathscr{H}$ in den Merkmalsraum $\mathscr{H} = \mathbb{R}^2$ definiert:

$$\Phi(x) = \big(\sin(\lambda_1 x), \cos(\lambda_2 x)\big)$$

mit den Parametern der Abbildung $\lambda_1, \lambda_2 \in \mathbb{R}$.
Hinweis: Verwechseln Sie bitte nicht die Parameter der Abbildung λ_1, λ_2 und die Parameter w_1, w_2 der Entscheidungsregel im Merkmalsraum (siehe unten).

a) Wie sieht die Menge der der Input-Muster $\mathscr{X}$ im Merkmalsraum $\mathscr{H}$ aus?

b) Zur Klassifikation der Muster in zwei Klassen wird eine lineare Entscheidungsregel im Merkmalsraum verwendet, d.h.

$$\langle \Phi(x), w \rangle \gtrless b$$

mit den Parametern $w = (w_1, w_2) \in \mathbb{R}^2$ und $b \in \mathbb{R}$. Gegeben sei eine Entscheidungsregel (d.h. ihre Parameter w und b). Wie sieht die Aufteilung des Input-Raumes $\mathscr{X}$ in die Klassen aus?

Aufgabe 3. Sei $k_1(x,x')$ und $k_2(x,x')$ zwei Kernel. Beweisen Sie, dass

a) $k_3(x,x') = k_1(x,x') + k_2(x,x')$ und

b) $k_3(x,x') = k_1(x,x') \cdot k_2(x,x')$ auch Kernel sind.

Sei $\mathscr{H}_1$, $\mathscr{H}_2$ und $\mathscr{H}_3$ die Merkmalsräume der Kernel k_1 bzw. k_2 und k_3. Wie ergeben sich die Dimensionen der jeweiligen Merkmalsräume $\mathscr{H}_3$ aus den Dimensionen der Merkmalsräume $\mathscr{H}_1$ und $\mathscr{H}_2$?

Aufgabe 4. Wir betrachten nochmals die „Exponential Family“-Modelle (siehe die entsprechende Vorlesung). Insbesondere diskutieren wir, wie diese Modelle diskriminativ gelernt werden können. Sei x eine beobachtbare Größe und y die Klasse. Die aposteriori Wahrscheinlichkeitsverteilung ist

$$p(y|x;w) = \frac{1}{Z(x,w)} \exp\Big[\big\langle \varphi(x,y), w \big\rangle\Big].$$

a) Weiterhin, sei die Entscheidungsstrategie die Maximum A-posteriori Entscheidung. Leiten Sie die Entscheidungsregel ab.

b) Gegeben sei eine klassifizierte Lernstichprobe $L = \big((x^l, y^l)\ldots\big)$. Gesucht wird der Parametervektor w der sie fehlerfrei klassifiziert (angenommen es existiert mindestens ein). Zeigen Sie wie man diese Aufgabe mithilfe des Perceptron-Algorithmus löst.