Hans U. Simon
Malte Darnstädt

Bochum, den 25.06.2013
Abgabe bis zum 03.07.2013

Übungen zur Vorlesung

# Theorie des Maschinellen Lernens

Sommer 2013

Übungsblatt 10

**Aufgabe 10.1**

Sei $N_n$ ein Netzwerk von linearen Schwellwertfunktionen über $X = \{0,1\}^n$ mit einem Ausgabeknoten und beliebig vielen Berechnungsknoten in einem Hidden Layer.

Zeige, dass man zu einer Stichprobe der Größe $m$ effizient (also in polynomieller Zeit in $n$ und $m$) eine konsistente Hypothese aus $N_n$ finden kann.
Warum gibt es trotzdem keinen effizienten Lerner im realisierbaren Modell für die Klasse $\cup_n N_n$?

**Aufgabe 10.2**

Sei $f : \mathbb{R}^n \to \{0,1\}$ eine Funktion die durch ein neuronales Netz $N_f$ mit $n$ Eingabeknoten, $k$ Berechnungsknoten in einem Hidden Layer und einem Ausgabeknoten gegeben ist.
Sei $b \in \{0,1\}^k$. Wir verändern das Netz $N_f$ folgendermaßen: Genau dann, wenn $b_i = 0$ ist, wird die Ausgabe von Berechnungsknoten $i$ gekippt (d.h. von 0 auf 1 gesetzt und umgekehrt).

Zeige: Für jedes $b$ kann man die Gewichte und den Schwellwert des Ausgabeknotens so anpassen, dass das veränderte Netz $N_{f,b}$ weiterhin die Funktion $f$ berechnet.

**Problem: Set-Splitting**
*Eingabe:* Eine Menge $S = \{x_1, \ldots, x_n\}$ und Teilmengen $A_1, \ldots, A_l$ von $S$.
*Frage:* Gibt es Teilmengen $S_1, S_2$ von $S$, so dass $S = S_1 \cup S_2$ gilt und kein $A_i$ vollständig in $S_1$ oder $S_2$ liegt?

**Definition: XOR-Netzwerk**
Sei $N^2_{\text{XOR}}$ die Klasse der Netzwerke mit zwei Berechnungsknoten (gewöhnliche lineare Schwellwerte) in einem Hidden Layer und einem Ausgabeknoten, der die Parität seiner zwei Eingabe-Bits berechnet, d.h. der Knoten gibt genau dann 1 aus wenn seine Eingabe $(0, 1)$ oder $(1, 0)$ ist.

## Aufgabe 10.3

Zeige, dass folgende Abbildung eine polynomielle Reduktion von Set-Splitting auf $N^2_{\text{XOR}}$-Consistency darstellt (weil Set-Splitting ein schwerstes Problem in NP ist, haben wir damit gezeigt, dass das Lernen von $N^2_{\text{XOR}}$ im Realizable Model hart ist):
Wir wählen $X = \{-1, 0, 1\}^n$. In der Stichprobe sind folgende Beispiele:

- $e_i$ mit Label 0 für $1 \leq i \leq n$.
- $-e_i$ mit Label 0 für $1 \leq i \leq n$.
- 0 (der Nullvektor) mit Label 1.
- $\sum_{x_j \in A_i} e_j$ mit Label 1 für $1 \leq i \leq l$.

Dabei bezeichnet $e_i$ den Vektor mit einer Eins in Position $i$ und Nullen in allen anderen Positionen.

## Aufgabe 10.4

Zeige mit Hilfe einer Reduktion auf eine mächtigere Klasse, die bekanntermaßen effizient lernbar ist, dass $N^2_{\text{XOR}}$ über $X_n = \{-1, 0, 1\}^n$ im realisierbaren Modell effizient lernbar ist (wenn man als Ausgabe des Lerners die Funktionen der mächtigeren Klasse erlaubt).