

Einführung in die Künstliche Intelligenz

Dr. Sebastian Rudolph
Sommersemester 2009

<http://semantic-web-grundlagen.de>

Übungsblatt 3: Maschinelles Lernen

Aufgabe 3.1 Erkläre die folgenden Termini:

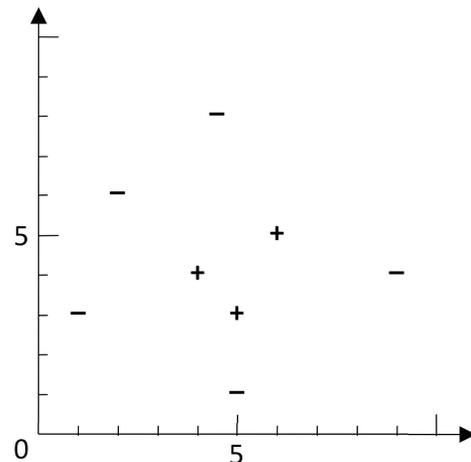
1. Begriffslernen,
2. Hypothesenraum,
3. Versionsraum,
4. induktiver Bias,
5. Overfitting.

Aufgabe 3.2 Maschinelles Lernen aus Beispielen lässt sich formalisieren als Aufgabe, eine Funktion f , basierend auf gegebenen Input-Output-Paaren $(x_1, f(x_1)), (x_2, f(x_2)), \dots (x_n, f(x_n))$ möglichst gut durch eine Funktion f' anzunähern. Erläutere diese Formalisierung an konkreten Problemstellungen. Beschreibe die Funktion f in diesen Fällen. Wie sehen die entsprechenden Input-Output-Paare aus?

Aufgabe 3.3 Betrachte das Beispiel zum Lernen des Begriffs *EnjoySport* aus der Vorlesung. Gib die Folge der Mengen S und G des Kandidateneliminationsalgorithmus für den Fall an, dass die Trainingsbeispiele in umgekehrter Reihenfolge abgearbeitet werden.

Trotzdem das Endresultat unabhängig von der Abarbeitungsreihenfolge ist, unterscheiden sich die Zwischenresultate. Schlage eine Strategie vor, die Trainingsbeispiele so anzuordnen, dass die Summe der Elemente von S und G während der Zwischenschritte möglichst klein wird.

Aufgabe 3.4 Bei einem Begriffslernproblem seien die Beispiele Paare ganzer Zahlen (x, y) , die als Punkte in einem zweidimensionalen Koordinatensystem aufgefasst werden sollen. Der Hypothesenraum H bestehe aus achsenparallelen Rechtecken in diesem Koordinatensystem mit ganzzahligen Eckpunktkoordinaten, d.h. die Hypothesen sind von der Form $a \leq x \leq b$, $c \leq y \leq d$ für beliebige ganzzahlige a, b, c, d . Die positiven (+) und negativen (-) Trainingsbeispiele sind aus nebenstehender Abbildung ersichtlich.



1. Beschreibe die Schranken S und G des Versionsraumes? Schreibe diese als Formel und zeichne Sie in das Diagramm ein.
2. Angenommen, der Lernalgorithmus kann nun einen “Lehrer” nach der Klassifikation (+ oder -) eines bestimmten Punktes (x, y) fragen. Schlage ein solches Wertepaar vor, dessen Klassifikation (egal ob positiv oder negativ) den Versionsraum garantiert einschränken wird.
3. Angenommen, Du wärst ein “Lehrer” und wolltest den Algorithmus den Begriff $3 \leq x \leq 5, 2 \leq y \leq 9$ lernen lassen. Was ist die kleinste Anzahl von Trainingsbeispielen, die dem Kandidateneliminationsalgorithmus präsentiert werden muss, damit der vorgegebene Begriff perfekt gelernt wird?

Aufgabe 3.5 Beschreibe ein einfaches Perzeptron mit zwei variablen Eingängen (zuzüglich eines weiteren Einganges der immer fest mit 1 belegt ist, siehe Skizze des Perzeptrons aus der Vorlesung) welches die boolesche Funktion $A \vee \neg B$ umsetzt. Beschreibe ein 2-layer Perzeptron, welches die XOR-Funktion realisiert.

Aufgabe 3.6 Beschreibe – möglicherweise unter Verwendung der Hinweise aus der Vorlesung – detailliert ein Verfahren basierend auf einem genetischen Algorithmus zur Lösung des Problems des Handlungsreisenden¹.

Aufgabe 3.7 Gegeben sei ein 2-layer Perzeptron mit 8 Input-Neuronen, 3 Neuronen im hidden layer und 8 Output-Neuronen. Die Trainingsbeispiele sind in nebenstehender Tabelle angegeben. Trainiere das Netzwerk mittels Backpropagation (verwende dabei tanh als Output-Funktion) und gib die resultierenden Gewichte nach dem Training an.

Input	Output
10000000	10000000
01000000	01000000
00100000	00100000
00010000	00010000
00001000	00001000
00000100	00000100
00000010	00000010
00000001	00000001

¹siehe z.B. http://de.wikipedia.org/wiki/Problem_des_Handlungsreisenden