2 Lernen im Mistake-Bound Modell

2.1 Konzeptklassen und Repräsentationen

Wir beschäftigen uns mit dem Erlernen von *Konzepten* aus einer gegebenen *Konzeptklasse*. Ein *Konzept* c ist eine Menge von Elementen aus einer beliebige Grundmenge X. Wir nennen X auch das *Universum*. Eine *Konzeptklasse* C über einem Universum X ist eine Menge von Konzepten. Das zu erlernende Konzept heißt auch *Zielkonzept*.

Wir schreiben oft auch einfach c für die *charakteristische Funktion* χ_c von c, wobei

$$\chi_c(x) = \begin{cases} 1 & \text{falls } x \in c \\ 0 & \text{falls } x \notin c \end{cases}$$

Ein *Beispiel* ist ein Element aus dem Universum X. Bezüglich einem gegebenen Zielkonzept c ist ein *positives Beispiel* (für c) ein Beispiel das in c enthalten ist. Entsprechend ist ein *negatives Beispiel* ein Beispiel das nicht in c enthalten ist. Ein *klassifiziertes Beispiel* ist Beispiel zusammen mit seiner Klassifikation durch c, also ein Paar $(x, c(x))$ mit $x \in X$.

Wir nehmen an, daß eine Konzeptklasse C stets durch eine *Repräsentationssprache* R beschrieben wird, d.h. es gibt eine surjektive *Sematikfunktion* $\Phi : R \to C$, die jeder Repräsentation $r \in R$ eindeutig ein Konzept $\Phi(r) \in C$ zuordnet.

Im einfachsten Fall können wir die Konzepte einer (höchstens abzählbaren) Konzeptklasse C einfach durchnumerieren. Dann ist $C = \{c_1, c_2, \dots\}$ und entsprechend $R \subseteq \mathbb{N}$. In den meisten Fällen wird die Repräsentationssprache R aus Wörtern über einem endlichen Alphabet Σ bestehen, wobei wir OBdA davon ausgehen können, daß $\Sigma = \{0, 1\}^*$. Wir werden aber auch Konzeptklassen betrachten, bei denen zur Beschreibung ihrer Konzepte reelle Zahlen benötigt werden. In diesem Fall nehmen wir zusätzlich an, daß unsere Lernalgorithmen reelle Zahlen handhaben können, obwohl diese ja im Allgemeinen nicht endlich beschrieben werden können.

In Beispiel 1.1 ist das Universum X die Menge aller Wörter über dem Alphabet $\Sigma = \{V, R, L, R\}$. Die Konzepte sind jeweils Sprachen über Σ, und werden durch einen endliche Automaten mit Alphabet Σ beschrieben. Die Konzeptklasse besteht also aus allen regulären Sprachen über dem Alphabet Σ. In Beispiel 1.2 ist das Universum $X = \mathbb{N}^2$. Konzepte sind Rechtecke, die jeweils durch die binär kodierten Koordinaten $\langle x_{min}, y_{min}, x_{max}, y_{max} \rangle$ zweier Eckpunkte repräsentiert werden können, und die Konzeptklasse ist die Menge aller Rechtecke.

Beispiel 2.1. Eine boolesche Formel F über n Variablen $x_1, \ldots, x_n$ repräsentiert das Konzept aller erfüllenden Belegungen $x \in \{0,1\}^n$ von F. Das Universum ist also $X = \{0,1\}^n$. Da jede Teilmengen von $\{0,1\}^n$ durch eine boolesche Formel beschrieben werden kann, besteht die entsprechende Konzeptklasse der booleschen Formeln über n Variablen aus alle Teilmengen von $\{0,1\}^n$.

2.2 Das Mistake-Bound Modell

Sei C eine Konzeptklasse über einem Universum X. Das Lernen im Mistake-Bound Modell erfolgt in einer Folge von *Versuchen* bzgl. einem Zielkonzept $c \in C$ und einer Beispielfolge $x_1, x_2, \ldots$ mit $x_i \in X$. Der Ablauf im i-ten Versuch ist wie folgt:

1. Der Lernalgorithmus erhält das Beispiel $x_i \in X$.
2. Der Lernalgorithmus trifft eine *Vorhersage* für die Klassifikation des Beispiels x_i durch das Zielkonzept c.
3. Der Lernalgorithmus erhält die (korrekte) Klassifikation $c(x_i)$.

Wir nenne einen solchen Lernalgorithmus auch *on-line* Lernalgorithmus.

Definition. *Eine Konzeptklasse C über einem Universum X ist mit k Vorhersagefehler im Mistake-Bound Modell* erlernbar*, falls es einen on-line Lernalgorithmus gibt, der für alle Konzepte $c \in C$ und für alle Beispielfolgen $x_1, x_2, \ldots$ mit $x_i \in X$ höchstens k Vorhersagefehler macht.*

2.3 Konjunktionen

Eine Konjunktion (von Literalen) über n Variablen hat im Allgemeinen die Form

$$\bigwedge_{i \in I} x_i \wedge \bigwedge_{j \in J} \overline{x}_j, \quad I, J \subseteq \{1, \ldots, n\}$$

Das von einer solchen Konjunktion repräsentierte Konzept ist die Menge aller erfüllenden Belegungen $x \in \{0,1\}^n$,

$$\{x = x_1 \ldots x_n \mid \forall i \in I : x_i = 1 \wedge \forall j \in J : x_j = 0\}$$

Das Universum ist also $X = \{0,1\}^n$.

Wir können die Konzeptklasse aller Konjunktionen über n Variablen mit dem folgenden on-line Lernalgorithmus im Mistake-Bound Modell erlernen.

- Initialisiere $h = x_1 \wedge \bar{x}_1 \wedge \ldots x_n \wedge \bar{x}_n$
- Im i-ten Versuch bei Vorlage des Beispiels $x_i \in \{0,1\}^n$
  - Sage x_i gemäß h voraus
  - Im Fall einer falschen Vorhersage entferne aus h alle Literale, die unter der Belegung x_i falsch sind

Lemma 2.2. *Kein Literal des Zielkonzepts wird aus h entfernt.*

Beweis. Sei h_{i-1} die Hypothese zu Beginn des i-ten Versuches, bei dem ein Literal l aus h_{i-1} entfernt wird. Das Literal l kann nur dann aus h_{i-1} entfernt werden, wenn ein Vorhersagefehler gemacht wird und l unter der Belegung x_i falsch ist. Da h_{i-1} das Literal l noch enthält, muß auch h_{i-1} unter der Belegung x_i falsch sein, und da ein Vorhersagefehler gemacht wird, ist das Zielkonzept c wahr unter der Belegung x_i. Dann kann l aber nicht in der Konjunktion c enthalten sein, da l unter der Belegung x_i falsch ist. □

Lemma 2.3. *Vorhersagefehler werden nur für positive Beispiele gemacht.*

Beweis. Da h stets alle Literale des Zielkonzepts c enthält, sind alle erfüllenden Belegungen von h auch erfüllende Belegungen von c. Also können Vorhersagefehler nur für erfüllende Belegungen von c gemacht werden. □

Lemma 2.4. *Bei jedem Vorhersagefehler wird wenigstens ein Literal aus h entfernt.*

Beweis. (indirekt) Nehmen wir an, daß im i-ten Versuch mit Beispiel x_i ein Vorhersagefehler gemacht wird, aber kein Literal entfernt wird. Dann erfüllt x_i alle Literale von h, und somit auch h. Da x_i einen Vorhersagefehler verursacht, muß x_i ein negatives Beispiel für das Zielkonzept c sein. Vorhersagefehler können aber nur für positive Beispiele gemacht werden. □

Theorem 2.5. *Konjunktionen über n Variablen sind mit höchstens $n + 1$ Vorhersagefehler im Mistake-Bound Modell erlernbar.*

Beweis. Für ein unerfüllbares Zielkonzept wird kein Vorhersagefehler gemacht, und für ein ein erfüllbares Zielkonzept werden beim ersten Vorhersagefehler n Literale aus h entfernt. Nach dem ersten Vorhersagefehler können höchstens n weitere Literale aus h entfernt werden. □