

Einführung in die Künstliche Intelligenz

Übungszettel 5

Prof. Dr. Claudia Schon

C.Schon@hochschule-trier.de

Fachbereich Informatik

Hochschule Trier

1 Entscheidungsbaum zu Bestellungen¹

Ein E-Commerce-Unternehmen möchte herausfinden, anhand welcher Merkmale sich vorhersagen lässt, ob eine Online-Bestellung von einem Kunden storniert wird oder nicht. Zu diesem Zweck wurden exemplarisch 14 vergangene Bestellungen untersucht. Für jede Bestellung wurden drei Merkmale erfasst:

- **TimeOfDay**: Zu welcher Tageszeit die Bestellung aufgegeben wurde (Morning, Afternoon, Evening, Night).
- **Device**: Über welches Gerät die Bestellung erfolgte (Desktop oder Mobile).
- **PaymentMethod**: Mit welcher Zahlungsmethode bezahlt wurde (CreditCard oder PayPal).

Zusätzlich wurde festgehalten, ob die Bestellung vom Kunden später storniert wurde (**Cancelled** = Yes/No). Die vollständigen Trainingsdaten sind in der unten abgebildeten Tabelle dargestellt.

Ziel dieser Aufgabe ist es, auf Basis der Daten einen Entscheidungsbaum zu erstellen, der die Stornierung von Bestellungen möglichst gut vorhersagt.

	TimeOfDay	Device	PaymentMethod	Cancelled
1	Morning	Desktop	CreditCard	No
2	Evening	Mobile	PayPal	Yes
3	Afternoon	Desktop	PayPal	No
4	Evening	Desktop	CreditCard	No
5	Morning	Mobile	CreditCard	Yes
6	Night	Desktop	PayPal	No
7	Night	Mobile	PayPal	Yes
8	Morning	Mobile	CreditCard	Yes
9	Afternoon	Mobile	CreditCard	No
10	Evening	Desktop	PayPal	No
11	Night	Mobile	CreditCard	Yes
12	Morning	Desktop	PayPal	No
13	Afternoon	Mobile	PayPal	No
14	Evening	Mobile	CreditCard	Yes

Bitte beachten Sie, dass die Nummerierung in der Tabelle nur als Hilfe für die Bearbeitung der Aufgabe gedacht ist und keine Attributwerte darstellt.

- (a) Bestimmen Sie das beste Attribut für die Wurzel des Entscheidungsbaums, indem Sie den *Information Gain* für jedes Attribut berechnen.

¹Die Aufgabe wurde mit Unterstützung von ChatGPT erstellt.

- (b) Konstruieren Sie schrittweise einen Entscheidungsbaum für die obigen Trainingsdaten, indem Sie den in der Vorlesung vorgestellten Algorithmus anwenden.
- (c) Auf Stud.IP finden Sie die Datei `online_orders.xlsx`, die die Trainingsdaten enthält. Berechnen Sie mit Hilfe von KNIME einen Entscheidungsbaum für das Beispiel. Wählen Sie den *Gain ratio* für die Auswahl des besten Attributs. Eine detaillierte Beschreibung der Erstellung eines Entscheidungsbaumes in KNIME finden Sie in den vorherigen beiden Übungsblättern.

2 Einfaches Perzeptron-Update

Wir betrachten ein Perzeptron mit zwei Eingängen.

Wir erweitern den Eingabevektor um einen zusätzlichen konstanten Wert $x_0 = 1$. Das Gewicht w_0 ist der Bias des Perzeptrons. Der vollständige (erweiterte) Eingabevektor ist also $\vec{x} = (x_0, x_1, x_2)$. Der Gewichtvektor hat die Komponenten w_0, w_1 und w_2 und ist gegeben als:

$$w_0 = 0,5; \quad w_1 = 1,0; \quad w_2 = -1,0$$

Die Ausgabe des Perzeptrons ergibt sich durch:

$$y = \varphi(\vec{x} \cdot \vec{w}) = \begin{cases} 1, & \text{wenn } \vec{x} \cdot \vec{w} > 0 \\ 0, & \text{sonst} \end{cases}$$

Zur Anpassung der Gewichte verwenden wir die folgende aus der Vorlesung bekannte Regel:

$$w_i^{\text{neu}} = w_i^{\text{alt}} + \Delta w_i \quad \text{mit} \quad \Delta w_i = \eta \cdot (t - y) \cdot x_i \quad \text{für } i = 0, \dots, n, \quad \eta > 0$$

Wobei t die gewünschte Ausgabe bezeichnet.

Das Gewicht w_0 wird wie alle anderen Gewichte mit dieser Regel angepasst. Es muss auch bei der Berechnung des Produkts $\vec{x} \cdot \vec{w}$ berücksichtigt werden, wobei x_0 konstant auf 1 gesetzt wird.²

1. Berechnen Sie für den Eingabevektor $\vec{x} = (1; 0; 1)$ die gewichtete Summe $\vec{x} \cdot \vec{w}$ und die Ausgabe $y = \varphi(\vec{x} \cdot \vec{w})$.
2. Gegeben sei die Zielausgabe $t = 1$ und die Lernrate $\eta = 1$. Bestimmen Sie die neuen Gewichtswerte nach einem Trainingsschritt mit der oben gegebenen Lernregel.

²Bitte beachten Sie, dass in der ursprünglichen Version der Folie 82 der Vorlesungsfolien der Wert x_0 fälschlicherweise nicht berücksichtigt wurde. Dieser Fehler wurde inzwischen korrigiert – die überarbeitete Version der Folie steht nun auf Stud.IP zur Verfügung.

3 Perzeptron-Training für die umgekehrte Implikation³

Ein Perzeptron soll die logische Funktion der **umgekehrten Implikation** $x_2 \rightarrow x_1$ erlernen. Die Wahrheitstabelle lautet:

x_1	x_2	$x_2 \rightarrow x_1$
0	0	1
0	1	0
1	0	1
1	1	1

Wir erweitern den Eingabevektor um einen zusätzlichen konstanten Wert $x_0 = 1$. Das Gewicht w_0 ist der Bias des Perzeptrons. Der vollständige (erweiterte) Eingabevektor ist also $\vec{x} = (x_0, x_1, x_2)$. Der Gewichtsvektor hat die Komponenten w_0, w_1 und w_2 .

Die Ausgabe des Perzeptrons ergibt sich durch:

$$y = \varphi(\vec{x} \cdot \vec{w}) = \begin{cases} 1, & \text{wenn } \vec{x} \cdot \vec{w} > 0 \\ 0, & \text{sonst} \end{cases}$$

Zur Anpassung der Gewichte verwenden wir die folgende aus der Vorlesung bekannte Regel:

$$w_i^{\text{neu}} = w_i^{\text{alt}} + \Delta w_i \quad \text{mit} \quad \Delta w_i = \eta \cdot (t - y) \cdot x_i \quad \text{für } i = 0, \dots, n, \quad \eta > 0$$

Wobei t die gewünschte Ausgabe bezeichnet (was der rechten Spalte in der obigen Wahrheitstabelle entspricht).

Das Gewicht w_0 wird wie alle anderen Gewichte mit dieser Regel angepasst. Es muss auch bei der Berechnung des Produkts $\vec{x} \cdot \vec{w}$ berücksichtigt werden, wobei x_0 konstant auf 1 gesetzt wird.⁴

Die Eingabedaten aus der obigen Wahrheitstabelle, ergänzt um $x_0 = 1$, sollen als Trainingsdaten verwendet werden.

1. Trainieren Sie mit diesen Daten ein Perzeptron.

Verwenden Sie die Lernrate $\eta = 1$ sowie die folgenden Startgewichte:

$$w_0 = -1,0; \quad w_1 = 1,0; \quad w_2 = -2,0$$

In der folgenden Tabelle ist der erste Trainingsschritt schon durchgeführt worden. Tragen Sie den weiteren Trainingsverlauf in der Tabelle ein:

³Die Aufgabe wurde mit Unterstützung von ChatGPT erstellt.

⁴Bitte beachten Sie, dass in der ursprünglichen Version der Folie 82 der Vorlesungsfolien der Wert x_0 fälschlicherweise nicht berücksichtigt wurde. Dieser Fehler wurde inzwischen korrigiert – die überarbeitete Version der Folie steht nun auf Stud.IP zur Verfügung.

	Erw. Eingabe ($x_0; x_1; x_2$)	Gewichte ($w_0; w_1; w_2$)	$x \cdot w$	Ausgabe y	Ziel- ausgabe t	$t - y$	Neue Gewichte
1	(1; 0; 0)	(-1,0; 1,0; -2,0)	-1,0	0	1	1	(0,0; 1,0; -2,0)
2	(1; 0; 1)	(0,0; 1,0; -2,0)					
3	(1; 1; 0)						
4	(1; 1; 1)						
1	(1; 0; 0)						
2	(1; 0; 1)						
3	(1; 1; 0)						
4	(1; 1; 1)						

- Welche Gewichte wurden nach zwei Runden Training gelernt?
- Stellen Sie die gelernten Gewichte graphisch als Entscheidungsgerade im (x_1, x_2) -Koordinatensystem dar.

4 Die Sigmoid Aktivierungsfunktion

In der Vorlesung haben wir die Sigmoid Aktivierungsfunktion kennengelernt:

$$\varphi(x) = \frac{1}{1 + e^{-x}}$$

a) Zeigen Sie, dass für die Sigmoidfunktion gilt:

$$\varphi(x)' = \varphi(x) \cdot (1 - \varphi(x))$$

Hinweis: Es kann hilfreich sein, $\varphi(x)'$ und auch $\varphi(x) \cdot (1 - \varphi(x))$ auszurechnen und zu vergleichen.

b) Zeigen Sie, dass die Sigmoidfunktion die folgende Eigenschaft hat:

$$\varphi(-x) = 1 - \varphi(x)$$