

# Einführung in die Künstliche Intelligenz

## Übungszettel 2

Prof. Dr. Claudia Schon

[C.Schon@hochschule-trier.de](mailto:C.Schon@hochschule-trier.de)

Fachbereich Informatik

Hochschule Trier

# 1 Definitionen

Definieren Sie kurz die folgenden Begriffe:

- (a) Hypothese
- (b) Hypothesenraum
- (c) Klassifikation
- (d) Regression
- (e) Ockhams Razor

## 2 Lernformen <sup>1</sup>

Finden Sie drei Beispiele aus Ihrem Alltag, wo maschinelles Lernen eingesetzt wird. Entscheiden oder recherchieren Sie jeweils, ob ein überwachter, unüberwachter oder verstärkender Lernansatz zugrunde liegt? Begründen Sie jeweils Ihre Entscheidung.

---

<sup>1</sup>Aufgabe wurde mit Unterstützung von ChatGPT erstellt.

### 3 Aufteilung von Daten <sup>2</sup>

Wir möchten ein maschinelles Lernverfahren trainieren, E-Mails in *Spam* und *Kein Spam* zu unterteilen. Dabei soll die Betreffzeile zur Vorhersage genutzt werden. Gegeben ist der folgende Datensatz:

Nr	Betreffzeile	Klasse
1	Jetzt Geld verdienen!	Spam
2	Ihre Rechnung ist da	Kein Spam
3	Sofort reich werden	Spam
4	Meeting morgen um 10 Uhr	Kein Spam
5	Gratis Urlaub gewinnen!	Spam
6	Projektabschluss heute	Kein Spam
7	Wichtige Informationen zur Steuer	Kein Spam
8	Update zur Lieferung	Kein Spam
9	Sie haben gewonnen!	Spam
10	Einladung zum Betriebsausflug	Kein Spam
11	Neue Termine für die Vorlesung	Kein Spam
12	Bestellung erfolgreich	Kein Spam

- (a) Wir betrachten die folgende Einteilung des Datensatzes in Trainings- und Testdaten:

**Trainingsdaten:**

- Beispiele 2, 3, 4, 6, 7, 8, 10, 11, 12

**Testdaten:**

- Beispiele 1, 5, 9

Ist diese Einteilung der Trainings- und Testdaten sinnvoll? Begründen Sie Ihre Antwort.

- (b) Der Datensatz wurde um zwei E-Mails erweitert:

Nr	Betreffzeile	Klasse
13	Sonderangebot für Sie!	Spam
14	Ihre Gehaltsabrechnung ist da	Kein Spam

Die Daten werden wie folgt aufgeteilt:

**Trainingsdaten:**

- Beispiele 2, 3, 4, 5, 6, 8, 9, 10, 11, 12, 13, 14

<sup>2</sup>Aufgabe wurde mit Unterstützung von ChatGPT erstellt.

**Testdaten:**

- Beispiele 1, 4, 7, 13

Ist diese Einteilung der Trainings- und Testdaten sinnvoll? Begründen Sie Ihre Antwort.

**Hinweis:** In der Praxis werden deutlich größere Datensätze verwendet. Der hier verwendete kleine Datensatz dient ausschließlich der Veranschaulichung.

## 4 Entscheidungsbäume anwenden

Wir betrachten den Entscheidungsbaum auf Vorlesungsfolie 26.

- (a) Wir wenden den Entscheidungsbaum auf neue Datensätze an. Entscheiden Sie für jeden der folgenden Datensätze, ob wir im Restaurant auf einen Tisch warten oder nicht.

<i>Alt</i>	<i>Bar</i>	<i>Fri</i>	<i>Hun</i>	<i>Pat</i>	<i>Price</i>	<i>Rain</i>	<i>Res</i>	<i>Type</i>	<i>Est</i>	<i>WillWait</i>
<i>Yes</i>	<i>No</i>	<i>No</i>	<i>Yes</i>	<i>Full</i>	<i>\$\$</i>	<i>Yes</i>	<i>No</i>	<i>Thai</i>	<i>30-60</i>	<i>?</i>
<i>No</i>	<i>Yes</i>	<i>Yes</i>	<i>Yes</i>	<i>Full</i>	<i>\$\$\$</i>	<i>No</i>	<i>No</i>	<i>Italian</i>	<i>10-30</i>	<i>?</i>
<i>Yes</i>	<i>No</i>	<i>No</i>	<i>Yes</i>	<i>Full</i>	<i>\$</i>	<i>Yes</i>	<i>Yes</i>	<i>Thai</i>	<i>0-10</i>	<i>?</i>

- (b) Geben Sie einen Beispieldatensatz an, bei dem die Entscheidung möglichst früh im Entscheidungsbaum getroffen wird.
- (c) Geben Sie einen Beispieldatensatz an, bei dem der Entscheidungsbaum bei der Entscheidung bis zur maximalen Tiefe betrachtet werden muss.

## 5 Entscheidungsbäume erstellen<sup>3</sup>

Wir betrachten Datensätze, mit den binären Attributen *Attribut 1*, *Attribut 2* und *Attribut 3*. Ausgabe ist eine binäre Variable *Klasse*.

- (a) Entwerfen Sie manuell für den folgenden Datensatz einen Entscheidungsbaum. Versuchen Sie einen möglichst kleinen Entscheidungsbaum zu konstruieren.

<i>Attribut 1</i>	<i>Attribut 2</i>	<i>Klasse</i>
Yes	Yes	No
Yes	No	Yes
No	Yes	No
No	No	No

- (b) Entwerfen Sie manuell für den folgenden Datensatz einen Entscheidungsbaum. Versuchen Sie einen möglichst kleinen Entscheidungsbaum zu konstruieren.

<i>Attribut 1</i>	<i>Attribut 2</i>	<i>Attribut 3</i>	<i>Klasse</i>
Yes	Yes	Yes	Yes
Yes	Yes	No	Yes
Yes	No	Yes	Yes
Yes	No	No	Yes
No	Yes	Yes	Yes
No	Yes	No	No
No	No	Yes	No
No	No	No	No

- (c) Entwerfen Sie manuell für den folgenden Datensatz einen Entscheidungsbaum. Versuchen Sie einen möglichst kleinen Entscheidungsbaum zu konstruieren.

<i>Attribut 1</i>	<i>Attribut 2</i>	<i>Klasse</i>
Yes	Yes	No
Yes	No	Yes
No	Yes	Yes
No	No	No

<sup>3</sup>Aufgabe basiert auf einer Aufgabe aus: Mitchell, T. M. (1997). *Machine learning (Vol. 1)*. McGraw-hill New York.

## 6 Entwurf von Entscheidungsbäumen<sup>4</sup>

Erstellen Sie einen Entscheidungsbaum für das Problem, ob Sie an einer Straßenkreuzung weiterfahren sollen, wenn die Ampel gerade auf Grün umgeschaltet hat.

Da keine Datensätze vorliegen, sollen Sie sich selbst sinnvolle Attribute (z.B. Verkehrssituation, Sichtverhältnisse, Verhalten anderer Verkehrsteilnehmer etc.) sowie jeweils passende Attributwerte überlegen, die bei dieser Entscheidung eine Rolle spielen könnten.

---

<sup>4</sup>Aufgabe basiert auf einer Aufgabe aus: *Stuart J. Russell, Peter Norvig. Artificial Intelligence: A Modern Approach (4th Edition). Pearson, 2020.* und wurde mit Unterstützung von ChatGPT erstellt.