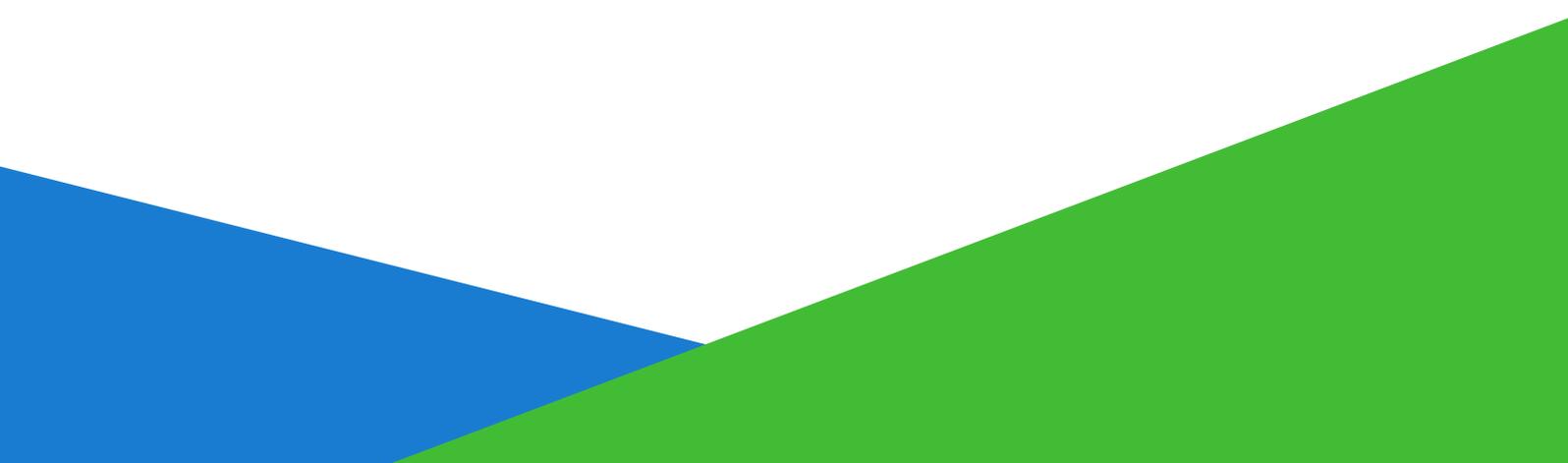


Workshopkonzept mit Kopiervorlagen

Klassifikationsaufgaben

Was Klassifikationsaufgaben sind und wie diese algorithmisch
gelöst werden können



Dr. Daniel Scholz
Braunschweig
daniel@taralino.de

Herausgeber

taralino e. V.
Braunschweig
post@taralino.de
www.taralino.de



Dieses Werk ist lizenziert unter einer Creative Commons Namensnennung, Nicht-kommerziell, Weitergabe unter gleichen Bedingungen 4.0 International Lizenz (CC BY-NC-SA 4.0).

<https://creativecommons.org/licenses/by-nc-sa/4.0>

Dies bedeutet insbesondere:

- Das Material darf in beliebigen Formaten (z.B. ausgedruckt oder als digitale Version) uneingeschränkt weiterverbreitet werden.
- Das Material darf verändert und weiterentwickelt werden. Ist anschließend eine erneute Veröffentlichung geplant, muss dafür jedoch auch eine Creative Commons Lizenz verwendet werden.
- Eine Nutzung des Materials für kommerzielle Zwecke ist nicht gestattet.

Stand des vorliegenden Dokuments: 14. September 2022

Vorwort

Der vorliegende Material bietet einen spielerischen Einstieg in das Themenfeld des maschinellen Lernens und damit einem Teilgebiet der Künstlichen Intelligenz. Insbesondere wird darauf eingegangen, wie Algorithmen anhand von Trainingsdaten Klassifikationsaufgaben lösen können.

- Gruppengröße: 10 bis 30 Teilnehmende (aufgeteilt in 3 bis 6 Gruppen)
- Zeitaufwand: etwa 30 Minuten
- Zielgruppe: Schulklassen ab Jahrgang 9
- Material: Eine Schere pro Gruppe, ggf. Stift und Zettel für Notizen

2
/
3

Die Lehreinheit besteht im Wesentlichen in einer Gruppenarbeit, wobei drei bis sechs Gruppen mit je drei bis fünf Teilnehmenden gebildet werden sollten. Der Zeitrahmen kann flexibel gestaltet werden: In jüngeren Jahrgängen sollten mindestens 30 Minuten eingeplant werden, in gut organisierten Gruppen lässt sich das Vorhaben aber auch in 20 Minuten durcharbeiten. Empfohlen werden die Inhalte in Schulklassen ab dem neunten Jahrgang.

Neben dem Ausbau der algorithmischen Problemlösekompetenz wird zudem Wert darauf gelegt, die Bedeutung von Daten in unserer digitalen Welt bewerten zu können.

Viel Spaß beim Lernen und Lehren

Daniel Scholz

April 2022



Inhaltsverzeichnis

1	Klassifikationsaufgaben	6
2	Ablauf der Lehreinheit	8
3	Lernziele	10

1

Klassifikationsaufgaben

Der Begriff der Künstlichen Intelligenz ist allgegenwärtig. Dabei ist es kaum möglich, eine allgemeine Definition von „Künstlicher Intelligenz“ anzugeben, da dies bereits für das Wort „Intelligenz“ schwierig ist. Verwendet wird der Begriff jedoch häufig im Zusammenhang mit einer algorithmischen Entscheidungsfindung.

Genauer lässt sich hingegen das maschinelle Lernen beschreiben: Dabei handelt es sich um Verfahren, die anhand von Erfahrungen in Form von Daten gewissen Entscheidungen oder Vorhersagen treffen können.



Abb. 1 Zur Einsortierung der Begriffe Künstliche Intelligenz, maschinelles Lernen und Klassifikationsaufgaben

Bei Klassifikationsaufgaben handelt es sich wiederum um ein Teilgebiet des maschinellen Lernens (Abb. 1). Um genauer zu sein: Eingangsdaten gehören jeweils einer Kategorie bzw. Klasse an und die Aufgabe besteht darin, anhand der Eingangsdaten die zugehörige Klasse zu bestimmen. Neben der Klassifikation von Bildern (z.B. handgeschriebene Ziffern oder Verkehrszeichen) finden sich klassische Anwendungsbeispiele in der Spamerkennung, der Bewertung einer Kreditwürdigkeit oder auch bei medizinischen Diagnosen.

Um Schülerinnen und Schüler an das Thema heranzuführen, wird eine Aufgabe zur Klassifikation von Früchten (Obst und Gemüse) vorgeschlagen. Wie bei allen Klassifikationsaufgaben werden dafür Trainingsdaten benötigt: Dabei handelt es sich um ein Set von Eingangsdaten mit jeweils bekanntem Klassifikationsergebnis. Ein Trainingsset bestehend aus drei Datensätzen zeigt Abb. 2. Die Eingangsdaten sind jeweils eine Frucht und das Klassifikationsergebnis wird anhand des Buchstabens im grauen Kreis veranschaulicht.

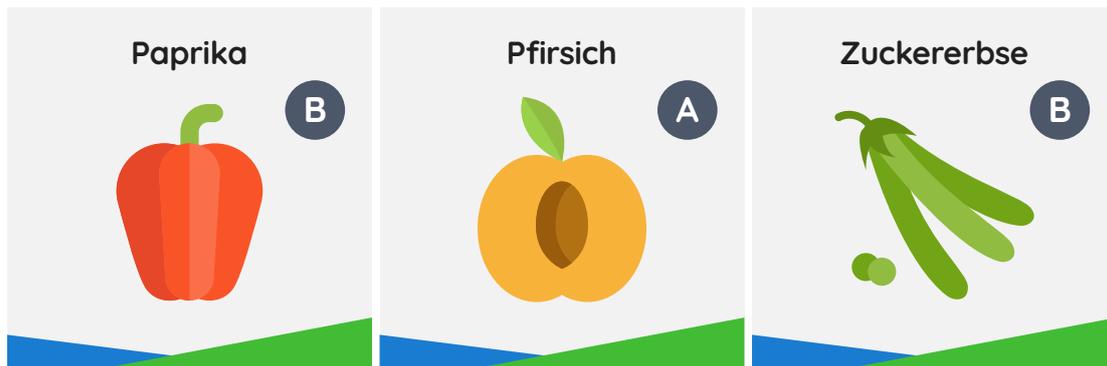


Abb. 2 Beispiel eines Trainingssets zur Klassifikation von Früchten

Das Klassifikationsergebnis hängt davon ab, welche Klassifikationsaufgabe gelöst werden soll. Falls die Früchte wie im Beispiels aus Abb. 2 in Obst und Gemüse zu klassifizieren sind, so steht der Buchstabe A für Obst und der Buchstabe B für Gemüse.

Ist ein Trainingsset mit einer ausreichend großen Anzahl an Datensätzen gegeben, kann anschließend ein Set an Testdaten ausgewertet werden. Dabei handelt es sich um Eingangsdaten, dessen Klassifikationsergebnis noch nicht bekannt ist. Anhand des Trainingssets soll nun eine Entscheidung darüber getroffen werden, welcher Klasse die Eingangsdaten des Testsets (wahrscheinlich) angehören.

Trainingssets beziehen sind stets auf eine explizite Klassifikationsaufgabe.

Wie beschrieben können Früchte beispielsweise in Obst und Gemüse klassifiziert werden. Weitere Möglichkeiten sind eine Klassifikation bezüglich der Größe oder bezüglich der Farbe der Früchte.

2

Ablauf der Lehreinheit

Die Lehreinheit beginnt, indem Kleingruppen bestehend aus jeweils drei bis fünf Teilnehmenden gebildet werden. Jeder Kleingruppe wird eines der drei möglichen Trainingssets überreicht (siehe Kopiervorlagen). In der gesamten Gruppe sollten dabei insgesamt alle drei Trainingssets zum Einsatz kommen. Nun wird jeder Kleingruppe folgende Aufgabe gestellt:



1. Zerschneidet das Trainingsset in die 15 abgebildeten Karten.
2. Sortiert die Karten bezüglich der jeweils abgebildeten Buchstaben, sodass die Früchte in zwei bzw. drei Kategorien aufgeteilt werden.
3. Diskutiert, welche Kategorien gebildet wurden, d.h., was die Früchte einer Kategorie gemeinsam haben.

Nach einer Arbeitsphase von ca. 10 Minuten kommt das Testset zum Einsatz. Der Moderierende zeigt eine Frucht bzw. Karte und fragt jede Kleingruppe reihum, in welche Kategorie die Frucht zu klassifizieren ist und warum (auf Grundlage des jeweils vorliegenden Trainingssets). Dies kann mit allen sechs Karten des Testsets wiederholt werden.

In der gesamten Gruppe sollte sich nun herausstellen, dass je Trainingsset unterschiedliche Klassifikationsaufgaben verfolgt wurden:

1. Klassifikation bezüglich der Farbe. (A: Gelb, B: Grün, C: Rot)
2. Klassifikation bezüglich Obst und Gemüse. (A: Obst, B: Gemüse)
3. Klassifikation bezüglich der Größe. (A: klein, B: groß, C: länglich)

Die Auswertung der einzelnen Karten des Testsets ist – je nach Klassifikationsaufgabe – bewusst nicht immer eindeutig: Heidelbeeren können beispielsweise nicht bezüglich der Farben klassifiziert werden, da das zugehörige Trainingsset nur die Farben Gelb, Grün und Rot berücksichtigt.

Abschließend sollte in der Gruppe darüber diskutiert werden, dass eine algorithmische Vorgehensweise zur Lösung von Klassifikationsaufgaben erarbeitet wurde. Zum Einstieg in die Diskussion bieten sich zum Beispiel die folgenden Fragestellungen an:

1. Beschreibe in eigenen Worten, was Klassifikationsaufgaben sind.
2. Welche Rolle spielen Trainingsdaten bei der algorithmischen Lösung von Klassifikationsaufgaben?
3. Welche Herausforderungen gibt es bei der algorithmischen Lösung von Klassifikationsaufgaben?
4. Welche Anwendungsfälle von Klassifikationsaufgaben fallen dir ein?

3

Lernziele

Die Lernziele der Lehreinheit sind ähnlich zu den Antworten der Diskussionsfragen zuvor. Als kleine Zusammenfassung:



1. Klassifikationsaufgaben bestehen darin, Eingangsdaten (z.B. Bilder) in zuvor festgelegte Klassen zu kategorisieren.
2. Trainingsdaten sind der wichtigste Baustein bei der algorithmischen Lösung von Klassifikationsaufgaben. Nur dadurch kann anschließend entschieden werden, welcher Klasse weitere Testdaten angehören (Prinzip Entscheidungen aus Erfahrungen treffen).
3. Wie gut eine Klassifikationsaufgabe gelöst werden kann, hängt zentral mit dem Umfang und der Güte der Trainingsdaten zusammen: Je geeigneter und je größer der Umfang des Trainingssets ist, desto besser kann die zugehörige Klassifikationsaufgabe gelöst werden.
4. Neben der Klassifikation von Bildern (z.B. handgeschriebene Ziffern oder Verkehrszeichen) finden sich klassische Anwendungsbeispiele in der Spamerkennung, der Bewertung einer Kreditwürdigkeit oder auch bei medizinischen Diagnosen.

In diesem Zusammenhang nicht zu vernachlässigen ist der allgemeine Umgang mit Daten (Datenschutz und Datenethik). Wir sollten uns bewusst sein, wie wir mit Daten umgehen wollen und wofür diese genutzt werden können.

Bildernachweise

Die folgende Liste dient als Nachweis über verwendete Fotomaterial (in ihrer jeweiligen Originalversion):

Seite 4: Photo by *Nathan Dumlao* on Unsplash

Alle Obst- und Gemüse-Vektorgrafiken wurden der Plattform www.svgrepo.com entnommen (abgerufen am 15. März 2022). Die verwendeten Grafiken werden (bzw. wurden zum Zeitpunkt des Downloads) unter der Creative Commons Public Domain 1.0 Universell License vertrieben:



www.creativecommons.org/publicdomain/zero/1.0/

10
/
11

Kopiervorlagen

Pfirsich



A

Paprika



C

Zuckererbse



B

Banane



A

Kirsche



C

Rosenkohl



B

Brokkoli



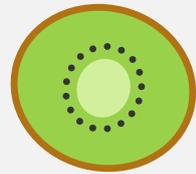
B

Möhre



A

Kiwi



B

Erdbeere



C

Ananas



A

Gurke



B

Peperoni



C

Apfel



C

Kürbis



A

Pfirsich



A

Paprika



B

Zuckererbse



B

Banane



A

Kirsche



A

Birne



A

Brokkoli



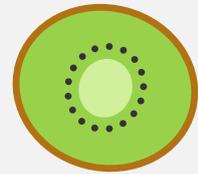
B

Möhre



B

Kiwi



A

Radieschen



B

Ananas



A

Gurke



B

Peperoni



B

Apfel



A

Kürbis



B

t

Himbeere



A

Apfel



B

Zuckererbse



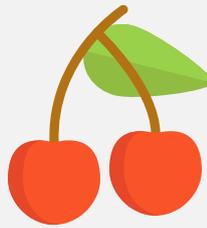
C

Banane



C

Kirsche



A

Birne



B

Brokkoli



B

Möhre



C

Rosenkohl



A

Radieschen



A

Ananas



B

Gurke



C

Peperoni



C

Erdbeere



A

Kürbis



B

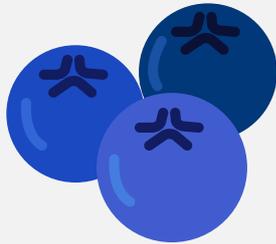
Mais



Tomate



Heidelbeeren



Zwiebel



Fliegenpilz



Croissant



t