

Pädagogische Hochschule Karlsruhe
Bachelorarbeit

Entwicklung und Evaluation eines
Workshop-Konzepts zur Einführung in das
überwachte Lernen

Institut für Informatik und digitale Bildung

Erstgutachter: Jun.-Prof. Dr. Bernhard Standl

Zweitgutachter: FSchR Vincenzo Fragapane

Verfasserin: **Sabine Britner**

Matrikelnummer: 3200735

E-Mail: sabine.britner@stud.ph-karlsruhe.de

Studiengang: Bachelor Lehramt Sekundarstufe 1

Fach 1: Informatik

Fach 2: Mathematik

Abgabedatum: 20.10.2022

Zusammenfassung

Obwohl maschinelles Lernen Bestandteil des alltäglichen Lebens ist, weisen Heranwachsende Defizite in den Kenntnissen über die zugrundeliegenden Konzepte auf. Das hat Konsequenzen für die Wahrnehmung solcher Technologien. Deshalb ist es notwendig Interventionen anzubieten, die Jugendliche über die Ideen hinter maschinellem Lernen aufklären. Die hier vorliegende Arbeit befasst sich mit der folgenden Forschungsfrage: Wie sieht eine geeignete Lernumgebung aus, um Lernende der Sekundarstufe I (ab Klasse 7) in die Grundideen überwachten Lernens, einer Teilkategorie maschinellen Lernens, einzuführen? Ziel ist dabei die Evaluierung bzw. Optimierung eines Workshop-Konzepts, das im Rahmen dieser Arbeit für die Einführung Lernender der Sekundarstufe I in das überwachte Lernen designt wurde. Um die Forschungsfrage zu beantworten, wurde das Workshop-Konzept einer quantitativen Einschätzung durch Expertinnen aus der Praxis mittels eines Fragebogens unterzogen. Ergänzt wurden die Antworten im Fragebogen durch eine qualitative Rückmeldung. Nach einer ersten Einschätzung durch die Expertinnen wurde der Workshop überarbeitet und erneut zur Bewertung freigegeben. Die Ergebnisse deuten darauf hin, dass das Workshop-Konzept nach Überarbeitung das Kriterium der Eignung aus der Forschungsfrage erfüllt.

Inhaltsverzeichnis

1 Einleitung	1
2 Fachsystematische Grundlagen	3
2.1 Künstliche Intelligenz	3
2.2 Maschinelles Lernen	4
2.2.1 Definition	4
2.2.2 Überwachtes Lernen	5
2.2.3 Unüberwachtes Lernen	9
2.2.4 Verstärkendes Lernen	12
2.2.5 Künstliche neuronale Netze	14
3 Workshop-Konzept	17
3.1 Themenauswahl	17
3.2 Voraussetzungen und Ziele	18
3.3 Bezug zum Bildungsplan Baden-Württemberg 2016	19
3.4 Aktivitäten und Materialien	19
4 Methode	21
4.1 Herangehensweise	21
4.2 Expertinnen	22
4.3 Instrument	22
4.4 Vorgehensweise	23
4.5 Analyse	23
5 Ergebnisse und Diskussion	25
5.1 Erste Befragung	25
5.2 Zwischendiskussion	27
5.3 Zweite Befragung	29
5.4 Vergleich beider Befragungen	31
5.5 Abschließende Diskussion	32
6 Fazit	34

Literaturverzeichnis	35
Abbildungsverzeichnis	40
Tabellenverzeichnis	42
Anhänge	43
A Ergänzende Informationen zur Fachsystematik	43
B Literatur zu maschinellen Lernalgorithmen	51
C Verlaufsplan und Workshop-Materialien	55
D Computer Science Lesson Analysis: Fragebogen	114
E Ausgefüllte Fragebögen und Rückmeldungen	118
F Re-Design: Verlaufsplan und Workshop-Materialien	128
Eidesstattliche Erklärung	187

Abkürzungsverzeichnis

Abb.	Abbildung
Abk.	Abkürzung
bzw.	beziehungsweise
Def.	Definition
engl.	englisch
FFNN	Feed Forward Neural Networks
ggf.	gegebenenfalls
IRA	Interrater Agreement
Kap.	Kapitel
KI	Künstliche Intelligenz
KNN	Künstliche neuronale Netze
ML	Maschinelles Lernen
Tab.	Tabelle
vgl.	vergleiche
z.B.	zum Beispiel

1. Einleitung

Das maschinelle Lernen (Abk. ML) stellt ein Teilgebiet der künstlichen Intelligenz (Abk. KI) dar [52], dessen Technologien mittlerweile Bestandteil unseres Alltags sind. Anwendungen wie Film-, Musik- oder Produktempfehlungen basieren auf Verfahren des maschinellen Lernens [24] und werden täglich von vielen Menschen genutzt. [8] schreibt dem maschinellen Lernen die Bedeutung zu, die wichtigste Universaltechnologie unserer Zeit zu sein.

Besonders für Jugendliche ist der Umgang mit Anwendungen von KI bzw. ML selbstverständlich. Der Großteil der Heranwachsenden besitzt ein Smartphone mit Internetzugang, das beispielsweise für den Zugriff auf Streaming-Dienste für Musik oder Filme genutzt wird [16]. Gleichzeitig hat eine Untersuchung ergeben, dass Schülerinnen und Schüler nur geringe Kenntnisse über die zugrunde liegenden Technologien von künstlicher Intelligenz verfügen. Zudem werden diese von vielen Jugendlichen mit Gefahr oder Angst assoziiert, was auch in Verbindung zu den mangelnden Kenntnissen steht. [31]. In einer weiteren Studie, die sich speziell auf das maschinelle Lernen bezieht, lassen sich zwei gegensätzliche Vorstellungen der Teilnehmenden zu ML finden. Während eine Gruppe der Lernenden davon ausgeht, dass Maschinen eigenständig lernen (Anteil des Menschen wird unterschätzt), nehmen andere Schülerinnen und Schüler den Lernprozess von Maschinen so wahr, dass dieser fast vollständig vom Menschen ausgeht (Anteil des Menschen wird überschätzt). [19]. Unter Berücksichtigung der Untersuchungsergebnisse und mit Blick auf die Entmystifikationsaufgabe des Informatikunterrichts [29] sollte die Aufklärung über die Konzepte hinter KI bzw. ML im Informatikunterricht eine Rolle spielen.

Die hier vorliegende Arbeit beschäftigt sich mit der Fragestellung, wie eine geeignete Lernumgebung aussieht, um Schülerinnen und Schüler der Sekundarstufe I (ab Klasse 7) in die Konzepte des überwachten Lernens einzuführen. Das überwachte Lernen ist eine von drei Hauptformen des maschinellen Lernens [28]. Den Ausgangspunkt der Untersuchung stellt das von der Verfasserin dieser Arbeit entwickelte Konzept eines Workshops dar, der sich an Lernende der Sekundarstufe I richtet und auf die Einführung in die Grundideen überwachten Lernens abzielt. Das Ziel der Arbeit ist die Evaluierung bzw. Optimierung des Workshop-Konzepts. In Anlehnung an die designbasierte Forschung wird die Eignung der Workshop-Planung durch Expertinnen aus der Praxis entlang eines Fragebogens eingeschätzt. Der Begriff der Eignung aus der Forschungsfrage wird in Kap. 4.3 operationalisiert.

In Kapitel zwei werden die fachsystematischen Grundlagen des maschinellen Lernens betrachtet. Danach wird in Kapitel drei das Workshop-Konzept vorgestellt, das im Rahmen der Arbeit evaluiert wird. Das vierte Kapitel gibt einen Einblick in die zur Bearbeitung der Forschungsfrage verwendete Methode. Darauf aufbauend werden in Kapitel fünf die Ergebnisse präsentiert und diskutiert. Die Arbeit schließt mit Kapitel sechs, dem Fazit, ab.

2. Fachsystematische Grundlagen

Dieses Kapitel behandelt die fachsystematischen Grundlagen maschinellen Lernens. Um das maschinelle Lernen kontextuell einordnen zu können, wird zunächst der Begriff der künstlichen Intelligenz eingeführt. Anschließend wird eine Begriffsdefinition von maschinellem Lernen vorgenommen. Der Themenbereich wird dann nach Lerntypen differenziert weiter vertieft. Künstliche neuronale Netze bilden den Abschluss des Kapitels.

2.1 Künstliche Intelligenz

Mit künstlicher Intelligenz (engl. *artificial intelligence*) wird ein Teilbereich der Informatik beschrieben [9, 52]. Eine Begriffserklärung liefert Def. 2.1. Diese Definition spiegelt die Thematik nur unscharf wider, da bereits eine eindeutige Begriffsdefinition von Intelligenz problematisch ist [5, 9]. Zudem ist es von der zeitlichen Einordnung abhängig, ob eine Technologie als künstliche Intelligenz wahrgenommen wird. So zählten Navigationssysteme vor einigen Jahren als künstliche Intelligenz. Heutzutage wird die Technologie als selbstverständlich angesehen. [24].

Definition 2.1 Der Begriff „künstliche Intelligenz“ umfasst diejenigen Computersysteme, die Aufgaben und Probleme lösen können, für die man Intelligenz benötigt [5, 24].

In der Literatur werden zwei Arten künstlicher Intelligenz unterschieden. Unter den Begriff der starken KI fallen Methoden, die versuchen Prozesse im menschlichen Gehirn nachzuahmen. Diese Kategorie ist durch Merkmale wie Empathie oder Bewusstsein charakterisiert. Die erfolgreiche Entwicklung einer starken KI ist aktuell nicht bekannt. Eine schwache KI hingegen wird gezielt für die Lösung klar definierter und abgegrenzter Problemstellungen eingesetzt. Schwache KI-Systeme sind inzwischen technisch umsetzbar. [9].

KI wird für unterschiedliche Aufgaben eingesetzt [52]. Dazu gehören beispielsweise Sprachverständnis, Bilderkennung, Hilfestellungen in der Medizin und autonomes Fahren [52] sowie Produkt- und Filmempfehlungen [24]. Dabei basieren die meisten KI-Systeme aktuell auf den Methoden des maschinellen Lernens [24].

2.2 Maschinelles Lernen

2.2.1 Definition

Bei maschinellem Lernen (engl. *machine learning*) handelt es sich um ein Teilgebiet der künstlichen Intelligenz [52]. In [23] wird maschinelles Lernen als der wichtigste KI-Bereich hervorgehoben. Die Beziehung der Bereiche Informatik, künstliche Intelligenz und maschinelles Lernen ist in Abb. 2.1 visualisiert. In der Literatur wird maschinelles Lernen häufig angelehnt an Samuel (vgl. [20, 52], siehe Def. 2.2) und/oder nach Mitchell (vgl. [9, 20, 47], siehe Def. 2.3) definiert.

Definition 2.2 Durch maschinelles Lernen können Algorithmen aus der Erfahrung bestimmte Verhaltensweisen erlernen, ohne dass eine explizite Programmierung dieser erfolgen muss [44].

Definition 2.3 Ein Computerprogramm lernt aus der Erfahrung eine bestimmte Aufgabe zu lösen, wenn sich seine Leistung in Bezug auf diese Aufgabe mit der Erfahrung verbessert [38].

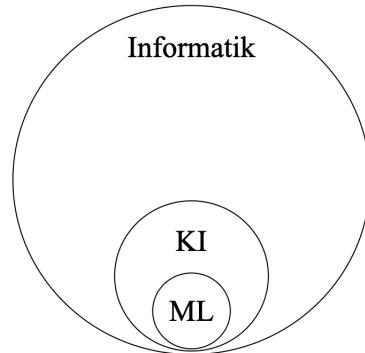


Abbildung 2.1: Abgrenzung und Verortung von KI und ML. In Anlehnung an [52], S. 56.

Man bezeichnet Algorithmen beim maschinellen Lernen als Lernalgorithmen [28]. Ein Lernalgorithmus kann beispielsweise die Fähigkeit erlangen, Hunde und Katzen auf Bildern zu erkennen, ohne dass typische Eigenschaften der Tiere beim Programmieren im Code hinterlegt werden müssen (vgl. Def. 2.2). Stattdessen lernt der Algorithmus das Erkennen der Tiere eigenständig. [9]. Damit dieses Verhalten funktionieren kann, benötigt ein Algorithmus Erfahrung (vgl. Def. 2.2 und Def. 2.3). Nach [5] kann das Lernen aus der Erfahrung beim maschinellen Lernen auf zwei Arten erfolgen. Dabei können drei Hauptformen des Lernens unterschieden werden:

- a. **Lernen aus Daten.** Algorithmen können aus Datensätzen lernen. In diese Kategorie gehören das überwachte und das unüberwachte Lernen (engl. *supervised* und *unsupervised learning*). [5]. Der Begriff der Daten wird in Anhang A.1 genauer beleuchtet.

b. Lernen durch Bestrafung oder Belohnung. Lernen Algorithmen über Belohnung oder Bestrafung eine Verhaltensstrategie zu optimieren, so spricht man von verstärkendem oder bestärkendem Lernen (engl. *reinforcement learning*) [5, 38, 47].

Um das überwachte und unüberwachte Lernen voneinander abgrenzen zu können, bedarf es einer weiteren Differenzierung, die in Kap. 2.2.2 und Kap. 2.2.3 vorgenommen wird.

2.2.2 Überwachtes Lernen

Damit überwachte Lernalgorithmen aus Daten lernen können, müssen Lösungen auf den Daten bereitgestellt werden [20, 28]. Ein Beispiel: Soll ein Algorithmus erlernen, Hunde und Katzen auf Bildern zu erkennen, so muss man dem Algorithmus Bilder zur Verfügung stellen, die mit „Hund“ oder „Katze“ beschriftet sind [9, 47]. Die Lösungen auf den Daten sind in diesem Fall die Beschriftungen der Bilder. Diese werden als Labels bezeichnet [20].

Mithilfe von gelabelten Daten können überwachte Lernalgorithmen Klassifikations- und Regressionsaufgaben lösen [23, 52]. Def. 2.4 liefert eine Begriffserklärung für die Klassifikation. Ein Beispiel ist in Abb. 2.2 dargestellt. Die Murmeln werden nach Farbe klassifiziert.

Definition 2.4 Eine Klassifikation beschreibt das Einordnen von Objekten in vorgegebene Kategorien, die Klassen genannt werden [23].

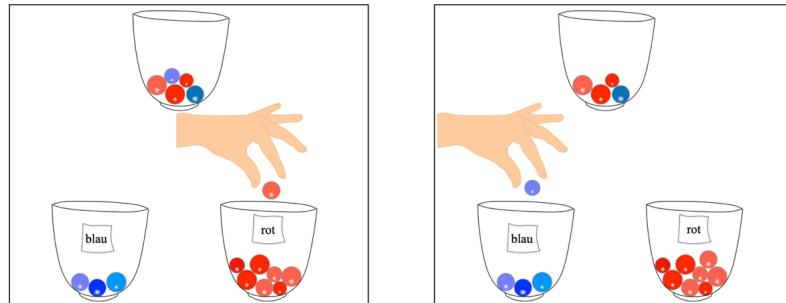
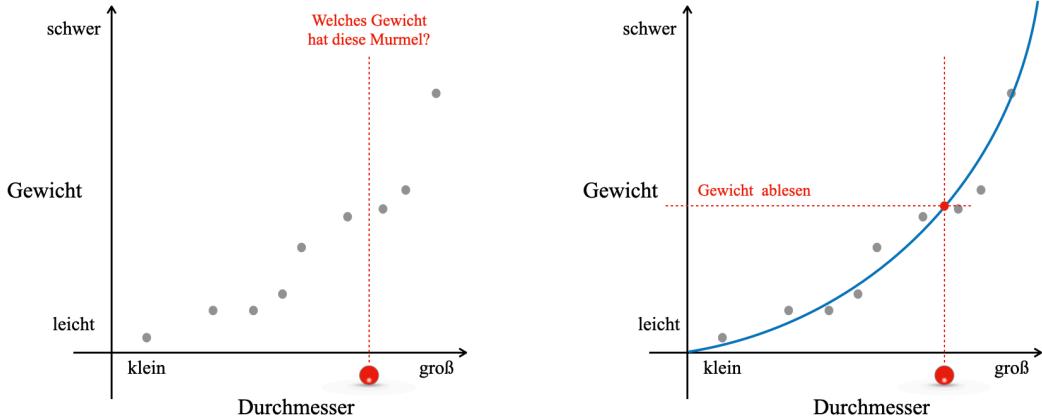


Abbildung 2.2: Murmeln werden den Klassen „rot“ und „blau“ zugeordnet. Eigene Darstellung.

Damit ein Algorithmus Daten klassifizieren kann, muss er die vorgegebenen Klassen erlernen. Dabei wird eine Klasse durch positive und negative Beispiele erlernt. Positive Beispiele sind Beispiele, die in die zu erlernende Klasse einzuordnen sind, während negative Beispiele nicht in diese Klasse gehören. Unter Erlernen versteht man in diesem Zusammenhang das Definieren von Eigenschaften, die alle positiven Beispiele gemeinsam haben, aber auf keines der negativen Beispiele zutreffen. [3]. Bezogen auf das Beispiel von oben: Wenn ein Lernalgorithmus erlernen soll Hunde- und Katzenfotos nach Tierart zu klassifizieren, dienen Bilder, die mit „Katze“ gelabelt sind, als positive Beispiele für die Klasse „Katze“ und als negative Beispiele für die Klasse „Hund“. Aufgabe des Algorithmus ist es, Merkmale zu finden, die die Klassen charakterisieren und voneinander abgrenzen [28].

Wie man Def. 2.5 entnehmen kann, geht es bei der Regression um die Vorhersage stetiger Zielgrößen aus Eingabegrößen. So kann ein Algorithmus etwa erlernen, aus Bildern von Katzen und Hunden das Tiergewicht zu bestimmen [52]. Durch die Verwendung von Eingabedaten mit bekannter Lösung (Werte für die Zielgröße sind bekannt) kann der Lernalgorithmus den Zusammenhang zwischen Eingabegröße und Zielgröße durch eine Funktion annähern [3]. So kann für jeden beliebigen Wert die Zielgröße bestimmt werden [26]. In Abb. 2.3 ist ein Beispiel veranschaulicht. Aus dem Durchmesser einer Glas-Murmel soll das Gewicht der Murmel vorhergesagt werden.

Definition 2.5 Die Regression beschäftigt sich mit der Vorhersage stetiger Zahlenwerte [23]. Die vorherzusagende Größe wird als Zielgröße bezeichnet [26]. Für seine Vorhersage werden Größen aus den Eingabedaten unter der Annahme gewählt, dass sie die Zielgröße beeinflussen [3].



(a) Bekannte Murmel-Gewichte werden über dem Murmel-Durchmesser aufgetragen. In Anlehnung an [26], S. 41.

(b) Schnittpunkt mit Regressionskurve liefert Vorhersage für das Murmel-Gewicht. In Anlehnung an [25], S. 65.

Abbildung 2.3: Vorhersage des Murmel-Gewichts mittels des Durchmessers einer Murmel.

Beim überwachten Lernen werden die Daten in Trainings- und Testdaten unterteilt. In der Praxis liegt das Verhältnis von Trainingsdaten zu Testdaten bei etwa 70:30. Zu Beginn werden einem überwachten Lernalgorithmus die Trainingsdaten bereitgestellt. Der Algorithmus untersucht die Daten auf Merkmale und Muster. Dieser Prozess wird als das Lernen im Lernalgorithmus verstanden. [28]. Anschließend wird die Fehlerquote [28] bzw. Treffergenauigkeit [5] des trainierten Algorithmus anhand von Testdaten geprüft. Die Treffergenauigkeit ist dabei der prozentuale Anteil der korrekt zugeordneten Testbeispiele an der Gesamtzahl der Testdaten [5]. Die Testdaten sind dem Lernalgorithmus noch unbekannt. Fällt der Test gut aus, kann der Algorithmus auf neuen Daten mit unbekannter Lösung angewendet werden. [51]. Wenn man mit dem Testergebnis des Lernalgorithmus nicht zufrieden ist, kann

man zum Training zurückkehren [45]. In Abb. 2.4 ist der Prozess vom Training zur Anwendung dargestellt. Will man also etwa, dass ein Lernalgorithmus die Klassifikation von Katzen und Hunden auf Fotos nach Tierart erlernt, so müssen die für diesen Zweck bereitgestellten Tierfotos in Trainings- und Testdaten aufgeteilt werden. Aus den Trainingsdaten erlernt der Algorithmus Eigenschaften, die die Klassen „Hund“ und „Katze“ auszeichnen und voneinander unterscheiden. Danach versucht der Lernalgorithmus die Testdaten der jeweils richtigen Klasse zuzuordnen. Ob eine Zuordnung richtig ist, wird über die zugehörigen Labels geprüft [28]. So kann die Fehlerquote bzw. die Treffergenauigkeit bestimmt werden. Bei der Bilderkennung wird eine Treffergenauigkeit von mindestens 90% [47] und damit eine Fehlerquote von maximal 10% angestrebt.

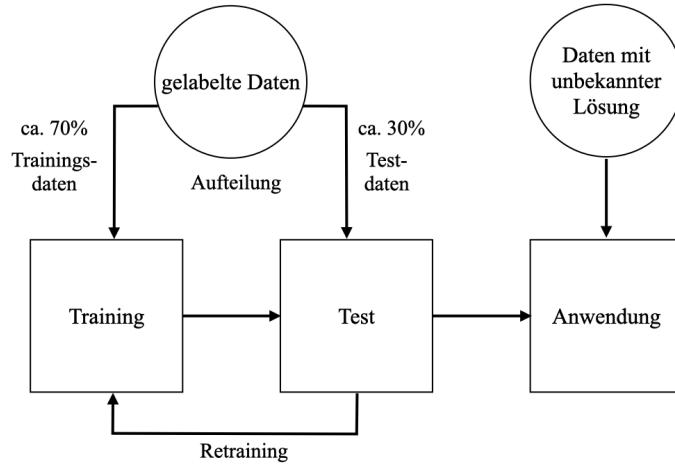


Abbildung 2.4: Vom Training zur Anwendung eines überwachten Lernalgorithmus. In Anlehnung an [45].

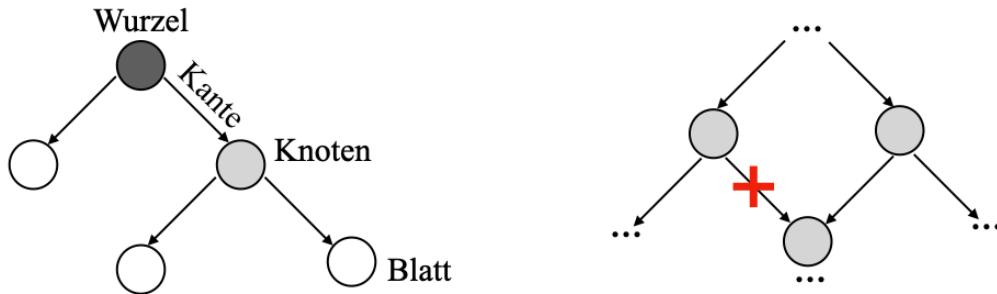
Dass ein Algorithmus beim Testen eine hohe Fehlerquote aufweist, kann verschiedene Ursachen haben. Beispielsweise spielt die Menge der bereitgestellten Trainingsdaten eine Rolle. Prinzipiell verbessert sich das Testergebnis mit Zunahme der Trainingsdatenmenge. In der Praxis liegt die Größenordnung im Bereich mehrerer Tausend Trainingsbeispiele. [28]. Häufig entstehen Fehler durch eine Überanpassung (engl. *overfitting*) [51]. Darunter versteht man, dass die vom Lernalgorithmus erlernten Merkmale und Muster sich zu stark an den Trainingsdaten orientieren und somit nicht für eine Verallgemeinerung geeignet sind. Die Verallgemeinerbarkeit eines Lernalgorithmus setzt auch voraus, dass die gewählten Trainingsdaten repräsentativ sind. [20]. Wenn ein Algorithmus Hunde auf Bildern erkennen soll, so ist es nicht ausreichend, dem Algorithmus nur Fotos einer Hunderasse vorzuführen.

Beispiele für überwachte Lernalgorithmen können Tab. 2.1 entnommen werden, wobei die Algorithmen nach Aufgabengebiet sortiert sind. Diese Sortierung ist nicht exklusiv. Vielmehr gibt es Regressionsalgorithmen, die auch für Klassifikationsprobleme geeignet sind und andersherum [20]. Literaturquellen zur Vertiefung aller in dieser Arbeit genannten Lernverfahren sind in Anhang B aufgelistet. Im Folgenden wird der Entscheidungsbaum genauer betrachtet.

Aufgabe	Lernalgorithmus
Klassifikation	Entscheidungsbäume k-Nächste Nachbarn Logistische Regression Naive Bayes Support Vector Machines
Regression	Lineare Regression

Tabelle 2.1: Wichtige überwachte Lernalgorithmen im Überblick. In Anlehnung an [20] und [23].

Der Entscheidungsbau ist ein beliebtes überwachtes Lernverfahren, das sich durch seine schrittweise Nachvollziehbarkeit auszeichnet [27]. Entscheidungsbäume werden grafisch durch eine Baumstruktur aus Knoten und Verzweigungen (Kanten) dargestellt. Der oberste Knoten in einem Baum wird als Wurzel bezeichnet. Ein Knoten führt über eine Kante zu einem weiteren Knoten. Man bezeichnet Knoten, von denen keine Kanten mehr ausgehen, als Blätter. [52]. Die Baumstruktur ist in Abb. 2.5a veranschaulicht. Im Allgemeinen erlauben es Bäume nicht, dass mehrere Kanten zum selben Knoten führen (vgl. Abb. 2.5b) [43].



(a) Komponenten einer Baumstruktur. Eigene Darstellung.

(b) Ein Knoten besitzt nur eine hinführende Kante. Eigene Darstellung.

Abbildung 2.5: Grafische Darstellung von Entscheidungsbäumen.

Bei Klassifikationsaufgaben repräsentieren die Blätter eines Baumes jeweils eine Klasse. Die übrigen Knoten stellen Attribute dar. Die von einem Knoten ausgehenden Kanten geben Werte (oder Wertintervalle) an, die das zugehörige Attribut annehmen kann. Ein Beispiel ist in Abb. 2.6 dargestellt. Mithilfe des Baumes soll entschieden werden, ob das Wetter zum Tennis spielen geeignet ist. Attribute sind „Himmel“, „Luftfeuchtigkeit“ und „Wind“. Die Kanten sind mit den zugehörigen Werten beschriftet. Die Blätter geben die Klassen „Tennis“ und „kein Tennis“ an. Will man einen Baum nutzen, um Objekte zu klassifizieren, so beginnt man an der Wurzel und läuft den Baum Knoten für Knoten ab, bis man bei einem Blatt landet. An jedem Knoten, der kein Blatt ist, wird das zu klassifizierende Objekt auf das zugehörige Attribut geprüft. Über die vom Knoten ausgehenden Kanten wählt man den zutreffenden Wert aus. So gelangt man über die ausgewählte Kante zum nächsten Knoten. [38].

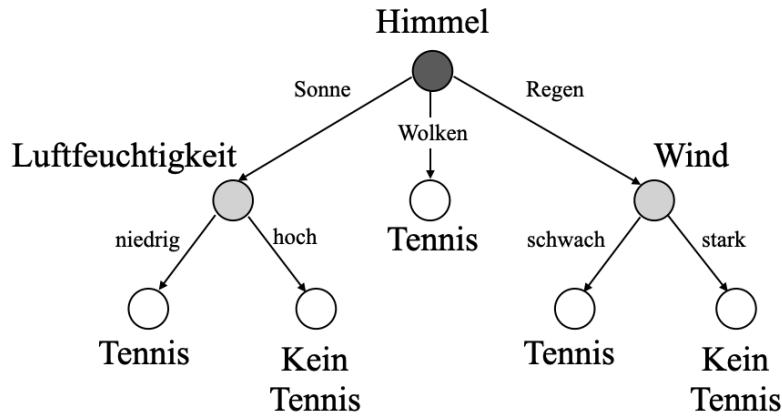


Abbildung 2.6: Entscheidungsbaum zur Vorhersage, ob sich das Wetter zum Tennisspielen eignet. In Anlehnung an [38], S. 53.

Um eine Regression in einem Entscheidungsbaum abbilden zu können, muss eine Diskretisierung der Zielgröße stattfinden, da Regressionen der Vorhersage kontinuierlicher Werte dienen. Dies kann etwa durch die Festlegung von Intervallen erfolgen, wobei jedes Blatt einem Intervall entspricht. Die Herausforderung ist dabei, die Anzahl der Intervalle festzulegen. [23].

Entscheidungsbäume werden mithilfe von Trainingsdaten erstellt. Das ist der Lernprozess im Entscheidungsbaum. [23]. Dabei ist die Reihenfolge, in der die Merkmale der Trainingsdaten im Baum angeordnet werden, entscheidend, da sie den Erfolg des Baumes bezüglich seiner Aufgabe beeinflusst [27]. Mithilfe des ID3-Algorithmus können Entscheidungsbäume ausgehend von der Wurzel konstruiert werden. Dabei wird an jedem Knoten das Merkmal ausgewählt, das an der jeweiligen Stelle im Baum die beste Klassifikation hervorbringt. Das Maß für die beste Klassifikation wird über einen statistischen Test definiert. [38]. Mehr Informationen zum ID3-Algorithmus sind in Anhang A.2 zu finden.

2.2.3 Unüberwachtes Lernen

Auch beim unüberwachten Lernen erlernen Algorithmen Verhaltensweisen aus Daten. Im Gegensatz zum überwachten Lernen sind die Daten hier aber nicht gelabelt [5, 20]. Ohne Labels kann ein Lernalgorithmus jedoch weder für Klassifikations- noch für Regressionsaufgaben eingesetzt werden. Stattdessen wird für unüberwachte Lernalgorithmen in der Literatur häufig die Clusteranalyse (Clustering) als Aufgabe angeführt (vgl. [5, 20, 23, 52]). In Def. 2.6 wird die Clusteranalyse genauer beschrieben. Weitere Aufgaben unüberwachter Lernalgorithmen sind die Dimensionsreduktion, Assoziationsanalysen und Anomaliedetektion [20]. Unüberwachte Lernalgorithmen zu den jeweiligen Aufgabengebieten sind in Tab. 2.2 aufgelistet.

Definition 2.6 Unter einer Clusteranalyse versteht man das Einteilen einer Gruppe von Objekten in Untergruppen [23]. Diese Untergruppen bezeichnet man als Cluster. Objekte mit ähnlichen Eigenschaften werden demselben Cluster zugeordnet. [1]. Darüber hinaus sollen sich die Objekte eines Clusters möglichst stark von den Objekten anderer Cluster unterscheiden [14].

Aufgabe	Lernalgorithmen
Anomaliedetektion	Isolation Forest
Assoziationsanalyse	Apriori-Algorithmus
Clustering	Density-based Clustering of Applications with Noise (DBSCAN) Erwartungsmaximierung (EM-Algorithmus) k-Means-Algorithmus
Dimensionsreduktion	Hauptkomponentenanalyse

Tabelle 2.2: Wichtige unüberwachte Lernalgorithmen im Überblick. In Anlehnung an [14] und [20].

Anders als bei der Klassifikation, bei der die Klassen vorgegeben sind, sind die Cluster bei der Clusteranalyse vorab nicht bekannt [5, 28]. Der Algorithmus muss die Einteilung selbstständig erlernen [5, 9]. Die Idee hinter der Clusteranalyse ist in Abb. 2.7 veranschaulicht. Dabei wird eine Gruppe Murmeln in zwei Cluster aufgeteilt. Jedes Cluster wird durch eine zufällig gewählte Murmel repräsentiert. Die übrigen Murmeln werden jeweils mit den Repräsentanten der Cluster, die als Cluster-Prototypen bezeichnet werden, verglichen. So kann die Zuordnung zu dem Cluster mit der größten Ähnlichkeit erfolgen. [23].

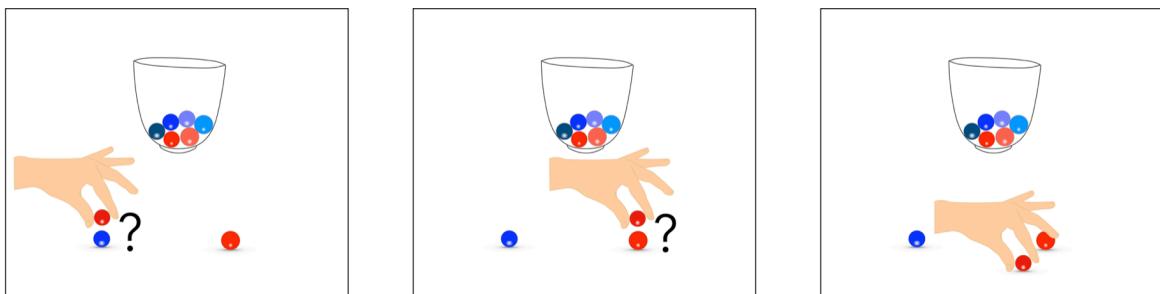


Abbildung 2.7: Murmel mit Cluster-Prototypen vergleichen. Eigene Darstellung.

Bei unüberwachten Lernalgorithmen erfolgt das Training direkt auf den Eingabedaten [51]. Lernalgorithmen untersuchen bei der Clusteranalyse die Eingabedaten auf Muster und nehmen darauf basierend eine Gruppierung der Daten vor [28, 47]. Dies birgt Chancen, aber auch Probleme: Es besteht die Gefahr, dass die Gruppierung der Daten keinen Sinn ergibt. So kann es beispielsweise passieren, dass ein Algorithmus Tierbilder nicht nach Tierart sondern nach Fellfarbe sortiert. [9].

Der k-Means-Algorithmus ist ein beliebter und einfacher Algorithmus für die Clusteranalyse [5], der in der Praxis häufig angewandt wird [52]. Er gehört zu den sogenannten partitio-

nierenden Verfahren. Dabei wird eine Gruppe von Objekten in k Cluster zerlegt, wobei kein Cluster leer sein darf. Und jedes Objekt wird genau einem Cluster zugeordnet. [14]. Für jedes Cluster wird ein Prototyp bestimmt, der die Objekte eines Clusters am besten repräsentiert. Anschaulich gesehen handelt es sich dabei um den Mittelpunkt eines Clusters. Ein Objekt wird dem Cluster zugeordnet, zu dessen Mittelpunkt es den kleinsten Abstand hat. [40]. Der Mittelpunkt eines Clusters wird als Centroid bezeichnet [14]. In Abb. 2.8 wird der Algorithmus grafisch veranschaulicht. Er benötigt als Eingabe einen Wert k , der die Anzahl der zu bildenden Cluster angibt [40, 52]. Man kann den Algorithmus in vier Schritten beschreiben:

1. Lege k Objekte der Gruppe zufällig als potenzielle Centroide fest.
2. Ordne alle Objekte jeweils dem nächsten Centroid zu, sodass Cluster entstehen.
3. Berechne für jedes Cluster den Centroid neu.
4. Wiederhole die Schritte 2 und 3 bis sich die Centroide unter Schritt 3 nicht mehr verändern. [40].

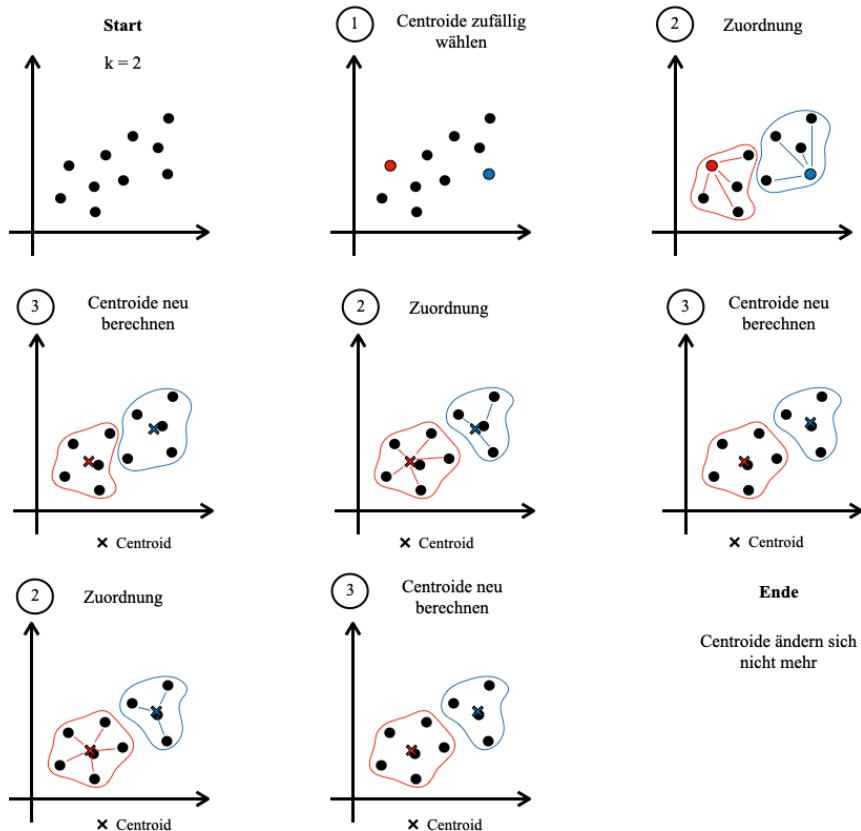


Abbildung 2.8: Grafische Darstellung des k-Means-Algorithmus. In Anlehnung an [14], S. 53.

2.2.4 Verstärkendes Lernen

Beim verstärkenden Lernen erlernt ein Computerprogramm, das als autonomer Agent bezeichnet wird, durch Rückmeldungen in Form von Belohnung oder Bestrafung eine bestimmte Verhaltensstrategie [38]. Die Idee zu dieser Art des Lernens stammt aus der Psychologie [47] und ist auch aus dem Alltag bekannt: Will man beispielsweise einen Hund erziehen oder ihm Tricks beibringen, so werden bestimmte Verhaltensweisen des Hundes etwa durch Leckerlis belohnt oder durch Schimpfen bestraft [21].

Dem Agenten steht eine Menge bestimmter Handlungen zur Verfügung, die er in seiner Umgebung ausführen kann. Dabei wird die Umgebung, in der sich der Agent befindet, über eine Menge von Zuständen definiert. Durch die Handlung des Agenten findet eine Zustandsänderung der Umgebung statt, die vom Agenten wahrgenommen wird. Abhängig von der Zustandsänderung erhält der Agent eine Belohnung in Form eines Zahlenwertes. Die Belohnung wird über eine Funktion festgelegt, die entweder im Computerprogramm implementiert oder einer außenstehenden Person bekannt ist. Mithilfe der Rückmeldungen soll der Agent eine Strategie (Policy) für die Auswahl einer Handlung entwickeln, die in einem maximalen Belohnungserhalt resultiert. [38]. Anstelle einer Belohnungsfunktion ist auch der Einsatz einer Kostenfunktion möglich, die durch die Policy des Agenten minimiert werden soll [5].

Als Beispiel für einen Agenten in einer Umgebung kann man sich einen Roboter auf einem Gitter vorstellen, dessen Aufgabe darin besteht, ein bestimmtes Zielfeld zu erreichen. Die ihm hierfür zur Verfügung stehenden Handlungen sind Bewegungen nach vorne, hinten, rechts und links. Über eine Handlung wechselt der Roboter das Gitterfeld, wobei jedes Gitterfeld einen Zustand der Umgebung darstellt. Das Erreichen des Zielfelds wird mit einem Wert von 100 belohnt. Für das Erreichen jedes anderen Zustands erhält der Roboter den Wert 0. [38].

In Abb. 2.9 ist die Interaktion des Roboters mit seiner Umgebung illustriert.

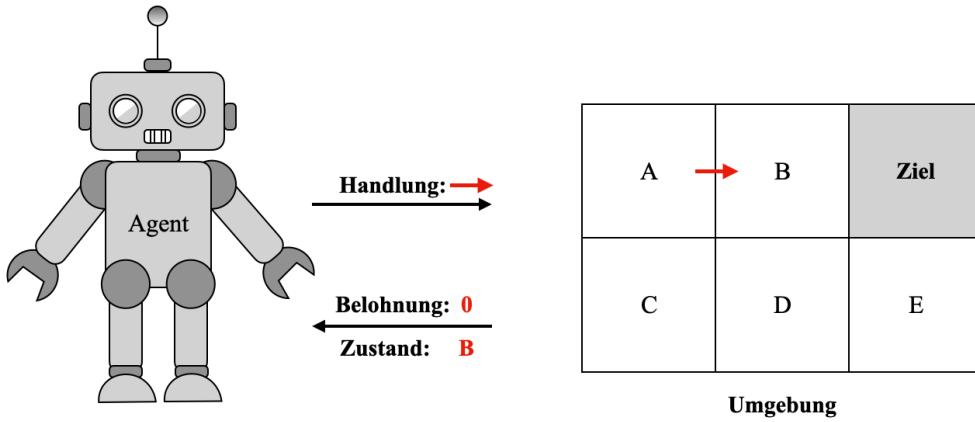


Abbildung 2.9: Ein autonomer Agent interagiert mit seiner Umgebung. In Anlehnung an [38], S. 368 und 372.

Das Q-Learning ist der beliebteste Algorithmus des verstärkenden Lernens [23]. Er basiert auf einer Tabelle, in der für jedes Paar aus Zustand und Handlung ein Wert (sogenannter \hat{Q} -

Wert) hinterlegt ist. Zunächst wird jedem Paar der Wert 0 zugeordnet, da der Agent noch keine Erfahrungen mit der Umgebung gemacht hat (vgl. Abb. 2.10). Der Agent führt wiederholt Handlungen in seiner Umgebung aus, um so die Tabelle zu aktualisieren. Für jede ausgeführte Handlung erhält der Agent eine Belohnung. Mithilfe der Gleichung in Abb. 2.11 wird der neue \hat{Q} -Wert berechnet und anschließend in der Tabelle an entsprechender Stelle abgespeichert. [38]. Die Herleitung und Erklärung der Gleichung kann in Anhang A.3 nachvollzogen werden. Nach [38] kann der Q-Learning Algorithmus wie folgt beschrieben werden:

1. Initialisiere Tabelle für alle Zustände und Handlungen. Setze alle \hat{Q} -Werte auf Null.
2. Beobachte den aktuellen Zustand.
3. Wiederhole unendlich oft:
 - Wähle eine Handlung und führe diese aus.
 - Erhalte eine Belohnung.
 - Beobachte den neuen Zustand.
 - Aktualisiere die Tabelle.

Durch iteratives Aktualisieren der Tabelle kommt der Agent einer optimalen Policy näher. Die Policy wird über die Tabelle bestimmt. Dabei wird für jeden Zustand die Handlung mit dem größten \hat{Q} -Wert ausgewählt. [20].

		Handlungen			
		↑	↓	→	←
		A	B	C	D
		0	0	0	0
		0	0	0	0
		0	0	0	0
		0	0	0	0
		0	0	0	0
		Ziel	0	0	0

Abbildung 2.10: Initialisierte Tabelle. In Anlehnung an [23], S. 369.

$$\hat{Q}(s, a) = r(s, a) + \gamma \cdot \max \hat{Q}(s', a')$$

Abbildung 2.11: Funktion für den Q-Learning Algorithmus. Nach [38], S. 375.

2.2.5 Künstliche neuronale Netze

Die Grundideen künstlicher neuronaler Netze (Abk. KNNs) sind vom menschlichen Gehirn inspiriert [5, 20]. In Anlehnung an das biologische Vorbild, bei dem viele Nervenzellen (Neuronen) miteinander verbunden sind, werden bei KNNs mehrere künstliche Neuronen verknüpft. Mithilfe von KNNs lassen sich Klassifikations-, Regressions- und Clusteringaufgaben lösen. [22]. Damit sind sowohl Problemstellungen des überwachten als auch des unüberwachten Lernens abgedeckt. Aber auch bestärkendes Lernen ist mit KNNs möglich. Ein Beispiel hierfür ist das bekannte Computerprogramm AlphaGo, das mithilfe eines künstlichen neuronalen Netzes eine Policy für das Spielen von Go erlernt hat [5].

Das menschliche Gehirn besteht aus schätzungsweise 10^{11} Neuronen, die in einem Netzwerk untereinander verbunden sind. Im Durchschnitt steht ein Neuron mit 10^4 weiteren Neuronen in Verbindung. [38]. Ein biologisches Neuron ist in Abb. 2.12 schematisch dargestellt. Dabei ist ein Teil des Neurons für die Aufnahme von Signalen vorhergehender Neuronen verantwortlich. Diese ankommenden Signale werden im Inneren des Neurons aufsummiert. Überschreitet die Summe eine bestimmte Schwelle, so wird ein Signal an anderer Stelle an anschließende Neuronen weitergeleitet. [22].

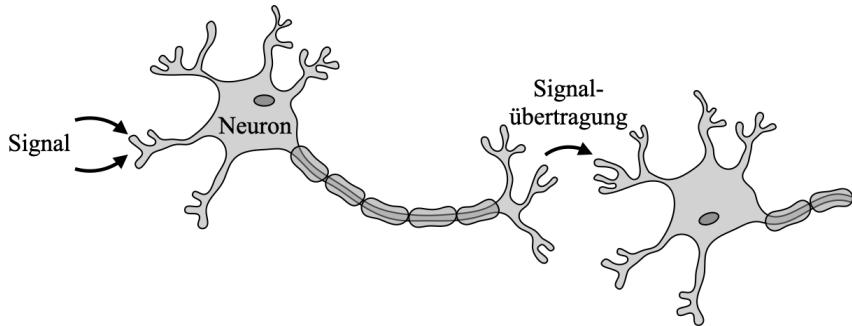


Abbildung 2.12: Ein Neuron empfängt Signale und leitet diese weiter. In Anlehnung an [22], S. 153.

Dieses Prinzip der Signalverarbeitung findet man bei künstlichen Neuronen wieder. Ein einzelnes künstliches Neuron wird als Perzeptron bezeichnet. [22]. Ein solches Perzeptron ist in Abb. 2.13 illustriert. Das Perzeptron nimmt reelle Werte x_i als Eingaben entgegen. Jeder Eingabewert wird mit einem ihm zugeordneten reellen Gewicht w_i multipliziert. Die Paare $x_i \cdot w_i$ werden anschließend im Perzeptron aufsummiert. [38]. Über die Gewichte wird somit der Einfluss bestimmt, den die Eingaben auf die Summe haben [22]. Wenn die gewichtete Summe größer als ein bestimmter Schwellenwert ist, so gibt das Perzeptron den Wert 1 aus, andernfalls -1 . Der Schwellenwert wird dabei durch den Term $-(x_0 \cdot w_0)$ ausgedrückt, sodass $x_1 w_1 + x_2 w_2 + \dots + x_n w_n > -x_0 w_0$ die Bedingung für das Überschreiten des Schwellenwerts ist. Durch Äquivalenzumformung ergibt sich $x_0 w_0 + x_1 w_1 + x_2 w_2 + \dots + x_n w_n > 0$. Im Perzeptron wird dies durch einen konstanten Eingabewert x_0 mit Gewicht w_0 realisiert. [38]. Die Übersetzung der gewichteten Summe in eine Ausgabe wird als Aktivierungsfunk-

tion bezeichnet [22]. Die hier vorgestellte Aktivierungsfunktion ist nur eine Möglichkeit, einen Übersetzungsprozess zu beschreiben. Sie wird Schwellenwertfunktion genannt. Daneben werden beispielsweise auch Sigmoid-Funktionen eingesetzt, die die Form $\frac{1}{e^{-\alpha \cdot y}}$ mit $y = \sum_{i=0}^n x_i w_i$ besitzen. [5].

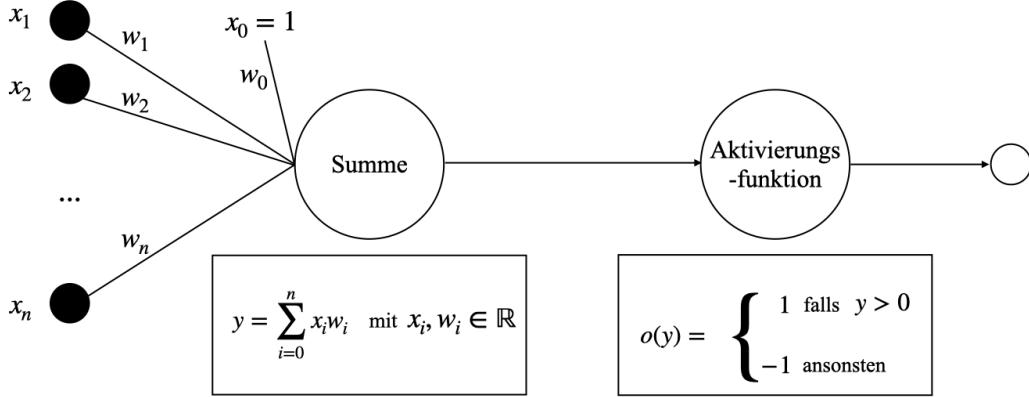


Abbildung 2.13: Ein Perzeptron mit gewichteter Summe und Schwellenwertfunktion. In Anlehnung an [5], S. 174 und [38], S. 87.

Ein einfaches Beispiel für die Anwendung eines Perzeptrons ist das logische UND. Dies kann über zwei Eingaben mit den Gewichten $w_1 = w_2 = 0,5$ realisiert werden. Der Schwellenwert kann etwa auf 0,8 gesetzt werden. Damit ist $-(x_0 \cdot w_0) = 0,8$, sodass daraus mit $x_0 = 1$ als konstanter Eingabewert $w_0 = -0,8$ folgt. [38]. In Abb. 2.14 ist das beschriebene Perzeptron dargestellt. Nur wenn es für beide Eingaben den Wert 1 erhält, wird der Schwellenwert überschritten und 1 ausgegeben.

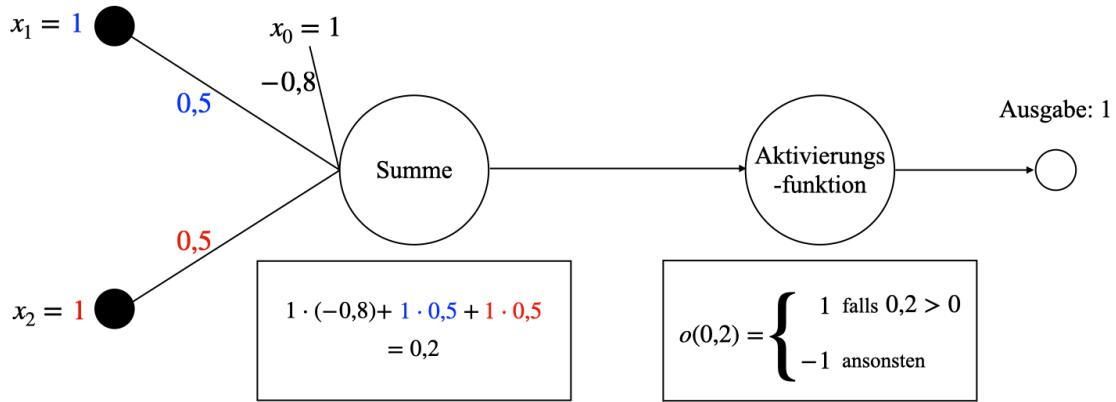


Abbildung 2.14: Ein Perzeptron, das das logische UND repräsentiert. In Anlehnung an [5], S. 174 und [38], S. 87.

Ein Perzeptron lernt über die Anpassung seiner Gewichte. Ausgangspunkt sind dabei zufällig festgelegte Gewichte. Mithilfe von Trainingsdaten werden beim überwachten Lernen die Gewichte so geändert, dass man für jedes Trainingsbeispiel die korrekte Ausgabe erhält. [22].

Zwei wichtige Algorithmen für die Bestimmung der Gewichte sind die Perzeptron-Regel und die Delta-Regel, die auf dem Gradientenabstiegsverfahren basiert. Beide Algorithmen sind fundamental für das Lernen von Netzen. [38]. Die Perzeptron-Regel wird in Anhang A.4 genauer betrachtet.

Der Zusammenschluss mehrerer künstlicher Neuronen zu Netzen kann durch unterschiedliche Architekturen erfolgen. Die aktuell häufigste Form sind sogenannte Feed Forward Neural Networks (Abk. FFNNs). Solche Netze bestehen aus Neuronen, die in Schichten vorliegen. Die Bezeichnung „Feed Forward“ kommt daher, dass Informationen von einer Schicht nach vorne in die nächste Schicht übermittelt werden. [22]. Die FFNN-Architektur ist in Abb. 2.15 dargestellt. Dabei werden einzelne Neuronen (repräsentiert durch Knoten) mit Neuronen anderer Schichten, aber nicht mit Neuronen derselben Schicht, verbunden. Häufig werden Verbindungen zwischen allen Knoten zweier benachbarter Schichten gezogen. Die erste Schicht wird als Input-Schicht oder Eingangsschicht bezeichnet. Über sie werden Informationen in das neuronale Netz eingegeben. Die letzte Schicht ist die Output- oder Ausgangsschicht. Sie dient der Ausgabe des Ergebnisses. Dazwischen können eine bis mehrere tausend Zwischenschichten (engl. *hidden layers*) liegen. Die Auswahl der Anzahl der Zwischenschichten und Neuronen pro Schicht ist dabei für den Erfolg eines Netzes entscheidend. Diese müssen vor dem Lernprozess des Netzes festgelegt werden. Allerdings gibt es noch keine eindeutigen Regeln, um die Anzahl der Zwischenschichten zu bestimmen. Vielmehr beruht die Auswahl auf Erfahrung. Bei neuronalen Netzen mit vielen Zwischenschichten spricht man von Deep Learning. [5].

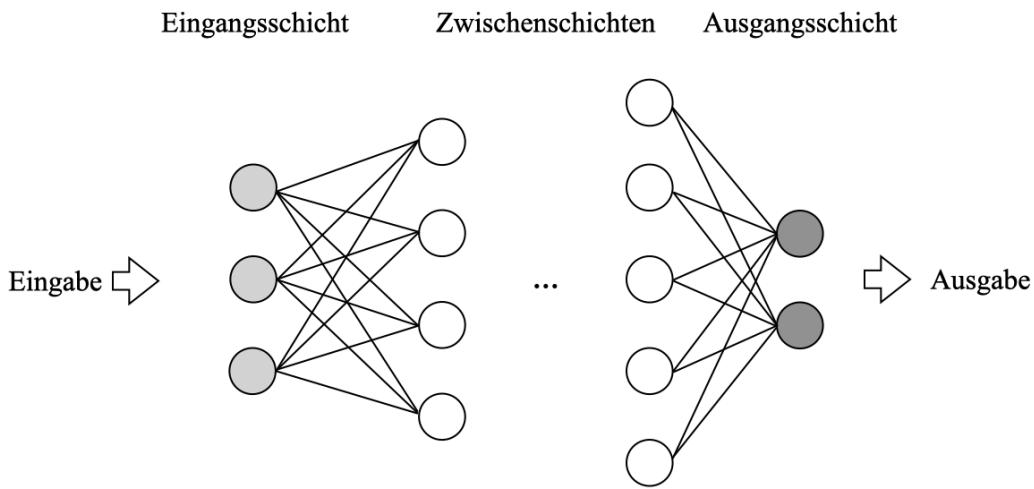


Abbildung 2.15: Die Schichten einer FFNN Architektur. In Anlehnung an [5], S. 164.

3. Workshop-Konzept

In diesem Kapitel wird das Workshop-Konzept betrachtet, das Gegenstand der Evaluation ist. Es werden Einblicke in die Themenwahl, die Lernzielentwicklung und die Gestaltung von Materialien und Aktivitäten gegeben sowie Bezüge zum aktuellen Bildungsplan von Baden-Württemberg hergestellt. Der Verlaufsplan und die Unterrichtsmaterialien sind in Anhang C angehängt.

3.1 Themenauswahl

Das verstärkende Lernen wird von [23] als eine fortgeschrittene Thematik des maschinellen Lernens eingestuft. Dennoch kann die Grundidee, Lernen auf Basis von Rückmeldungen, gut vermittelt werden, was ein Grund dafür sein könnte, warum es im deutschsprachigen Raum bereits mehrere Aktivitäten für den Unterricht zum verstärkenden Lernen gibt (vgl. [35, 46, 49] sowie das „Mensch, Maschine!“-Spiel¹). Das überwachte Lernen ist diejenige Form des maschinellen Lernens, die in der Fachliteratur häufig zuerst eingeführt wird (vgl. [3, 5, 9, 28, 47, 52]). Sie ist beliebter als das unüberwachte Lernen, was womöglich darauf zurückzuführen ist, dass unüberwachtes Lernen mit weniger Vorgaben arbeitet [5]. Aus der Studie von [33] geht hervor, dass englischsprachige Lehreinheiten häufig das überwachte Lernen thematisieren. Im deutschsprachigen Raum gibt es beispielsweise eine Aktivität von [46] und den „Workshop Maschinelles Lernen Unplugged“².

Unter Berücksichtigung bereits bestehender Aktivitäten und Kurse fiel die Wahl für den hier vorgestellten Workshop auf das überwachte Lernen. Entlang von Entscheidungsbäumen werden Grundideen zu Labels, Klassifikation sowie Trainings- und Testdaten eingeführt. Entscheidungsbäume sind das Lernverfahren der Wahl, weil sie für Menschen intuitiv [5] und schrittweise [27] nachvollziehbar und damit gut für eine erste Begegnung mit dem überwachten Lernen geeignet sind. Es wurde für den Einführungsworkshop eine Entscheidung gegen künstliche neuronale Netze getroffen, da diese in der Fachliteratur als Blackbox bezeichnet werden (vgl. [5, 52]). Damit ist gemeint, dass Menschen nicht mehr einfach nachvollziehen können, wie ein Netz zu einer Ausgabe gelangt [5, 52].

¹ProDaBi-Team der Universität Paderborn: <https://www.prodabi.de/mensch-maschine-spiel>

²Universität Osnabrück: <https://inf-didaktik.rz.uos.de/phycos/ki.php>

3.2 Voraussetzungen und Ziele

Zielgruppe des Workshops sind Lernende der Sekundarstufe I ab Klasse 7. Wie bereits in Kap. 2.2 beschrieben, bilden Algorithmen eine wichtige Grundlage für das maschinelle Lernen. Somit wird ein bereits vorhandenes Begriffsverständnis von Algorithmen für den Workshop vorausgesetzt.

Die Lernziele für den Workshop sind von Richtlinien abgeleitet, die von der AI for K-12 entwickelt wurden. Bei AI for K-12 handelt es sich um eine von der Association for the Advancement of Artificial Intelligence und Computer Science Teachers Association ins Leben gerufene Arbeitsgruppe, deren Aufgabe in der Definition von Richtlinien für die Vermittlung von künstlicher Intelligenz vom Kindergarten bis zur Klassenstufe 12 besteht. Diese Richtlinien werden entlang fünf großer Ideen organisiert, die der Arbeitsgruppe zufolge jedem und jeder Lernenden bekannt sein sollten. Dabei umfasst Idee Nummer drei („Computers can learn from data“) die Konzepte des maschinellen Lernens. [50]. Auf der Webseite der Arbeitsgruppe kann die aktuelle Version der Richtlinien für Idee Nummer drei heruntergeladen werden (vgl. [2]). Diese sind nach Klassenstufen und Konzepten strukturiert. Unter dem Konzept „Nature of Learning“ findet man für die Klassenstufen 6 bis 8 das Lernziel „Train and evaluate a classification or prediction model using machine learning on a tabular dataset.“ ([2], S. 3), welches den Ausgangspunkt für die Entwicklung der Lernziele für den hier präsentierten Workshop darstellte. Diese lauten wie folgt:

Übergeordnetes Lernziel (Kernkompetenz)

Im Workshop lernen die Schülerinnen und Schüler das überwachte Lernen anhand von Entscheidungsbäumen kennen. Nach der Teilnahme können die Schülerinnen und Schüler Entscheidungsbäume erstellen, testen und anwenden.

Feinlernziele (Teilkompetenzen)

Die Schülerinnen und Schüler können ...

... aus tabellarischen Datensätzen einen Entscheidungsbaum erstellen, indem sie Attribute aus der Tabelle als Knoten in der richtigen Reihenfolge anordnen und über Kanten, beschriftet mit Werten, in geeigneter Weise verbinden (TK1).

... einen Entscheidungsbaum auf tabellarischen Datensätzen korrekt anwenden (TK2).

... einen Entscheidungsbaum mithilfe tabellarischer Testdatensätze bewerten, indem sie den Baum korrekt auf den Testdaten anwenden und die Ergebnisse des Baumes mit der Lösung der Testdaten vergleichen (TK3).

Unter Berücksichtigung von [32] wurden jeder Teilkompetenz Kriterien hinzugefügt, anhand welcher bestimmt werden kann, ob eine Schülerin oder ein Schüler die jeweilige Teilkompetenz erworben hat. Mithilfe dieser soll das Erreichen der gesetzten Kompetenzziele evaluiert

werden. Inspiriert durch [48] wird im Rahmen dieses Workshops eine Arbeitsblatt-Analyse als Evaluationsinstrument eingesetzt. Auf dem Arbeitsblatt 2 (vgl. Anhang C.10) werden alle drei oben genannten Teilkompetenzen adressiert. Damit entstehen durch das Bearbeiten des Blattes durch die Lernenden Produkte, die hinsichtlich der Kompetenzzielerreichung entlang der beschriebenen Kriterien unter Zuhilfenahme einer Musterlösung (vgl. Anhang C.11) analysiert werden können.

3.3 Bezug zum Bildungsplan Baden-Württemberg 2016

Das maschinelle Lernen ist ein extracurriculares Themengebiet aus der Informatik. Gleichwohl lassen sich Bezüge zum aktuellen Bildungsplan der Informatik in Baden-Württemberg herstellen (vgl. [36] und [37]). Hinsichtlich der inhaltsbezogenen Kompetenzen werden die Konzepte „Algorithmen“ und „Daten und Codierung“ tangiert. Darüber hinaus werden die prozessbezogenen Kompetenzbereiche „Modellieren und Implementieren“ sowie „Kommunizieren und Kooperieren“ angesprochen:

Algorithmen

In der Erarbeitungsphase entwickeln Lernende einen Algorithmus, um die vorgegebene Problemstellung zu lösen.

Daten und Codierung

Im Rahmen des Workshops setzen sich die Schülerinnen und Schüler mit der Baumstruktur am Beispiel von Entscheidungsbäumen auseinander.

Modellieren und Implementieren

Die Lernenden üben im Workshop aus einem gegebenen tabellarischen Datensatz einen Entscheidungsbaum zu modellieren.

Kommunizieren und Kooperieren

Im Workshop gibt es Aktivitäten, die in Zweiergruppen durchgeführt werden. Dabei lösen die Lernenden die Aufgabenstellungen gemeinsam, indem sie Lösungsideen sowie Ergebnisse diskutieren.

3.4 Aktivitäten und Materialien

Unter Unplugged-Aktivitäten versteht man im Kontext der Informatik-Didaktik Aktivitäten zur Vermittlung informatischer Konzepte ohne den Einsatz von Computern. Ausgangspunkt ist dabei eine Problemstellung oder Herausforderung, die von den Lernenden gelöst werden soll. [6]. Solche Unplugged-Aktivitäten werden auch bei der Vermittlung von maschinellem Lernen eingesetzt (vgl. [33, 35, 46] sowie das „Mensch, Maschine!“-Spiel und der „Workshop

Maschinelles Lernen Unplugged“). In [46] wird argumentiert, dass mithilfe von Unplugged-Aktivitäten das Themenfeld künstliche Intelligenz in seiner Komplexität heruntergebrochen werden kann ohne dabei die Anschlussfähigkeit für Vertiefungen (auch Softwareeinsatz) zu verlieren. Angesichts der Beliebtheit von Unplugged-Aktivitäten und Argumente dafür wurde für den hier geplanten Workshop ebenfalls mit Unplugged-Aktivitäten gearbeitet. Inspiriert durch die Aktivität zu Entscheidungsbäumen aus [46]³ wurde für die Erarbeitungsphase eine Aktivität geplant, bei der Lernende in die Rolle eines überwachten Lernalgorithmus schlüpfen. Hierbei wurde die Idee übernommen, aus Bildern Merkmale zu extrahieren, um daraus einen Klassifikationsalgorithmus zu entwickeln. Bei [46] geht es um die Klassifikation von Bildern von Affen danach, ob die Tiere beißen oder nicht. Die im Rahmen des hier vorgestellten Workshops entwickelte Aktivität versucht einen Bezug zur Lebenswelt der Lernenden herzustellen und stellt sie vor das Problem, einen Lügendetektor zu entwickeln, um damit das Rätsel um ein verschwundenes Handy zu lösen (vgl. Anhang C.4). Die Lernenden untersuchen dabei Bilder von Personen auf Merkmale, die lügende von nicht lügenden Personen unterscheiden und konstruieren davon ausgehend einen eigenen Algorithmus zur Klassifikation.

Für die Vermittlung von überwachtem Lernen gibt es Online-Tools wie etwa die Google Teachable Machine⁴ oder Machine Learning For Kids⁵. Die Teachable Machine von Google ist ein einfach zu bedienendes Tool, das Lernenden eine praktische Auseinandersetzung mit maschinellem Lernen ermöglicht. Die Grundlage bildet allerdings ein künstliches neuronales Netz (Deep Learning). [39]. Deshalb ist das Tool hier nicht passend. Mit Machine Learning For Kids können Entscheidungsbäume trainiert, getestet und visualisiert werden. Allerdings wird das Lernziel TK1 mit dem Tool nicht adressiert, da der Computer das Erstellen des Entscheidungsbaumes übernimmt. Darüber hinaus ist nach Auffassung der Verfasserin dieser Arbeit die Darstellung der Entscheidungsbäume im Tool vergleichsweise abstrakt und deshalb weniger gut für eine erste Auseinandersetzung geeignet.

Der Studie von [33] kann entnommen werden, dass Videos, Arbeitsblätter und Präsentationsfolien beliebte Materialien in englischsprachigen Lehreinheiten zu maschinellem Lernen sind. Auch für den hier präsentierten Workshop sind dies die Unterrichtsmaterialien der Wahl. Bei der Ausgestaltung dieser wurde in Orientierung an [4] auf ein konsistentes Design, die Auswahl und Gestaltung ansprechender und altersgerechter Bilder, eine alters- und gendergerechte Sprache, einen Lebensweltbezug, den Einsatz unterschiedlicher Sozialformen sowie auf eine Adressierung der gesetzten Lernziele geachtet. Darüber hinaus wurde eine Begleitbroschüre entwickelt, die Informationen zur Fachsystematik, zu den Lernzielen und Materialien für Lehrkräfte bereitstellt, (vgl. Anhang C.12).

³Zugehörige Materialien erhältlich unter <https://www.aiunplugged.org>

⁴<https://teachablemachine.withgoogle.com>

⁵<https://machinelearningforkids.co.uk>

4. Methode

Dieses Kapitel beschäftigt sich mit der zur Untersuchung der Forschungsfrage verwendeten Methode. Zu Beginn wird die Herangehensweise betrachtet, die auf dem designbasierten Forschungsansatz aufbaut. Da das Workshop-Konzept einer Einschätzung durch Expertinnen unterzogen wird, erfolgt eine Charakterisierung dieser. Anschließend werden Instrument und Vorgehensweise vorgestellt sowie die Analyse der Untersuchungsergebnisse dargelegt.

4.1 Herangehensweise

Bei der designbasierten Forschung handelt es sich um einen Forschungsansatz, der darauf abzielt, durch das Design von Interventionen zu aktuellen Problemstellungen aus der Bildung einen Beitrag zur Entwicklung innovativer Lernumgebungen zu leisten und theoretische Erkenntnisse abzuleiten. Charakteristisch für diesen Ansatz ist die enge Zusammenarbeit mit der Praxis. In die Entwicklung von Interventionen fließen neben wissenschaftlichen Theorien Erfahrungen aus der Praxis ein. Auch für die Erprobung und Evaluation werden Praktikerinnen und Praktiker herangezogen. Die Ergebnisse der designbasierten Forschung sollen einen Nutzen für die Praxis bringen. [15]. Der Forschungsprozess kann als fortlaufender Zyklus aus Design, Evaluation und Re-Design aufgefasst werden (vgl. Abb. 4.1) [12, 13].

In Kap. 3 wird dargelegt, wie unter Berücksichtigung bereits vorhandener nationaler und internationaler Literatur ein Workshop-Konzept für das überwachte Lernen entwickelt wurde. Dies entspricht gerade der Design-Phase aus Abb. 4.1. Um der Forschungsfrage dieser Arbeit nachzugehen, wurde das entworfene Konzept in Anlehnung an das Prozessmodell aus Abb. 4.1 evaluiert, überarbeitet und erneut bewertet.

Nach dem Vorbild des Software-Engineerings unterscheidet man in der designbasierten Forschung drei Evaluationsformen, abhängig vom Fokus der Evaluation: Alpha-, Beta- und Gamma-Tests. Alpha-Tests untersuchen die interne Konsistenz und praktische Durchführbarkeit einer Intervention. Dabei kommen hauptsächlich Einschätzungen durch Expertinnen und Experten etwa durch Interviews oder Fragebögen zum Einsatz. [34].

Daran angelehnt wurde die interne Konsistenz des im Rahmen dieser Arbeit entworfenen Workshop-Konzepts durch eine quantitative Befragung von Expertinnen aus der Praxis untersucht. Dieser wurde um eine qualitative Rückmeldung ergänzt, mit dem Ziel die Gründe hinter den Antworten im Fragebogen zu beleuchten.

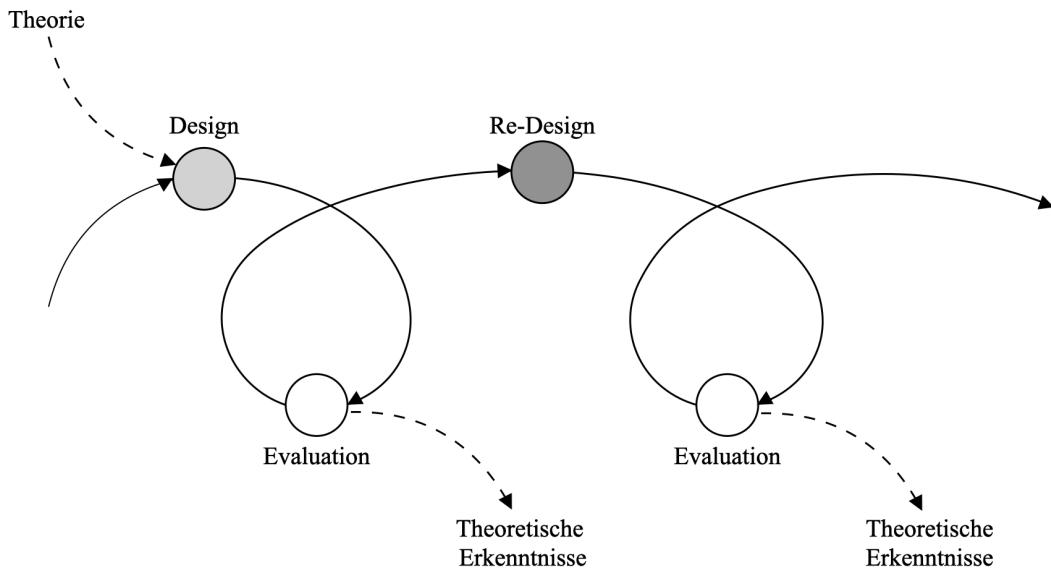


Abbildung 4.1: Prozessmodell zur designbasierten Forschung. In Anlehnung an [17].

4.2 Expertinnen

Bei den für die Befragung ausgewählten Expertinnen handelt es sich um zwei Mitarbeiterinnen aus dem Lehr-Lern-Labor Informatik der Pädagogischen Hochschule Karlsruhe. Beide stammen aus dem Bereich der Informatik und bringen durch ihre Tätigkeit Workshop-Erfahrung mit. Sie zeichnen sich zusätzlich dadurch aus, dass sie im Lehr-Lern-Labor promovieren.

4.3 Instrument

Durch den Fragebogen wurde der Begriff der Eignung aus der Forschungsfrage operationalisiert. Im Zentrum des Fragebogens steht der Verlaufsplan des Workshops. Entlang der Bereiche „Unterrichtsphasen“, „Kompetenzen“, „Methoden“ und „Assessment (Evaluation)“ soll dieser mit Berücksichtigung der entworfenen Materialien auf seine Eignung untersucht werden. Die Fragebogen-Items sind durch die genannten vier Bereiche gegliedert. Eine Grundlage für die Entwicklung der Items war das 5e-Modell nach [10] für die Erfassung einer geeigneten Abfolge der Unterrichtsphasen. Zur Untersuchung der Kommunikation, methodischen Adressierung und Feststellung der Kompetenzziele wurden das Analyse-Tool für kompetenzorientierte Unterrichtspläne von [41] sowie der Leitfaden zur Evaluierung standardbasierter Unterrichtspläne nach [11] und das Constructive Alignment von [7] berücksichtigt. Insgesamt wurden 23 Items formuliert. Diese sind in Tab. 4.1 aufgelistet. Jedes Item wurde über eine Likert-Skala mit fünf Stufen (von „trifft überhaupt nicht zu“ bis „trifft sehr zu“) abgefragt. Der Fragebogen ist in Anhang D beigefügt.

Bereich	Item-Nr.	Beschreibung
Unterrichtsphasen	1 2 3 4 5 6 7 8	Ersichtlichkeit der Phasen klare Beschreibung Aktivitäten Engagement-Phase Exploration-Phase Explanation-Phase Elaboration-Phase Evaluation-Phase Angemessenheit der Zeiträume
Kompetenzen	9 10 11 12 13 14 15 16	Ersichtlichkeit der Kompetenzziele zeitliche Angemessenheit der Kompetenzziele Adressierung der Kompetenzziele Bezug der Kompetenzziele zu Thema und Inhalt Kompetenzziel-Komponente Bedingung Kompetenzziel-Komponente Handlung Kompetenzziel-Komponente Inhalt Kompetenzziel-Komponente Kriterium
Methoden	17 18 19 20 21	Ersichtlichkeit der Methoden zielgerichteter Einsatz individuelle Auseinandersetzungen kooperative Auseinandersetzungen angemessenes Verhältnis frontaler Anteile und Lernenden-Aktivitäten
Assessment (Evaluation)	22 23	Erfassung Kompetenzziele Messung individueller Kompetenzziel-Erreichung

Tabelle 4.1: Fragebogen-Items nach Bereichen organisiert.

4.4 Vorgehensweise

Zunächst erhielten die Expertinnen den Fragebogen für eine erste Befragung gemeinsam mit dem Verlaufsplan des Workshops sowie den Unterrichtsmaterialien. Die Kommunikation fand für beide Expertinnen auf demselben Weg statt (Nachrichten via E-Mail und Microsoft Teams). Die Expertinnen schätzten unabhängig voneinander das Workshop-Konzept entlang des Fragebogens ein und formulierten eine schriftliche Rückmeldung. Danach送eten sie den Fragebogen und die inhaltliche Rückmeldung zurück. Anschließend wurden die Ergebnisse der ersten Befragung analysiert und genutzt, um das Konzept des Workshops zu überarbeiten (Re-Design). Schließlich fand eine zweite Befragung statt. Dabei schätzten die Expertinnen das überarbeitete Konzept nochmals anhand des Fragebogens ein und verfassten ein inhaltliches Feedback.

4.5 Analyse

Um den Grad der Übereinstimmung zwischen den Bewertungen der Expertinnen innerhalb einer Befragung quantitativ zu bestimmen, wurden mithilfe von Microsoft Excel Indizes des

Interrater Agreement (Abk. IRA) berechnet. Unter Berücksichtigung von [30] und [42] wurden zwei IRA-Indizes ($r_{wg(j)}^*$ und $AD_{M(J)}$) aus verschiedenen Indexfamilien ausgewählt. Andere Indizes wurden ausgeschlossen (z.B. $r_{wg(j)}$ angesichts der Nichtlinearität, $a_{wg(j)}$ aufgrund der Anzahl benötigter Bewertenden und CV_{wg} mangels eindeutiger Interpretation).

Excel wurde auch dazu genutzt, die Antworten in Diagrammen zu veranschaulichen. Mithilfe vertikaler Liniendiagramme wurden die Antworten der Expertinnen innerhalb einer Befragung gegenübergestellt. So konnten die individuellen Antworten, aber auch Gemeinsamkeiten und Unterschiede übersichtlich dargestellt werden.

Aus den schriftlichen Rückmeldungen wurden die Kernaussagen notiert, um Rückschlüsse auf die Antworten im Fragebogen ziehen zu können und Anhaltspunkte für eine Überarbeitung des Workshop-Konzepts zu erhalten. Diese wurden tabellarisch festgehalten.

Um zu untersuchen, ob die Überarbeitung des Workshop-Konzepts zu einer Optimierung geführt hat, wurden die Mittelwerte der ersten und zweiten Befragung in einer Tabelle gegenübergestellt. Voraussetzung hierfür war, dass die IRA-Indizes eine hohe Übereinstimmung der Expertinnen anzeigen, sodass die individuellen Ergebnisse durch die Mittelwerte nicht verzerrt würden [42].

5. Ergebnisse und Diskussion

In diesem Kapitel werden die Ergebnisse der Einschätzung durch die Expertinnen berichtet. Zusätzlich zu den Antworten im Fragebogen werden die Ergebnisse der schriftlichen Rückmeldungen vorgestellt. Die ausgefüllten Fragebögen und Rückmeldungen befinden sich in Anhang E. Nach Vorstellung der Ergebnisse der ersten Befragung erfolgt eine Zwischendiskussion, an die die Präsentation der Ergebnisse der zweiten Befragung anschließt. Auf den Vergleich der ersten und zweiten Befragung folgt dann die abschließende Diskussion.

5.1 Erste Befragung

Die Antworten der ersten Befragung sind in Abb. 5.1, Abb. 5.2 und Abb. 5.3 nach Bereichen strukturiert dargestellt. Auf der y-Achse sind die einzelnen Items des jeweiligen Bereichs angeordnet. Von der x-Achse können die Stufen der Likert-Skala (1 = „trifft überhaupt nicht zu“; 5 = „trifft sehr zu“) abgelesen werden. In Tab. 5.1 sind die Kernpunkte der schriftlichen Rückmeldungen zusammengefasst.

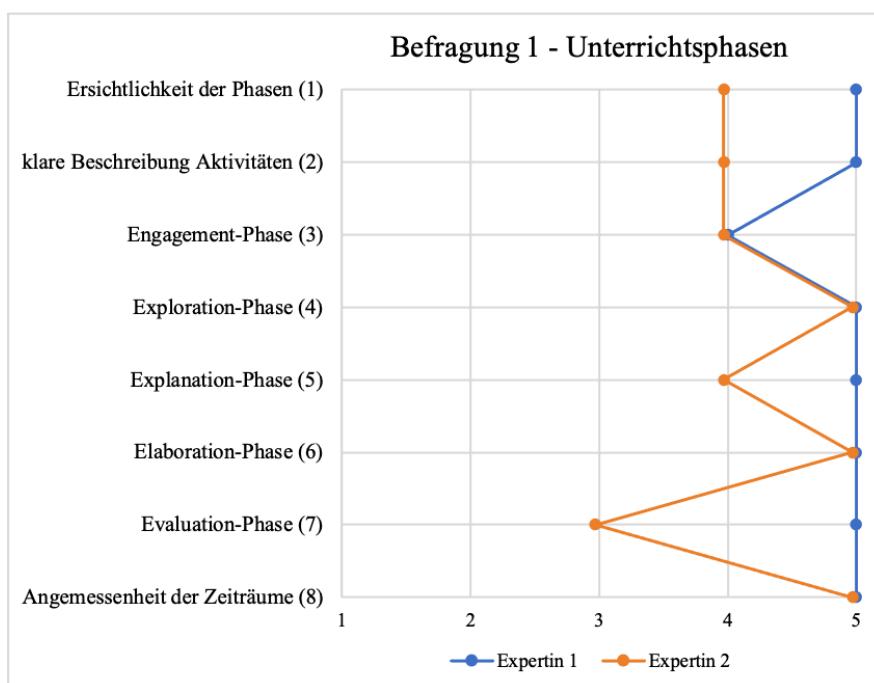


Abbildung 5.1: Befragungsergebnisse von Expertin 1 und 2 zu den Unterrichtsphasen.

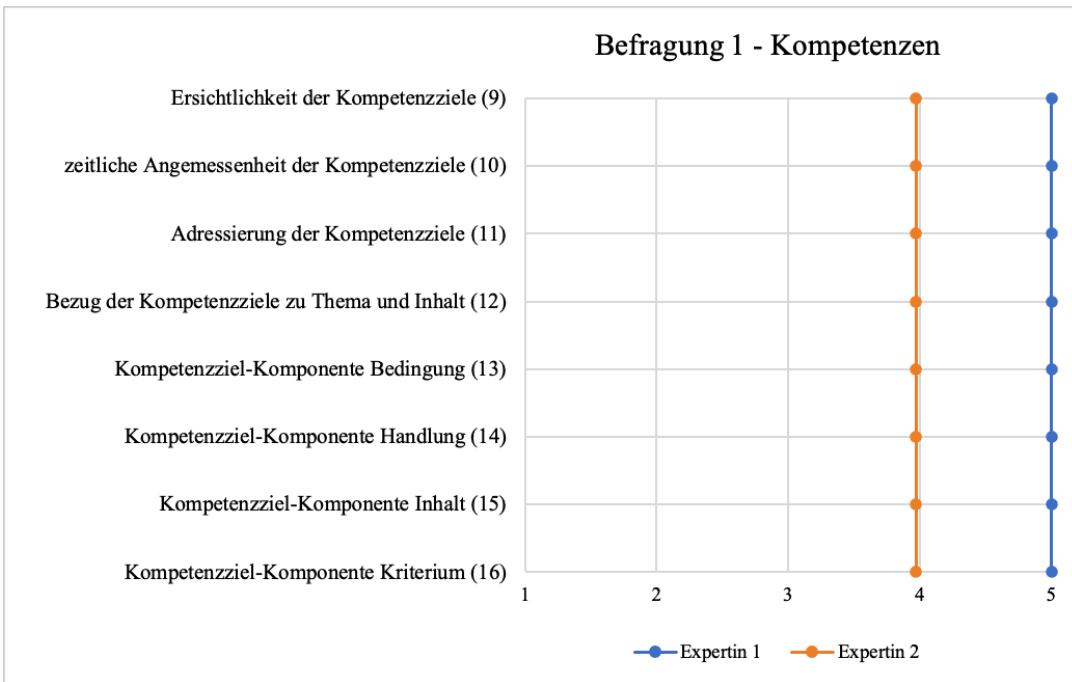


Abbildung 5.2: Befragungsergebnisse von Expertin 1 und 2 zu den Kompetenzen.

Insgesamt stimmte Expertin 1 bei der ersten Befragung 22 von 23 Items vollständig zu. Für Item 3 (Engagement-Phase) vergab sie 4 von 5 Punkten auf der Likert-Skala (vgl. Abb. 5.1). In ihrer schriftlichen Rückmeldung bezog sie sich auf die Ausgestaltung der Einführung (vgl. Tab. 5.1). Dabei wies die Expertin zum einen darauf hin, dass der Begriff KI, der in der Quizfrage vorkommt, die die Lernenden motivieren soll, noch unbekannt sein könnte, zum anderen meldete sie zurück, dass Sprache und Visualisierungen auf den Folien stellenweise zu abstrakt sind und zu viele Fachbegriffe eingeführt werden.

Nr. Rückmeldung		
Expertin 1	1	Einführung: Quizfrage zur Motivation der Lernenden nur bedingt geeignet, da der Begriff KI Lernenden noch unbekannt sein könnte; teilweise zu abstrakte Formulierungen und Grafiken in den Folien; zu viele neue Fachbegriffe
Expertin 2	2	Unterrichtsphasen, die mehrere Sozialformen besitzen, aufteilen
	3	Komponenten der Kompetenzen ausführen
	4	Evaluation der Arbeitsblätter unklar
	5	Mehr Binnendifferenzierung

Tabelle 5.1: Kernaussagen aus den schriftlichen Rückmeldungen der beiden Expertinnen.

In Abb. 5.3 sieht man, dass Expertin 2 die Items 22 und 23 (Assessment) nicht beantwortet hat. Den übrigen Items stimmte sie vollständig oder teilweise zu, ausgenommen Item 7 (Evaluation-Phase), das von ihr mit 3 von 5 Punkten auf der Likert-Skala bewertet wurde (vgl. Abb. 5.1). Bereits im Fragebogen wurden einzelne Items mit Kommentaren versehen (vgl. Anhang E.1). In ihrem schriftlichen Feedback führte sie vier Punkte an. Dazu gehören

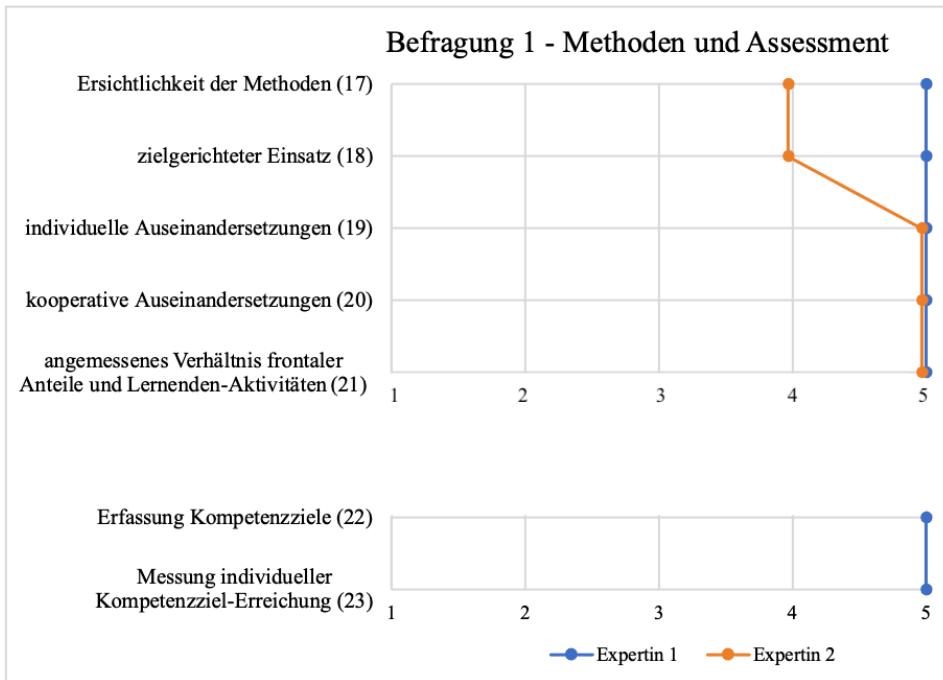


Abbildung 5.3: Befragungsergebnisse von Expertin 1 und 2 zu Methoden und Assessment.

das Aufteilen von Unterrichtsphasen mit mehreren Sozialformen, die Ausführung der Komponenten der Kompetenzen sowie die Unklarheit über die Evaluation der Arbeitsblätter (vgl. Tab. 5.1). Darüber hinaus merkte die Expertin an, dass ihr Angebote zur Binnendifferenzierung fehlen.

Vergleicht man die Einschätzungen beider Expertinnen, so fällt auf, dass die Items 3, 4, 6, 8 (Unterrichtsphasen) sowie 19, 20, 21 (Methoden) gleich bewertet wurden. Allen genannten Items wurde vollständig zugestimmt, bis auf Item 3 (Engagement-Phase), bei dem auf der Likert-Skala 4 von 5 Punkten vergeben wurden. Bei den Items 1, 2, 5 (Unterrichtsphasen), 9 bis 16 (Kompetenzen) und 17, 18 (Methoden) unterscheiden sich die Bewertungen um einen Punkt. Während Expertin 2 für diese Items 4 Punkte vergab, stimmte Expertin 1 uneingeschränkt zu (5 Punkte). Der größte Unterschied zeichnet sich bei Item 7 ab (vgl. Abb. 5.1). Dieses bezieht sich auf die Evaluation-Phase. Im Gegensatz zu Expertin 1, die die volle Punktzahl vergab, entschied sich Expertin 2 hier für die Mittelkategorie.

Für die IRA-Indizes wurden die Werte $r_{wg(j)}^* = 0,899$ und $AD_{M(J)} = 0,357$ ermittelt. Da Expertin 2 die Items 22 und 23 nicht beantwortet hat, sind diese beiden Items bei der Berechnung nicht berücksichtigt.

5.2 Zwischendiskussion

Die Indizes des Interrater Agreement zeigen für die erste Befragung eine hohe Übereinstimmung in der Bewertung durch die Expertinnen an.

Im Folgenden sollen die Rückmeldungen (vgl. Tab. 5.1) und Antworten im Fragebogen (vgl. Abb. 5.1, Abb. 5.2 und Abb. 5.3) der Expertinnen gemeinsam betrachtet werden. Führt man die qualitative Rückmeldung von Expertin 1 mit ihren Antworten im Fragebogen zusammen, so fällt auf, dass sich die Rückmeldung in der quantitativen Bewertung widerspiegelt. Dass die Expertin für das Item 3 (Engagement-Phase) keine volle Punktzahl auf der Likert-Skala vergab, stimmt mit ihrer Rückmeldung überein, in der sie potenzielle Probleme hinsichtlich der Gestaltung der Einführungsphase nannte (Rückmeldung 1).

Auch kann eine Verbindung zwischen den Antworten von Expertin 2 und ihrer schriftlichen Rückmeldung hergestellt werden. Die Rückmeldung der Expertin stellt eine Zusammenfassung ihrer im Fragebogen notierten Kommentare dar. Die Kommentare können dabei als Orientierungshilfe für die Zuordnung von Rückmeldungen zu Items angesehen werden. So könnte man einen Zusammenhang zwischen der Einschätzung der Items 1 und 2 (Unterrichtsphasen) mit 4 von 5 Punkten und ihrer Rückmeldung über das Aufteilen von Unterrichtsphasen mit mehreren Sozialformen (Rückmeldung 2) annehmen. Ihre Rückmeldung über die Ausformulierung der Komponenten der Kompetenzen (Rückmeldung 3) passt damit zusammen, dass sie den Items des Bereichs „Kompetenzen“ nur teilweise zustimmte. Bei Betrachtung ihres Feedbacks lässt sich außerdem vermuten, weshalb sie die Items 22 und 23 nicht beantwortet hat. Für Expertin 2 war die Arbeitsblatt-Analyse als Instrument unklar, um die Kompetenzziel-Erreichung der Lernenden zu messen (Rückmeldung 4), sodass sie sich womöglich nicht in der Lage gesehen hat, eine Bewertung abzugeben. Darüber hinaus könnte Rückmeldung 4 der Hintergrund für ihre Wahl der Mittelkategorie bei Item 7 (Evaluation-Phase) sein. Das Fehlen von Angeboten zur Binnendifferenzierung (Rückmeldung 5), auf das die Expertin ebenfalls verwies, geht über den Fragebogen hinaus.

Diejenigen Rückmeldungen aus Tab. 5.1, die sich auf den Fragebogen beziehen, bildeten den Ausgangspunkt für die Überarbeitung des Workshop-Konzepts (Rückmeldungen 1 bis 4). Das Ergebnis des Re-Designs ist in Anhang F zu finden. Auf Rückmeldung 3 und 4 wurde eingegangen, indem Informationstexte zu den Komponenten der Lernziele und der Evaluation der Lernzielerreichung in der Begleitbroschüre ergänzt wurden (vgl. Anhang F.12). Aus Rückmeldung 1 resultierte, dass die Unterrichtsphase „Einführung“ aus dem Verlaufsplan entfernt wurde (vgl. Anhang F.1). Stattdessen steigen die Lernenden über die Problemstellung zum Lügendetektor in die Thematik ein. Der Begriff des überwachten Lernens wird erst bei der Besprechung der Ergebnisse eingeführt. Diejenigen Fachbegriffe, die zwar einen Überblick geben, aber nicht unmittelbar zum überwachten Lernen gehören, wurden aus dem Workshop ausgelagert. Um der rückgemeldeten Abstraktheit entgegenzuwirken, wurde für das überwachte Lernen ein Beispiel aus dem Alltag hinzugefügt (vgl. Anhang F.2). Schließlich erfolgte eine Aufspaltung der Transferphase in zwei Unterrichtsphasen entlang der Sozialformen (Bearbeitung der Transferaufgabe und Besprechung der Transferaufgabe), um an Rückmeldung 2 anzuknüpfen (vgl. Anhang F.1).

5.3 Zweite Befragung

Abb. 5.4, Abb. 5.5 und Abb. 5.6 zeigen die Antworten der zweiten Befragung. Die Hauptpunkte der schriftlichen Rückmeldungen können Tab. 5.2 entnommen werden.

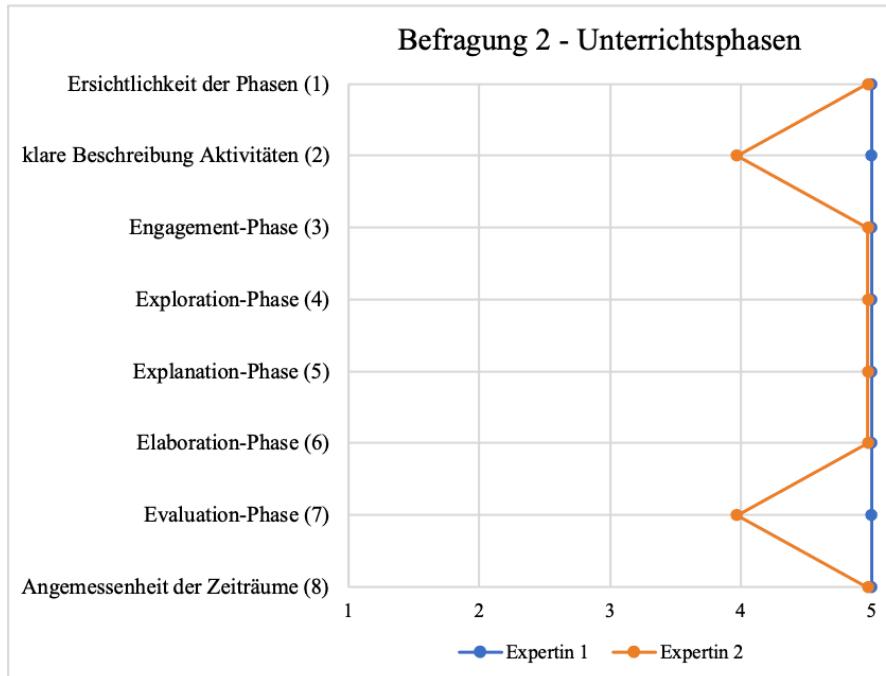


Abbildung 5.4: Befragungsergebnisse von Expertin 1 und 2 nach Re-Design zu den Unterrichtsphasen.

Nr.	Rückmeldung
Expertin 1	- keine Änderungsvorschläge
Expertin 2	1 Unterrichtsphasen mit mehreren Sozialformen vorhanden 2 Bewertungsschema für die Evaluation der Kompetenzziele fehlt 3 Kompetenziel-Komponente „Kriterium“ konkreter ausformulieren

Tabelle 5.2: Kernaussagen aus den schriftlichen Rückmeldungen der beiden Expertinnen nach Re-Design des Workshop-Konzepts.

Bei der Einschätzung des überarbeiteten Workshop-Konzepts stimmte Expertin 1 allen Items uneingeschränkt zu. In ihrer schriftlichen Rückmeldung wurden keine Überarbeitungspunkte angeführt (vgl. Tab. 5.2).

Expertin 2 vergab für das re-designte Workshop-Konzept eine vollständige Zustimmung für 20 Items. Den übrigen Items 2,7 (Unterrichtsphasen) sowie 16 (Kompetenzen) stimmte sie teilweise zu. Im Fragebogen formulierte sie Kommentare (vgl. Anhang E.2). Diese sind in Tab. 5.2 inhaltlich wiedergegeben. Dabei äußerte die Expertin, dass im Workshop Unterrichtsphasen mit mehreren Sozialformen existieren. Auch wies sie auf das Fehlen eines Bewertungsschemas für die Evaluation der Kompetenzziele hin und schlug eine konkretere Ausführung der Kompetenziel-Komponente „Kriterium“ vor.

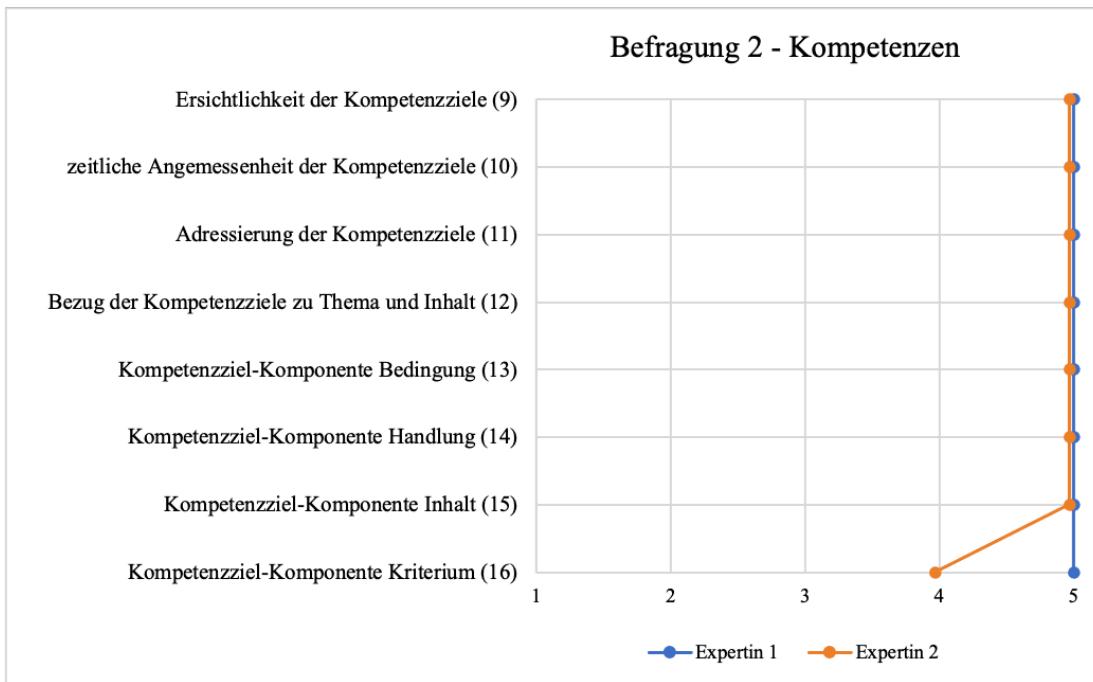


Abbildung 5.5: Befragungsergebnisse von Expertin 1 und 2 nach Re-Design zu den Kompetenzen.

Die Gegenüberstellung der Antworten beider Expertinnen zeigt eine Übereinstimmung in 20 Items. Dabei schätzten sie die Items 1, 3, 4, 5, 6, 8 (Unterrichtsphasen) sowie 9 bis 15 (Kompetenzen) und 17 bis 23 (Methoden und Assessment) mit voller Zustimmung ein. Den Items 2, 7 und 16 stimmte Expertin 1 vollständig und Expertin 2 teilweise zu.

Für die IRA-Indizes wurden die Werte $r_{wg(j)}^* = 0,984$ und $AD_{M(J)} = 0,065$ berechnet.

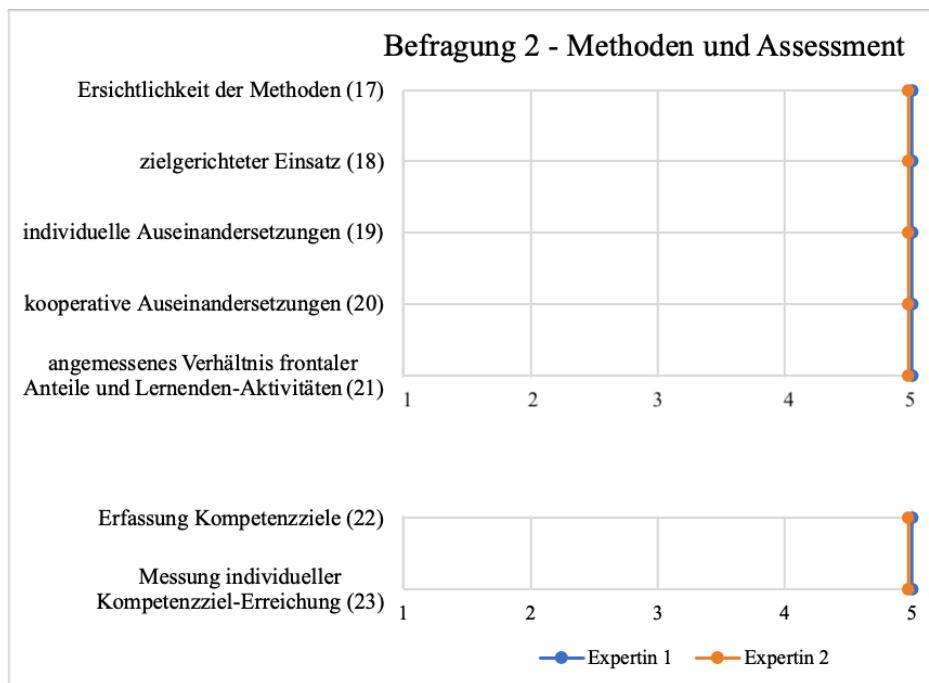


Abbildung 5.6: Befragungsergebnisse von Expertin 1 und 2 nach Re-Design zu Methoden und Assessment.

5.4 Vergleich beider Befragungen

In Tab. 5.3 sind die durchschnittlichen Bewertungen der Items innerhalb der ersten und zweiten Befragung gegenübergestellt. Die Items wurden mit einer fünfstufigen Likert-Skala abgefragt. So liegt der maximal zu erreichende Wert bei 5 („trifft sehr zu“) und der kleinste Wert bei 1 („trifft überhaupt nicht zu“).

Bereich	Item	Befragung 1	Befragung 2
		\bar{x}	\bar{x}
Unterrichtsphasen	1 Ersichtlichkeit der Phasen	4,5	5
	2 klare Beschreibung Aktivitäten	4,5	4,5
	3 Engagement-Phase	4	5
	4 Exploration-Phase	5	5
	5 Explanations-Phase	4,5	5
	6 Elaboration-Phase	5	5
	7 Evaluation-Phase	4	4,5
	8 Angemessenheit der Zeiträume	5	5
Kompetenzen	9 Ersichtlichkeit der Kompetenzziele	4,5	5
	10 zeitliche Angemessenheit der Kompetenzziele	4,5	5
	11 Adressierung der Kompetenzziele	4,5	5
	12 Bezug der Kompetenzziele zu Thema und Inhalt	4,5	5
	13 Kompetenzziel-Komponente Bedingung	4,5	5
	14 Kompetenzziel-Komponente Handlung	4,5	5
	15 Kompetenzziel-Komponente Inhalt	4,5	5
	16 Kompetenzziel-Komponente Kriterium	4,5	4,5
Methoden	17 Ersichtlichkeit der Methoden	4,5	5
	18 zielgerichteter Einsatz	4,5	5
	19 individuelle Auseinandersetzungen	5	5
	20 kooperative Auseinandersetzungen	5	5
	21 angemessenes Verhältnis frontaler Anteile und Lernenden-Aktivitäten	5	5
Assessment (Evaluation)	22 Erfassung Kompetenzziele	5*	5
	23 Messung individueller Kompetenzziel-Erreichung	5*	5
Mittelwert		4,6	4,9
$r_{wg(j)}^*$		0,899	0,984
$AD_{M(J)}$		0,357	0,065

Tabelle 5.3: Durchschnittliche Bewertung der Fragebogen-Items auf der Likert-Skala sowie IRA-Indizes. Bewertungen, die nur von einer Expertin abgegeben wurden, sind mit einem * gekennzeichnet.

Bei der ersten Befragung wurden im Mittel 4,6 Punkte auf der Likert-Skala vergeben. Ein Blick auf die Einzelbewertung der Expertinnen zeigt, dass kein Item abgelehnt wurde. Die Mittelkategorie wurde einmal vergeben. Bei der zweiten Befragung, die nach dem Re-Design des Workshop-Konzepts stattfand, wurde eine durchschnittliche Bewertung von 4,9 Punkten erreicht. Die niedrigste absolute Bewertung lag dabei bei 4 Punkten.

Vergleicht man die Ergebnisse der ersten und zweiten Befragung, so fällt auf, dass kein Item bei der zweiten Befragung eine geringere durchschnittliche Bewertung als bei der ersten Befragung erhalten hat. Mehr als die Hälfte der Items wurde bei der zweiten Befragung im Durchschnitt mit einer höheren Punktzahl bewertet. Zudem wurde bei der zweiten Befragung ein höherer $r_{wg(j)}^*$ - und ein niedrigerer $AD_{M(J)}$ -Wert erzielt.

5.5 Abschließende Diskussion

Auch für die zweite Befragung geben die IRA-Indizes eine hohe Übereinstimmung der Expertinnen in der Einschätzung des überarbeiteten Workshop-Konzepts an.

Betrachtet man die Ergebnisse der zweiten Befragung, so kann auch hier ein Zusammenhang zwischen den qualitativen Rückmeldungen (vgl. Tab. 5.2) und den Antworten im Fragebogen (vgl. Abb. 5.4, Abb. 5.5 und Abb. 5.6) hergestellt werden. Bei Expertin 1 kann man beobachten, dass ihre vollständige Zustimmung zu allen Items mit dem Fehlen weiterer Änderungsvorschläge zusammenpasst.

Da Expertin 2 ihr schriftliches Feedback in Form von Kommentaren unmittelbar im Fragebogen platzierte, war die Zuordnung zu den jeweiligen Items eindeutig abzulesen. Ihre Einschätzung von Item 2 (klare Beschreibung Aktivitäten) begründete sie damit, dass Unterrichtsphasen mit mehreren Sozialformen existieren (Rückmeldung 1). Für Item 7 (Evaluation-Phase) vergab sie 4 Punkte auf der Likert-Skala mangels eines Bewertungsschemas für die Evaluation der Kompetenzziele (Rückmeldung 2). Eine nach Auffassung von Expertin 2 zu wenig konkrete Beschreibung der Kompetenzziel-Komponente Kriterium (Rückmeldung 3) resultierte in einer Bewertung von Item 16 (Kompetenzziel-Komponente Kriterium) mit 4 Punkten. Insgesamt sind ihre Rückmeldungen bei der zweiten Befragung ähnlich zu denen der ersten Befragung ausgefallen.

Aus dem Vergleich der ersten und zweiten Befragung geht hervor, dass in der zweiten Befragung sowohl der Grad der Übereinstimmung als auch die Bewertung der Expertinnen höher ausfiel. Dieses Ergebnis konnte erzielt werden, indem für das Re-Design die Rückmeldungen aus der ersten Befragung, die das Verbesserungspotenzial des Workshop-Konzepts anzeigen, berücksichtigt wurden. Die Steigerung der Ergebnisse von der ersten zur zweiten Befragung lässt den Schluss zu, dass das zu Beginn der Arbeit formulierte Ziel der Optimierung des Workshop-Konzepts erreicht wurde.

Durch den zeitlich begrenzten Rahmen dieser Arbeit konnte nur die interne Konsistenz des Workshops als Evaluationsfokus in den Blick genommen werden. In [34] werden weitere Evaluationsfokusse in der designbasierten Forschung wie etwa die Durchführbarkeit im Klassenzimmer (Beta-Test) oder Wirksamkeit der Intervention (Gamma-Test) angeführt. Auch wird in dieser Arbeit nur ein Teil des Themenfeldes „Maschinelles Lernen“ abgedeckt.

Gemäß [34] könnten für eine zukünftige Optimierung des Workshop-Konzepts andere Evaluationsfokusse etwa in Pilotversuchen oder Erprobungen im Feld mit Schülerinnen und

Schülern anvisiert werden. Darüber hinaus ließe sich die Forschungsfrage auf andere Bereiche des maschinellen Lernens wie beispielsweise das unüberwachte Lernen ausweiten.

6. Fazit

Diese Arbeit ging der Frage nach, wie eine geeignete Lernumgebung aussieht, um Lernende der Sekundarstufe I (ab Klasse 7) in die Konzepte des überwachten Lernens einzuführen.

Im Rahmen der Arbeit wurde von der Verfasserin ein Workshop-Konzept entwickelt, das aus einem Unterrichtsverlaufsplan, einer Begleitbroschüre, Präsentationsfolien, Videos, Arbeits- sowie Lösungsblättern besteht. Im Workshop erkunden Lernende das überwachte Lernen anhand einer Unplugged-Aktivität zu Entscheidungsbäumen, die sich um die Konstruktion eines Lügendetektors dreht. Die Lernziele des Workshops knüpfen an die Richtlinien der Arbeitsgruppe AI for K-12 zur Vermittlung von KI an und stellen Bezüge zum aktuellen Bildungsplan der Informatik in Baden-Württemberg her.

Um die Forschungsfrage zu beantworten, wurde das Workshop-Konzept durch die Befragung von zwei Expertinnen mit Workshop-Erfahrung aus dem Bereich der Informatik eingeschätzt. Auf Grundlage der designbasierten Forschung wurde die Befragung nach Überarbeitung des Workshop-Konzepts wiederholt. Als quantitatives Befragungsinstrument wurde ein Fragebogen eingesetzt. Zusätzlich formulierten die Expertinnen qualitative Rückmeldungen bei jeder Befragung.

Nach der Überarbeitung wurde das Konzept von den Expertinnen bei sehr starker Übereinstimmung hoch bewertet. Dabei entsprach die Bewertung der Fragebogen-Items im Durchschnitt 4,9 von 5 Punkten auf der Likert-Skala. Vor dem Hintergrund dieser Ergebnisse kann das überarbeitete Workshop-Konzept im Kontext der Einführung Lernender der Sekundarstufe I (ab Klasse 7) in die Grundideen überwachten Lernens empfohlen werden.

Als Einschränkung der Forschung ist zu nennen, dass die Evaluation des Workshop-Konzepts in dieser Arbeit auf die interne Konsistenz (Alpha-Test) begrenzt ist. Um den Reifegrad des Workshop-Konzepts zu erhöhen, sind weitere Untersuchungen mit anderen Evaluationsfokus notwendig. Da die Forschungsfrage nur eine Teilkategorie des maschinellen Lernens umfasst, ließe sich diese auf andere ML-Bereiche übertragen.

Literaturverzeichnis

- [1] Aberham, J. und F. Kuruc: *Clusteranalyse*. In: Kersting, K., C. Lampert und C. Rothkopf (Hrsg.): *Wie Maschinen lernen. Künstliche Intelligenz verständlich erklärt*, S. 53–57. Springer, Wiesbaden, 2019.
- [2] AI for K-12: *Draft Big Idea 3 - Progression Chart*, November 2020.
<https://ai4k12.org/wp-content/uploads/2021/01/AI4K12-Big-Idea-3-Progression-Chart-Working-Draft-of-Big-Idea-3-v.11.19.2020.pdf>, besucht: 15.08.2022.
- [3] Alpaydin, E.: *Maschinelles Lernen*. De Gruyter Oldenbourg, 2. Aufl., 2019.
- [4] Asbrand, B. und G. Lang-Wojtasik: *Qualitätskriterien für Unterrichtsmaterialien entwicklungsbezogener Bildungsarbeit*. ZEP: Zeitschrift für internationale Bildungsforschung und Entwicklungspädagogik, 32(2):8–13, 2009.
- [5] Aust, H.: *Das Zeitalter der Daten. Was Sie über Grundlagen, Algorithmen und Anwendungen wissen sollten*. Springer, Berlin und Heidelberg, 2021.
- [6] Bell, T., J. Alexander, I. Freeman und M. Grimley: *Computer Science Unplugged: School students doing real computing without computers*. The New Zealand Journal of Applied Computing and Information Technology, 13(1):20–29, 2009.
- [7] Biggs, J.: *Enhancing teaching through constructive alignment*. Higher education, 32(3):347–364, 1996.
- [8] Brynjolfsson, E. und A. McAfee: *The Business of Artificial Intelligence—What It Can and Cannot Do for Your Organization*, 2017. <https://www.predictiveanalyticsworld.com/machinelearningtimes/business-artificial-intelligence-can-not-organization/8990/>, besucht: 04.10.2022.
- [9] Buxmann, P. und H. Schmidt: *Grundlagen der Künstlichen Intelligenz und des Maschinellen Lernens*. In: Buxmann, P. und H. Schmidt (Hrsg.): *Künstliche Intelligenz. Mit Algorithmen zum wirtschaftlichen Erfolg*, S. 3–25. Springer Gabler, Berlin und Heidelberg, 2021.

- [10] Bybee, R. W.: *The BSCS 5E instructional model: Personal reflections and contemporary implications*. Science and Children, 51(8):10–13, 2014.
- [11] Carpinelli, J., H. Kimmel, L. Hirsch, L. Burr-Alexander, R. Rockland und M. OShea: *A Rubric To Evaluate Standards Based Lesson Plans And Students' Achievement Of The Standards*. In: *ASEE Annual Conference and Exposition, Conference Proceedings*, S. 13.98.1–13.98.10, 2008.
- [12] Design-Based Research Collective: *Design-based research: An emerging paradigm for educational inquiry*. Educational researcher, 32(1):5–8, 2003.
- [13] Dolmans, D. H. und D. Tigelaar: *Building bridges between theory and practice in medical education using a design-based research approach: AMEE Guide No. 60*. Medical teacher, 34(1):1–10, 2012.
- [14] Ester, M. und J. Sander: *Knowledge Discovery in Databases. Techniken und Anwendungen*. Springer, Berlin und Heidelberg, 2020.
- [15] Euler, D.: *Wirkungs-vs. Gestaltungsforschung – eine feindliche Koexistenz?* Zeitschrift für Berufs-und Wirtschaftspädagogik, 107(4):520–542, 2011.
- [16] Feierabend, S., T. Rathgeb, H. Kheredmand und S. Glöckler: *JIM-Studie 2020: Jugend, Information, Medien: Basisuntersuchung zum Medienumgang 12-bis 19-Jähriger*. Medienpädagogischer Forschungsverbund Südwest, 2020.
- [17] Fraefel, U.: *Professionalization of pre-service teachers through university-school partnerships*. In: *Conference Proceedings of WERA Focal Meeting, Edinburgh*, 2014.
- [18] Gilch, A. und T. Schüler: *Daten*. In: Kersting, K., C. Lampert und C. Rothkopf (Hrsg.): *Wie Maschinen lernen. Künstliche Intelligenz verständlich erklärt*, S. 29–37. Springer, Wiesbaden, 2019.
- [19] Große-Böling, G. und A. Mühling: *Students Perception of the Inner Workings of Learning Machines*. In: *International Conference on Learning and Teaching in Computing and Engineering (LaTICE)*. Ho-Chi-Minh-City, Vietnam, 2020.
- [20] Géron, A.: *Praxiseinstieg Machine Learning mit Scikit-Learn, Keras und TensorFlow. Konzepte, Tools und Techniken für intelligente Systeme*. O'Reilly, 2020.
- [21] Herrmann, T. und L. F. Peiss: *Verstärkendes Lernen*. In: Kersting, K., C. Lampert und C. Rothkopf (Hrsg.): *Wie Maschinen lernen. Künstliche Intelligenz verständlich erklärt*, S. 203–212. Springer, Wiesbaden, 2019.

- [22] Hetzel, L. und F. Wangelik: *Künstliche neuronale Netze*. In: Kersting, K., C. Lampert und C. Rothkopf (Hrsg.): *Wie Maschinen lernen. Künstliche Intelligenz verständlich erklärt*, S. 149–161. Springer, Wiesbaden, 2019.
- [23] Jo, T.: *Machine Learning Foundations. Supervised, Unsupervised, and Advanced Learning*. Springer, Cham, 2021.
- [24] Kossen, J., F. Kuruc und M. E. Müller: *Einleitung*. In: Kersting, K., C. Lampert und C. Rothkopf (Hrsg.): *Wie Maschinen lernen. Künstliche Intelligenz verständlich erklärt*, S. 3–10. Springer, Wiesbaden, 2019.
- [25] Kossen, J. und M. E. Müller: *Lineare Regression*. In: Kersting, K., C. Lampert und C. Rothkopf (Hrsg.): *Wie Maschinen lernen. Künstliche Intelligenz verständlich erklärt*, S. 61–67. Springer, Wiesbaden, 2019.
- [26] Kossen, J. und M. E. Müller: *Regression*. In: Kersting, K., C. Lampert und C. Rothkopf (Hrsg.): *Wie Maschinen lernen. Künstliche Intelligenz verständlich erklärt*, S. 39–43. Springer, Wiesbaden, 2019.
- [27] Kossen, J., M. E. Müller und M. Ruckriegel: *Entscheidungsbäume*. In: Kersting, K., C. Lampert und C. Rothkopf (Hrsg.): *Wie Maschinen lernen. Künstliche Intelligenz verständlich erklärt*, S. 111–123. Springer, Wiesbaden, 2019.
- [28] Krause, M. und E. Natterer: *Maschinelles Lernen*. In: Kersting, K., C. Lampert und C. Rothkopf (Hrsg.): *Wie Maschinen lernen. Künstliche Intelligenz verständlich erklärt*, S. 21–27. Springer, Wiesbaden, 2019.
- [29] Laubach, U.: *Informatik für alle! Ein Plädoyer*, 2018. <https://gi.de/meldung/informatik-fuer-alle-ein-plaedyer>, besucht: 03.08.2022.
- [30] LeBreton, J. und J. Senter: *Answers to 20 Questions About Interrater Reliability and Interrater Agreement*. *Organizational Research Methods*, 11:815–852, Okt. 2008.
- [31] Lindner, A., M. Berges und M. Lechner: *KI im Toaster? Schüler: innenvorstellungen zu künstlicher Intelligenz*. INFOS 2021–19. GI-Fachtagung Informatik und Schule, S. 133–142, 2021.
- [32] Mager, R. F.: *Preparing Instructional Objectives*. David S. Lake, Belmont (Kalifornien), 2. Aufl., 1984.
- [33] Marques, L. S., C. Gresse von Wangenheim und J. C. Hauck: *Teaching Machine Learning in School: A Systematic Mapping of the State of the Art*. *Informatics in Education*, 19(2):283–321, 2020.

- [34] McKenney, S. und T. C. Reeves: *Conducting Educational Design Research*. Routledge, London and New York, 2012.
- [35] Michaeli, T., S. Seegerer und S. Jatzlau: *Kein Hexenwerk–Ideen des maschinellen Lernens in SNAP!* LOG IN: Vol. 40, No. 1, 2020.
- [36] Ministerium für Kultus, Jugend und Sport Baden-Württemberg: *GEMEINSAMER BILDUNGSPLAN DER SEKUNDARSTUFE I. Aufbaukurs Informatik*. http://bildungsplaene-bw.de/site/bildungsplan/get/documents/lsbw/export-pdf/depot-pdf/ALLG/BP2016BW_ALLG_SEK1_INF7.pdf, besucht: 15.08.2022.
- [37] Ministerium für Kultus, Jugend und Sport Baden-Württemberg: *GEMEINSAMER BILDUNGSPLAN DER SEKUNDARSTUFE I. Wahlfach Informatik an der Hauptschule, Werkrealschule und Realschule*. http://bildungsplaene-bw.de/site/bildungsplan/get/documents/lsbw/export-pdf/depot-pdf/ALLG/BP2016BW_ALLG_SEK1_INFWF.pdf, besucht: 15.08.2022.
- [38] Mitchell, T. M.: *Machine Learning*. McGraw Hill Education, internationale Aufl., 1997.
- [39] MSV, J.: *Teachable Machine From Google Makes It Easy To Train And Deploy ML Models*, November 2020. <https://www.forbes.com/sites/janakirammsv/2020/11/29/teachable-machine-from-google-makes-it-easy-to-train-and-deploy-ml-models/>, besucht: 17.08.2022.
- [40] Müller, D.: *k-Means-Algorithmus*. In: Kersting, K., C. Lampert und C. Rothkopf (Hrsg.): *Wie Maschinen lernen. Künstliche Intelligenz verständlich erklärt*, S. 81–88. Springer, Wiesbaden, 2019.
- [41] Ndihokubwayo, K., C. Byukusenge, E. Byusa, H. T. Habiyaremye, A. Mbonyiryivuze und J. Mukagihana: *Lesson plan analysis protocol (LPAP): A useful tool for researchers and educational evaluators*. Heliyon, 8(1):e08730, 2022.
- [42] O'Neill, T. A.: *An Overview of Interrater Agreement on Likert Scales for Researchers and Practitioners*. Frontiers in Psychology, 8, 2017.
- [43] Rosebrock, S.: *Geometrische Gruppentheorie. Ein Einstieg mit dem Computer: Basiswissen für Studium und Mathematikunterricht*. Vieweg, Wiesbaden, 2004.
- [44] Samuel, A. L.: *Some Studies in Machine Learning Using the Game of Checkers*. IBM Journal of Research and Development 44(1/2), S. 210–229, 1959.
- [45] SAP: *Was ist maschinelles Lernen?* <https://www.sap.com/germany/insights/what-is-machine-learning.html>, besucht: 08.08.2022.

- [46] Seegerer, S., A. Lindner und R. Romeike: *AI Unplugged–Wir ziehen Künstlicher Intelligenz den Stecker*. In: Pasternak, A. (Hrsg.): *Informatik für alle*, S. 325–334. Gesellschaft für Informatik, Bonn, 2019.
- [47] Seegerer, S., T. Michaeli und R. Romeike: *So lernen Maschinen!* LOG IN: Vol.40, No. 1, S. 27–31, 2020.
- [48] Standl, B.: *Solving everyday challenges in a computational way of thinking*. In: *International Conference on Informatics in Schools: Situation, Evolution, and Perspectives*, S. 180–191. Springer, 2017.
- [49] Strecker, K. und E. Modrow: *Eine Unterrichtssequenz zum Einstieg in Konzepte des maschinellen Lernens*. In: Pasternak, A. (Hrsg.): *Informatik für alle*, S. 335–344. Gesellschaft für Informatik, Bonn, 2019.
- [50] Touretzky, D., C. Gardner-McCune, F. Martin und D. Seehorn: *Envisioning AI for K-12: What Should Every Child Know about AI?* In: *Proceedings of the AAAI conference on artificial intelligence*, Bd. 33, S. 9795–9799, 2019.
- [51] Wuttke, L.: *Training-, Validierung- und Testdatensatz*, 2020. <https://datasolut.com/wiki/trainingsdaten-und-testdaten-machine-learning/>, besucht: 09.08.2022.
- [52] Wuttke, L.: *Praxisleitfaden für Künstliche Intelligenz in Marketing und Vertrieb. Beispiele, Konzepte und Anwendungsfälle*. Springer Gabler, Wiesbaden, 2021.

Abbildungsverzeichnis

2.1	Abgrenzung und Verortung von KI und ML. In Anlehnung an [52], S. 56.	4
2.2	Murmeln werden den Klassen „rot“ und „blau“ zugeordnet. Eigene Darstellung.	5
2.3	Vorhersage des Murmel-Gewichts mittels des Durchmessers einer Murmel.	6
2.4	Vom Training zur Anwendung eines überwachten Lernalgorithmus. In Anlehnung an [45].	7
2.5	Grafische Darstellung von Entscheidungsbäumen.	8
2.6	Entscheidungsbaum zur Vorhersage, ob sich das Wetter zum Tennisspielen eignet. In Anlehnung an [38], S. 53.	9
2.7	Murmel mit Cluster-Prototypen vergleichen. Eigene Darstellung.	10
2.8	Grafische Darstellung des k-Means-Algorithmus. In Anlehnung an [14], S. 53.	11
2.9	Ein autonomer Agent interagiert mit seiner Umgebung. In Anlehnung an [38], S. 368 und 372.	12
2.10	Initialisierte Tabelle. In Anlehnung an [23], S. 369.	13
2.11	Funktion für den Q-Learning Algorithmus. Nach [38], S. 375.	13
2.12	Ein Neuron empfängt Signale und leitet diese weiter. In Anlehnung an [22], S. 153.	14
2.13	Ein Perzeptron mit gewichteter Summe und Schwellenwertfunktion. In Anlehnung an [5], S. 174 und [38], S. 87.	15
2.14	Ein Perzeptron, das das logische UND repräsentiert. In Anlehnung an [5], S. 174 und [38], S. 87.	15
2.15	Die Schichten einer FFNN Architektur. In Anlehnung an [5], S. 164.	16
4.1	Prozessmodell zur designbasierten Forschung. In Anlehnung an [17].	22
5.1	Befragungsergebnisse von Expertin 1 und 2 zu den Unterrichtsphasen.	25
5.2	Befragungsergebnisse von Expertin 1 und 2 zu den Kompetenzen.	26
5.3	Befragungsergebnisse von Expertin 1 und 2 zu Methoden und Assessment.	27
5.4	Befragungsergebnisse von Expertin 1 und 2 nach Re-Design zu den Unterrichtsphasen.	29
5.5	Befragungsergebnisse von Expertin 1 und 2 nach Re-Design zu den Kompetenzen.	30

5.6 Befragungsergebnisse von Expertin 1 und 2 nach Re-Design zu Methoden und Assessment.	30
--	----

Anhang

A.1 Entropiekurve einer Klassifikation mit zwei Klassen. Nach [38], S. 57.	44
A.2 Erster Konstruktionsschritt eines Entscheidungsbaums nach dem ID3-Algorithmus. In Anlehnung an [38], S. 61.	46
A.3 Sofortige und kumulative Belohnung im Vergleich. In Anlehnung an [38], S. 372.	47
A.4 Aktualisieren des \hat{Q} -Wertes für den Zustand A. Der schwarze Pfeil gibt die sofortige Belohnung an. Die roten Pfeile repräsentieren die \hat{Q} -Werte für Zustand B. Der Agent führt eine Bewegung nach rechts aus, sodass $\delta(s_A, \rightarrow) = s_B$. In Anlehnung an [38], S. 376.	48
A.5 Trainieren eines neuronalen Netzes nach der Perzeptron-Regel. In Anlehnung an [49], S. 341.	50

Tabellenverzeichnis

2.1	Wichtige überwachte Lernalgorithmen im Überblick. In Anlehnung an [20] und [23].	8
2.2	Wichtige unüberwachte Lernalgorithmen im Überblick. In Anlehnung an [14] und [20].	10
4.1	Fragebogen-Items nach Bereichen organisiert.	23
5.1	Kernaussagen aus den schriftlichen Rückmeldungen der beiden Expertinnen.	26
5.2	Kernaussagen aus den schriftlichen Rückmeldungen der beiden Expertinnen nach Re-Design des Workshop-Konzepts.	29
5.3	Durchschnittliche Bewertung der Fragebogen-Items auf der Likert-Skala sowie IRA-Indizes. Bewertungen, die nur von einer Expertin abgegeben wurden, sind mit einem * gekennzeichnet.	31
Anhang		
A.1	Tabellarischer Trainingsdatensatz. In Anlehnung an [38], S. 59.	45

Anhang A Ergänzende Informationen zur Fachsystematik

A.1 Daten

Daten stellen eine wichtige Grundlage für das maschinelle Lernen dar. Nach [18] können beim maschinellen Lernen im Wesentlichen vier verschiedene Datentypen unterschieden werden:

- a. **Kategoriale Daten** sind Variablen, die bestimmte Ausprägungen annehmen können. Dabei wird jeder möglichen Ausprägung der Variable eine Zahl zugeordnet, die die jeweilige Ausprägung repräsentiert. Die zugeordnete Zahl enthält keine Informationen über die Ausprägung und wird zufällig ausgewählt. Beispiele sind Geschlecht und Bundesland.
- b. **Ordinalen Daten** unterscheiden sich von kategorischen Daten darin, dass die Ausprägungen einer Variable in einer Reihenfolge angeordnet sind. So kann man beispielsweise ein Bewertungssystem mit sechs Stufen durch die Zahlen 0, 1, 2, 3, 4 und 5 kodieren.
- c. **Diskrete Daten** sind keine Codes um Ausprägungen darzustellen, sondern bilden tatsächliche Zahlenwerte ab. Deshalb können auf diskreten Daten Rechenoperationen durchgeführt werden. Zu den diskreten Daten gehört beispielsweise die Anzahl der Nutzerinnen und Nutzer, die einen Film angesehen haben.
- d. **Stetige Daten** zeichnen sich dadurch aus, dass sie im Gegensatz zu diskreten Daten auch Kommazahlen sein können. Beispielsweise könnte man mithilfe von stetigen Daten die genaue Zeitspanne bis zum (vorzeitigen) Beenden eines Filmes durch eine Nutzerin oder einen Nutzer festhalten.

A.2 ID3-Algorithmus

Dieses Kapitel befasst sich mit dem ID3-Algorithmus nach [38]. Mithilfe des Algorithmus können Entscheidungsbäume mittels Trainingsdaten von der Wurzel ausgehend konstruiert werden. Der Algorithmus beginnt mit der Frage nach dem besten Attribut für den Wurzelknoten. Dieses wird bestimmt und der Wurzel zugeordnet. Für jeden möglichen Wert, den das Attribut annehmen kann, wird eine Kante angelegt, die zu einem Kindknoten führt. Die Trainingsdaten werden entlang der Kanten jeweils dem entsprechenden Kindknoten zugeordnet. Der für die Wurzel geschilderte Ablauf wird dann für jeden Kindknoten mithilfe der ihm zugehörigen Trainingsdaten wiederholt.

Über den sogenannten Informationsgewinn wird das beste Attribut für die jeweilige Position im Baum identifiziert. In diesem Kontext ist dasjenige Attribut am besten, das den höchsten

Informationsgewinn besitzt. Um den Informationsgewinn quantitativ beschreiben zu können, wird die Entropie als Größe benötigt. Der Begriff stammt aus der Informationstheorie und stellt das Maß für die (Un-)Reinheit einer Menge von Beispielen dar¹. Beim maschinellem Lernen kann die Entropie als die kleinste Anzahl an Bits aufgefasst werden, die nötig ist, um die Klassifikation eines beliebigen Elements in einer Menge von Beispielen zu kodieren. Gehören etwa alle Beispiele in einer Menge S derselben Klasse c_1 an, so werden 0 Bit benötigt. Da klar ist, welcher Klasse ein Element zugeordnet ist, muss keine Nachricht kodiert werden. Die Entropie ist dann gleich null. Sind die Beispiele genau zur Hälfte auf zwei Klassen c_1 und c_2 aufgeteilt, so ist 1 Bit notwendig, um die korrekte Klasse eines Elements zu kodieren. In diesem Fall hat die Entropie den Wert 1. Das Schaubild in Abb. A.1 stellt diesen Sachverhalt dar. Der dort abgebildete Graph kann durch die Funktion $Entropie(S) = -p_1 \cdot \log_2 p_1 - (1 - p_1) \cdot \log_2 (1 - p_1)$ mit $p_1 + p_2 = 1 \Leftrightarrow p_2 = 1 - p_1$ beschrieben werden. Dabei gibt p_1 den Anteil der Elemente an, die der Klasse c_1 zugeordnet sind. Entsprechend beschreibt p_2 den Anteil der Elemente, die in Klasse c_2 gehören. Die Basis 2 des Logarithmus kann auf die Kodierung der Klassifikation in Bits (0 oder 1) zurückgeführt werden. Ein Beispiel: Gegeben ist eine Menge S aus 14 Beispielen. Neun davon werden der Klasse c_1 zugeordnet und die restlichen fünf gehören in Klasse c_2 . Für die Entropie von S ergibt sich hieraus: $Entropie(S) = -\frac{9}{14} \cdot \log_2(\frac{9}{14}) - \frac{5}{14} \cdot \log_2(\frac{5}{14}) = 0,940$. Der Wert 0,940 kann so interpretiert werden, dass im Durchschnitt weniger als 1 Bit für die Kodierung notwendig sind. Will man die Entropie einer Menge S beschreiben, deren Elemente mehr als zwei Klassen, etwa n verschiedenen Klassen, zugeordnet sind, so kann die Formel verallgemeinert werden: $Entropie(S) = \sum_{i=1}^n -p_i \cdot \log_2 p_i$.

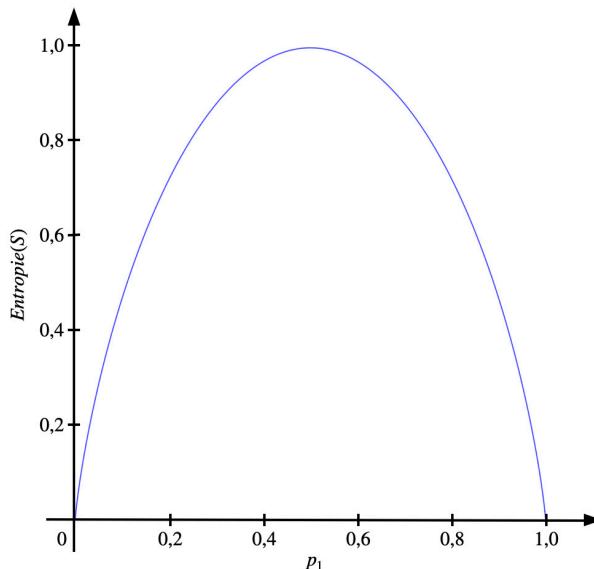


Abbildung A.1: Entropiekurve einer Klassifikation mit zwei Klassen. Nach [38], S. 57.

¹Mehr Informationen zur Informationstheorie und Entropie: Weitz, E.: *Konkrete Mathematik (nicht nur) für Informatiker. Mit vielen Grafiken und Algorithmen in Python*. Springer Spektrum, Wiesbaden, 2018. S. 865 ff.

Mithilfe der Entropie kann nun der Informationsgewinn beschrieben werden. Unter dem Informationsgewinn eines Attributs A versteht man die erwartete Entropie-Abnahme, die durch die Zerlegung der Trainingsdaten in einer Menge S entlang dieses Attributs erfolgt: $Gewinn(S, A) = Entropie(S) - \sum_{w \in Werte(A)} \frac{|S_w|}{|S|} \cdot Entropie(S_w)$. Dabei ist $Werte(A)$ die Menge aller möglichen Werte, die das Attribut A annehmen kann. Für jeden Wert w von A gibt es eine Teilmenge S_w , die alle Elemente aus S enthält, die für A den Wert w besitzen. Der Minuend gibt die Entropie der ursprünglichen Gesamtmenge S an. Der Subtrahend stellt die erwartete Entropie nach Zerlegung von S entlang von A dar. Diese setzt sich aus der gewichteten Summe aller Entropie-Werte der zugehörigen Teilmengen S_w zusammen. Die Differenz von Entropie der Gesamtmenge und erwarteter Entropie nach Zerlegung entspricht gerade der erwarteten Entropie-Abnahme.

Der Einsatz des Informationsgewinns im Zusammenhang mit dem ID3-Algorithmus soll an einem Beispiel veranschaulicht werden. In Tab. A.1 ist ein tabellarischer Trainingsdatensatz mit den Attributen „Himmel“, „Luftfeuchtigkeit“ und „Wind“ abgebildet. Mithilfe dieser soll ein Entscheidungsbaum konstruiert werden, mit dem man entscheiden kann, ob das Wetter zum Tennisspielen geeignet ist. Nach dem ID3-Algorithmus muss zuerst das beste Attribut für den Wurzelknoten identifiziert werden. Um dieses zu bestimmen, muss für jedes Attribut der Informationsgewinn berechnet werden. Anschließend wird das Attribut mit dem höchsten Wert ausgewählt.

Tag	Himmel	Luftfeuchtigkeit	Wind	Tennis
T1	Sonne	hoch	schwach	Nein
T2	Sonne	hoch	stark	Nein
T3	Wolken	hoch	schwach	Ja
T4	Regen	hoch	schwach	Ja
T5	Regen	normal	schwach	Ja
T6	Regen	normal	stark	Nein
T7	Wolken	normal	stark	Ja
T8	Sonne	hoch	schwach	Nein
T9	Sonne	normal	schwach	Ja
T10	Regen	normal	schwach	Ja
T11	Sonne	normal	stark	Ja
T12	Wolken	hoch	stark	Ja
T13	Wolken	normal	schwach	Ja
T14	Regen	hoch	stark	Nein

Tabelle A.1: Tabellarischer Trainingsdatensatz. In Anlehnung an [38], S. 59.

Betrachtet man die 14 Trainingsbeispiele aus Tab. A.1, so erkennt man, dass zwei verschiedene Klassen („Tennis“ und „Kein Tennis“) existieren. Dabei gehören neun Beispiele der Klasse „Tennis“ an und fünf der Klasse „Kein Tennis“. So ergibt sich für die Entropie der Gesamtmenge: $Entropie(S) = 0,940$ (siehe oben). Das Attribut „Wind“ kann zwei verschiedene Werte („schwach“ und „stark“) annehmen. Durch eine Aufteilung entlang dieses Attributs erhält man die beiden Teilmengen $S_{schwach} = \{T1, T3, T4, T5, T8, T9, T10, T13\}$ und $S_{stark} = \{T2, T6, T7, T11, T12, T14\}$. Da die Trainingsbeispiele in S_{stark} genau zur Hälfte den Klassen „Tennis“ und „Kein Tennis“ zugeordnet sind, ist $Entropie(S_{stark}) = 1$ (siehe oben). In $S_{schwach}$ liegt das Verhältnis von „Tennis“ zu „Kein Tennis“ bei 6:2. Hieraus berechnet sich für die Entropie: $Entropie(S_{schwach}) = -\frac{6}{8} \cdot log_2(\frac{6}{8}) - \frac{2}{8} \cdot log_2(\frac{2}{8}) = 0,811$. Daraus folgt für den Informationsgewinn dieses Attributs: $Gewinn(S, Wind) = Entropie(S) - (\frac{|S_{schwach}|}{|S|} \cdot Entropie(S_{schwach}) + \frac{|S_{stark}|}{|S|} \cdot Entropie(S_{stark})) = 0,940 - (\frac{8}{14} \cdot 0,811 + \frac{6}{14} \cdot 1) = 0,048$. Der Informationsgewinn der anderen Attribute beträgt $Gewinn(S, Himmel) = 0,246$ und $Gewinn(S, Luftfeuchtigkeit) = 0,151$. Da das Attribut „Himmel“ den größten Informationsgewinn besitzt, wird es dem Wurzelknoten zugeordnet. Da es die Werte „Sonne“, „Wolken“ und „Regen“ annehmen kann, wird für jedes dieser Werte eine Kante angelegt, die auf einen Kindknoten zeigt (vgl. Abb. A.2). Durch das Attribut wird die Menge S in die Teilmengen S_{Sonne} , S_{Wolken} und S_{Regen} zerlegt. Jede Teilmenge wird entlang der Kanten dem entsprechenden Kindknoten zugeordnet. Da S_{Wolken} nur Beispiele der Klasse „Tennis“ enthält, wird der zugehörige Kindknoten zum Blatt. An jedem übrigen Kindknoten wiederholt sich der Prozess, wobei das Attribut mit dem höchsten Informationsgewinn bezüglich der dem Knoten zugeordneten Teilmenge bestimmt wird.

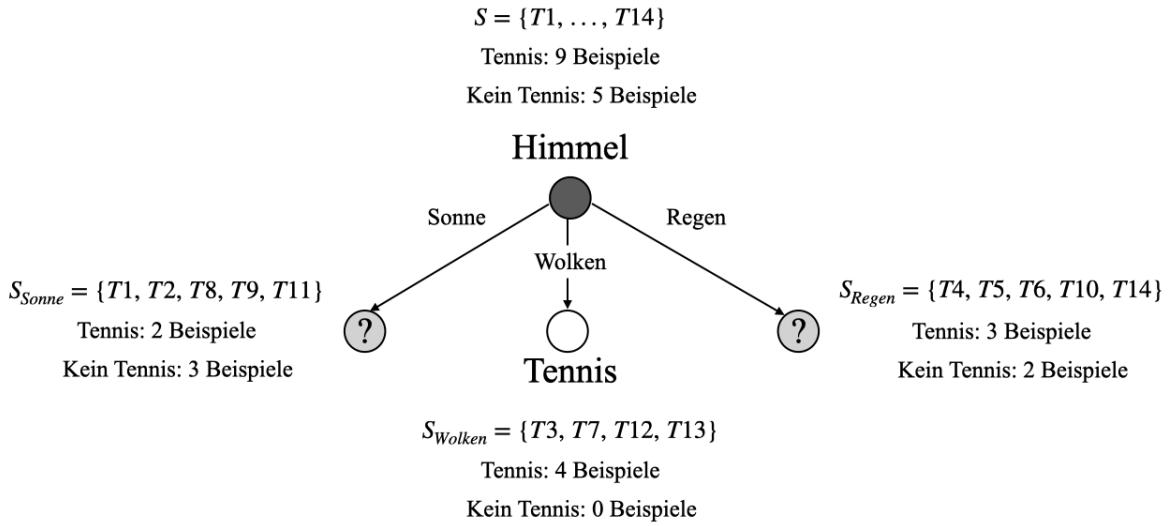


Abbildung A.2: Erster Konstruktionsschritt eines Entscheidungsbaums nach dem ID3-Algorithmus. In Anlehnung an [38], S. 61.

A.3 Herleitung Q-Funktion

In diesem Kapitel wird die Q-Funktion nach [38] hergeleitet und erläutert. Gegeben sind die Mengen $A = \{a_1, a_2, \dots, a_n\}$ und $S = \{s_1, s_2, \dots, s_k\}$, wobei A die möglichen Handlungen (engl. *actions*) des Agenten und S die Zustände (engl. *states*) seiner Umgebung beschreibt. Die Aufgabe des autonomen Agenten besteht in der Entwicklung einer Policy π für die Auswahl einer Handlung. Die Policy wird als $\pi(s) = a$ mit $a \in A$ und $s \in S$ aufgefasst, sodass jedem Zustand eine Handlung zugeordnet ist, die ausgeführt werden soll. Für jede ausgeführte Handlung erhält der Agent eine Belohnung r , wobei r abhängig von Zustand und Handlung ist ($r(s, a)$ mit $a \in A$ und $s \in S$). So werden mit jedem Schritt Belohnungen gesammelt. Diese kumulative Belohnung für einen beliebigen Anfangszustand s_t wird durch die Funktion $V(s_t) = r_t + \gamma \cdot r_{t+1} + \gamma^2 \cdot r_{t+2} + \dots = \sum_{i=0}^{\infty} \gamma^i \cdot r_{t+i}$ beschrieben. Bei γ handelt es sich um eine Konstante mit $0 \leq \gamma < 1$, die das Verhältnis zwischen sofortiger und zeitlich verzögter (zukünftiger) Belohnungen ausdrückt. Da die zu entwickelnde Policy in einem maximalen Belohnungserhalt resultieren soll, folgt daraus die Maximierung von V . Diese optimale Policy wird mit π^* gekennzeichnet und die zugehörige maximale kumulative Belohnung mit V^* .

Ein Beispiel hierfür ist in Abb. A.3 dargestellt. Wie schon im Beispiel aus Kap. 2.2.4, handelt es sich hierbei um ein Gitter, in dem jedes Feld einen anderen Zustand repräsentiert und Bewegungen nach vorne, hinten, rechts und links als Handlungen möglich sind. In Abb. A.3b sieht man, dass die kumulative Belohnung für das rot umrandete Feld 90 beträgt. Wie kommt dieser Wert zustande? Um von diesem Feld aus das Ziel mit einer optimalen Policy zu erreichen, werden zwei Bewegungen benötigt (erst nach rechts, dann nach oben oder umgekehrt). Für die erste Bewegung erhält der Agent sofort den Belohnungswert 0 (vgl. Abb. A.3a). Da er mit der zweiten Bewegung das Ziel erreicht, wird der Agent dann (zeitlich verzögert) mit 100 belohnt. Mit $\gamma = 0,9$ ergibt sich $V^* = 0 + 0,9 \cdot 100 = 90$.

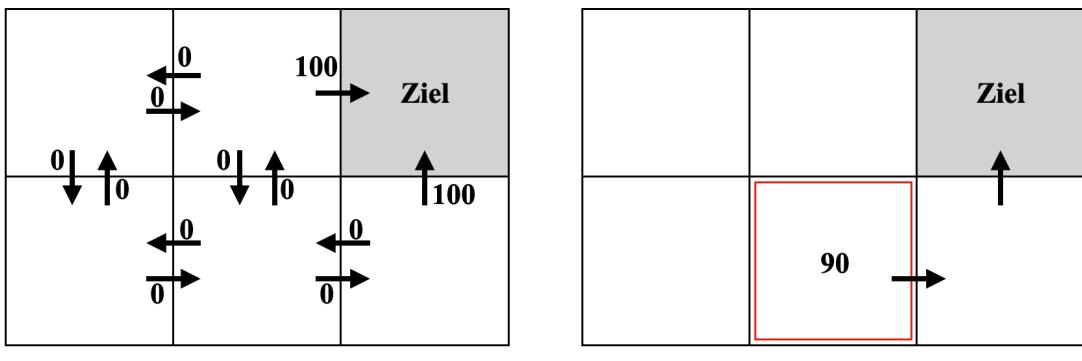


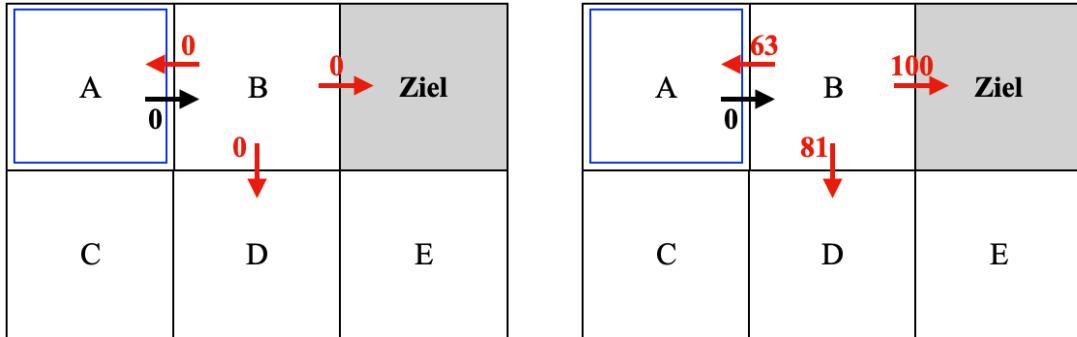
Abbildung A.3: Sofortige und kumulative Belohnung im Vergleich. In Anlehnung an [38], S. 372.

Die Q-Funktion $Q(s, a)$ wird in [38] über den Funktionswert definiert. Dieser soll die maximale kumulative Belohnung angeben, die ausgehend von einem Zustand s bei Ausführung

einer Handlung a erzielt werden kann: $Q(s, a) = r(s, a) + \gamma \cdot V^*(\delta(s, a))$ mit $\delta(s, a) = s'$. $r(s, a)$ beschreibt dabei die sofortige Belohnung, die der Agent für Ausführung der Handlung a im Zustand s erhält. Durch die Handlung geht die Umgebung in den Zustand s' über ($\delta(s, a) = s'$). Von diesem Zustand aus wird die optimale Policy ausgeführt, was durch $V^*(s')$ ausgedrückt wird. Durch den Faktor γ wird die zeitliche Verögerung berücksichtigt. Betrachtet man den Zusammenhang zwischen Q und V^* , so fällt auf, dass $V^*(s) = \max Q(s, a')$ gilt. Dieser Zusammenhang kann genutzt werden, um die Q-Funktion rekursiv zu definieren: $Q(s, a) = r(s, a) + \gamma \cdot \max Q(\delta(s, a), a')$.

Beim Q-Learning besteht die Aufgabe des Agenten darin, die Q-Funktion zu erlernen, um so eine optimale Policy zu erhalten. Hierbei wird die Q-Funktion iterativ angenähert. Eine Beschreibung des Q-Learning-Algorithmus befindet sich in Kap. 2.2.4. Im Algorithmus wird die Tabelle bei jedem Schleifendurchlauf mithilfe der Funktion $\hat{Q}(s, a) = r(s, a) + \gamma \cdot \max \hat{Q}(s', a')$ aktualisiert. Dabei ist \hat{Q} die aktuelle Schätzung der Q-Funktion, die in der Tabelle repräsentiert wird.

Das Q-Learning soll am Beispiel aus Abb. A.3a veranschaulicht werden. Der Agent wird zufällig auf einem Gitterfeld platziert. Nach dem Zufallsprinzip werden Handlungen ausgewählt, die auf der Umgebung ausgeführt werden. Solange alle \hat{Q} -Werte in der Tabelle Null sind und der Agent das Zielfeld nicht erreicht, finden keine Änderungen in der Tabelle statt (vgl. Abb. A.4a).



$$\begin{aligned}\hat{Q}(s_A, \rightarrow) &= r(s_A, \rightarrow) + \gamma \cdot \max \hat{Q}(s_B, a') \\ &= 0 + 0,9 \cdot \max \{0,0,0\} \\ &= 0 + 0,9 \cdot 0 \\ &= 0\end{aligned}$$

$$\begin{aligned}\hat{Q}(s_A, \rightarrow) &= r(s_A, \rightarrow) + \gamma \cdot \max \hat{Q}(s_B, a') \\ &= 0 + 0,9 \cdot \max \{63,81,100\} \\ &= 0 + 0,9 \cdot 100 \\ &= 90\end{aligned}$$

(a) Berechnen des neuen \hat{Q} -Wertes für Zustand A nach Initialisierung der Tabelle.

(b) Berechnen des neuen \hat{Q} -Wertes für Zustand A nach einigen Durchläufen.

Abbildung A.4: Aktualisieren des \hat{Q} -Wertes für den Zustand A. Der schwarze Pfeil gibt die sofortige Belohnung an. Die roten Pfeile repräsentieren die \hat{Q} -Werte für Zustand B. Der Agent führt eine Bewegung nach rechts aus, sodass $\delta(s_A, \rightarrow) = s_B$. In Anlehnung an [38], S. 376.

Mit dem Erreichen des Ziels erhält der Agent zum ersten Mal eine Belohnung, die größer Null ist und somit eine tatsächliche Aktualisierung der Tabelle bewirkt. Nachdem der Agent

das Ziel erreicht hat, wird der Durchgang beendet und es beginnt ein neuer Durchgang, indem der Agent erneut auf ein zufälliges Feld bewegt wird. Die aktualisierte Tabelle beeinflusst die Auswahl einer Handlung, sofern der Agent auf seinem Weg ins Ziel das entsprechende Feld betritt, was wiederum eine Aktualisierung weiterer \hat{Q} -Werte nach sich zieht. Ein Beispiel für das Aktualisieren einer Tabelle nach einigen Durchläufen zeigt Abb. A.4b.

A.4 Perzeptron-Regel

Die Perzeptron-Regel ist ein aus der Biologie motivierter Algorithmus für das Trainieren von Perzeptrons. Die sogenannte hebbische Regel besagt, dass die Verbindung zwischen zwei biologischen Neuronen dann verstärkt wird, wenn beide gleichzeitig Signale übertragen („feuern“). Diese Regel kann auf Perzeptrons transferiert werden. [20]. Hierbei findet eine Verstärkung durch die Erhöhung des Gewichts der Verbindung zweier Perzeptrons statt, wenn diese simultan aktiv sind [49].

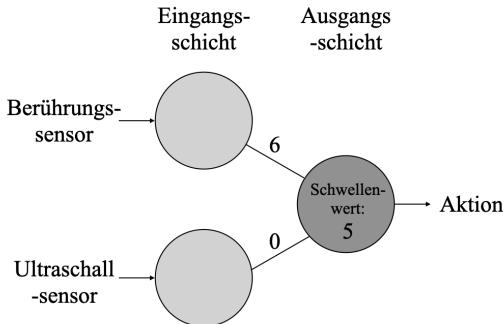
Die Perzeptron-Regel soll an einem einfachen Beispiel veranschaulicht werden. Dieser und alle nachfolgenden Absätze nehmen Bezug auf [49]. Man stelle sich einen Roboter vor, der sich in einem Raum bewegt. Der Roboter verfügt über einen Berührungssensor und einen Ultraschallsensor, die eine Wahrnehmung seiner Umgebung ermöglichen. Aktionen, die der Roboter in seiner Umgebung ausführen kann, sind „Geradeaus Fahren“ und „Umdrehen“. Seine Lernaufgabe besteht darin, umzudrehen bevor er gegen eine Wand stößt.

Das neuronale Netz, das für diese Aufgabe trainiert werden soll, ist in Abb. A.5a illustriert. Dabei werden die einzelnen Perzeptrons vereinfacht durch Knoten dargestellt. Die Eingangsschicht besteht aus je einem Eingabe-Neuron für jeden Sensor. Sie werden in der Ausgangsschicht mit einem Neuron verbunden, das feuert, wenn der Schwellenwert überschritten ist. Zwischenschichten gibt es in diesem Beispiel nicht.

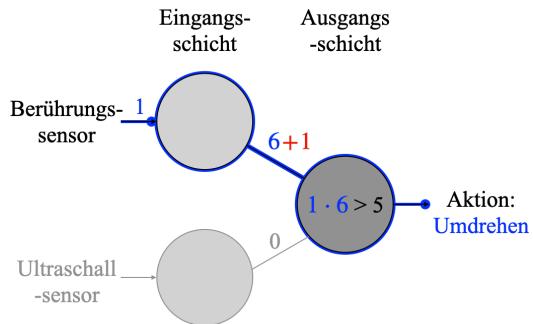
Zu Beginn lässt sich beobachten, dass der Roboter umdreht, nachdem er gegen eine Wand gefahren ist. Dies kommt dadurch zustande, dass der Berührungssensor bei Zusammenstoß mit der Wand ein Signal aussendet. Dadurch wird über das entsprechende Eingabe-Neuron eine 1 in das Netz eingegeben. Durch Multiplikation mit dem Gewicht wird der Schwellenwert überschritten und das Perzeptron gibt die Aktion „Umdrehen“ aus (vgl. A.5b). Da das Eingabe-Neuron des Berührungssensors und das Ausgabe-Neuron gemeinsam feuern, kommt hier die Perzeptron-Regel zum Tragen und das Gewicht zwischen beiden Neuronen wird erhöht.

Nach mehreren Zusammenstößen verändert der Roboter jedoch sein Verhalten. Er dreht bereits um, bevor es zu einer Kollision mit der Wand kommen kann. Dies liegt daran, dass über den Ultraschallsensor eine Eingabe in das Netz gelangt, sobald sich der Roboter einer Wand nähert. Zunächst kann der Ultraschallsensor das Verhalten des Roboters nicht beeinflussen, da das Gewicht zwischen Eingabe-Neuron und Ausgabe-Neuron zu gering ist, um den Schwellenwert zu überschreiten (vgl. Abb. A.5c), sodass das Ausgabe-Neuron nicht feuert. Bei ei-

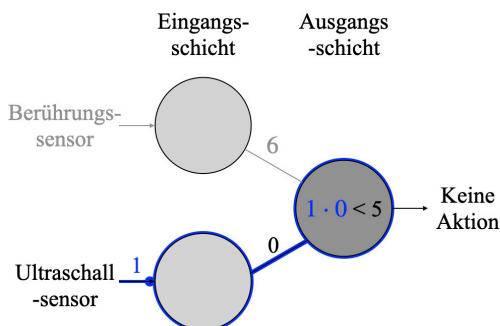
nem Zusammenstoß wird die Nähe zur Wand über den Ultraschallsensor detektiert und das zugehörige Eingabe-Neuron gibt den Wert 1 ein. Damit feuert es simultan mit den anderen beiden Perzeptrons. So wird auch das Gewicht zwischen Eingabe-Neuron des Ultraschallsensors und Ausgabe-Neuron bei einer Kollision mit der Wand gemäß der Perzeptron-Regel erhöht (vgl. Abb. A.5d). Schließlich ist das Gewicht nach mehreren Kollisionen ausreichend groß, um den Schwellenwert zu überschreiten und der Roboter dreht um, sobald er sich der Wand nähert.



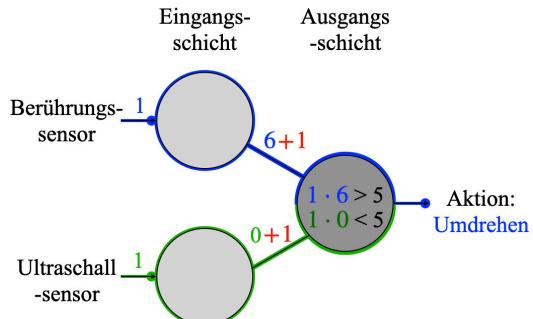
(a) Neuronales Netz, das trainiert werden soll. Die Verbindungen zwischen Eingabe-Neuronen und Ausgabe-Neuron sind unterschiedlich gewichtet.



(b) Berührungssensor löst Umdrehen des Roboters aus. Nach der Perzeptron-Regel wird das Gewicht erhöht.



(c) Das Ausgabe-Neuron feuert zunächst nicht, wenn der Ultraschallsensor ein Signal sendet.



(d) Beim Zusammenstoß sendet auch der Ultraschallsensor ein Signal. Die Gewichte beider Verbindungen zum Ausgabe-Neuron nehmen zu.

Abbildung A.5: Trainieren eines neuronalen Netzes nach der Perzeptron-Regel. In Anlehnung an [49], S. 341.

Anhang B Literatur zu maschinellen Lernalgorithmen

B.1 Überwachte Lernalgorithmen

Entscheidungsbäume

Alpaydin, E.: *Maschinelles Lernen*. De Gruyter Oldenbourg, 2. Aufl., 2019. S. 217 ff.

Aust, H.: *Das Zeitalter der Daten. Was Sie über Grundlagen, Algorithmen und Anwendungen wissen sollten*. Springer, Berlin und Heidelberg, 2021. S. 40 ff.

Géron, A.: *Praxiseinstieg Machine Learning mit ScikitLearn, Keras und TensorFlow. Konzepte, Tools und Techniken für intelligente Systeme*. O'Reilly, 2020. S. 177 ff.

Jo, T.: *Machine Learning Foundations. Supervised, Unsupervised, and Advanced Learning*. Springer, Cham, 2021. S. 141 ff.

Kossen, J., M. E. Müller und M. Ruckriegel: *Entscheidungsbäume*. In: Kersting, K., C. Lampert und C. Rothkopf (Hrsg.): *Wie Maschinen lernen. Künstliche Intelligenz verständlich erklärt*, S. 111–123. Springer, Wiesbaden, 2019.

Mitchell, T. M.: *Machine Learning*. McGraw Hill Education, internationale Aufl., 1997. S. 52 ff.

Wuttke, L.: *Praxisleitfaden für Künstliche Intelligenz in Marketing und Vertrieb. Beispiele, Konzepte und Anwendungsfälle*. Springer Gabler, Wiesbaden, 2021. S. 63.

k-Nächste Nachbarn

Jo, T.: *Machine Learning Foundations. Supervised, Unsupervised, and Advanced Learning*. Springer, Cham, 2021. S. 99 ff.

Mitchell, T. M.: *Machine Learning*. McGraw Hill Education, internationale Aufl., 1997. S. 231 ff.

Neumann, M.: *k-Nächste-Nachbarn*. In: Kersting, K., C. Lampert und C. Rothkopf (Hrsg.): *Wie Maschinen lernen. Künstliche Intelligenz verständlich erklärt*, S. 73–79. Springer, Wiesbaden, 2019.

Logistische Regression

Géron, A.: *Praxiseinstieg Machine Learning mit ScikitLearn, Keras und TensorFlow. Konzepte, Tools und Techniken für intelligente Systeme*. O'Reilly, 2020. S. 144 ff.

Schüler, T.: *Logistische Regression*. In: Kersting, K., C. Lampert und C. Rothkopf (Hrsg.): *Wie Maschinen lernen. Künstliche Intelligenz verständlich erklärt*, S. 105-110. Springer, Wiesbaden, 2019.

Naive Bayes

Jo, T.: *Machine Learning Foundations. Supervised, Unsupervised, and Advanced Learning*. Springer, Cham, 2021. S. 144 ff.

ning. Springer, Cham, 2021. S. 124 ff.

Fehrling, J. und M. Krause: *Bayesregel*. In: Kersting, K., C. Lampert und C. Rothkopf (Hrsg.): *Wie Maschinen lernen. Künstliche Intelligenz verständlich erklärt*, S. 185–194. Springer, Wiesbaden, 2019.

Mitchell, T. M.: *Machine Learning*. McGraw Hill Education, internationale Aufl., 1997. S. 154 ff.

Support Vector Machines

Aberham, J. und F. Kuruc: *Support Vector Machine*. In: Kersting, K., C. Lampert und C. Rothkopf (Hrsg.): *Wie Maschinen lernen. Künstliche Intelligenz verständlich erklärt*, S. 95–103. Springer, Wiesbaden, 2019.

Géron, A.: *Praxiseinstieg Machine Learning mit ScikitLearn, Keras und TensorFlow. Konzepte, Tools und Techniken für intelligente Systeme*. O'Reilly, 2020. S. 155 ff.

Jo, T.: *Machine Learning Foundations. Supervised, Unsupervised, and Advanced Learning*. Springer, Cham, 2021. S. 167 ff.

Lineare Regression

Aust, H.: *Das Zeitalter der Daten. Was Sie über Grundlagen, Algorithmen und Anwendungen wissen sollten*. Springer, Berlin und Heidelberg, 2021. S. 35 ff.

Kossen, J. und M. E. Müller: *Lineare Regression*. In: Kersting, K., C. Lampert und C. Rothkopf (Hrsg.): *Wie Maschinen lernen. Künstliche Intelligenz verständlich erklärt*, S. 61–67. Springer, Wiesbaden, 2019.

B.2 Unüberwachte Lernalgorithmen

Isolation Forest

Géron, A.: *Praxiseinstieg Machine Learning mit ScikitLearn, Keras und TensorFlow. Konzepte, Tools und Techniken für intelligente Systeme*. O'Reilly, 2020. S. 276.

Apriori-Algorithmus

Alpaydin, E.: *Maschinelles Lernen*. De Gruyter Oldenbourg, 2. Aufl., 2019. S. 60 ff.

Géron, A.: *Praxiseinstieg Machine Learning mit ScikitLearn, Keras und TensorFlow. Konzepte, Tools und Techniken für intelligente Systeme*. O'Reilly, 2020. S. 273.

Densitybased Clustering of Applications with Noise (DBSCAN)

Ester, M. und J. Sander: *Knowledge Discovery in Databases. Techniken und Anwendungen*. Springer, Berlin und Heidelberg, 2020. S. 72 ff.

Géron, A.: *Praxiseinstieg Machine Learning mit ScikitLearn, Keras und TensorFlow. Konzepte, Tools und Techniken für intelligente Systeme*. O'Reilly, 2020. S. 257 ff.

Erwartungsmaximierung (EM-Algorithmus)

Ester, M. und J. Sander: *Knowledge Discovery in Databases. Techniken und Anwendungen*. Springer, Berlin und Heidelberg, 2020. S. 59 ff.

Mitchell, T. M.: *Machine Learning*. McGraw Hill Education, internationale Aufl., 1997. S. 191 ff.

Jo, T.: *Machine Learning Foundations. Supervised, Unsupervised, and Advanced Learning*. Springer, Cham, 2021. S. 241 ff.

k-Means-Algorithmus

Alpaydin, E.: *Maschinelles Lernen*. De Gruyter Oldenbourg, 2. Aufl., 2019. S. 167 ff.

Aust, H.: *Das Zeitalter der Daten. Was Sie über Grundlagen, Algorithmen und Anwendungen wissen sollten*. Springer, Berlin und Heidelberg, 2021. S. 67 ff.

Ester, M. und J. Sander: *Knowledge Discovery in Databases. Techniken und Anwendungen*. Springer, Berlin und Heidelberg, 2020. S. 54.

Géron, A.: *Praxiseinstieg Machine Learning mit ScikitLearn, Keras und TensorFlow. Konzepte, Tools und Techniken für intelligente Systeme*. O'Reilly, 2020. S. 243 f.

Jo, T.: *Machine Learning Foundations. Supervised, Unsupervised, and Advanced Learning*. Springer, Cham, 2021. S. 217 ff.

Müller, D.: *k-Means-Algorithmus*. In: Kersting, K., C. Lampert und C. Rothkopf (Hrsg.): *Wie Maschinen lernen. Künstliche Intelligenz verständlich erklärt*, S. 81-88. Springer, Wiesbaden, 2019.

Wuttke, L.: *Praxisleitfaden für Künstliche Intelligenz in Marketing und Vertrieb. Beispiele, Konzepte und Anwendungsfälle*. Springer Gabler, Wiesbaden, 2021. S. 65 f.

Hauptkomponentenanalyse

Alpaydin, E.: *Maschinelles Lernen*. De Gruyter Oldenbourg, 2. Aufl., 2019. S. 125 ff.

Aust, H.: *Das Zeitalter der Daten. Was Sie über Grundlagen, Algorithmen und Anwendungen wissen sollten*. Springer, Berlin und Heidelberg, 2021. S. 65 f.

Géron, A.: *Praxiseinstieg Machine Learning mit ScikitLearn, Keras und TensorFlow. Konzepte, Tools und Techniken für intelligente Systeme*. O'Reilly, 2020. S. 221 ff.

Hölzer, C.: *Hauptkomponentenanalyse*. In: Kersting, K., C. Lampert und C. Rothkopf (Hrsg.): *Wie Maschinen lernen. Künstliche Intelligenz verständlich erklärt*, S. 125-133. Springer, Wiesbaden, 2019.

Jo, T.: *Machine Learning Foundations. Supervised, Unsupervised, and Advanced Learning*. Springer, Cham, 2021. S. 43 ff.

B.3 Algorithmen verstrkenden Lernens

Q-Learning

Geron, A.: *Praxiseinstieg Machine Learning mit ScikitLearn, Keras und TensorFlow. Konzepte, Tools und Techniken fr intelligente Systeme*. O'Reilly, 2020. S. 632 ff.

Jo, T.: *Machine Learning Foundations. Supervised, Unsupervised, and Advanced Learning*. Springer, Cham, 2021. S. 368 ff.

Mitchell, T. M.: *Machine Learning*. McGraw Hill Education, internationale Aufl., 1997. S. 373 ff.

B.4 Knstliche neuronale Netze

Perzeptron-Regel

Mitchell, T. M.: *Machine Learning*. McGraw Hill Education, internationale Aufl., 1997. S. 88 f.

Geron, A.: *Praxiseinstieg Machine Learning mit ScikitLearn, Keras und TensorFlow. Konzepte, Tools und Techniken fr intelligente Systeme*. O'Reilly, 2020. S. 288.

Delta-Regel und Gradientenabstiegsverfahren

Alpaydin, E.: *Maschinelles Lernen*. De Gruyter Oldenbourg, 2. Aufl., 2019. S. 283 ff.

Aust, H.: *Das Zeitalter der Daten. Was Sie ber Grundlagen, Algorithmen und Anwendungen wissen sollten*. Springer, Berlin und Heidelberg, 2021. S. 178 ff.

Bttcher, W., C. Bunne und J. von Stetten: *Gradientenabstiegsverfahren*. In: Kersting, K., C. Lampert und C. Rothkopf (Hrsg.): *Wie Maschinen lernen. Knstliche Intelligenz verstndlich erklrt*, S. 171-180. Springer, Wiesbaden, 2019.

Mitchell, T. M.: *Machine Learning*. McGraw Hill Education, internationale Aufl., 1997. S. 89 ff.

Anhang C Verlaufsplan und Workshop-Materialien

C.1 Verlaufsplan

Zeitrahmen	90 Minuten
Übergeordnete Idee	Im Workshop lernen die Schülerinnen und Schüler das überwachte Lernen anhand von Entscheidungsbäumen kennen.
Konzepte der Informatik	Nature of Learning (nach AI for K12)
Verortung Bildungsplan	Sekundarstufe I – Informatik ab Klasse 7
Prozessorientierte Kompetenzen	Modellieren und Implementieren, Kommunizieren und Kooperieren
Inhaltsbezogene Kompetenzen	Daten und Codierung, Algorithmen
Groblerziele	Entscheidungsbäume mithilfe tabellarischer Datensätze erstellen, testen und anwenden.
Feinlernziele	Die Schülerinnen und Schüler können... <ul style="list-style-type: none">... aus tabellarischen Datensätzen einen Entscheidungsbäum erstellen, indem sie Attribute aus der Tabelle als Knoten in der richtigen Reihenfolge anordnen und über Kanten, beschriftet mit Werten, in geeigneter Weise verbinden (TK1).... einen Entscheidungsbäum auf tabellarischen Datensätzen korrekt anwenden (TK2).... einen Entscheidungsbäum mithilfe tabellarischer Testdatensätze bewerten, indem sie den Baum korrekt auf den Testdaten anwenden und die Ergebnisse des Baumes mit der Lösung der Testdaten vergleichen (TK3).

Zeit / Phase	Lernziel	Inhalt / Thema	Methode	Transfer	Sonstiges
<i>5 Min Begrüßung</i>		Die Lehrperson begrüßt die Lernenden und stellt Workshopthema sowie Agenda vor.	Plenum, Folie 1-2		Transparenz schaffen und Orientierung geben.
<i>5 Min Einführung</i>		Die Einführung beginnt mit einer Quizfrage. Die Lehrperson führt die Begriffe „künstliche Intelligenz“, „maschinelles Lernen“ und „überwachtes Lernen“ ein.	Lehrvortrag, Ampel-Methode ¹ , Ampel-Kärtchen ² , Folie 3-9		Motivation und Aktivierung von Vorwissen. Kontextuelle Einordnung der Thematik.
<i>5 Min Problemstellung</i>		Die Lehrperson stellt die Problemstellung (Video) vor, gibt den Arbeitsauftrag, nimmt die Gruppeninteilung vor und teilt das Arbeitsblatt aus.	Lehrvortrag, Video Problemstellung, Arbeitsblatt 1, Folie 10-11		
<i>20 Min Erarbeitung</i>	Kommunikation und Kooperation der Lernenden	Die Lernenden bearbeiten das Arbeitsblatt. Während der Bearbeitung steht die Lehrperson den Schülerinnen und Schülern bei Fragen und Problemen zur Seite.	Partnerarbeit, Arbeitsblatt 1, Tablet/ Smartphone, Kopfhörer, Lösungskarten, (Bunt-)Stifte, (Lineal)		Ziel: Unterrichtsthematik erkunden und Problemstellung aus dem Video lösen.

<p><i>20 Min</i></p> <p><i>Besprechung</i></p> <p><i>Ergebnisse</i></p>	<p>Die Lehrperson ruft Lernende auf, die ihre Teillösungen vorstellen. Die Lernenden können vorgestellte Ergebnisse ergänzen oder Fragen dazu stellen. Präsentierte Algorithmus-Skizzen sowie der Sinn des Tests werden diskutiert. Die Lehrperson führt den Entscheidungsbaum sowie den Prozess vom Training zur Anwendung eines überwachten Lernalgorithmus ein.</p>	<p>Präsentation der Lernenden, Klassengespräch, Lehrvortrag, Folie 12-20</p> <p>Auf Grundlage der von den Lernenden entwickelten Skizzen, wird der Entscheidungsbaum eingeführt. Bei Durchführung der Aktivität durchlaufen die Lernenden bereits die Phasen Training, Test und Anwendung. Diese sollen nun für das überwachte Lernen verallgemeinert werden.</p>
<p><i>10 Min</i></p> <p><i>Transfer</i></p>	<p>Modellierung Entscheidungsbaum: TK1 wird adressiert</p>	<p>Die Lehrperson stellt den Lernenden eine Transferaufgabe. Die Lernenden bearbeiten die Aufgabe. Die Lösung wird gemeinsam besprochen.</p>

<p><i>10 Min Sicherung</i></p>	<p>Die Sicherung wird von der Lehrperson ammodiert. Die Lernenden erhalten von der Lehrperson den Lückentext. Sie besprechen und füllen den Lückentext gemeinsam mit der Lehrperson aus.</p>	<p>Klassengespräch, Sicherungsblatt, Stifte, Folie 24-31</p>	<p>Zusammenfassung und schriftliches Festhalten der neuen Erkenntnisse.</p>
<p><i>10 Min Übung/ Evaluation</i></p>	<p>TK1, TK2 und TK3 werden adressiert. Die Lehrperson moderiert die Phase an. Sie teilt das Arbeitsblatt aus. Die Lernenden bearbeiten das Blatt. Die Lehrperson sammelt das Blatt am Ende der Phase ein.</p>	<p>Einzelarbeit, Arbeitsblatt 2, (Bunt-)Stifte, Folie 32</p>	<p>Evaluation: Die eingesammelten Arbeitsblätter werden einer Arbeitsblatt-Analyse unterzogen.</p>
<p><i>5 Min Verabschiedung</i></p>	<p>Die Lehrperson verabschiedet sich.</p>	<p>Plenum, Folie 33</p>	

¹ Die Lernenden erhalten von der Lehrperson jeweils ein grünes, gelbes und rotes Kärtchen. Jedes Kärtchen repräsentiert eine Antwortmöglichkeit. Zuerst nennt die Lehrperson die Quizfrage, danach die Antwortmöglichkeiten mit der zugeordneten Farbe. Quizfrage und Antworten werden auf der Folie gezeigt. Die Lernenden denken kurz über die Frage nach. Dann gibt die Lehrperson ein Zeichen (z.B. Runterzählen, ein Klatschen, ein Pfiff, usw.) und alle Lernenden halten gleichzeitig das Kärtchen ihrer Wahl in die Höhe. Anschließend zeigt die Lehrperson den Lernenden die richtige Lösung auf der Folie und erläutert diese kurz. In Anlehnung an Wahl, D. Lernumgebungen erfolgreich gestalten: Vom tragen Wissen zum kompetenten Handeln, Bad Heilbrunn: Julius Klinkhardt, 2013.

² Die Kärtchen bereits vor Beginn des Workshops verteilen, um eine Unterbrechung zu vermeiden.

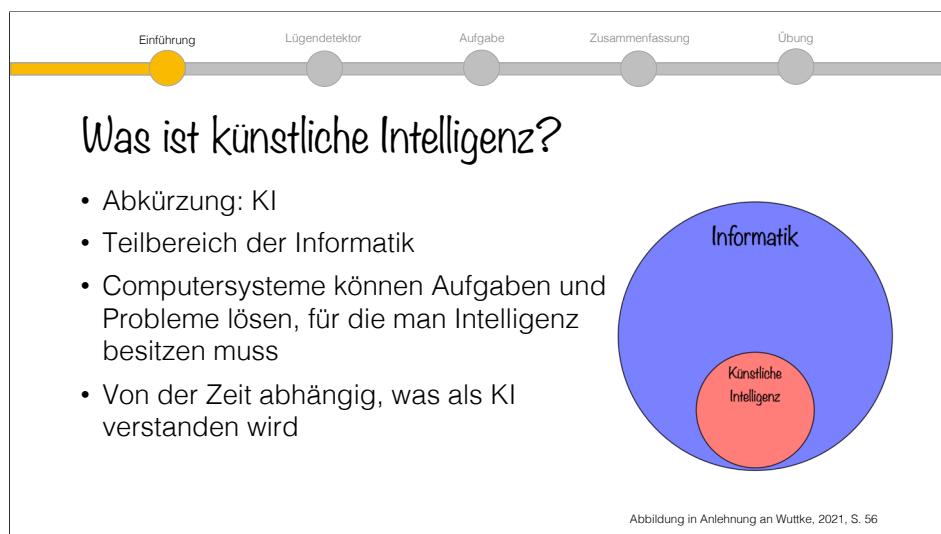
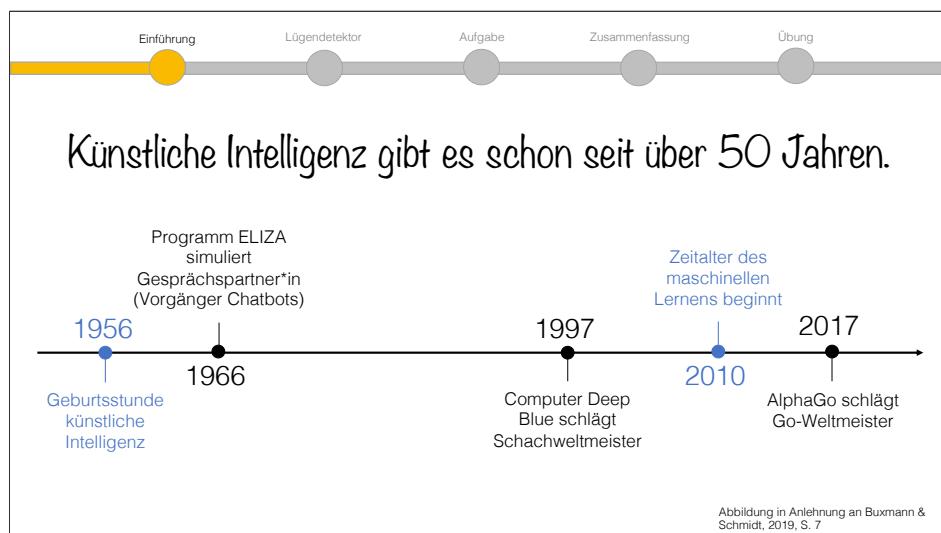
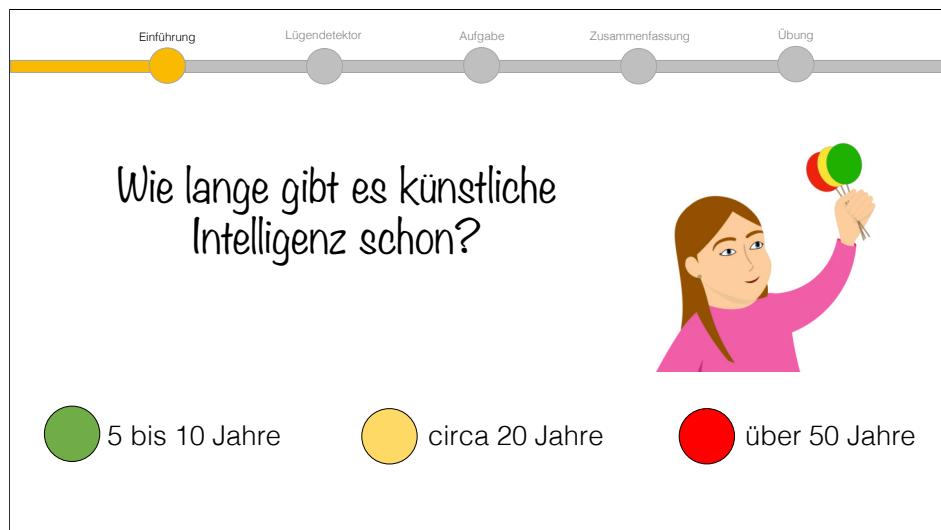
C.2 Präsentationsfolien



Agenda

- Einführung
- Der Lügendetektor
- Aufgabe
- Zusammenfassung
- Übung

A horizontal timeline diagram with five grey circles labeled "Einführung", "Lügendetektor", "Aufgabe", "Zusammenfassung", and "Übung". The first circle, "Einführung", is highlighted with a yellow background. Below the timeline, the word "Einführung" is written in a large, black, hand-drawn style font. To the right of the text is an illustration of two people, a girl with brown hair and a boy with black hair, facing each other with speech bubbles above them. They are positioned within a light green circle.



Einführung Lügendetektor Aufgabe Zusammenfassung Übung

Aufgaben künstlicher Intelligenz

Die meisten KI-Systeme basieren heutzutage auf den Methoden des **maschinellen Lernens**.

Quelle: pixabay.com

Einführung Lügendetektor Aufgabe Zusammenfassung Übung

Was ist maschinelles Lernen?

- Abkürzung: ML
- Teilbereich künstlicher Intelligenz
- Algorithmen können aus der **Erfahrung** bestimmte Verhaltensweisen erlernen (**Lernalgorithmus**)

Abbildung in Anlehnung an Wuttke, 2021, S. 56

Einführung Lügendetektor Aufgabe Zusammenfassung Übung

Lernen aus Erfahrung

- Bilder
- Texte
- Töne
- Zahlen
- ...

Überwachtes Lernen	Unüberwachtes Lernen	Verstärkendes Lernen
--------------------	----------------------	----------------------

Einführung Lügendetektor Aufgabe Zusammenfassung Übung

Der Lügendetektor

Eine Aktivität zum überwachten Lernen



 Problemstellung.mp4

Einführung Lügendetektor Aufgabe Zusammenfassung Übung

Jetzt seid ihr an der Reihe!

Helft Lisa dabei, Tims Handy wiederzufinden. Bearbeitet hierfür die Aufgaben auf dem Arbeitsblatt 1.



 20 Min  Zweiergruppe

Einführung Lügendetektor Aufgabe Zusammenfassung Übung

Ergebnisse

Aufgabe 1a)

Welche Merkmale habt ihr gefunden?

Mund:

Schülerinnen und Schüler, die lachen oder lächeln und dabei ihre Zähne zeigen, findet man nur in der Kategorie „lügt“.



Blickrichtung und Augenbraue:

Den Blick nach vorne in Kombination mit der hochgezogenen Augenbraue gibt es nur in der Kategorie „lügt“.



Einführung Lügendetektor Aufgabe Zusammenfassung Übung

Ergebnisse

Aufgabe 1b)

Welchen Algorithmus habt ihr euch überlegt?
Wie sieht eure Skizze aus?

Entscheidungsbaum

```

graph TD
    Mund((Mund)) -- "Lachen mit Zähnen" --> luegt1[lügt]
    Mund -- "Keine Zähne sichtbar" --> Blickrichtung((Blickrichtung))
    Blickrichtung -- "nach vorn" --> Augenbraue1((Augenbraue))
    Blickrichtung -- "zur Seite" --> luegt2["lügt nicht"]
    Augenbraue1 -- "hochgezogen" --> luegt3[lügt]
    Augenbraue1 -- "normal" --> luegt4["lügt nicht"]
  
```

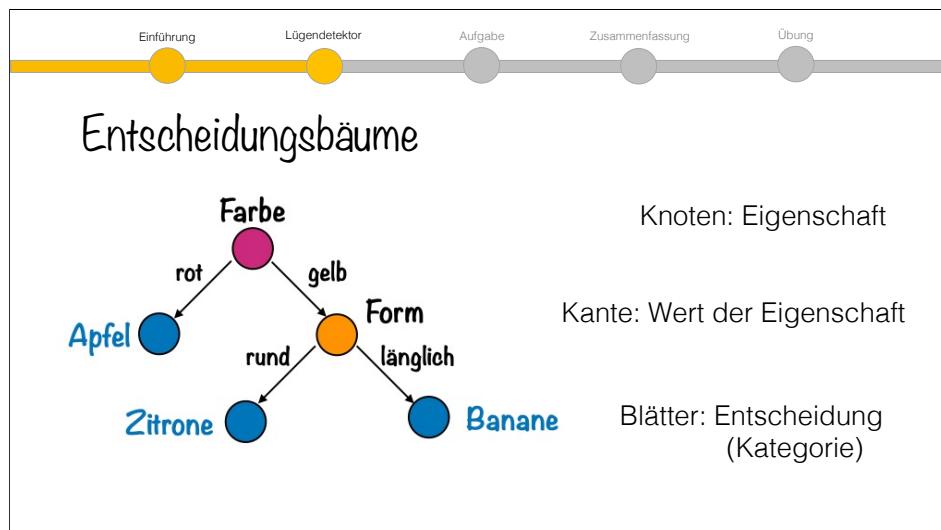
Einführung Lügendetektor Aufgabe Zusammenfassung Übung

Entscheidungsbäume

- Verfahren des überwachten Lernens
- Objekte in vorgegebene Kategorien einordnen (**Klassifikation**)
- Werden grafisch durch eine **Baumstruktur** dargestellt
- Weg zur Entscheidung **Schritt für Schritt** nachvollziehbar

Einführung Lügendetektor Aufgabe Zusammenfassung Übung

Entscheidungsbäume



Einführung Lügendetektor Aufgabe Zusammenfassung Übung

Ergebnisse

Aufgabe 1c)
Welche Personen lügen und welche nicht?

<input type="checkbox"/> lügt <input checked="" type="checkbox"/> lügt nicht	<input checked="" type="checkbox"/> lügt <input type="checkbox"/> lügt nicht	<input type="checkbox"/> lügt <input checked="" type="checkbox"/> lügt nicht

<input checked="" type="checkbox"/> lügt nicht	<input type="checkbox"/> lügt nicht	<input type="checkbox"/> lügt <input checked="" type="checkbox"/> lügt nicht

Wie ist der Test für euren Algorithmus ausgefallen?
Welchen Sinn hat der Test?

Einführung Lügendetektor Aufgabe Zusammenfassung Übung

Training, Test und Anwendung

Training: A robot examines training data (Hund, Katze) and identifies features.

- Trainingsdaten auf Merkmale untersuchen

Test: The robot tests its knowledge on test data (Hund, Katze).

- Wie viele Fehler macht der Algorithmus? (Testdaten)

Anwendung: The robot applies its learned knowledge to new data (Katz, Hund).

- Algorithmus kann auf neuen Daten mit unbekannten Lösungen angewendet werden

Beim überwachten Lernen müssen Trainings- und Testdaten **beschriftet** sein!

Quelle Tierotos: pixabay.com

Einführung Lügendetektor Aufgabe Zusammenfassung Übung

Ergebnisse

Aufgabe 1d)

Wer lügt und wer sagt die Wahrheit?

Elif	Elias	Anna

Elif, Elias und Anna sagen die Wahrheit.

Lukas	Alex	Erik

Lukas, Alex und Erik lügen.

Einführung Lügendetektor Aufgabe Zusammenfassung Übung

Ergebnisse

Aufgabe 1d)

Was ist mit Tims Handy passiert?

Anna hat Tims Handy auf der Fensterbank gesehen und wollte es ins Fundbüro bringen. Doch sie wurde abgelenkt. Stattdessen haben Lukas und Alex Tims Handy versteckt, um ihm einen Streich zu spielen. Elias hat die beiden gehört als sie damit auf dem Pausenhof angegeben haben.

Einführung Lügendetektor Aufgabe Zusammenfassung Übung

Aufgabe

Erstelle aus den Trainingsdaten einen Entscheidungsbaum, mit dem man vorhersagen kann, wann eine Schülerin/ein Schüler Nachmittagsunterricht hat.

Klasse	AG	Nachmittagsschule
7b	Französisch	Mittwoch
7a	Robotik	Montag
7a	Französisch	Dienstag
7b	Robotik	Mittwoch

5 Min

Zweiergruppe
(Sitznachbar/
Sitznachbarin)

Einführung Lügendetektor Aufgabe Zusammenfassung Übung

Lösung

Klasse	AG	Nachmittagsschule
7b	Französisch	Mittwoch
7b	Robotik	Mittwoch

In der 7b findet der Nachmittagsunterricht, unabhängig von der AG, am Mittwoch statt.

Klasse	AG	Nachmittagsschule
7a	Robotik	Montag
7a	Französisch	Dienstag

In der 7a findet der Nachmittagsunterricht, montags statt, wenn man die Robotik-AG besucht, sonst dienstags.

Einführung Lügendetektor Aufgabe Zusammenfassung Übung

Lösung

Klasse

```

graph TD
    Klasse((Klasse)) --> 7b((7b))
    Klasse --> 7a((7a))
    7b --> Mittwoch((Mittwoch))
    7a --> AG((AG))
    7a --> Franzo["Französisch"]
    7a --> Montag((Montag))
    7a --> Dienstag((Dienstag))
    AG --- Franzo
    AG --- Montag
    AG --- Dienstag
  
```

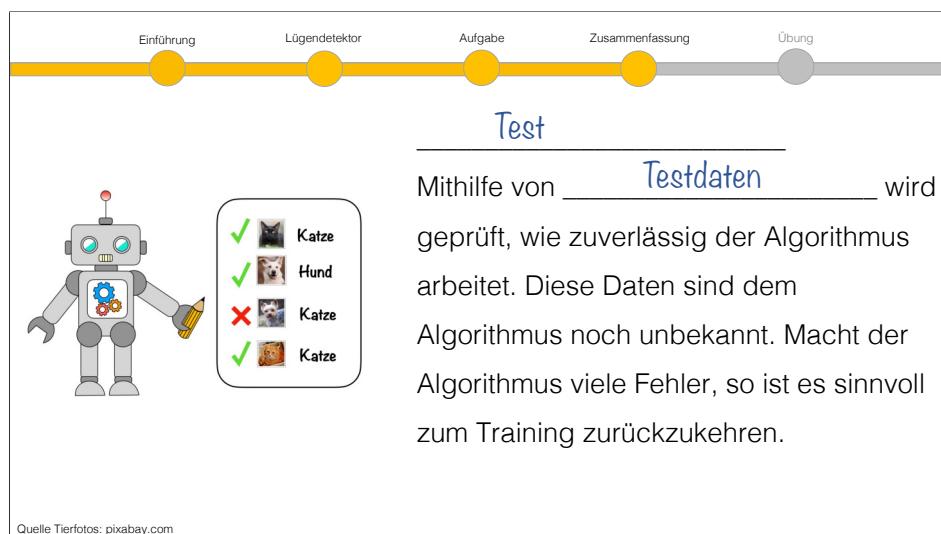
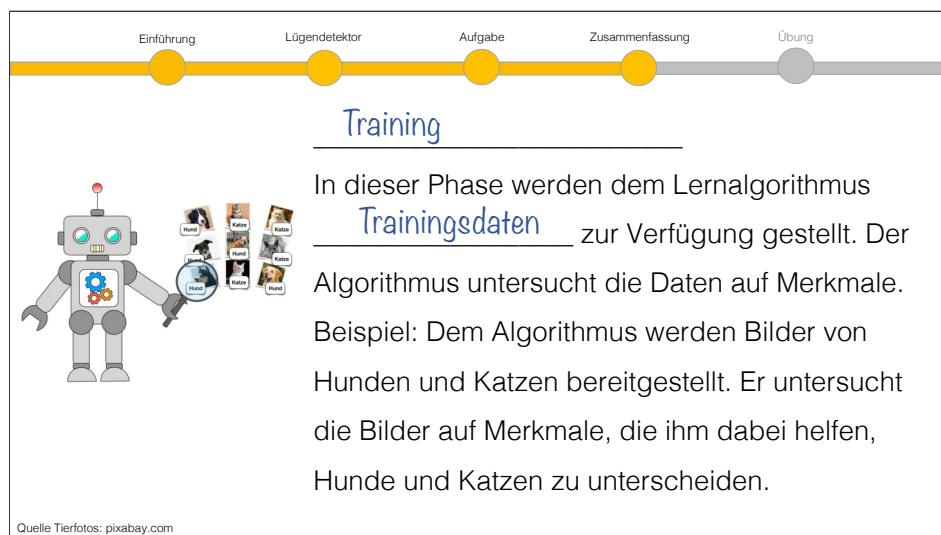
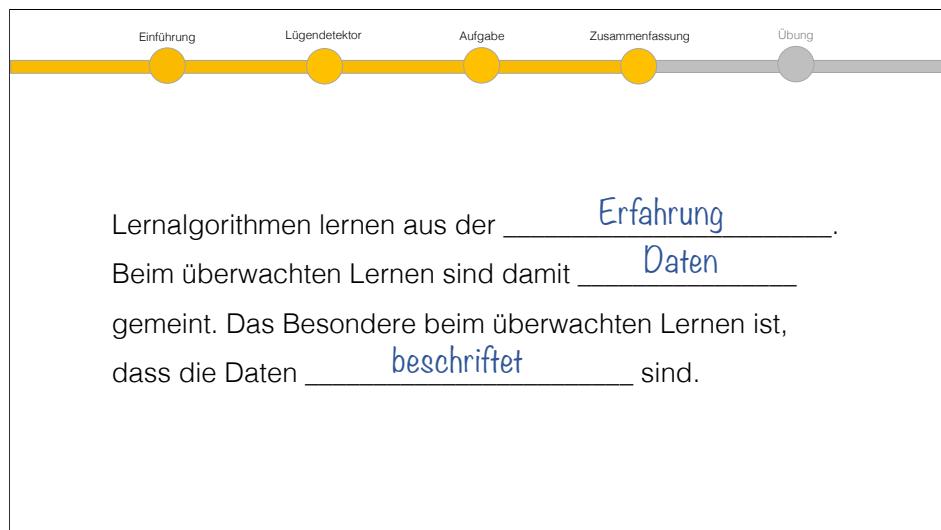
Reihenfolge der Knoten beachten!

Einführung Lügendetektor Aufgabe Zusammenfassung Übung

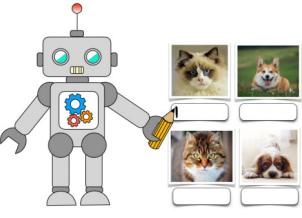
Zusammenfassung

Lückentext ausfüllen,

Hand wurde mit Ressourcen von pixabay.com erstellt.



Einführung Lügendetektor Aufgabe Zusammenfassung Übung



Anwendung

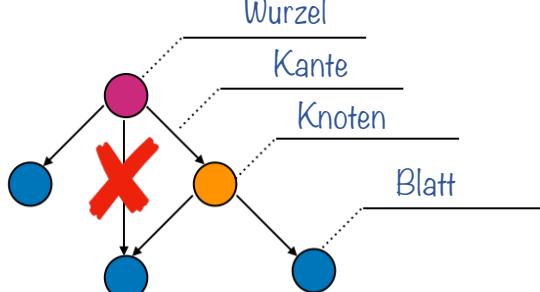
Wenn der Test gut ausgefallen ist (viele richtige Ergebnisse), kann der Lernalgorithmus auf neuen Daten mit unbekannter Beschriftung eingesetzt werden.

Quelle Tierfotos: pixabay.com

Einführung Lügendetektor Aufgabe Zusammenfassung Übung

Entscheidungsbäume sind ein Verfahren des überwachten Lernens. Man kann sie verwenden, um Objekte in vorgegebene Kategorien einzurichten (Klassifikation). Sie werden grafisch durch eine Baumstruktur dargestellt:

Einführung Lügendetektor Aufgabe Zusammenfassung Übung



Einführung Lügendetektor Aufgabe Zusammenfassung Übung

Beispiel:
 Eigenschaft: Farbe
 Wert : rot
 Entscheidung: Apfel

Einführung Lügendetektor Aufgabe Zusammenfassung Übung

Übung

Bearbeitet die Aufgaben auf dem Arbeitsblatt 2.

10 Min Einzelarbeit

Einführung Lügendetektor Aufgabe Zusammenfassung Übung

Gut gemacht!

Literaturquellen

- E. Alpaydin, *Maschinelles Lernen*, Berlin/Boston: De Gruyter Oldenbourg, 2019.
- H. Aust, *Das Zeitalter der Daten: Was Sie über Grundlagen, Algorithmen und Anwendungen wissen sollten*, Berlin und Heidelberg: Springer, 2021.
- P. Buxmann und H. Schmidt, „Grundlagen der Künstlichen Intelligenz und des Maschinellen Lernens,“ in *Künstliche Intelligenz: Mit Algorithmen zum wirtschaftlichen Erfolg*, Berlin und Heidelberg, Springer Gabler, 2021, pp. 3-25.
- A. Géron, *Praxiseinstieg Machine Learning mit Scikit-Learn, Keras und TensorFlow: Konzepte, Tools und Techniken für intelligente Systeme*, Heidelberg: O'Reilly, 2020.

Literaturquellen

- T. Jo, *Machine Learning Foundations: Supervised, Unsupervised, and Advanced Learning*, Cham: Springer, 2021.
- I. Kalder, „Eine kurze Geschichte der künstlichen Intelligenz,“ in *Wie Maschinen lernen: Künstliche Intelligenz verständlich erklärt*, Wiesbaden, Springer, 2019, pp. 135-140.
- J. Kossen, F. Kuruc und M. E. Müller, „Einleitung,“ in *Wie Maschinen lernen: Künstliche Intelligenz verständlich erklärt*, Wiesbaden, Springer, 2019, pp. 3-10.
- J. Kossen, M. E. Müller und M. Ruckriegel, „Entscheidungsbäume,“ in *Wie Maschinen lernen: Künstliche Intelligenz verständlich erklärt*, Wiesbaden, Springer, 2019, pp. 111-118.

Literaturquellen

- M. Krause und E. Natterer, „Maschinelles Lernen,“ in *Wie Maschinen lernen: Künstliche Intelligenz verständlich erklärt*, Wiesbaden, Springer, 2019, pp. 21-27.
- T. M. Mitchell, *Machine Learning*, Singapore: McGraw Hill Education, 1997.
- S. Rosebrock, *Geometrische Gruppentheorie: Ein Einstieg mit dem Computer. Basiswissen für Studium und Mathematikunterricht*, Wiesbaden: Vieweg, 2004.
- SAP, „Was ist maschinelles Lernen?,“ SAP, [Online]. Available: <https://www.sap.com/germany/insights/what-is-machine-learning.html>. [Zugriff am 30 März 2022].

Literaturquellen

- S. Seegerer, T. Michaeli und R. Romeike, „So lernen Maschinen!“, *LOG IN!*: Vol. 40, No. 1, pp. 27-31, 2020.
- L. Wuttke, Praxisleitfaden für Künstliche Intelligenz in Marketing und Vertrieb: Beispiele, Konzepte und Anwendungsfälle, Wiesbaden: Springer Gabler, 2021.
- L. Wuttke, „Training-, Validierung- und Testdatensatz,“ datasolut, 24 Juli 2020. [Online]. Available: <https://datasolut.com/wiki/trainingsdaten-und-testdaten-machine-learning/>. [Zugriff am 30 März 2022].

C.3 Video Problemstellung

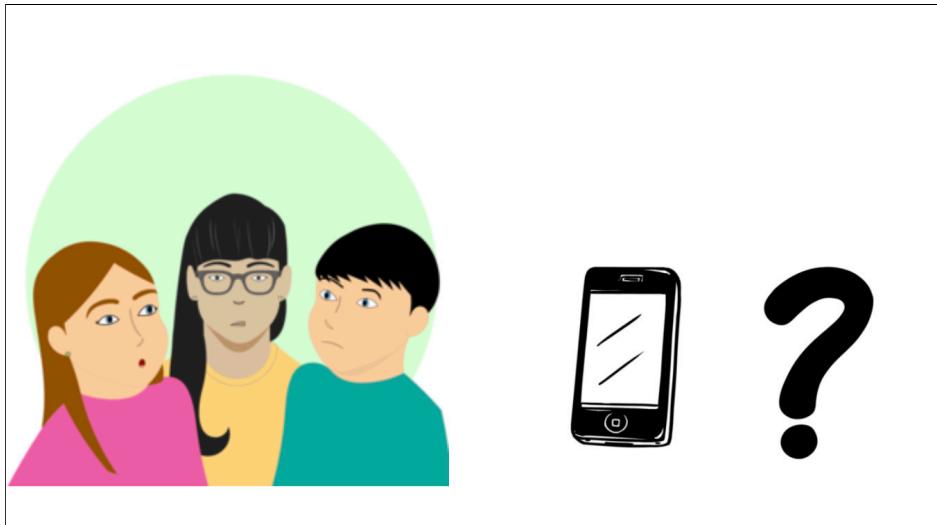


Der Lügendetektor

Sprechertext: Der Lügendetektor.



Sprechertext: Das hier sind Lisa und Tim. Die beiden besuchen die Klasse 7b. Sie sind beste Freunde.



Sprechertext: Heute ist die ganze 7b in Aufregung. Tims Handy ist verschwunden! Wurde es womöglich gestohlen?



Sprechertext: Um herauszufinden was passiert ist, möchte Lisa ihre Mitschülerinnen und Mitschüler befragen. Da sie nicht davon ausgehen kann, dass alle die Wahrheit sagen, muss sie sich etwas einfallen lassen.

überwachtes Lernen



Flaticon.com

Sprechertext: Sie beschließt mithilfe von überwachtem Lernen einen Lügendetektor zu entwickeln. Sie hat auch schon eine Idee, wie das funktionieren könnte.



**Was ist mit
Tims Handy
passiert?**

Sprechertext: Helft Lisa dabei, einen Lügendetektor zu entwickeln und Tims Handy wiederzufinden, indem ihr die Aufgaben auf dem Arbeitsblatt 1 löst.

Create your simpleshow now!



videomaker.simpleshow.com

C.4 Arbeitsblatt 1



Arbeitsblatt 1

Der Lügendetektor

Die Klasse 7b ist in Aufregung. Tims Handy ist verschwunden. Wurde es womöglich gestohlen? Lisa ist Tims beste Freundin. Darum möchte sie den Fall unbedingt lösen. Um herauszufinden, was geschehen ist, will sie ihre Mitschülerinnen und Mitschüler befragen. Da sie aber nicht davon ausgehen kann, dass alle die Wahrheit sagen, muss sie sich etwas einfallen lassen. Sie beschließt mithilfe von überwachtem Lernen einen Lügendetektor-Algorithmus zu entwickeln.



- Scannt den QR-Code mit eurem Tablet oder Smartphone ab und seht euch das Video an:



Arbeitsauftrag:

Seht euch die Schülerinnen und Schüler der Klasse 7a in Abbildung 1 genau an. Findet Merkmale, mit denen man entscheiden kann, ob eine Person in die Kategorie „lügt“ oder „lügt nicht“ eingeteilt wird.

Gefundene Merkmale:

- Nehmt die Merkmale aus Aufgabenteil a) und überlegt euch einen Algorithmus, der Lisa dabei hilft zu entscheiden, ob eine Person lügt oder nicht. Haltet den Algorithmus in einer **Skizze** fest.

Platz für eure Skizze

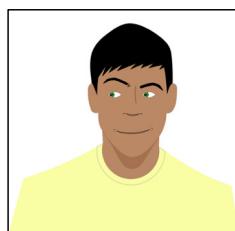
Das Icon in der Kopfzeile stammt von Flaticon.com.

- c) Scannt den QR-Code mit eurem Tablet oder Smartphone ab und seht euch das Video an:

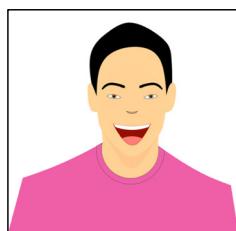


Arbeitsauftrag:

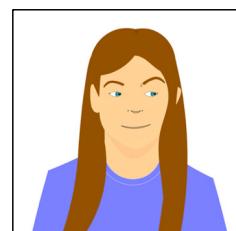
Ordnet die Schülerinnen und Schüler in Abbildung 2 mithilfe des Algorithmus in die Kategorien „lügt“ und „lügt nicht“ ein. Überprüft euer Ergebnis mit der Lösungskarte 1. Ist euer Algorithmus geeignet, einen Lügner/eine Lügnerin zu entlarven?



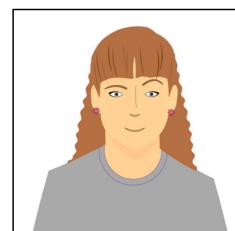
- lügt
 lügt nicht



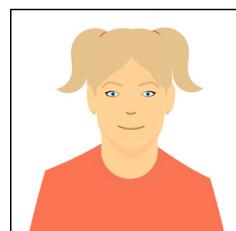
- lügt
 lügt nicht



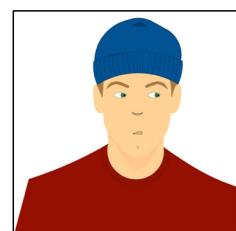
- lügt
 lügt nicht



- lügt
 lügt nicht



- lügt
 lügt nicht



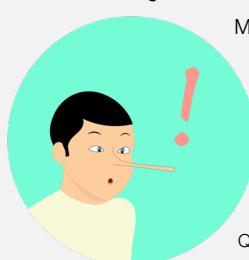
- lügt
 lügt nicht

Abbildung 2

- d) Lisa befragt nun die Schülerinnen und Schüler aus der 7b, die am „Tatort“ anwesend waren.

Verwendet euren Algorithmus, um herauszufinden, welche Schülerinnen und Schüler aus der 7b die Wahrheit sagen und welche nicht (Abbildung 3). Findet heraus, was mit Tims Handy passiert ist und schreibt es in das vorgegebene Feld.

Lügen



Menschen, die Lügen erzählen, können sich durch ihren Gesichtsausdruck oder ihre Körpersprache verraten. Aber auch körperliche Stressanzeichen wie Schwitzen oder Herzklagen können einen Lügner oder eine Lügnerin überführen. 1902 erfand der Herzspezialist James Mackenzie den *Polygraphen*, einen Lügendetektor, der die körperlichen Stressanzeichen misst. Allerdings gibt es verschiedene Möglichkeiten diesen Lügendetektor zu überlisten. Heutzutage wird maschinelles Lernen in der Forschung zur Lügenerkennung eingesetzt. Noch gibt es aber kein sicheres Mittel, um einen Lügner oder eine Lügnerin zu entlarven.

Quellen:

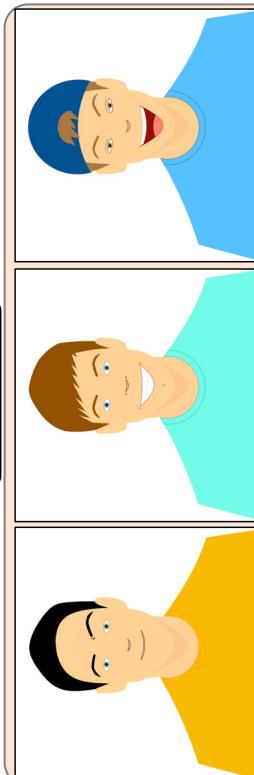
- <https://www.planet-wissen.de/gesellschaft/kommunikation/koerpersprache/pwieluegen100.html>
- <https://www.frontiersin.org/articles/10.3389/frobt.2019.00064/full>



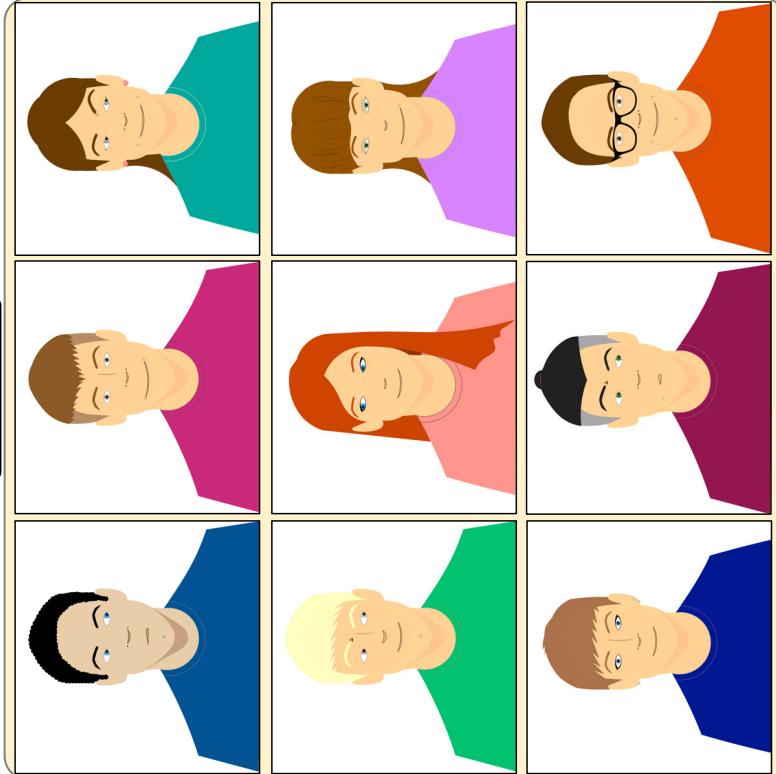
Arbeitsblatt 1

Abbildung 1

Lügt



Lügt nicht



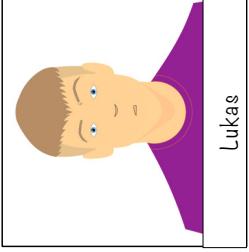
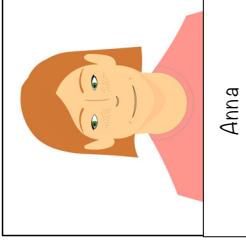
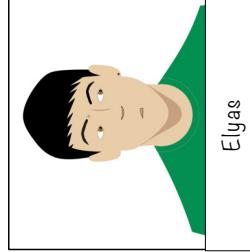
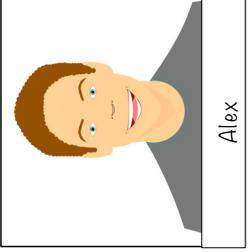
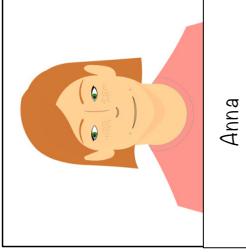
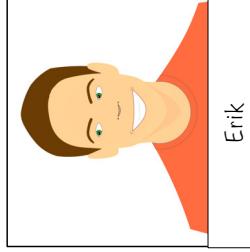
Das Icon in der Kopfzeile stammt von Flaticon.com.

Arbeitsblatt 1



Abbildung 3

Das Icon in der Kopfzeile stammt von Flaticon.com.

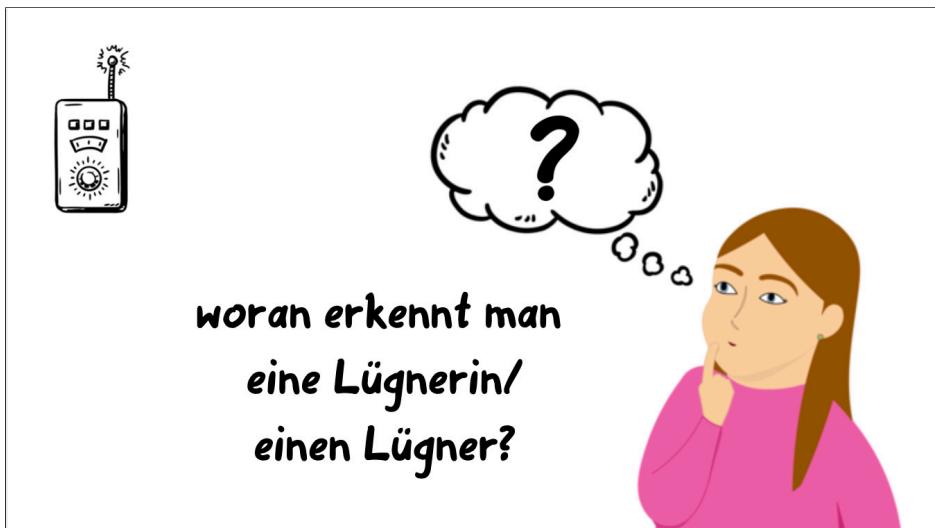
 <p>Lukas</p> <p>Ich habe das Handy zuletzt auf der Fensterbank gesehen.</p> <p><input type="checkbox"/> lügt <input type="checkbox"/> lügt nicht</p>	 <p>Anna</p> <p>Anna hat ein Handy auf der Fensterbank gefunden. Sie hat mir gesagt, dass sie es ins Fundbüro bringen möchte, weil sie nicht weiß, wem es gehört.</p> <p><input type="checkbox"/> lügt <input type="checkbox"/> lügt nicht</p>	 <p>Elias</p> <p>Vorhin haben ein paar Jungs auf dem Pausenhof damit angegeben das Handy versteckt zu haben, um Tim einen Streich zu spielen.</p> <p><input type="checkbox"/> lügt <input type="checkbox"/> lügt nicht</p>
 <p>Alex</p> <p>Anna hat das Handy von der Fensterbank genommen und in ihren Rucksack gepackt.</p> <p><input type="checkbox"/> lügt <input type="checkbox"/> lügt nicht</p>	 <p>If</p> <p>Ich wollte das Handy auf der Fensterbank ins Fundbüro bringen, aber dann wurde ich abgelenkt und habe es liegen lassen.</p> <p><input type="checkbox"/> lügt <input type="checkbox"/> lügt nicht</p>	 <p>Erik</p> <p>Lukas und Alex haben Tims Handy nicht versteckt. Elias lügt.</p> <p><input type="checkbox"/> lügt <input type="checkbox"/> lügt nicht</p>

Das ist mit Tims Handy passiert:

C.5 Verlinktes Video zu Aufgabenteil a) auf Arbeitsblatt 1



Sprechertext: Merkmale finden.



Sprechertext: Damit Lisa einen Lügendetektor entwickeln kann, muss sie ja erst einmal wissen, woran man einen Lügner oder eine Lügnerin erkennt.

Klasse 7a



Sprechertext: Um das herauszufinden, bittet sie die Klasse 7a um Hilfe. Lisa wählt zunächst 18 Schülerinnen und Schüler aus der 7a aus. Auf die restlichen sechs wird sie später noch zurückgreifen.

TRAININGSDATEN

lügt



lügt nicht



Sprechertext: Die ausgewählten Schülerinnen und Schüler werden von ihr in zwei Gruppen eingeteilt: die eine Gruppe erzählt Lisa eine Lüge, die andere Gruppe eine wahre Geschichte. Beim überwachten Lernen spricht man in diesem Zusammenhang von Trainingsdaten. Wichtig ist, dass die Trainingsdaten beschriftet sind. In unserem Fall bedeutet beschriftet, dass für jede Person bekannt ist, in welche Kategorie sie gehört.



Sprechertext: Nun kann Lisa die Trainingsdaten auf Merkmale untersuchen, die lügende Personen von nicht lügenden unterscheiden. Deshalb ist auch die Beschriftung notwendig.

ARBEITSAUFTAG

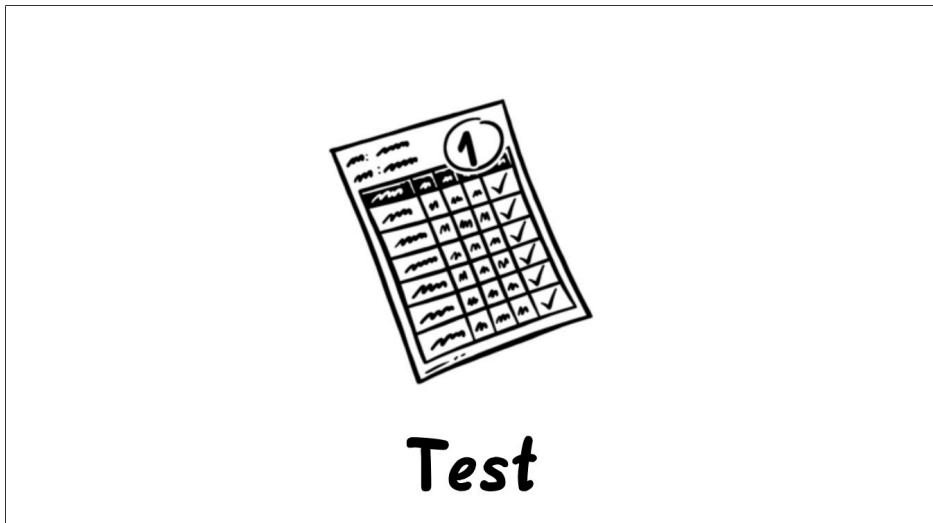
- Schüler*innen in Abb. 1 ansehen
- Merkmale finden, die lügende und nicht lügende Personen unterscheiden

Sprechertext: Arbeitsauftrag: Seht euch die Schülerinnen und Schüler der Klasse 7a in Abbildung 1 genau an. Findet Merkmale, mit denen man entscheiden kann, ob eine Person lügt oder die Wahrheit sagt.

[Create your simpleshow now!](#)

videomaker.simpleshow.com

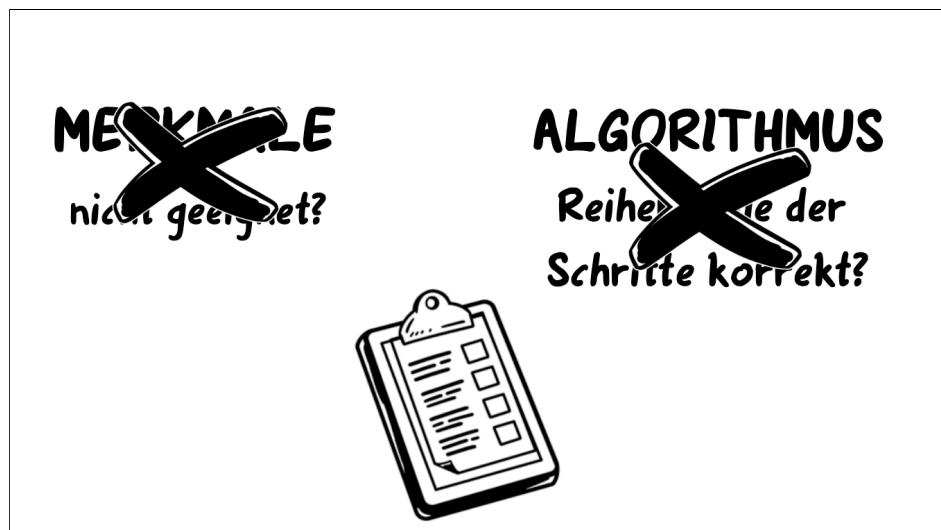
C.6 Verlinktes Video zu Aufgabenteil c) auf Arbeitsblatt 1



Sprechertext: Den Algorithmus testen.



Sprechertext: Nun da Lisa einen Algorithmus entwickelt hat, möchte sie sichergehen, dass er auch wirklich Lügnerinnen und Lügner entlarven kann.



Sprechertext: Es könnte ja sein, dass die Merkmale, die sie entdeckt hat, garnicht geeignet sind, um lügende Personen von nicht lügenden zu unterscheiden. Oder vielleicht hat sie ja einen Fehler beim Erstellen des Algorithmus gemacht? Um das auszuschließen, will sie also die Zuverlässigkeit des Algorithmus testen.



Sprechertext: Dafür zieht Lisa Beispiele heran, die dem Algorithmus noch unbekannt sind. Nun kommen die Schülerinnen und Schüler der Klasse 7a ins Spiel, die noch übrig sind. Beim überwachten Lernen werden Daten, die für den Test eingesetzt werden, als Testdaten bezeichnet.

The diagram shows six people in boxes, each with two options below them:

- Top row:
 - Person 1: lügt lügt nicht
 - Person 2: lügt lügt nicht
 - Person 3: lügt lügt nicht
- Bottom row:
 - Person 4: lügt lügt nicht
 - Person 5: lügt lügt nicht
 - Person 6: lügt lügt nicht

A black pen is shown pointing to the bottom-left box (Person 4). To the right is a blue circle with a yellow lightbulb icon. Below it is the text "Lösungskarte 1" and a downward arrow. Below the arrow is the question "Kann der Algorithmus Lügen entlarven?"

Sprechertext: Die Schülerinnen und Schüler erzählen Lisa eine Lüge oder eine wahre Geschichte. Und Lisa versucht mithilfe des Algorithmus zu bestimmen, wer lügt und wer nicht. Nachdem der Algorithmus alle Personen eingeordnet hat, sieht sie sich die Lösung an, um die Zuverlässigkeit des Algorithmus zu bewerten.

ARBEITSAUFTAG

- Bestimmen welche Schüler*innen in Abb. 2 lügen und welche nicht
- Ergebnis mit Lösungskarte 1 vergleichen
- Zuverlässigkeit bewerten

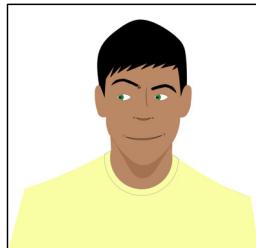
Sprechertext: Arbeitsauftrag: Bestimmt mithilfe des Algorithmus, welche Schülerinnen und Schüler in Abbildung 2 lügen und welche nicht. Überprüft eure Lösung mit Lösungskarte 1. Überlegt euch, ob euer Algorithmus zuverlässig genug arbeitet.

Create your simpleshow now!



videomaker.simpleshow.com

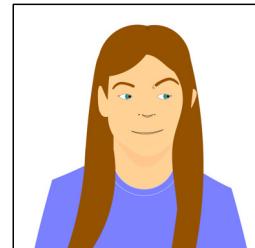
C.7 Lösungskarten zu Arbeitsblatt 1



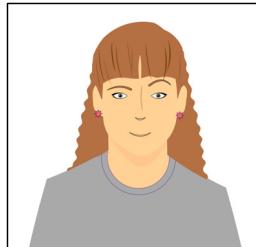
lügt
 lügt nicht



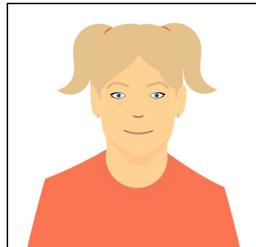
lügt
 lügt nicht



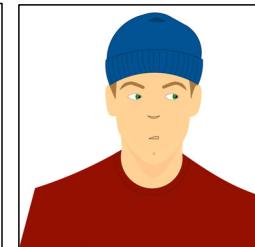
lügt
 lügt nicht



lügt
 lügt nicht



lügt
 lügt nicht



lügt
 lügt nicht

Fragen zur Lösung? Dann werft einen Blick auf Lösungskarte 2.

Merkmale:

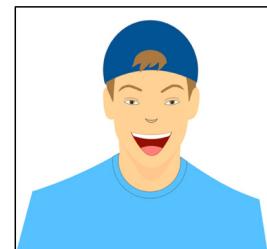
– **Mund:**

Schülerinnen und Schüler, die lachen oder lächeln und dabei ihre Zähne zeigen, findet man nur in der Kategorie „lügt“.

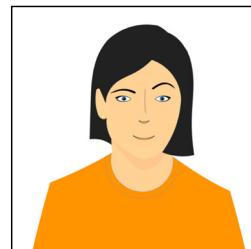
– **Blickrichtung und Augenbraue:**

Dass Schülerinnen und Schüler nach vorne blicken, ist in beiden Kategorien zu finden. Allerdings findet man die Kombination dieser Blickrichtung mit einer hochgezogenen Augenbraue nur bei den lügenden Personen.

Beispiele



Das Zeigen der Zähne beim Lachen ist ein Kriterium, um eine Person in die Klasse „lügt“ einzuteilen.



Der Blick nach vorn und die hochgezogene Augenbraue sind der Grund, warum diese Schülerin zur Klasse „lügt“ gehört.



Lösungskarte 1



Lösungskarte 2

C.8 Lösung zu Arbeitsblatt 1



Lösung zu Arbeitsblatt 1

Der Lügendetektor

Die Klasse 7b ist in Aufregung. Tims Handy ist verschwunden. Wurde es womöglich gestohlen? Lisa ist Tims beste Freundin. Darum möchte sie den Fall unbedingt lösen. Um herauszufinden, was geschehen ist, will sie ihre Mitschülerinnen und Mitschüler befragen. Da sie aber nicht davon ausgehen kann, dass alle die Wahrheit sagen, muss sie sich etwas einfallen lassen. Sie beschließt mithilfe von maschinellem Lernen einen Lügendetektor-Algorithmus zu entwickeln.

- a) Scannt den QR-Code mit eurem Tablet oder Smartphone ab und seht euch das Video an:



Arbeitsauftrag:

Seht euch die Schülerinnen und Schüler der Klasse 7a in Abbildung 1 genau an. Mithilfe welcher Merkmale kann man entscheiden, ob eine Person in die Kategorie „lügt“ oder „lügt nicht“ eingeteilt wird?



Lösung

Die Schülerinnen und Schüler nehmen die Rolle eines Lernalgorithmus ein und begeben sich auf die Suche nach Merkmalen in den Daten, die die Klassen „lügt“ und „lügt nicht“ charakterisieren und voneinander abgrenzen. Die Daten sind hierbei durch die Bilder der 18 Schülerinnen und Schüler in Abbildung 1 gegeben. Das Erkennen von Mustern und Merkmalen in den Daten ist Bestandteil der Trainingsphase im maschinellen Lernprozess. Somit stellen die Daten in Abbildung 1 die Trainingsdaten dar.

Die zu extrahierenden Merkmale sind:

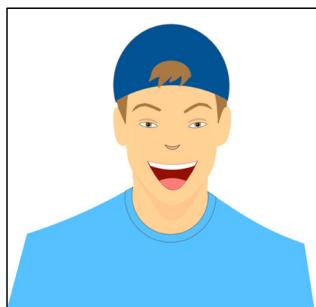
– **Mund:**

Schülerinnen und Schüler, die lachen oder lächeln und dabei ihre Zähne zeigen, findet man nur in der Klasse „lügt“.

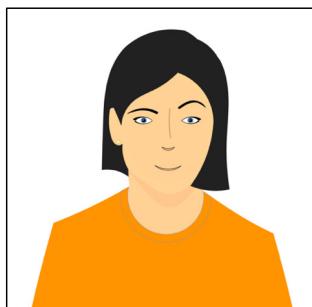
– **Blickrichtung und Augenbraue:**

Dass Schülerinnen und Schüler nach vorne blicken, ist in beiden Klassen zu finden. Allerdings findet man die Kombination dieser Blickrichtung mit einer hochgezogenen Augenbraue nur bei den lügenden Personen.

Beispiele:



Das Zeigen der Zähne beim Lachen ist ein Kriterium, um eine Person in die Klasse „lügt“ einzuteilen.



Der Blick nach vorn und die hochgezogene Augenbraue sind der Grund, warum diese Schülerin zur Klasse „lügt“ gehört.

Das Icon in der Kopfzeile wurde unter Verwendung von Ressourcen von Flaticon.com erstellt.

1



Lösung zu Arbeitsblatt 1

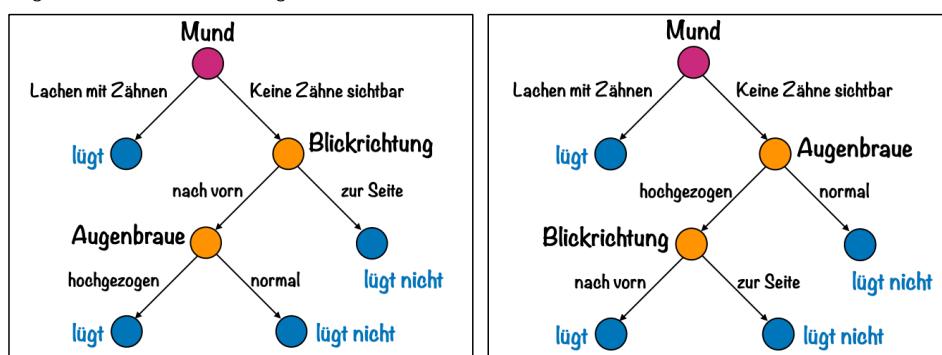
- b) Nehmt die Merkmale aus Aufgabenteil a) und überlegt euch einen Algorithmus, der Lisa dabei hilft zu entscheiden, ob eine Person lügt oder nicht. Haltet den Algorithmus in einer **Skizze** fest.

Lösung

Mit den aus Aufgabenteil a) extrahierten Merkmalen soll ein Algorithmus entwickelt werden, mit dem Personen in die Kategorien „lügt“ und „lügt nicht“ eingeteilt werden können. Der Algorithmus soll grafisch veranschaulicht werden.

Die Schülerinnen und Schüler sollen so eine eigene Darstellung entwickeln, die ggf. Elementen eines Entscheidungsbaumes nahekommen, wie etwa die Verwendung von Pfeilen. Die entwickelten Grafiken können später als Grundlage genutzt werden, um Entscheidungsbäume einzuführen.

Für den Entscheidungsbaum sind zwei Lösungen denkbar, da die Reihenfolge der Merkmale Augenbraue und Blickrichtung vertauscht werden kann:



- c) Scannet den QR-Code mit eurem Tablet oder Smartphone ab und seht euch das Video an:



Arbeitsauftrag:

Ordnet die Schülerinnen und Schüler in Abbildung 2 mithilfe des Algorithmus in die Kategorien „lügt“ und „lügt nicht“ ein. Überprüft euer Ergebnis mit der Lösungskarte 1. Ist euer Algorithmus geeignet, einen Lügner/eine Lügnerin zu entlarven?

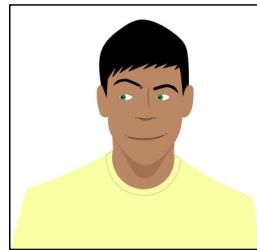
Lösung

Die Lernenden nutzen ihren Algorithmus, um zu entscheiden, ob eine Person lügt oder nicht. Dieser Aufgabenteil entspricht der Testphase im überwachten Lernprozess. Anhand von Daten (Testdaten), deren Lösung bereits bekannt ist (überwachtes Lernen), wird der Algorithmus auf seine Zuverlässigkeit geprüft. Sollten die Testdaten falsch zugeordnet werden, dann ist das ein Hinweis darauf, dass entweder die Merkmale im Algorithmus in der falschen Reihenfolge abgefragt werden oder gar nicht erst die richtigen Merkmale in a) extrahiert wurden. Die korrekte Zuordnung der Testdaten ist auf der nächsten Seite dargestellt. Die Schülerinnen und Schüler erhalten die korrekte Lösung der Testdaten über Lösungskarte 1. Sie sollen bewerten, ob ihr Algorithmus für die Problemlösung geeignet ist. Falls die Lernenden in der Testphase nicht zu der richtigen Lösung gelangen, gibt ihnen Lösungskarte 2 Informationen und Hinweise.

Das Icon in der Kopfzeile wurde unter Verwendung von Ressourcen von Flaticon.com erstellt.



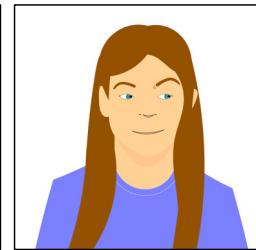
Lösung zu Arbeitsblatt 1



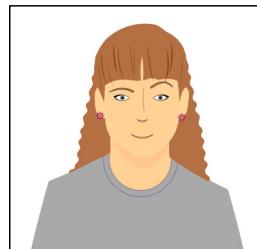
lügt
 lügt nicht



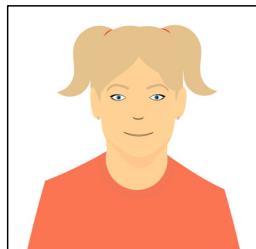
lügt
 lügt nicht



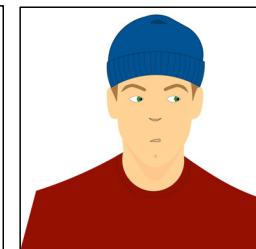
lügt
 lügt nicht



lügt
 lügt nicht



lügt
 lügt nicht



lügt
 lügt nicht

- d) Lisa befragt nun die Schülerinnen und Schüler aus der 7b, die am „Tatort“ anwesend waren.

Verwendet euren Algorithmus, um herauszufinden, welche Schülerinnen und Schüler aus der 7b die Wahrheit sagen und welche nicht (Abbildung 3). Findet heraus, was mit Tims Handy passiert ist und schreibt es in das vorgegebene Feld.

Lösung

Dieser Aufgabenteil entspricht der Anwendungsphase, die nach abgeschlossenem Training und Test erfolgt. In der Anwendungsphase wird der Algorithmus auf Daten angewandt, für die keine Lösungen vorliegen. Wie in Aufgabenteil c) verwenden die Schülerinnen und Schüler hier den Algorithmus, um die Personen zu klassifizieren. Nach der Klassifikation können sie das Rätsel um Tims Handy lösen. Es ergibt sich:

- Elif, Elyas und Anna sagen die Wahrheit.
- Lukas, Alex und Erik lügen.
- Anna hat Tims Handy auf der Fensterbank gesehen und wollte es ins Fundbüro bringen. Doch sie wurde abgelenkt. Stattdessen haben Lukas und Alex Tims Handy versteckt, um ihm einen Streich zu spielen. Elyas hat die beiden gehört als sie damit auf dem Pausenhof angegeben haben.

Das Icon in der Kopfzeile wurde unter Verwendung von Ressourcen von Flaticon.com erstellt.

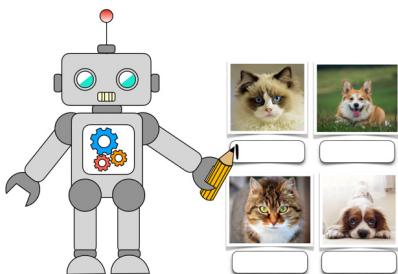
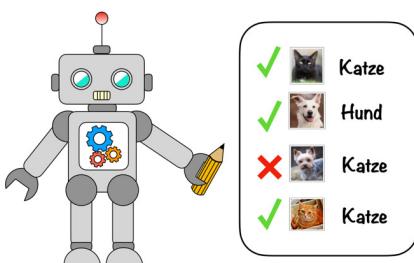
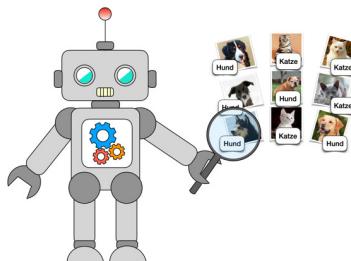
C.9 Sicherungsblatt



Unüberwachtes Lernen

Überwachtes Lernen

Lernalgorithmen lernen aus der _____. Beim überwachten Lernen sind damit _____ gemeint. Das Besondere beim überwachten Lernen ist, dass die Daten _____ sind.



In dieser Phase werden dem Lernalgorithmus _____ zur Verfügung gestellt. Der Algorithmus untersucht die Daten auf Merkmale.

Beispiel: Dem Algorithmus werden Bilder von Hunden und Katzen bereitgestellt. Er untersucht die Bilder auf Merkmale, die ihm dabei helfen, Hunde und Katzen zu unterscheiden.

Mithilfe von _____ wird geprüft, wie zuverlässig der Algorithmus arbeitet. Diese Daten sind dem Algorithmus noch unbekannt. Macht der Algorithmus viele Fehler, so ist es sinnvoll zum Training zurückzukehren.

Wenn der Test gut ausgefallen ist (viele richtige Ergebnisse), kann der Lernalgorithmus auf neuen Daten mit unbekannter Beschriftung eingesetzt werden.

Anwendung – Erfahrung – Testdaten – Daten – beschriftet – Trainingsdaten – Test – beschriftet – Training

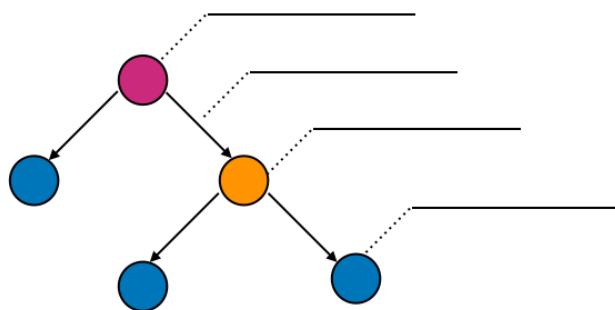
Das Icon in der Kopfzeile stammt von Flaticon.com
Quelle Tierbilder: pixabay.com



Unüberwachtes Lernen

Entscheidungsbäume

Entscheidungsbäume sind ein Verfahren des _____.
_____. Man kann sie verwenden, um Objekte in vorgegebene Kategorien einzuordnen (Klassifikation). Sie werden grafisch durch eine Baumstruktur dargestellt:



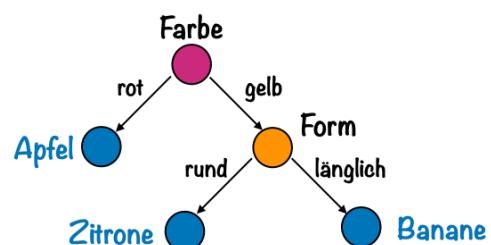
Die Blätter stellen die Entscheidungen darüber dar, in welche Kategorie ein Objekt eingeteilt wird. Jeder andere Knoten steht für eine Eigenschaft. Die Kanten, die von einem Knoten ausgehen, beschreiben jeweils eine mögliche Werte, die die zugehörige Eigenschaft annehmen kann.

Beispiel

Eigenschaft: _____

Wert: _____

Entscheidung: _____



Will man einen Baum nutzen, um eine Entscheidung zu treffen, so beginnt man an der Wurzel und läuft den Baum Knoten für Knoten ab.

Apfel – Blatt – überwachten Lernens – Knoten – Farbe – Wurzel – rot – Kante

C.10 Arbeitsblatt 2



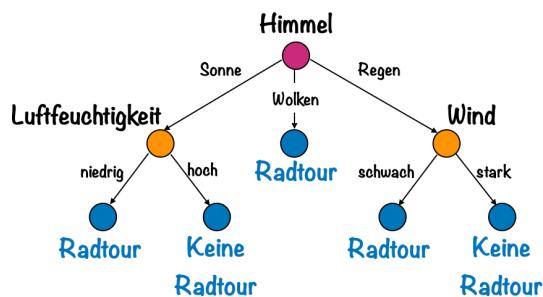
Arbeitsblatt 2

Aufgabe 1

Lisa und Tim möchten am Wochenende eine Radtour machen. Tim überprüft auf seinem Smartphone die Wettervorhersage. Welcher Tag ist für die Radtour geeignet? Markiere die Wege im Baum, die zu deiner Entscheidung geführt haben.

Lösung: Die Radtour findet am _____ statt.

	Himmel	Wind	Luftfeuchtigkeit
Samstag		stark	hoch
Sonntag		schwach	gering



Aufgabe 2

Elyas und Tim wollen sich am Freitag das Spiel ihres Lieblingsvereins ansehen und das Team anfeuern. Erstelle mithilfe der abgebildeten Trainingsdaten einen Entscheidungsbaum, mit dem man vorhersagen kann, auf welcher Tribüne ein Fan einen Platz zugewiesen bekommt.

Trainingsdaten:

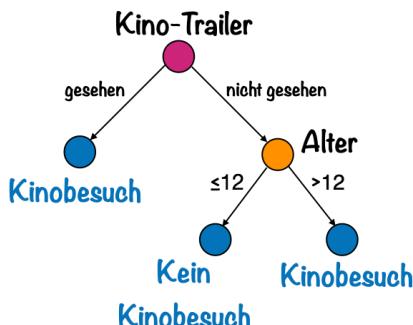
Ankunft	Fan von Verein ...	Tribüne
Nach 17 Uhr	Victoria XY-Stadt	C
Vor 17 Uhr	Fortuna ABC-Hausen	A
Nach 17 Uhr	Fortuna ABC-Hausen	C
Vor 17 Uhr	Victoria XY-Stadt	B

Zeichnung:

Aufgabe 3

Mit folgendem Baum soll vorhergesagt werden, welche Personen den neuen Superhelden-Film im Kino ansehen werden.

- a) Überprüfe den Baum mithilfe der Testdaten. Eignet sich der Baum für die Anwendung? Begründe deine Entscheidung.



Testdaten:

Alter	Kino-Trailer	Kinobesuch	Testergebnis?
13	nicht gesehen	Kein Kinobesuch	
14	gesehen	Kinobesuch	
11	gesehen	Kein Kinobesuch	

Eignet sich der Baum für die Anwendung?

Ja Nein

Begründung: _____

- b) Die unten abgebildeten Trainingsdaten wurden für den Entscheidungsbaum verwendet. Sieh dir die Trainingsdaten an und zeichne den zugehörigen Entscheidungsbaum. Wie unterscheidet er sich vom Baum in Aufgabestellung a)?

Trainingsdaten:

Alter	Kino-Trailer	Kinobesuch
9	gesehen	Kein Kinobesuch
11	gesehen	Kein Kinobesuch
12	gesehen	Kinobesuch
11	nicht gesehen	Kein Kinobesuch
13	gesehen	Kinobesuch
10	nicht gesehen	Kein Kinobesuch
12	nicht gesehen	Kein Kinobesuch
14	nicht gesehen	Kein Kinobesuch

Zeichnung:

Wie unterscheidet er sich vom Baum in Aufgabestellung a)?

C.11 Lösung zu Arbeitsblatt 2



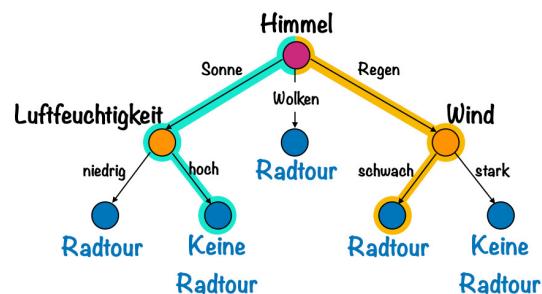
Lösung zu Arbeitsblatt 2

Aufgabe 1

Lisa und Tim möchten am Wochenende eine Radtour machen. Tim überprüft auf seinem Smartphone die Wettervorhersage. Welcher Tag ist für die Radtour geeignet? Markiere die Wege im Baum, die zu deiner Entscheidung geführt haben.

Lösung: Die Radtour findet am **Sonntag** statt.

	Himmel	Wind	Luftfeuchtigkeit
Samstag		stark	hoch
Sonntag		schwach	gering



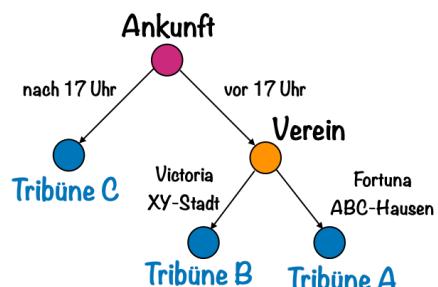
Aufgabe 2

Elyas und Tim wollen sich am Freitag das Spiel ihres Lieblingsvereins ansehen und das Team anfeuern. Erstelle mithilfe der abgebildeten Trainingsdaten einen Entscheidungsbaum, mit dem man vorhersagen kann, auf welcher Tribüne ein Fan einen Platz zugewiesen bekommt.

Trainingsdaten:

Ankunft	Fan von Verein ...	Tribüne
nach 17 Uhr	Victoria XY-Stadt	C
vor 17 Uhr	Fortuna ABC-Hausen	A
nach 17 Uhr	Fortuna ABC-Hausen	C
vor 17 Uhr	Victoria XY-Stadt	B

Zeichnung:



Das Icon in der Kopfzeile wurde unter Verwendung von Ressourcen von Flaticon.com erstellt.

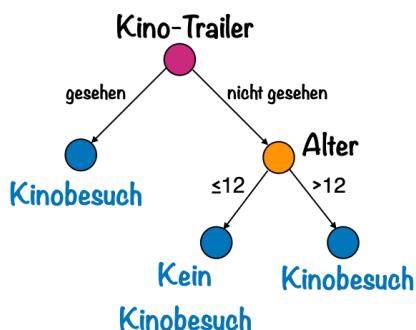


Lösung zu Arbeitsblatt 2

Aufgabe 3

Mit folgendem Baum soll vorhergesagt werden, welche Personen den neuen Superhelden-Film im Kino ansehen werden.

- a) Überprüfe den Baum mithilfe der Testdaten. Eignet sich der Baum für die Anwendung? Begründe deine Entscheidung.



Alter	Kino-Trailer	Kinobesuch	Testergebnis?
13	nicht gesehen	Kein Kinobesuch	falsch
14	gesehen	Kinobesuch	richtig
11	gesehen	Kein Kinobesuch	falsch

Eignet sich der Baum für die Anwendung?

Ja Nein

Begründung: Weil 2 von 3 Testdaten falsch zugeordnet wurden, ist die Genauigkeit des Baumes nicht gut.

- b) Die unten abgebildeten Trainingsdaten wurden für den Entscheidungsbaum verwendet. Sieh dir die Trainingsdaten an und zeichne den zugehörigen Entscheidungsbaum. Wie unterscheidet er sich vom Baum in Aufgabestellung a)?

Trainingsdaten:

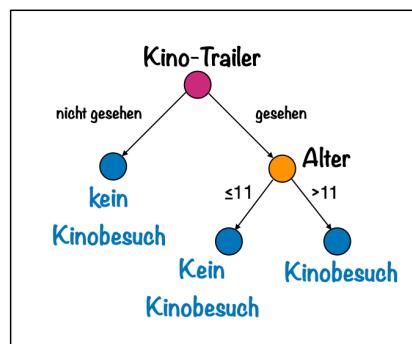
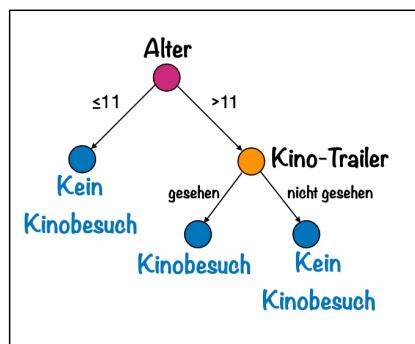
Alter	Kino-Trailer	Kinobesuch
9	gesehen	Kein Kinobesuch
11	gesehen	Kein Kinobesuch
12	gesehen	Kinobesuch
11	nicht gesehen	Kein Kinobesuch
13	gesehen	Kinobesuch
10	nicht gesehen	Kein Kinobesuch
12	nicht gesehen	Kein Kinobesuch
14	nicht gesehen	Kein Kinobesuch

Das Icon in der Kopfzeile wurde unter Verwendung von Ressourcen von Flaticon.com erstellt.



Lösung zu Arbeitsblatt 2

Zeichnung (2 mögliche Lösungen):



Wie unterscheidet er sich vom Baum in Aufgabestellung a)?

Die Altersgrenze ist falsch gesetzt. (Reihenfolge der Knoten ist vertauscht).

Bemerkung

Der Aufgabenteil b) von Aufgabe 3 ist die schwierigste Aufgabe auf dem Arbeitsblatt. Bisher mussten nur eindeutige Werte identifiziert werden, wie etwa ein Fußball-Verein, oder die Wertintervalle waren bereits vorgegeben, wie zum Beispiel in Aufgabe 2 „vor 17 Uhr“ und „nach 17 Uhr“. Bei dieser Aufgabe müssen die Lernenden die **Wertintervalle** (bis 11 Jahre und ab 12 Jahren) selbst ermitteln. Die Aufgabe muss nicht von allen Lernenden gelöst werden. Um das Erreichen der angestrebten Kompetenzen zu evaluieren, sind die Aufgaben 1 bis 3a) ausreichend. Aufgabe 3b) ist primär für leistungsstarke Lernende als Zusatzaufgabe gedacht.

C.12 Begleitbroschüre für Lehrkräfte

Wie lernen Maschinen?

Ein Workshop zur Einführung in die Konzepte
überwachten Lernens



Begleitbroschüre

Quelle Icon: Flaticon.com

Sabine Britner

Inhaltsverzeichnis

<u>EINLEITUNG</u>	2
<u>HINTERGRUND</u>	3
KÜNSTLICHE INTELLIGENZ	3
DEFINITION	3
GESCHICHTE	3
AUFGABEN	4
MASCHINELLES LERNEN	4
DEFINITION	4
LERNTYPEN	5
ÜBERWACHTES LERNEN	5
ENTSCHEIDUNGSBÄUME	7
<u>WORKSHOP</u>	9
ZIELE UND VORAUSSETZUNGEN	9
ZIELGRUPPE UND VORAUSSETZUNGEN	9
BEZUG ZUM BILDUNGSPLAN BADEN-WÜRTTEMBERG 2016	9
BEZUG ZU DEN RICHTLINIEN VON AI FOR K-12	9
LERNZIELE	10
MATERIALIEN	10
BEGLEITBROSCHÜRE	10
POWERPOINT-FOLIEN	10
VIDEOS	10
ARBEITSBLÄTTER	11
LÖSUNGSBLÄTTER	11
SICHERUNGSBLATT	11
VERLAUFSPLAN	12
<u>LITERATURVERZEICHNIS</u>	14

"We are entering a new world. The technologies of machine learning, speech recognition, and natural language understanding are reaching a nexus of capability. The end result is that we'll soon have artificially intelligent assistants to help us in every aspect of our lives." - Amy Stapleton [1]

Einleitung

Ob Bilderkennung, Sprachassistenten [2] oder Film-, Musik- und Produktempfehlungen [3] – besonders bei Jugendlichen ist heutzutage der Umgang mit künstlicher Intelligenz (Abk. KI) alltäglich. Gleichzeitig hat eine Untersuchung ergeben, dass Schülerinnen und Schüler nur wenig über die zugrunde liegenden Technologien von künstlicher Intelligenz wissen. Darüber hinaus werden diese von vielen Jugendlichen mit Gefahr oder Angst assoziiert, was auch mit dem mangelnden Wissen in Verbindung gebracht werden kann. [4]. Gerade deshalb und mit Blick auf die Entmystifikationsaufgabe des Informatikunterrichts [5] sollte die Aufklärung über die Konzepte hinter der künstlichen Intelligenz im Informatikunterricht erfolgen.

Das hier vorgestellte Workshop-Konzept richtet sich an Schülerinnen und Schüler der Sekundarstufe I (ab Klasse 7) und zielt darauf ab, sie in die grundlegenden Ideen des überwachten Lernens einzuführen. Nach einer kurzen Einführung und Verortung der Begriffe „maschinelles Lernen“ und „künstliche Intelligenz“ lernen die Jugendlichen gemeinsam mit den Hauptfiguren Lisa und Tim (siehe Abb.1) das überwachte Lernen anhand einer Unplugged-Aktivität mit Realitätsbezug kennen.

Überwachtes Lernen ist ein Typ des maschinellen Lernens (Abk. ML) [2, 6]. Das maschinelle Lernen stellt einen Teilbereich der künstlichen Intelligenz dar [2], auf den sich die Entwicklung der KI in den letzten Jahren stark fokussiert hat [7]. Aktuell basieren die meisten KI-Systeme auf Verfahren des maschinellen Lernens [3]. Deshalb ist es naheliegend Konzepte des maschinellen Lernens im Unterricht zu thematisieren.

Auf den folgenden Seiten erfahren Sie mehr über die fachlichen Hintergründe, die Lernziele und -voraussetzungen sowie die Workshop-Struktur.



Abbildung 1: Gemeinsam mit Lisa und Tim lernen Schülerinnen und Schüler das überwachte Lernen kennen. Eigene Darstellung.

"The most important general-purpose technology of our era is artificial intelligence, particularly machine learning (ML) — that is the machine's ability to keep improving its performance without humans having to explain exactly how to accomplish all the tasks it's given." – Erik Brynjolfsson & Andrew McAfee [8]

Hintergrund

Künstliche Intelligenz

Definition

Künstliche Intelligenz (engl. *artificial intelligence*) beschreibt einen Teilbereich der Informatik [2, 7], der diejenigen Computersysteme umfasst, die Aufgaben und Probleme lösen können, für die man Intelligenz benötigt [3, 9]. Diese Definition greift den Umfang der Thematik nur grob auf, was unter anderem dadurch erklärt werden kann, dass bereits die Begriffsdefinition von Intelligenz problematisch ist [7, 9]. Darüber hinaus ist das Verständnis künstlicher Intelligenz von der zeitlichen Einordnung abhängig. Während vor einigen Jahren Navigationssysteme als künstliche Intelligenz zählten, wird die Technologie heute als selbstverständlich angesehen. [3].

Man kann zwei Arten künstlicher Intelligenz unterscheiden. Mit **starker KI** sind Methoden gemeint, die versuchen Abläufe im menschlichen Gehirn nachzuahmen. Hierzu zählen Merkmale wie Empathie oder Bewusstsein. Die erfolgreiche Entwicklung einer starken KI ist aktuell nicht bekannt. Eine **schwache KI** hingegen wird gezielt für die Lösung klar definierter und abgegrenzter Problemstellungen eingesetzt. Schwache KI-Systeme sind inzwischen technisch umsetzbar. [7].

Geschichte

Bereits im Jahre 1956 wurde der Begriff der künstlichen Intelligenz eingeführt. In diesem Jahr fand die sechswöchige Konferenz „Summer Research Project on Artificial Intelligence“ am Dartmouth College in den Vereinigten Staaten statt, die als die Geburtsstunde der künstlichen Intelligenz angesehen wird. [7, 10].

Von da an ging es mit der KI-Forschung bergauf. Schnellere und günstigere Computer sowie eine Zunahme der Speicherkapazitäten von Daten ermöglichten Fortschritte auf dem Gebiet der künstlichen Intelligenz. Ein Beispiel für einen Erfolg aus dieser Zeit ist das von Joseph Weizenbaum entwickelte Programm ELIZA, das eine(n) Gesprächspartner(in) simulieren konnte. Die Erfolge resultierten jedoch in einer zu optimistischen Erwartungshaltung an die KI-Entwicklung. Das Ausbleiben der erwarteten Erfolge führte zum sogenannten (ersten) KI-Winter. [7]. In dieser Phase ging das Interesse an der KI-Forschung und ihre finanzielle Förderung zurück [9, 10].

In den 1980er Jahren fand der erste KI-Winter ein Ende und man fokussierte sich verstärkt auf die Entwicklung sogenannter Expertensysteme. Hierunter versteht man Computerprogramme, die Menschen beim Lösen spezieller Probleme unterstützen können. Allerdings stellte sich heraus, dass Expertensysteme nur in besonderen Fällen von wirtschaftlichem Nutzen waren. [9]. Und so kam es erneut zu einem KI-Winter, welcher aber nicht lange andauerte [10].

Im Jahre 1997 gelang es dem Schachcomputer Deep Blue von IBM den damaligen Schachweltmeister Garri Kasparow zu besiegen [7, 9, 10]. Dieses Ereignis erregte große öffentliche Aufmerksamkeit, besonders da Schach jahrzehntelang als ultimativer Test für künstliche Intelligenz angesehen wurde [11].

In den letzten Jahren wurde der Fokus in der KI-Forschung auf das maschinelle Lernen gelegt [7]. So bilden heutzutage Verfahren des maschinellen Lernens zumeist die Grundlage für KI [3]. Die Wissenschaftler Erik Brynjolfsson und Andrew McAfeeheben das maschinelle Lernen als wichtigste Universalttechnologie unserer Zeit hervor [8].

Aufgaben

Künstliche Intelligenz wird für unterschiedliche Aufgaben eingesetzt [2]. Dazu gehören beispielsweise Sprachverständnis, Bilderkennung, Hilfestellungen in der Medizin und autonomes Fahren [2] sowie Produkt- und Filmempfehlungen [3].

Maschinelles Lernen

Definition

Bei maschinellem Lernen (engl. *machine learning*) handelt es sich um ein Teilgebiet der künstlichen Intelligenz [2]. Die Beziehung der Bereiche Informatik, künstliche Intelligenz und maschinelles Lernen ist in Abb.2 veranschaulicht. In der Literatur werden häufig die folgenden beiden Definitionen angeführt:

- Definition nach Arthur Samuel:

Durch maschinelles Lernen können Algorithmen aus der Erfahrung bestimmte Verhaltensweisen erlernen, ohne dass eine explizite Programmierung dieser erfolgen muss [12].

- Definition nach Tom Mitchell:

Ein Computerprogramm lernt aus der Erfahrung eine bestimmte Aufgabe zu lösen, wenn sich seine Leistung in Bezug auf diese Aufgabe mit der Erfahrung verbessert [13].

Algorithmen werden beim maschinellen Lernen als **Lernalgorithmen** bezeichnet [6]. Ein Algorithmus kann beispielsweise erlernen, Hunde und Katzen auf Bildern zu erkennen, ohne dass typische Eigenschaften der Tiere beim Programmieren im Code hinterlegt werden müssen. Stattdessen lernt der Algorithmus das Erkennen der Tiere eigenständig. [7]. Damit dieses Verhalten funktionieren kann, benötigt ein Algorithmus Erfahrung.

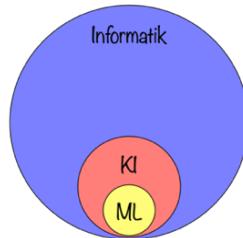


Abbildung 2: Abgrenzung und Verortung von KI und ML. Abbildung in Anlehnung an [2], S.56.

Lerntypen

Wie können Algorithmen aus der Erfahrung lernen? Beim maschinellen Lernen kann das auf zwei Arten erfolgen:

- Lernen aus Daten:

Algorithmen können aus Daten wie Bildern, Texten, Tönen oder Zahlen lernen. In diese Kategorie gehören das **überwachte** und das **unüberwachte Lernen** (engl. *supervised* und *unsupervised learning*) [9].

Lernen durch Bestrafung und Belohnung:

Lernen Algorithmen über Belohnung und Bestrafung eine Verhaltensstrategie zu optimieren, so spricht man von **verstärkendem** oder **bestärkendem Lernen** (engl. reinforcement learning) [9, 13, 14].

Überwachtes Lernen

Das überwachte Lernen wird im Wesentlichen für **Klassifikations-** und **Regressionsaufgaben** angewandt [2, 15, 16].

- Klassifikation:

Unter einer Klassifikation versteht man das Einordnen von Objekten in vorgegebene Kategorien, die **Klassen** genannt werden [15]. Ein Beispiel für eine Klassifikationsaufgabe ist in Abb.3 zu sehen. Die Murmeln werden in die Klassen „rot“ und „blau“ eingeteilt. Beispiele für Lernverfahren, die für die Klassifikation eingesetzt werden, sind Entscheidungsbäume, logistische Regression und Naive Bayes [17].

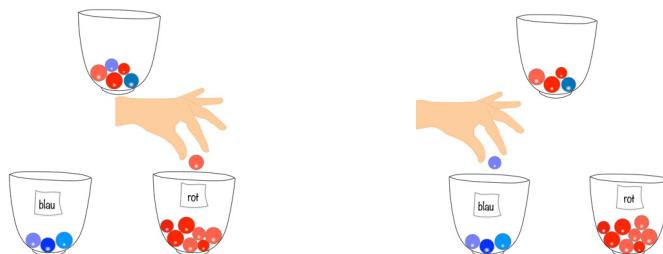


Abbildung 3: Die Murmeln werden in die beiden vorgegebenen Behälter mit der Aufschrift „rot“ bzw. „blau“ eingeordnet. Eigene Darstellung.

- Regression:

Die Regression beschäftigt sich mit der Vorhersage kontinuierlicher Zahlenwerte [2, 15]. Die vorherzusagende Größe wird als Zielgröße bezeichnet [18]. Für seine Vorhersage werden Größen aus den Eingabedaten unter der Annahme gewählt, dass sie die Zielgröße beeinflussen. Durch die Verwendung von Eingabedaten mit bekannter Lösung (Werte für Zielgröße sind bekannt), kann der Lernalgorithmus den Zusammenhang zwischen Eingabegröße und Zielgröße durch eine Funktion annähern. [16]. So kann für jeden beliebigen Wert die Zielgröße bestimmt werden [18]. Ein Beispiel ist in Abb.4 zu sehen. Mithilfe des Durchmessers einer Glas-Murmel soll das Gewicht der Murmel bestimmt werden. Das bekannteste Regressionsverfahren ist die lineare Regression [17].

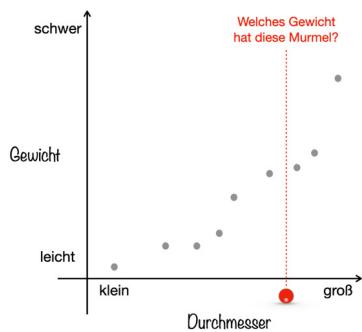


Abbildung 4a: Bekannte Murmel-Gewichte werden über dem Murmel-Durchmesser aufgetragen. In Anlehnung an [18], S. 41.

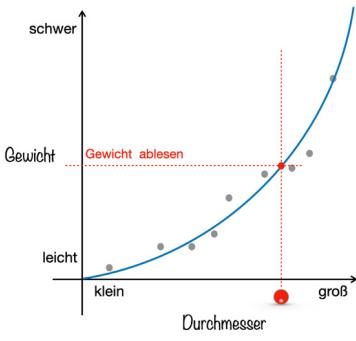


Abbildung 4b: Schnittpunkt mit Regressionskurve liefert Vorhersage für das Murmel-Gewicht. In Anlehnung an [30], S. 65.

Wie bereits geschildert, lernen Algorithmen beim überwachten Lernen aus Daten. Das überwachte Lernen zeichnet sich dadurch aus, dass die Lösung der Daten bereits bekannt ist [19, 6]. Ein Beispiel: Soll ein Algorithmus erlernen, Hunde und Katzen auf Bildern zu erkennen, so muss man dem Algorithmus Bilder zur Verfügung stellen, die mit „Hund“ oder „Katze“ beschriftet sind [7, 14]. Die Beschriftung der Bilder ist hier die bekannte Lösung. Die Beschriftungen der Daten werden als **Labels** bezeichnet [19].

Darüber hinaus werden die Daten beim überwachten Lernen in **Trainings-** und **Testdaten** unterteilt. Der Grund für diese Unterteilung wird in den folgenden Abschnitten erläutert. In der Praxis liegt das Verhältnis von Trainingsdaten zu Testdaten bei etwa 70:30. Abb.5 zeigt ein Beispiel für eine solche Aufteilung. [6].



Abbildung 5: Einteilung beschrifteter Tierbilder in Trainings- und Testdaten im Verhältnis von ca. 70:30 (Quelle Tierfotos: pixabay.com).

Wie kann man die Daten nutzen, damit ein Algorithmus ein bestimmtes Verhalten erlernt? In der ersten Phase durchläuft ein Algorithmus das sogenannte **Training**. In der Trainingsphase werden die Trainingsdaten zur Verfügung gestellt. [7]. Der Algorithmus untersucht die Daten und begibt sich auf die Suche nach Merkmalen und Mustern [6, 20]. Auf unser Beispiel bezogen: Dem Algorithmus werden Trainingsbilder von Hunden

und Katzen bereitgestellt [14]. Er untersucht die Bilder auf Merkmale, die die Klassen „Katze“ bzw. „Hund“ auszeichnen und voneinander unterscheiden. Dieser Prozess wird als das Lernen im Lernalgorithmus verstanden [7, 6].

Anschließend wird der trainierte Algorithmus einem **Test** unterzogen [7]. Anhand von Testdaten wird die Fehlerquote [6] bzw. die Treffergenauigkeit [9] des Algorithmus bestimmt. Die Treffergenauigkeit ist dabei der prozentuale Anteil der korrekt zugeordneten Testbeispiele an der Gesamtzahl der Testdaten [9]. Die Testdaten sind dem Lernalgorithmus noch unbekannt [20]. Für unser Beispiel bedeutet dies, dass der Algorithmus versucht die Tierbilder, die für den Test zurückgelegt wurden, jeweils in die richtige Klasse einzurichten. Ob eine Zuordnung richtig ist, wird über die zugehörigen Labels geprüft. Fällt der Test gut aus, kann der Algorithmus auf neuen Daten mit unbekannter Lösung angewendet werden [6, 20]. Ist man mit dem Testergebnis nicht zufrieden, so kann man zum Training zurückkehren [21]. In Abb.6 ist der Prozess vom Training zur Anwendung dargestellt. Bei der Bilderkennung wird eine Treffergenauigkeit von mindestens 90% [14] und damit eine Fehlerquote von maximal 10% angestrebt.

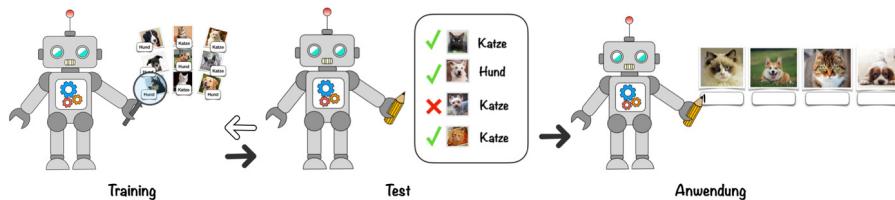


Abbildung 6: Vom Training zur Anwendung eines Lernalgorithmus (Quelle Tierfotos: pixabay.com).

Eine hohe Fehlerquote kann unterschiedliche Gründe haben. Beispielsweise kann eine zu geringe Datenmenge ursächlich für ein schlechtes Testergebnis sein. Grundsätzlich gilt, dass die Fehlerquote mit Zunahme der Trainingsdatenmenge abnimmt. In der Praxis liegt die Größenordnung im Bereich mehrerer Tausend Trainingsbeispiele. [6]. Eine weitere häufige Ursache ist die Überanpassung (engl. *overfitting*). Darunter versteht man, dass die vom Lernalgorithmus erlernten Merkmale und Muster zu stark an den Trainingsdaten orientiert und der Algorithmus somit nicht verallgemeinert werden kann. Auch muss man bei der Wahl der Trainingsdaten darauf achten, dass diese repräsentativ sind. [19]. So ist es beispielsweise nicht zielführend einem Algorithmus nur Bilder einer Hunderasse vorzuführen, wenn dieser erlernen soll Hunde auf Bildern zu erkennen.

Entscheidungsbäume

Im Workshop setzen sich Schülerinnen und Schüler mit dem überwachten Lernen auseinander. Entscheidungsbäume können schrittweise [22] und intuitiv [9] nachvollzogen werden, weshalb sie für die erste Auseinandersetzung im Workshop besonders geeignet sind.

Entscheidungsbäume werden grafisch durch eine Baumstruktur aus **Knoten** und Verzweigungen (**Kanten**) dargestellt. Der oberste Knoten in einem Baum wird als Wurzel bezeichnet. Ein Knoten führt über eine Kante zu einem weiteren Knoten. [2]. Die Baumstruktur ist in Abb.7a veranschaulicht. Im Allgemeinen erlauben es Bäume nicht,

dass mehrere Kanten zum selben Knoten führen (siehe Abb.7b) [23]. Knoten, von denen keine Kanten mehr ausgehen, werden als **Blätter** bezeichnet [2, 9].

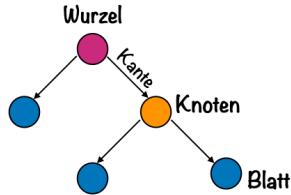


Abbildung 7a: Bestandteile einer Baumstruktur.
Eigene Darstellung.

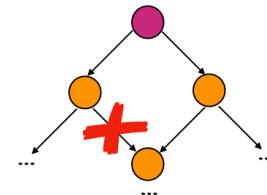


Abbildung 6b: Ein Knoten besitzt nur eine hinführende Kante. Eigene Darstellung.

Bei Klassifikationsaufgaben liefern Entscheidungsbäume Vorhersagen darüber, in welche Klasse ein Objekt eingeteilt wird. Jedes Blatt eines Baumes repräsentiert eine Klasse. Die restlichen Knoten stehen jeweils für ein **Attribut** (eine Eigenschaft [22]). Die von einem Knoten ausgehenden Kanten sind mit **Werten** oder Wertintervallen beschriftet, die ein Attribut annehmen kann. [15]. Ein Beispielbaum mit Attributen und zugehörigen Werten ist in Abb.8 dargestellt. Will man einen Baum für eine Vorhersage nutzen, beginnt man an der Wurzel und läuft den Baum Knoten für Knoten ab, bis man bei einem Blatt landet. An jedem Knoten, der kein Blatt ist, wird das zu klassifizierende Objekt auf das zugehörige Attribut geprüft. Über die vom Knoten ausgehenden Kanten wählt man den zutreffenden Wert aus. So gelangt man über die ausgewählte Kante zum nächsten Knoten. [13].

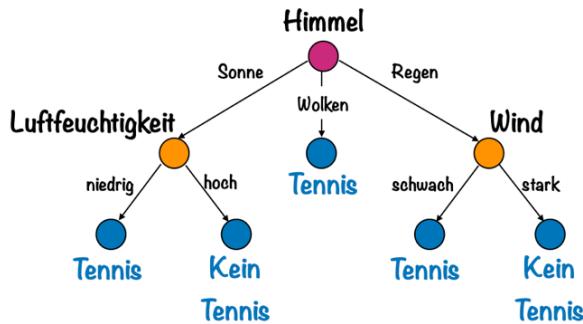


Abbildung 8: Entscheidungsbaum zur Vorhersage, ob das Wetter zum Tennisspielen geeignet ist. Attribute sind „Himmel“, „Luftfeuchtigkeit“ und „Wind“. Die Kanten sind mit den zugehörigen Werten beschriftet. Die Blätter repräsentieren die vorgegebenen Klassen „Tennis“ und „kein Tennis“. Abbildung in Anlehnung an [13], S.53.

Wie oben erwähnt, müssen beim überwachten Lernen vor der Anwendung eines Lernverfahrens Training und Test absolviert werden. Die Trainingsdaten werden dazu genutzt, Entscheidungsbäume zu erstellen. Dies ist der Lernprozess im Entscheidungsbau. [15]. Dabei spielt die Reihenfolge, in der die Merkmale der Trainingsdaten im Baum angeordnet werden, eine wichtige Rolle. Die Wahl der Position im Baum, an der auf ein bestimmtes Attribut geprüft wird, wirkt sich auf den Erfolg des Baumes hinsichtlich der Problemstellung aus. [22].

*"Artificial Intelligence, deep learning,
machine learning – whatever you're doing
if you don't understand it – learn it.
Because otherwise you're going to be a
dinosaur within 3 years." – Marc Cuban
[24]*

Workshop

Ziele und Voraussetzungen

Zielgruppe und Voraussetzungen

Der hier vorgestellte Workshop wurde für Schülerinnen und Schüler der Sekundarstufe I ab Klasse 7 konzipiert.

Um sich mit maschinellem Lernen auseinanderzusetzen zu können, wird ein Begriffsverständnis von **Algorithmen** vorausgesetzt. Das vorangegangene Kapitel zeigt auf, dass (Lern-)Algorithmen ein grundlegender Bestandteil maschinellen Lernens sind.

Bezug zum Bildungsplan Baden-Württemberg 2016 [25, 26]

Inhaltsbezogene Kompetenzen

- Daten und Codierung:
Im Rahmen des Workshops setzen sich die Schülerinnen und Schüler mit der Datenstruktur Baum am Beispiel von Entscheidungsbäumen auseinander.
- Algorithmen:
In der Erarbeitungsphase entwickeln Lernende einen Algorithmus, um die vorgegebene Problemstellung zu lösen.

Prozessbezogene Kompetenzen

- Modellieren und Implementieren:
Die Lernenden üben im Workshop aus einem gegebenen tabellarischen Datensatz einen Entscheidungsbaum modellieren.
- Kommunizieren und Kooperieren:
Im Workshop gibt es Aktivitäten, die in Zweiergruppen durchgeführt werden. Dabei lösen die Lernenden die Aufgabenstellungen gemeinsam, indem sie Lösungsideen sowie Ergebnisse diskutieren.

Bezug zu den Richtlinien von AI for K-12

AI for K-12 ist eine von der Association for the Advancement of Artificial Intelligence und Computer Science Teachers Association ins Leben gerufene Arbeitsgruppe, deren Aufgabe in der Definition von Richtlinien für die Vermittlung von künstlicher Intelligenz vom Kindergarten bis zur Klassenstufe 12 besteht. Diese Richtlinien werden durch fünf Ideen strukturiert, die der Arbeitsgruppe zufolge jeder Schülerin und jedem Schüler bekannt sein sollte, wobei sich hinter Idee Nummer drei („Computers can learn from data“) die Konzepte des maschinellen Lernens verbergen. [27].

Auf der Webseite der Arbeitsgruppe können die Richtlinien heruntergeladen werden (siehe [28]). Die Richtlinien für Idee Nummer drei sind nach Klassenstufen und Konzepten aufgeteilt. Unter dem Konzept „Nature of Learning“ ist für die Klassenstufen 6 bis 8 das Lernziel „Train and evaluate a classification or prediction model using machine learning on a tabular dataset.“ ([28], S. 3) aufgeführt. Die dem Workshop zugrunde liegenden Lernziele sind daran angelehnt.

Lernziele

Übergeordnetes Lernziel (Kernkompetenz)

Im Workshop lernen die Schülerinnen und Schüler das überwachte Lernen anhand von Entscheidungsbäumen kennen. Nach der Teilnahme können die Schülerinnen und Schüler Entscheidungsbäume erstellen, testen und anwenden.

Feinlernziele (Teilkompetenzen)

Die Schülerinnen und Schüler können ...

- ... aus tabellarischen Datensätzen einen Entscheidungsbaum erstellen, indem sie Attribute aus der Tabelle als Knoten in der richtigen Reihenfolge anordnen und über Kanten, beschriftet mit Werten, in geeigneter Weise verbinden (TK1).
- ... einen Entscheidungsbaum auf tabellarischen Datensätzen korrekt anwenden (TK2).
- ... einen Entscheidungsbaum mithilfe tabellarischer Testdatensätze bewerten, indem sie den Baum korrekt auf den Testdaten anwenden und die Ergebnisse des Baumes mit der Lösung der Testdaten vergleichen (TK3).

Materialien

Begleitbroschüre

- Begleitbroschüre.pdf

Die hier vorliegende Begleitbroschüre enthält Informationen über das Workshop-Konzept. Neben einem Kapitel, das Lehrpersonen in die relevanten fachsystematischen Hintergründe künstlicher Intelligenz und maschinellen Lernens einführt, werden hier die Lernziele und -voraussetzungen sowie wichtige Informationen zur Workshop-Durchführung präsentiert.

PowerPoint-Folien

- Folien.pptx

Zur Workshop-Durchführung werden animierte PowerPoint-Folien zur Verfügung gestellt. Die Folien sind zur besseren Orientierung mit Präsentationsnotizen versehen.

Videos

- Problemstellung.mp4
- Verlinkte Videos

Mithilfe des Videos Problemstellung.mp4 sollen die Lernenden in die Problemstellung für die Erarbeitungsphase eingeführt werden. Die Videos, die auf dem Arbeitsblatt 1 über

einen QR-Code verlinkt sind, stellen den Lernenden Informationen bereit, die sie zur Bearbeitung der Teilaufgaben benötigen.

Arbeitsblätter

- AB1.pdf
- AB2.pdf

Das Arbeitsblatt AB1.pdf wurde für die Erarbeitungsphase konzipiert. Das zweite Arbeitsblatt ist für die Übungsphase gedacht und wird zum Zwecke der Evaluation nach dem Workshop einer Arbeitsblatt-Analyse unterzogen.

Lösungsblätter

- LB1.pdf
- LB2.pdf

Die Lösungsblätter liefern die Lösungen zu den in den Arbeitsblättern beschriebenen Arbeitsaufträgen. Sie sind ausschließlich für die Lehrpersonen konzipiert. In der Erarbeitungsphase erhalten die Schülerinnen und Schüler Lösungskarten (Lösungskarten.pdf), um eine Zwischenlösung zu kontrollieren.

Sicherungsblatt

- Sicherungsblatt.pdf

Um die erarbeiteten Inhalte des Workshops schriftlich festzuhalten, gibt es ein Blatt für die Lernenden mit Lücken, das in der Sicherungsphase gemeinsam mit der Lehrperson ausgefüllt wird.

Verlaufsplan

Zeit	Lernziel	Inhalt / Thema	Methode	Transfer	Sonstiges
5 Min Begrüßung		Die Lehrperson begrüßt die Lernenden und stellt Workshopthema sowie Agenda vor.	Plenum, Folie 1-2		Transparenz schaffen und Orientierung geben.
5 Min Einführung		Die Einführung beginnt mit einer Quizfrage. Die Lehrperson führt die Begriffe „künstliche Intelligenz“, „maschinelles Lernen“ und „überwachtes Lernen“ ein.	Lehrvortrag, Ampel-Methode ¹ , Ampel-Kärtchen ² Folie 3-9		Motivation und Aktivierung von Vorwissen. Kontextuelle Einordnung der Thematik.
5 Min Problemstellung		Die Lehrperson stellt die Problemstellung (Video) vor, gibt den Arbeitsauftrag, nimmt die Gruppeninteilung vor und teilt das Arbeitsblatt aus.	Lehrvortrag, Arbeitsblatt 1, Folie 10-11,		
20 Min Erarbeitung	Kommunikation und Kooperation der Lernenden	Die Lernenden bearbeiten das Arbeitsblatt. Während der Bearbeitung steht die Lehrperson den Schülerinnen und Schülern bei Fragen und Problemen zur Seite.	Partnerarbeit, Arbeitsblatt 1, Tablet/ Smartphone, Kopfhörer, Lösungskarten, (Bunt-)Stifte, (Lineal)		Ziel: Unterrichtsthematik erkunden und Problemstellung aus dem Video lösen.
20 Min Besprechung Ergebnisse		Die Lehrperson ruft Lernende auf, die ihre Teillösungen vorstellen. Die Lernenden können vorgestellte Ergebnisse ergänzen oder Fragen dazu stellen. Präsentierte Algorithmus-Skizzen sowie der Sinn des Tests werden diskutiert.	Präsentation der Lernenden, Klassengespräch, Lehrvortrag, Folie 12-20		Auf Grundlage der von den Lernenden entwickelten Skizzen, wird der Entscheidungsbaum eingeführt. Bei

¹ Die Lernenden erhalten von der Lehrperson jeweils ein grünes, gelbes und rotes Kärtchen. Jedes Kärtchen repräsentiert eine Antwortmöglichkeit. Zuerst nennt die Lehrperson die Quizfrage, danach die Antwortmöglichkeiten mit der zugehörigen Farbe. Quizfrage und Antworten werden auf der Folie gezeigt. Die Lernenden denken kurz über die Frage nach. Dann gibt die Lehrperson ein Zeichen (z.B. Runterzählen, ein Klatschen, ein Piff, usw.) und alle Lerner den halten gleichzeitig das Kärtchen ihrer Wahl in die Höhe. Anschließend zeigt die Lehrperson den Lernenden die richtige Lösung auf der Folie und erläutert diese kurz. In Anlehnung an Wahl, D. Lernumgebungen erfolgreich gestalten: Von tragen Wissen zum kompetenten Handeln, Bad Heilbrunn: Julius Klinkhardt, 2013.

² Die Kärtchen bereits vor Beginn des Workshops verteilen, um eine Unterbrechung zu vermeiden.

		Die Lehrperson führt den Entscheidungsbaum sowie den Prozess vom Training zur Anwendung eines überwachten Lernalgorithmus ein.	Durchführung der Aktivität durchlaufen die Lernenden bereits die Phasen Training, Test und Anwendung. Diese sollen nun für das überwachte Lernen verallgemeinert werden.
10 Min Transfer	Modellierung Entscheidungsbaum: TK1 wird adressiert	Die Lehrperson stellt den Lernenden eine Transferaufgabe. Die Lernenden bearbeiten die Aufgabe. Die Lösung wird gemeinsam besprochen.	Partnerarbeit (Sitznachbarin/Sitznachbar), Klassengespräch, Papier, (Bunt-)Stifte, Folie 21-23
10 Min Sicherung		Die Sicherung wird von der Lehrperson anmoderiert. Die Lernenden erhalten von der Lehrperson den Lückentext. Sie besprechen und füllen den Lückentext gemeinsam mit der Lehrperson aus.	Klassengespräch, Sicherungsblatt, Stifte, Folie 24-31
10 Min Übung/ Evaluation	TK1, TK2 und TK3 werden adressiert	Die Lehrperson moderiert die Phase an. Sie teilt das Arbeitsblatt aus. Die Lernenden bearbeiten das Blatt. Die Lehrperson sammelt das Blatt am Ende der Phase ein.	Einzelarbeit, Arbeitsblatt 2, (Bunt-)Stifte, Folie 32
5 Min Verabschiedung		Die Lehrperson verabschiedet sich.	Folie 33

Literaturverzeichnis

- [1] A. Stapleton, „Virtual Personal Assistants: Future Gatekeeper to Your Attention?“, Opus Research, 17 November 2015. [Online]. Available: <https://opusresearch.net/wordpress/2015/11/17/virtual-personal-assistants-future-gatekeeper-to-your-attention/>. [Zugriff am 24 März 2022].
- [2] L. Wuttke, Praxisleitfaden für Künstliche Intelligenz in Marketing und Vertrieb: Beispiele, Konzepte und Anwendungsfälle, Wiesbaden: Springer Gabler, 2021.
- [3] J. Kossen, F. Kuruc und M. E. Müller, „Einleitung,“ in *Wie Maschinen lernen: Künstliche Intelligenz verständlich erklärt*, Wiesbaden, Springer, 2019, pp. 3-10.
- [4] A. Lindner, M. Berges und M. Lechner, „KI im Toaster? Schüler:innenvorstellungen zu künstlicher Intelligenz,“ in *INFOS 2021 – 19. GI-Fachtagung Informatik und Schule*, Bonn, Gesellschaft für Informatik, 2021, pp. 133-142.
- [5] U. Lautebach, „Informatik für alle! Ein Plädoyer.,“ Gesellschaft für Informatik, 1 Februar 2018. [Online]. Available: <https://gi.de/meldung/informatik-fuer-alle-ein-plaedyoer>. [Zugriff am 24 März 2022].
- [6] M. Krause und E. Natterer, „Maschinelles Lernen,“ in *Wie Maschinen lernen: Künstliche Intelligenz verständlich erklärt*, Wiesbaden, Springer, 2019, pp. 21-27.
- [7] P. Buxmann und H. Schmidt, „Grundlagen der Künstlichen Intelligenz und des Maschinellen Lernens,“ in *Künstliche Intelligenz: Mit Algorithmen zum wirtschaftlichen Erfolg*, Berlin und Heidelberg, Springer Gabler, 2021, pp. 3-25.
- [8] E. Brynjolfsson und A. McAfee, „The Business of Artificial Intelligence—What It Can and Cannot Do for Your Organization,“ Machine Learning Times, 9 Oktober 2017. [Online]. Available: <https://www.predictiveanalyticsworld.com/machinelarningtimes/business-artificial-intelligence-can-not-organization/8990/>. [Zugriff am 5 Mai 2022].
- [9] H. Aust, Das Zeitalter der Daten: Was Sie über Grundlagen, Algorithmen und Anwendungen wissen sollten, Berlin und Heidelberg: Springer, 2021.
- [10] I. Kalder, „Eine kurze Geschichte der künstlichen Intelligenz,“ in *Wie Maschinen lernen: Künstliche Intelligenz verständlich erklärt*, Wiesbaden, Springer, 2019, pp. 135-140.
- [11] A. Rogers, „What Deep Blue And AlphaGo Can Teach Us About Explainable AI,“ Forbes, 9 Mai 2019. [Online]. Available: <https://www.forbes.com/sites/forbestechcouncil/2019/05/09/what-deep-blue-and-alphago-can-teach-us-about-explainable-ai/?sh=20c1750552fd>. [Zugriff am 28 März 2022].
- [12] A. L. Samuel, „Some Studies in Machine Learning Using the Game of Checkers,“ *IBM Journal of Research and Development* 44(1/2), pp. 210-229, 1959.
- [13] T. M. Mitchell, Machine Learning, Singapore: McGraw Hill Education, 1997.
- [14] S. Seegerer, T. Michaeli und R. Romeike, „So lernen Maschinen!,“ *LOG IN: Vol. 40, No. 1*, pp. 27-31, 2020.

- [15] T. Jo, Machine Learning Foundations: Supervised, Unsupervised, and Advanced Learning, Cham: Springer, 2021.
- [16] E. Alpaydin, Maschinelles Lernen, Berlin/Boston: De Gruyter Oldenbourg, 2019.
- [17] N. Castle, „Regression vs. Classification Algorithms,“ Oracle, 8 März 2018. [Online]. Available: <https://blogs.oracle.com/ai-and-datasience/post/regression-vs-classification-algorithms>. [Zugriff am 29 März 2022].
- [18] J. Kossen und M. E. Müller, „Regression,“ in *Wie Maschinen lernen: Künstliche Intelligenz verständlich erklärt*, Wiesbaden, Springer, 2019, pp. 39-43.
- [19] A. Géron, Praxiseinstieg Machine Learning mit Scikit-Learn, Keras und TensorFlow: Konzepte, Tools und Techniken für intelligente Systeme, Heidelberg: O'Reilly, 2020.
- [20] L. Wuttke, „Training-, Validierung- und Testdatensatz,“ datasolut, 24 Juli 2020. [Online]. Available: <https://datasolut.com/wiki/trainingsdaten-und-testdaten-machine-learning/>. [Zugriff am 30 März 2022].
- [21] SAP, „Was ist maschinelles Lernen?,“ SAP, [Online]. Available: <https://www.sap.com/germany/insights/what-is-machine-learning.html>. [Zugriff am 30 März 2022].
- [22] J. Kossen, M. E. Müller und M. Ruckriegel, „Entscheidungsbäume,“ in *Wie Maschinen lernen: Künstliche Intelligenz verständlich erklärt*, Wiesbaden, Springer, 2019, pp. 111-118.
- [23] S. Rosebrock, Geometrische Gruppentheorie: Ein Einstieg mit dem Computer. Basiswissen für Studium und Mathematikunterricht, Wiesbaden: Vieweg, 2004.
- [24] Upfront Ventures, „Mark Cuban Interview by Jason Hirschhorn | Upfront Summit 2017,“ 08 Februar 2022. [Online]. Available: https://www.youtube.com/watch?v=RtZ_H_aSTCI&list=PLavJa4-6tbplIWvLGw-LphcRA_2coZuRx. [Zugriff am 04 April 2022].
- [25] Ministerium für Kultus, Jugend und Sport Baden-Württemberg, „GEMEINSAMER BILDUNGSPLAN DER SEKUNDARSTUFE I. Aufbaukurs Informatik,“ 31 Mai 2018. [Online]. Available: http://bildungsplaene-bw.de/site/bildungsplan/get/documents/lbw/export-pdf/depot-pdf/ALLG/BP2016BW_ALLG_SEK1_INF7.pdf. [Zugriff am 21 April 2022].
- [26] Ministerium für Kultus, Jugend und Sport Baden-Württemberg, „GEMEINSAMER BILDUNGSPLAN DER SEKUNDARSTUFE I. Wahlfach Informatik an der Hauptschule, Werkrealschule und Realschule,“ 28 Juli 2018. [Online]. Available: http://bildungsplaene-bw.de/site/bildungsplan/get/documents/lbw/export-pdf/depot-pdf/ALLG/BP2016BW_ALLG_SEK1_INFWF.pdf. [Zugriff am 21 April 2022].
- [27] D. Touretzky, C. Gardner-McCune, F. Martin und D. Seehorn, „Envisioning AI for K-12: What Should Every Child Know about AI?,“ *Proceedings of the AAAI Conference on Artificial Intelligence*. vol. 33, p. 9795–9799, 2019.
- [28] AI for K-12, „Draft Big Idea 3 - Progression Chart,“ 2019 November 2020. [Online]. Available: <https://ai4k12.org/wp-content/uploads/2021/01/AI4K12-Big-Idea-3-Progression-Chart-Working-Draft-of-Big-Idea-3-v.11.19.2020.pdf>. [Zugriff am 16 Mai 2022].

- [29] D. Wahl, Lernumgebungen erfolgreich gestalten: Vom tragen Wissen zum kompetenten Handeln, Bad Heilbrunn: Julius Klinkhardt, 2013.
- [30] J. Kossen und M. E. Müller, „Lineare Regression,“ in *Wie Maschinen lernen. Künstliche Intelligenz verständlich erklärt*, Wiesbaden, Springer, 2019, pp. 61-67.

Anhang D Computer Science Lesson Analysis: Fragebogen

A. Unterrichtsphasen

	trifft überhaupt nicht zu	trifft nicht zu	weder/ noch	trifft zu	trifft sehr zu
Die Unterrichtsphasen sind ersichtlich.					
Aus dem Verlaufsplan geht klar hervor, was Lehrperson und Lernende in den einzelnen Unterrichtsphasen tun.					
Die Lernenden werden zu Beginn für die Auseinandersetzung mit der Unterrichtsthematik motiviert (Engagement-Phase).					
Die Lernenden erkunden anhand praktischer Aktivitäten die Unterrichtsthematik (Exploration-Phase).					
Auf Grundlage der Erfahrungen/ Ergebnisse der Lernenden aus der Erkundungsaktivität werden für die Unterrichtsthematik relevante Begriffe und Konzepte eingeführt (Explanation-Phase).					
Die im Unterricht erworbenen Kenntnisse und Fähigkeiten werden auf neue Situationen übertragen (Elaboration-Phase).					
Das Erreichen der Kompetenzziele wird durch die Lehrperson evaluiert (Evaluation-Phase).					
Die jeweils eingeplanten Zeiträume für die Unterrichtsphasen sind angemessen.					

B. Kompetenzen (Bitte Hinweis beachten)

	trifft überhaupt nicht zu	trifft nicht zu	weder/ noch	trifft zu	trifft sehr zu
Die Kompetenzziele sind ersichtlich.					
Die Kompetenzziele sind für den Zeitrahmen angemessen.					
Aus dem Verlaufsplan geht klar hervor, an welchen Stellen die Kompetenzen adressiert werden.					
Die Kompetenzziele stehen in Bezug zu Thema und Inhalt der Unterrichtsstunde.					

Die Kompetenzziele besitzen die Komponente...	trifft überhaupt nicht zu	trifft nicht zu	weder/ noch	trifft zu	trifft sehr zu
... Bedingung.					
... Handlung.					
... Inhalt.					
... Kriterium.					

Hinweis:

Ein Lernziel setzt sich aus fünf Komponenten zusammen: Wer, Bedingung, Handlung (vgl. Operatoren im Bildungsplan¹ oder Bloom Taxonomie), Inhalt und Kriterium. Beispiel: „Mit Hilfe von Venn-Diagrammen sind die Lernenden in der Lage, die Wahrscheinlichkeit eines Ereignisses korrekt zu bestimmen“. Wer = die Lernenden; Bedingung = Mit Hilfe von Venn-Diagrammen; Handlung = bestimmen; Inhalt = Wahrscheinlichkeit eines Ereignisses; Kriterium = korrekt;²

¹https://bildungsplaene-bw.de/_Lde/LS/BP2016BW/ALLG/SEK1/INF7/OP

²vgl. Ndihokubwayo et al. (2022)

C. Methoden

	trifft überhaupt nicht zu	trifft nicht zu	weder/ noch	trifft zu	trifft sehr zu
Die Unterrichtsmethoden sind ersichtlich.					
Die Unterrichtsmethoden werden zielgerichtet eingesetzt.					
Lernende setzen sich individuell mit den Unterrichtsinhalten auseinander.					
Lernende setzen sich kooperativ mit den Unterrichtsinhalten auseinander.					
Das Verhältnis frontaler Unterrichtsanteile zu Lernenden-Aktivitäten ist angemessen.					

D. Assessment (Evaluation)

	trifft überhaupt nicht zu	trifft nicht zu	weder/ noch	trifft zu	trifft sehr zu
Die Kompetenzziele werden durch die verwendeten Evaluationsinstrumente erfasst.					
Die Evaluationsinstrumente sind dazu geeignet, die individuelle Kompetenzziel-Erreichung der Lernenden zu messen.					

Literatur

Biggs, J.: *Enhancing teaching through constructive alignment*. Higher education, 32(3):347-364, 1996.

Bybee, R. W.: *The BSCS 5E instructional model: Personal reflections and contemporary implications*. Science and Children, 51(8):10–13, 2014.

Carpinelli, J., H. Kimmel, L. Hirsch, L. BurrAlexander, R. Rockland und M. OShea: *A Rubric To Evaluate Standards Based Lesson Plans And Students' Achievement Of The Standards*. In: ASEE Annual Conference and Exposition, Conference Proceedings, S. 13.98.1–13.98.10, 2008.

Ndihokubwayo, K., C. Byukusenge, E. Byusa, H. T. Habiyaremye, A. Mbonyiryivuze und J. Mukagihana: *Lesson plan analysis protocol (LPAP): A useful tool for researchers and educational evaluators*. *Heliyon*, 8(1):e08730, 2022.

Anhang E Ausgefüllte Fragebögen und Rückmeldungen

E.1 Befragung 1

Ausgefüllter Fragebogen Expertin 1

Bachelorarbeit
Sabine Britner

Computer Science Lesson Analysis: Fragebogen

A. Unterrichtsphasen

	trifft überhaupt nicht zu	trifft nicht zu	weder/ noch	trifft zu	trifft sehr zu
Die Unterrichtsphasen sind ersichtlich.					X
Aus dem Verlaufsplan geht klar hervor, was Lehrperson und Lernende in den einzelnen Unterrichtsphasen tun.					X
Die Lernenden werden zu Beginn für die Auseinandersetzung mit der Unterrichtsthematik motiviert (Engagement-Phase).				X	
Die Lernenden erkunden anhand praktischer Aktivitäten die Unterrichtsthematik (Exploration-Phase).					X
Auf Grundlage der Erfahrungen/ Ergebnisse der Lernenden aus der Erkundungsaktivität werden für die Unterrichtsthematik relevante Begriffe und Konzepte eingeführt (Explanation-Phase).					X
Die im Unterricht erworbenen Kenntnisse und Fähigkeiten werden auf neue Situationen übertragen (Elaboration-Phase).					X
Das Erreichen der Kompetenzziele wird durch die Lehrperson evaluiert (Evaluation-Phase).					X
Die jeweils eingeplanten Zeiträume für die Unterrichtsphasen sind angemessen.					X

B. Kompetenzen (Bitte Hinweis beachten)

	trifft überhaupt nicht zu	trifft nicht zu	weder/ noch	trifft zu	trifft sehr zu
Die Kompetenzziele sind ersichtlich.					X
Die Kompetenzziele sind für den Zeitrahmen angemessen.					X
Aus dem Verlaufsplan geht klar hervor, an welchen Stellen die Kompetenzen adressiert werden.					X
Die Kompetenzziele stehen in Bezug zu Thema und Inhalt der Unterrichtsstunde					X

Die Kompetenzziele besitzen die Komponente...	trifft überhaupt nicht zu	trifft nicht zu	weder/noch	trifft zu	trifft sehr zu
... Bedingung.					X
... Handlung.					X
... Inhalt.					X
... Kriterium.					X

Hinweis:

Ein Lernziel setzt sich aus fünf Komponenten zusammen: Wer, Bedingung, Handlung (vgl. Operatoren im Bildungsplan¹ oder Bloom Taxonomie), Inhalt und Kriterium. Beispiel: „Mit Hilfe von Venn-Diagrammen sind die Lernenden in der Lage, die Wahrscheinlichkeit eines Ereignisses korrekt zu bestimmen“.

Wer = die Lernenden; Bedingung = Mit Hilfe von Venn-Diagrammen; Handlung = bestimmen; Inhalt = Wahrscheinlichkeit eines Ereignisses; Kriterium = korrekt;²

C. Methoden

	trifft überhaupt nicht zu	trifft nicht zu	weder/noch	trifft zu	trifft sehr zu
Die Unterrichtsmethoden sind ersichtlich.					X
Die Unterrichtsmethoden werden zielgerichtet eingesetzt.					X
Lernende setzen sich individuell mit den Unterrichtsinhalten auseinander.					X
Lernende setzen sich kooperativ mit den Unterrichtsinhalten auseinander.					X
Das Verhältnis frontaler Unterrichtsanteile zu Lernenden-Aktivitäten ist angemessen.					X

D. Assessment

	trifft überhaupt nicht zu	trifft nicht zu	weder/noch	trifft zu	trifft sehr zu
Die Kompetenzziele werden durch die verwendeten Evaluationsinstrumente erfasst.					X
Die Evaluationsinstrumente sind dazu geeignet, die individuelle Kompetenzziel-Erreichung der Lernenden zu messen.					X

¹ https://bildungplaene-bw.de/_Lde/LS/BP2016BW/ALLG/SEK1/INF7/OP

² vgl. Ndihokubwayo et al. (2022)

Schriftliche Rückmeldung Expertin 1

Folie 4

Sie beginnen auf der ersten Seite mit der Frage: „Wie lange gibt es künstliche Intelligenz schon“. Dabei ist den SuS KI vielleicht noch kein Begriff. Vielleicht wäre guter Einstieg zu fragen „wo findet ihr in eurem Alltag KI?“, „was stellt ihr euch unter KI vor?“

Folie 6

Vielen SuS in der 7. Klasse sind wahrscheinlich die Begriffe „Teilbereich der Informatik „und „Computersysteme können Aufgaben lösen“ zu abstrakt. Ich denke, Beispiele aus dem Alltag der SuS wären hier noch hilfreich.

Folie 7

Schöne Folie

Folie 8

Sie haben sich ein spannendes Thema ausgesucht, das nicht leicht zu erklären ist. Ich finde die Grafik mit den Kreisen sehr schön, könnte mir aber vorstellen, dass manche jüngeren SuS innerlich abschalten, wenn zu viele „Informatik“ Wörter fallen, die sie noch nicht verinnerlicht haben. Sie haben bereits sehr schön das Beispiel mit Katze und Hund, aber ich würde ehrlich gesagt hier noch eine Folie einbauen, die das Beispiel erklärt und dabei den Dialog mit den SuS suchen, um sie für die neuen Inhalte zu aktivieren. Zum Beispiel „was meint ihr“, „wo in dem Kreis kann man xx einordnen“, „wie wäre das mit Hunden und Katzen ...“

Folie 9

Der Roboter ist sehr ansprechend. Haben sie an ein Beispiel aus der Schülerwelt gedacht. Wann lernen SuS aus Daten oder durch Bestrafung / Belohnung?

Folie 10

Super Video!

Arbeitsblatt 1

Schön gestaltet und eine tolle Idee. Ich finde die Kombination zwischen Logik und Personalisierung gut. Spannend wäre jetzt herauszufinden, ob die SuS damit klarkommen.

Folie 13

Hier ist ein kleiner Formatierungsfehler in meiner Version. Aber die Folie ist super.

Folie 14 / 17

Sehr schön.

Folie 18

Hier ist auch noch ein Formatierungsfehler

Ausgefüllter Fragebogen Expertin 2

Bachelorarbeit
Sabine Britner

Computer Science Lesson Analysis: Fragebogen

A. Unterrichtsphasen

	trifft überhaupt nicht zu	trifft nicht zu	weder/noch	trifft zu	trifft sehr zu	(Tabelle)
Die Unterrichtsphasen sind ersichtlich.				X →		Ich würde viele Worte herausstellen, ob Schüler oder Lehrer aktiv sind
Aus dem Verlaufsplan geht klar hervor, was Lehrperson und Lernende in den einzelnen Unterrichtsphasen tun.				X →		
Die Lernenden werden zu Beginn für die Auseinandersetzung mit der Unterrichtsthematik motiviert (Engagement-Phase).				X Würde über nichts ausklingen	✗	Stark herausarbeiten! → Thema ergibt sich dann!
Die Lernenden erkunden anhand praktischer Aktivitäten die Unterrichtsthematik (Exploration-Phase).					X	
Auf Grundlage der Erfahrungen/ Ergebnisse der Lernenden aus der Erkundungsaktivität werden für die Unterrichtsthematik relevante Begriffe und Konzepte eingeführt (Explanation-Phase).				X → schon vorgegeben in Präsentation → passt auf -		
Die im Unterricht erworbenen Kenntnisse und Fähigkeiten werden auf neue Situationen übertragen (Elaboration-Phase).					X	
Das Erreichen der Kompetenzziele wird durch die Lehrperson evaluiert (Evaluation-Phase).				X? Vorstellen wie genau sieht die vorletzte		Pläne nicht Evaluation aus?
Die jeweils eingeplanten Zeiträume für die Unterrichtsphasen sind angemessen.					X	

B. Kompetenzen (Bitte Hinweis beachten)

	trifft überhaupt nicht zu	trifft nicht zu	weder/noch	trifft zu	trifft sehr zu
Die Kompetenzziele sind ersichtlich.				X	
Die Kompetenzziele sind für den Zeitrahmen angemessen.				X	
Aus dem Verlaufsplan geht klar hervor, an welchen Stellen die Kompetenzen adressiert werden.				X	
Die Kompetenzziele stehen in Bezug zu Thema und Inhalt der Unterrichtsstunde				X	

TK2 + TK3 kommen oft in den letzten 10min
⇒ immer etwas heikel

Die Kompetenzziele besitzen die Komponente...	trifft überhaupt nicht zu	trifft nicht zu	weder/noch	trifft zu	trifft sehr zu
... Bedingung. → etwas unbew.				X	
... Handlung. → klar				X	
... Inhalt. → klar				X	
... Kriterium. → korrekt?				X	

} aus-
füllen!

Hinweis:

Ein Lernziel setzt sich aus fünf Komponenten zusammen: Wer, Bedingung, Handlung (vgl. Operatoren im Bildungsplan¹ oder Bloom Taxonomie), Inhalt und Kriterium. Beispiel: „Mit Hilfe von Venn-Diagrammen sind die Lernenden in der Lage, die Wahrscheinlichkeit eines Ereignisses korrekt zu bestimmen“. Wer = die Lernenden; Bedingung = Mit Hilfe von Venn-Diagrammen; Handlung = bestimmen; Inhalt = Wahrscheinlichkeit eines Ereignisses; Kriterium = korrekt;²

C. Methoden

	trifft überhaupt nicht zu	trifft nicht zu	weder/noch	trifft zu	trifft sehr zu
Die Unterrichtsmethoden sind ersichtlich.				X	
Die Unterrichtsmethoden werden zielgerichtet eingesetzt.				X	
Lernende setzen sich individuell mit den Unterrichtsinhalten auseinander.				X	X
Lernende setzen sich kooperativ mit den Unterrichtsinhalten auseinander.					X
Das Verhältnis frontaler Unterrichtsanteile zu Lernenden-Aktivitäten ist angemessen.					X

④

D. Assessment

	trifft überhaupt nicht zu	trifft nicht zu	weder/noch	trifft zu	trifft sehr zu
Die Kompetenzziele werden durch die verwendeten Evaluationsinstrumente erfasst.					
Die Evaluationsinstrumente sind dazu geeignet, die individuelle Kompetenzziel-Erreichung der Lernenden zu messen.					

} unbew.

⑤ aber in einer Phase zwei Methoden
ist unübersichtlich! Klar trennen!
Bsp. 10 Min Transfer

¹ <https://bildungsplaene-bw.de/Lde/LS/BP2016BW/ALLG/SEK1/INF7/OP>

² vgl. Ndihokubwayo et al. (2022)

Schriftliche Rückmeldung Expertin 2

Insbesondere denke ich, sollten Sie aus meiner Sicht noch folgendes stärker ausformulieren:

1. IN den einzelnen Unterrichtsphasen gibt es Phasen, die sowohl kooperative Lernformen haben (Partnerarbeit) als auch zentrale. Das würde ich trennen und ggf. 2 Phasen daraus machen, dann wird der methodische Aufbau klarer und man weiß genau, wenn kooperativ gearbeitet wird und wann nicht.
2. Sie sollten ausführen, dass die Kompetenzziele Bedingung, Handlung, Inhalt und Kriterium besitzen (im Verlauf als Bemerkung oder in Ihrer Arbeit dann, da hatte ich etwas Probleme).
3. Assessment ist mir unklar, wie sie die Arbeitsblätter evaluieren.
4. Mir fehlt ein wenig die Binnendifferenzierung.... Könnte gut passieren, dass schwächere Schüler auf dem ersten Arbeitsblatt bei 1a und 1b festsitzen.... eventuell scaffolding-Angebote oder dort schon einmal sichern und dann anwenden.... Binnendifferenzierung ist jedenfalls im gesamten Entwurf nicht richtig ersichtlich.

E.2 Befragung 2

Ausgefüllter Fragebogen Expertin 1

Bachelorarbeit
Sabine Britner

Computer Science Lesson Analysis: Fragebogen

A. Unterrichtsphasen

	trifft überhaupt nicht zu	trifft nicht zu	weder/ noch	trifft zu	trifft sehr zu
Die Unterrichtsphasen sind ersichtlich.					x
Aus dem Verlaufsplan geht klar hervor, was Lehrperson und Lernende in den einzelnen Unterrichtsphasen tun.					x
Die Lernenden werden zu Beginn für die Auseinandersetzung mit der Unterrichtsthematik motiviert (Engagement-Phase).					x
Die Lernenden erkunden anhand praktischer Aktivitäten die Unterrichtsthematik (Exploration-Phase).					x
Auf Grundlage der Erfahrungen/ Ergebnisse der Lernenden aus der Erkundungsaktivität werden für die Unterrichtsthematik relevante Begriffe und Konzepte eingeführt (Explanation-Phase).					x
Die im Unterricht erworbenen Kenntnisse und Fähigkeiten werden auf neue Situationen übertragen (Elaboration-Phase).					x
Das Erreichen der Kompetenzziele wird durch die Lehrperson evaluiert (Evaluation-Phase).					x
Die jeweils eingeplanten Zeiträume für die Unterrichtsphasen sind angemessen.					x

B. Kompetenzen (Bitte Hinweis beachten)

	trifft überhaupt nicht zu	trifft nicht zu	weder/ noch	trifft zu	trifft sehr zu
Die Kompetenzziele sind ersichtlich.					x
Die Kompetenzziele sind für den Zeitrahmen angemessen.					x
Aus dem Verlaufsplan geht klar hervor, an welchen Stellen die Kompetenzen adressiert werden.					x
Die Kompetenzziele stehen in Bezug zu Thema und Inhalt der Unterrichtsstunde					x

Die Kompetenzziele besitzen die Komponente...	trifft überhaupt nicht zu	trifft nicht zu	weder/noch	trifft zu	trifft sehr zu
... Bedingung.					x
... Handlung.					x
... Inhalt.					x
... Kriterium.					x

Hinweis:

Ein Lernziel setzt sich aus fünf Komponenten zusammen: Wer, Bedingung, Handlung (vgl. Operatoren im Bildungsplan¹ oder Bloom Taxonomie), Inhalt und Kriterium. Beispiel: „Mit Hilfe von Venn-Diagrammen sind die Lernenden in der Lage, die Wahrscheinlichkeit eines Ereignisses korrekt zu bestimmen“.

Wer = die Lernenden; Bedingung = Mit Hilfe von Venn-Diagrammen; Handlung = bestimmen; Inhalt = Wahrscheinlichkeit eines Ereignisses; Kriterium = korrekt;²

C. Methoden

	trifft überhaupt nicht zu	trifft nicht zu	weder/noch	trifft zu	trifft sehr zu
Die Unterrichtsmethoden sind ersichtlich.					x
Die Unterrichtsmethoden werden zielgerichtet eingesetzt.					x
Lernende setzen sich individuell mit den Unterrichtsinhalten auseinander.					x
Lernende setzen sich kooperativ mit den Unterrichtsinhalten auseinander.					x
Das Verhältnis frontaler Unterrichtsanteile zu Lernenden-Aktivitäten ist angemessen.					x

D. Assessment

	trifft überhaupt nicht zu	trifft nicht zu	weder/noch	trifft zu	trifft sehr zu
Die Kompetenzziele werden durch die verwendeten Evaluationsinstrumente erfasst.					x
Die Evaluationsinstrumente sind dazu geeignet, die individuelle Kompetenzziel-Erreichung der Lernenden zu messen.					x

¹ https://bildungplaene-bw.de/_Lde/LS/BP2016BW/ALLG/SEK1/INF7/OP

² vgl. Ndihokubwayo et al. (2022)

Schriftliche Rückmeldung Expertin 1

Ich finde der Workshop ist sehr gut gelungen. Ich habe keine Änderungsvorschläge mehr.

Ausgefüllter Fragebogen Expertin 2

Bachelorarbeit
Sabine Britner

Computer Science Lesson Analysis: Fragebogen

A. Unterrichtsphasen

	trifft überhaupt nicht zu	trifft nicht zu	weder/noch	trifft zu	trifft sehr zu
Die Unterrichtsphasen sind ersichtlich.					x
Aus dem Verlaufsplan geht klar hervor, was Lehrperson und Lernende in den einzelnen Unterrichtsphasen tun.				x	teilweise noch Phasen mit Schülerpräsentation und Lehrvortrag.... (bspw. Phase 4, aber durch die Erläuterungen im Ablaufplan dann ersichtlich)
Die Lernenden werden zu Beginn für die Auseinandersetzung mit der Unterrichtsthematik motiviert (Engagement-Phase).					x
Die Lernenden erkunden anhand praktischer Aktivitäten die Unterrichtsthematik (Exploration-Phase).					x
Auf Grundlage der Erfahrungen/ Ergebnisse der Lernenden aus der Erkundungsaktivität werden für die Unterrichtsthematik relevante Begriffe und Konzepte eingeführt (Explanation-Phase).					x
Die im Unterricht erworbenen Kenntnisse und Fähigkeiten werden auf neue Situationen übertragen (Elaboration-Phase).					x
Das Erreichen der Kompetenzziele wird durch die Lehrperson evaluiert (Evaluation-Phase).	Bewertungsschema hierfür noch unklar: Punkteskala, Teilerreichung der Ziele? ja/nein? Sonst klar, wie evaluiert wird.			x	
Die jeweils eingeplanten Zeiträume für die Unterrichtsphasen sind angemessen.					x

B. Kompetenzen (Bitte Hinweis beachten)

	trifft überhaupt nicht zu	trifft nicht zu	weder/noch	trifft zu	trifft sehr zu
Die Kompetenzziele sind ersichtlich.					x
Die Kompetenzziele sind für den Zeitrahmen angemessen.					x
Aus dem Verlaufsplan geht klar hervor, an welchen Stellen die Kompetenzen adressiert werden.					x
Die Kompetenzziele stehen in Bezug zu Thema und Inhalt der Unterrichtsstunde					x

Die Kompetenzziele besitzen die Komponente...	trifft überhaupt nicht zu	trifft nicht zu	weder/ noch	trifft zu	trifft sehr zu
... Bedingung.					x
... Handlung.					x
... Inhalt.					x
... Kriterium.				x	Noch konkreter wäre super!

Hinweis:

Ein Lernziel setzt sich aus fünf Komponenten zusammen: Wer, Bedingung, Handlung (vgl. Operatoren im Bildungsplan¹ oder Bloom Taxonomie), Inhalt und Kriterium. Beispiel: „Mit Hilfe von Venn-Diagrammen sind die Lernenden in der Lage, die Wahrscheinlichkeit eines Ereignisses korrekt zu bestimmen“.

Wer = die Lernenden; Bedingung = Mit Hilfe von Venn-Diagrammen; Handlung = bestimmen; Inhalt = Wahrscheinlichkeit eines Ereignisses; Kriterium = korrekt;²

C. Methoden

	trifft überhaupt nicht zu	trifft nicht zu	weder/ noch	trifft zu	trifft sehr zu
Die Unterrichtsmethoden sind ersichtlich.					x
Die Unterrichtsmethoden werden zielgerichtet eingesetzt.					x
Lernende setzen sich individuell mit den Unterrichtsinhalten auseinander.					x
Lernende setzen sich kooperativ mit den Unterrichtsinhalten auseinander.					x
Das Verhältnis frontaler Unterrichtsanteile zu Lernenden-Aktivitäten ist angemessen.					x

D. Assessment

	trifft überhaupt nicht zu	trifft nicht zu	weder/ noch	trifft zu	trifft sehr zu
Die Kompetenzziele werden durch die verwendeten Evaluationsinstrumente erfasst.					x
Die Evaluationsinstrumente sind dazu geeignet, die individuelle Kompetenzziel-Erreichung der Lernenden zu messen.					x

¹ https://bildungplaene-bw.de/_Lde/LS/BP2016BW/ALLG/SEK1/INF7/OP

² vgl. Ndihokubwayo et al. (2022)

Schriftliche Rückmeldung Expertin 2

Siehe Kommentare im Fragebogen.

Anhang F Re-Design: Verlaufsplan und Workshop-Materialien

F.1 Verlaufsplan

Zeitrahmen	90 Minuten
Übergeordnete Idee	Im Workshop lernen die Schülerinnen und Schüler das überwachte Lernen anhand von Entscheidungsbäumen kennen.
Konzepte der Informatik	Nature of Learning (nach AI for K12)
Verortung Bildungsplan	Sekundarstufe I – Informatik ab Klasse 7
Prozessorientierte Kompetenzen	Modellieren und Implementieren, Kommunizieren und Kooperieren
Inhaltsbezogene Kompetenzen	Daten und Codierung, Algorithmen
Groblerziele	Entscheidungsbäume mithilfe tabellarischer Datensätze erstellen, testen und anwenden.
Feinlernziele	Die Schülerinnen und Schüler können... <ul style="list-style-type: none">... aus tabellarischen Datensätzen einen Entscheidungsbäum erstellen, indem sie Attribute aus der Tabelle als Knoten in der richtigen Reihenfolge anordnen und über Kanten, beschriftet mit Werten, in geeigneter Weise verbinden (TK1).<ul style="list-style-type: none">... einen Entscheidungsbäum auf tabellarischen Datensätzen korrekt anwenden (TK2).... einen Entscheidungsbäum mithilfe tabellarischer Testdatensätze bewerten, indem sie den Baum korrekt auf den Testdaten anwenden und die Ergebnisse des Baumes mit der Lösung der Testdaten vergleichen (TK3).

Zeit / Phase	Lernziel	Inhalt / Thema	Methode	Transfer	Sonstiges
<i>5 Min</i> <i>Begrüßung</i>		Die Lehrperson begrüßt die Lernenden und stellt Workshopthema sowie Agenda vor.	Plenum, Folie 1-2		Transparenz schaffen und Orientierung geben.
<i>5 Min</i> <i>Problemstellung</i>		Die Lehrperson stellt die Problemstellung (Video) vor, gibt den Arbeitsauftrag, nimmt die Gruppenenteilung vor und teilt das Arbeitsblatt aus.	Lehrvortrag, Video Problemstellung, Arbeitsblatt 1, Folie 3-4		Motivation der Lernenden
<i>20 Min</i> <i>Erarbeitung</i>	Kommunikation und Kooperation der Lernenden	Die Lernenden bearbeiten das Arbeitsblatt. Während der Bearbeitung steht die Lehrperson den Schülerinnen und Schülern bei Fragen und Problemen zur Seite.	Partnerarbeit, Arbeitsblatt 1, Tablet/ Smartphone, Kopfhörer, Lösungskarten, (Bunt-)Stifte, (Lineal)		Ziel: Unterrichtsthematik erkunden und Problemstellung aus dem Video lösen.

<p><i>25 Min</i></p> <p><i>Besprechung</i></p> <p><i>Ergebnisse</i></p>	<p>Die Lehrperson ruft Lernende auf, die ihre Teillösungen vorstellen. Die Lernenden können vorgestellte Ergebnisse ergänzen oder Fragen dazu stellen. Präsentierte Algorithmus-Skizzen sowie der Sinn des Tests werden diskutiert. Die Lehrperson führt die Begriffe „Entscheidungsbaum“, „überwachtes Lernen“ und „Lernalgorithmus“ sowie den Prozess vom Training zur Anwendung eines überwachten Lernalgorithmus ein.</p>	<p>Präsentation der Lernenden, Lehrvortrag, Folie 5-16</p> <p>Auf Grundlage der von den Lernenden entwickelten Skizzien, wird der Entscheidungsbaum eingeführt. Bei Durchführung der Aktivität durchlaufen die Lernenden bereits die Phasen Training, Test und Anwendung. Diese sollen nun für das überwachte Lernen verallgemeinert werden.</p>
<p><i>5 Min</i></p> <p><i>Transferaufgabe</i></p>	<p>Modellierung Entscheidungsbaum: TK1 wird adressiert</p>	<p>Transfer von Bild- auf tabellarische Datensätze. Werden Attribute aus der Tabelle als Knoten identifiziert und in korrekter Reihenfolge angeordnet? Werden die Knoten über Kanten verbunden, die mit den korrekten Werten beschriftet sind?</p>

<p><i>5 Min Besprechung Transferaufgabe</i></p>	<p>Die Lehrperson ruft eine Lernende/einen Lernenden auf, die/der einen Lösungsvorschlag präsentiert. Die anderen Lernenden können ergänzen oder Fragen stellen. Die Lehrperson ergänzt ggf. die vorgestellte Lösung.</p>	<p>Präsentation Lernende/ Lernender, (Lehrvortrag), Folie 18-19</p>	<p>Zusammenfassung und schriftliches Festhalten der neuen Erkenntnisse.</p>
<p><i>10 Min Sicherung</i></p>	<p>Die Sicherung wird von der Lehrperson amoderiert. Die Lernenden erhalten von der Lehrperson den Lückentext. Sie besprechen und füllen den Lückentext gemeinsam mit der Lehrperson aus.</p>	<p>Klassengespräch, Sicherungsblatt, Stifte, Folie 20-27</p>	
<p><i>10 Min Übung/ Evaluation</i></p>	<p>TK1, TK2 und TK3 werden adressiert.</p>	<p>Die Lehrperson moderiert die Phase an. Sie teilt das Arbeitsblatt aus. Die Lernenden bearbeiten das Blatt. Die Lehrperson sammelt das Blatt am Ende der Phase ein.</p>	<p>Einzelarbeit, Arbeitsblatt 2, (Bunt-)Stifte, Folie 28</p> <p>Evaluation: Die eingesammelten Arbeitsblätter werden einer Arbeitsblatt-Analyse¹ unterzogen.</p>
<p><i>5 Min Verabschiedung</i></p>	<p>Die Lehrperson verabschiedet sich.</p>	<p>Plenum, Folie 29</p>	

¹ Auf dem Arbeitsblatt 2 werden alle Teilkompetenzen (TK1, TK2, TK3) adressiert. Deshalb entstehen durch die Bearbeitung des Blattes durch die Lernenden Produkte, die hinsichtlich der Kompetenzielerreichung analysiert werden können. Die für jedes Lernziel formulierte Komponente „Kriterium“ gibt Orientierung, um die Erfüllung des jeweiligen Lernziels zu evaluieren. Als Hilfestellung dient die in den Materialien bereitgestellte Musterlösung zu Arbeitsblatt 2 (Lösungsblatt 2).

F.2 Präsentationsfolien

Wie lernen Maschinen?



Ein Workshop zur Einführung in die Konzepte
überwachten Lernens

Quelle: Flaticon.com

Sabine Britner

Agenda

- Der Lügendetektor
- Ergebnis-Besprechung
- Aufgabe
- Zusammenfassung
- Übung

Lügendetektor Ergebnis-Besprechung Aufgabe Zusammenfassung Übung

Der Lügendetektor



Problemstellung.mp4

Lügendetektor Ergebnis-Besprechung Aufgabe Zusammenfassung Übung

Jetzt seid ihr an der Reihe!

Helft Lisa dabei, Tims Handy wiederzufinden. Bearbeitet hierfür die Aufgaben auf dem Arbeitsblatt 1.

20 Min Zweiergruppe

Lügendetektor Ergebnis-Besprechung Aufgabe Zusammenfassung Übung

Ergebnisse

Aufgabe 1a)

Welche Merkmale habt ihr gefunden?

Mund:

Schülerinnen und Schüler, die lachen oder lächeln und dabei ihre Zähne zeigen, findet man nur in der Kategorie „lügt“.

Blickrichtung und Augenbraue:

Den Blick nach vorne in Kombination mit der hochgezogenen Augenbraue gibt es nur in der Kategorie „lügt“.

Lügendetektor Ergebnis-Besprechung Aufgabe Zusammenfassung Übung

Ergebnisse

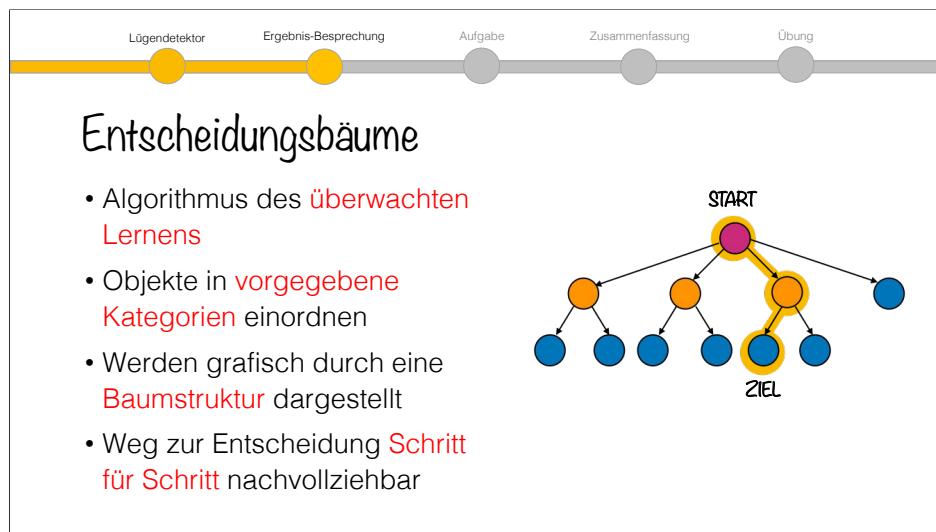
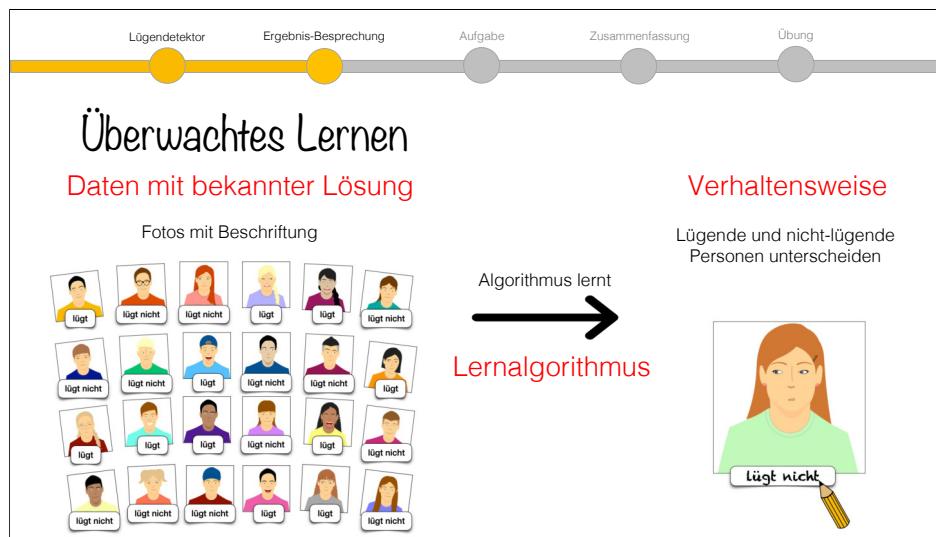
Aufgabe 1b)

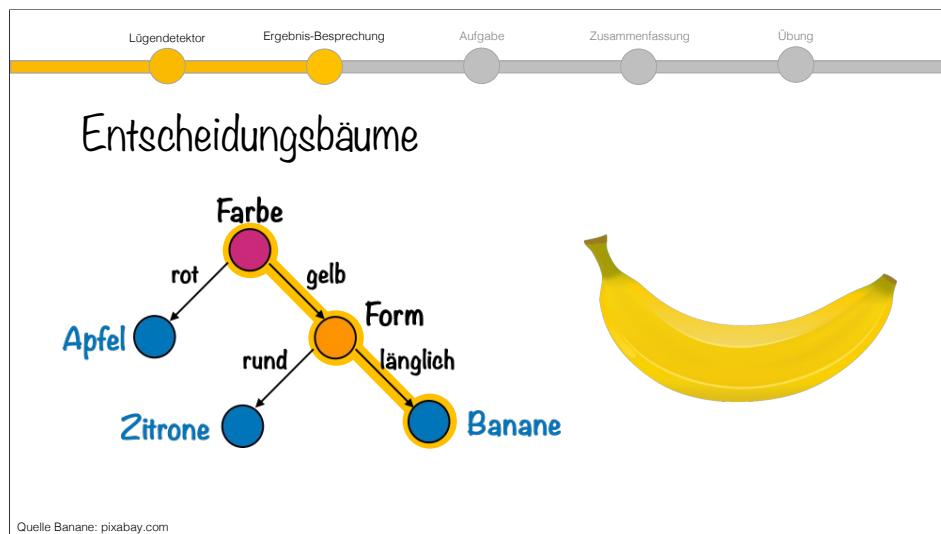
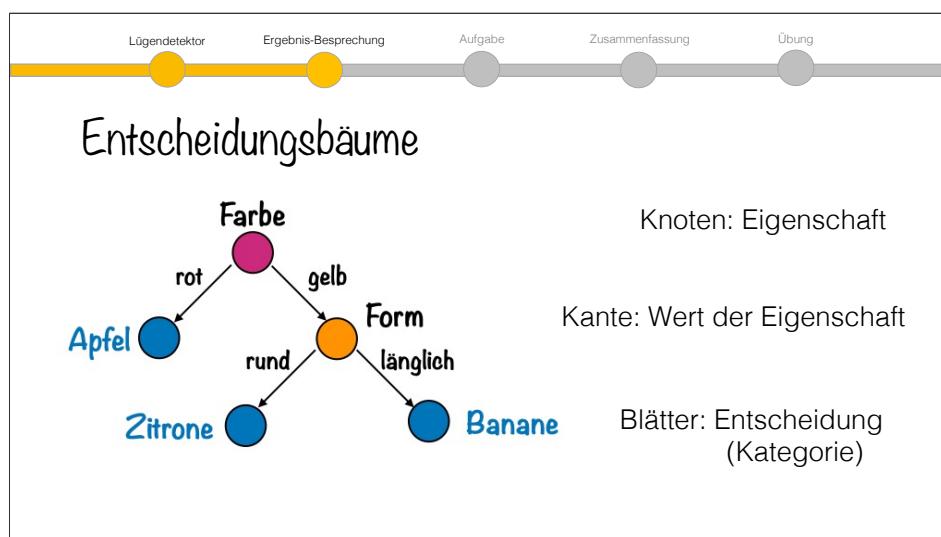
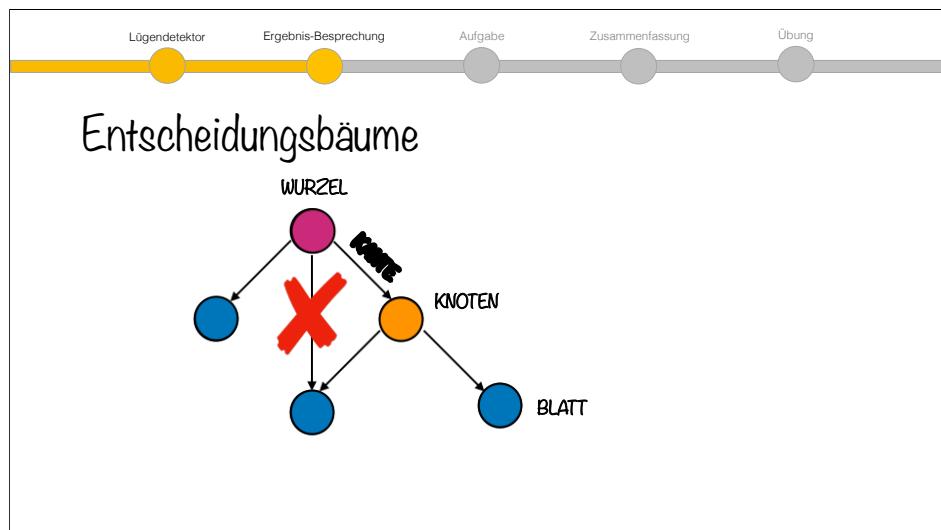
Welchen Algorithmus habt ihr euch überlegt?
Wie sieht eure Skizze aus?

Entscheidungsbaum

```

graph TD
    Mund((Mund)) -- "Lachen mit Zähnen" --> Lügen1((lügt))
    Mund -- "Keine Zähne sichtbar" --> Blickrichtung((Blickrichtung))
    Blickrichtung -- "nach vorn" --> Augenbraue((Augenbraue))
    Blickrichtung -- "zur Seite" --> Lügen2((lügt nicht))
    Augenbraue -- "hochgezogen" --> Lügen3((lügt))
    Augenbraue -- "normal" --> Lügen4((lügt nicht))
  
```





Lügendetektor Ergebnis-Besprechung Aufgabe Zusammenfassung Übung

Ergebnisse

Aufgabe 1c)

Welche Personen lügen und welche nicht?

<input type="checkbox"/> lügt <input checked="" type="checkbox"/> lügt nicht	<input type="checkbox"/> lügt <input checked="" type="checkbox"/> lügt nicht	<input type="checkbox"/> lügt <input checked="" type="checkbox"/> lügt nicht

<input checked="" type="checkbox"/> lügt nicht <input type="checkbox"/> lügt	<input checked="" type="checkbox"/> lügt nicht <input type="checkbox"/> lügt	<input type="checkbox"/> lügt <input checked="" type="checkbox"/> lügt nicht

Wie ist der Test für euren Algorithmus ausgefallen?

Welchen Sinn hat der Test?

Lügendetektor Ergebnis-Besprechung Aufgabe Zusammenfassung Übung

Training, Test und Anwendung

Training: Trainingsdaten auf Merkmale untersuchen

Test: Wie viele Fehler macht der Algorithmus? (Testdaten)

Anwendung: Algorithmus auf neuen Daten mit unbekannten Lösungen anwenden

Quelle Tierfotos: pixabay.com

Lügendetektor Ergebnis-Besprechung Aufgabe Zusammenfassung Übung

Ergebnisse

Aufgabe 1d)

Wer lügt und wer sagt die Wahrheit?

Elif	Elias	Anna

Lukas	Alex	Erik

Elif, Elias und Anna sagen die Wahrheit.

Lukas, Alex und Erik lügen.

Lügendetektor Ergebnis-Besprechung Aufgabe Zusammenfassung Übung

Ergebnisse

Aufgabe 1d)

Was ist mit Tims Handy passiert?

Anna hat Tims Handy auf der Fensterbank gesehen und wollte es ins Fundbüro bringen. Doch sie wurde abgelenkt. Stattdessen haben Lukas und Alex Tims Handy versteckt, um ihm einen Streich zu spielen. Elyas hat die beiden gehört als sie damit auf dem Pausenhof angegeben haben.

Lügendetektor Ergebnis-Besprechung Aufgabe Zusammenfassung Übung

Aufgabe

Erstelle aus den Trainingsdaten einen Entscheidungsbaum, mit dem man vorhersagen kann, wann eine Schülerin/ein Schüler Nachmittagsunterricht hat.

Klasse	AG	Nachmittagsschule
7b	Französisch	Mittwoch
7a	Robotik	Montag
7a	Französisch	Dienstag
7b	Robotik	Mittwoch

5 Min

Zweiergruppe
(Sitznachbar/
Sitznachbarin)

Lügendetektor Ergebnis-Besprechung Aufgabe Zusammenfassung Übung

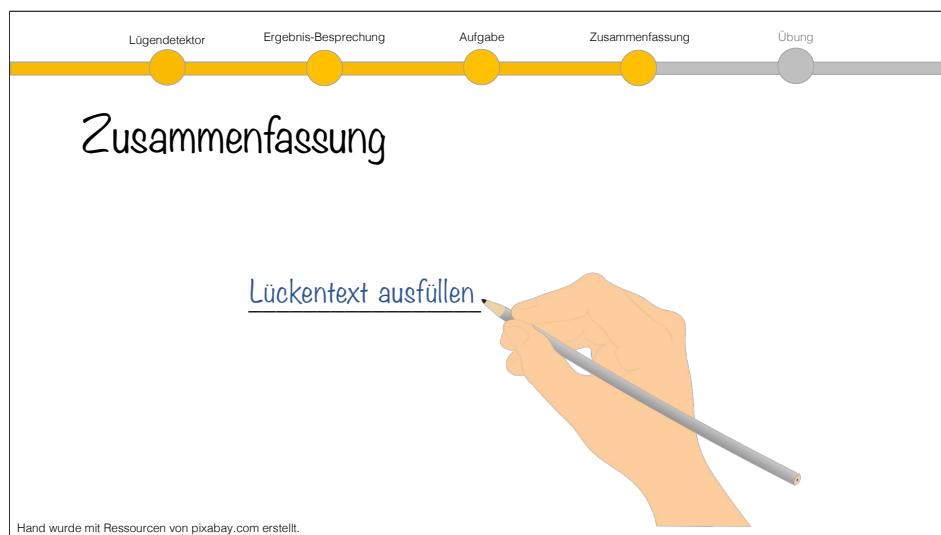
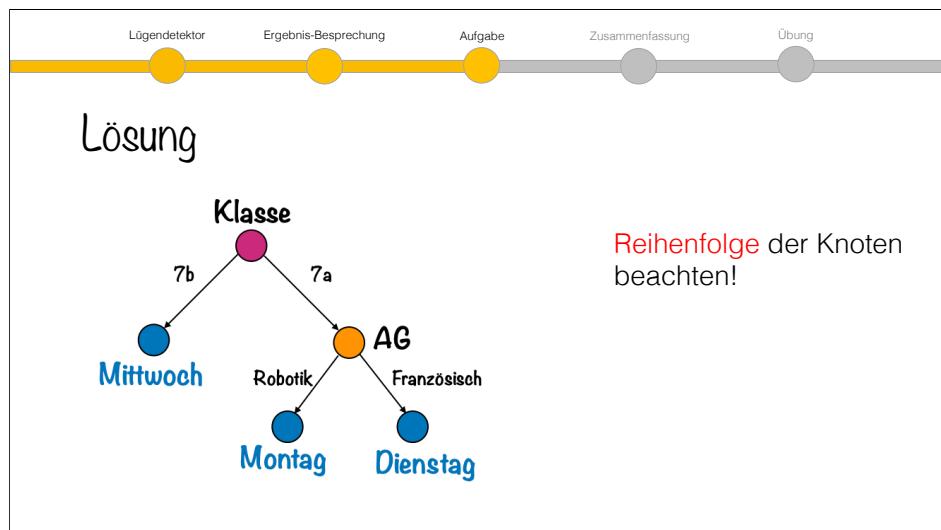
Lösung

Klasse	AG	Nachmittagsschule
7b	Französisch	Mittwoch
7b	Robotik	Mittwoch

In der 7b findet der Nachmittagsunterricht, unabhängig von der AG, am Mittwoch statt.

Klasse	AG	Nachmittagsschule
7a	Robotik	Montag
7a	Französisch	Dienstag

In der 7a findet der Nachmittagsunterricht, montags statt, wenn man die Robotik-AG besucht, sonst dienstags.

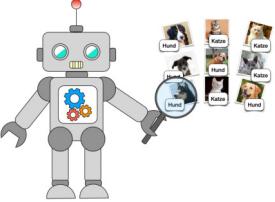


Lügendetektor Ergebnis-Besprechung Aufgabe Zusammenfassung Übung

Beim überwachten Lernen erlernt ein Algorithmus aus Daten mit bekannter Lösung eine bestimmte Verhaltensweise. Deshalb spricht man auch von Lernalgorithmen.

Lügendetektor Ergebnis-Besprechung Aufgabe Zusammenfassung Übung

Training



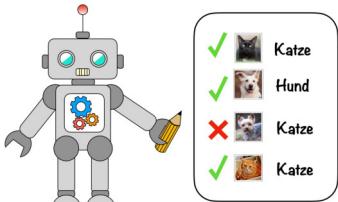
In dieser Phase werden dem Lernalgorithmus Trainingsdaten zur Verfügung gestellt. Der Algorithmus untersucht die Daten auf Merkmale.

Beispiel: Dem Algorithmus werden beschriftete Bilder von Hunden und Katzen bereitgestellt. Er untersucht die Bilder auf Merkmale, die ihm dabei helfen, Hunde und Katzen zu unterscheiden.

Quelle Tierfotos: pixabay.com

Lügendetektor Ergebnis-Besprechung Aufgabe Zusammenfassung Übung

Test

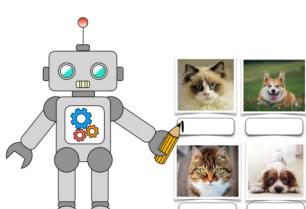


Mithilfe von Testdaten wird geprüft, wie zuverlässig der Algorithmus arbeitet. Diese Daten sind dem Algorithmus noch unbekannt. Macht der Algorithmus viele Fehler, so ist es sinnvoll zum Training zurückzukehren.

Quelle Tierfotos: pixabay.com

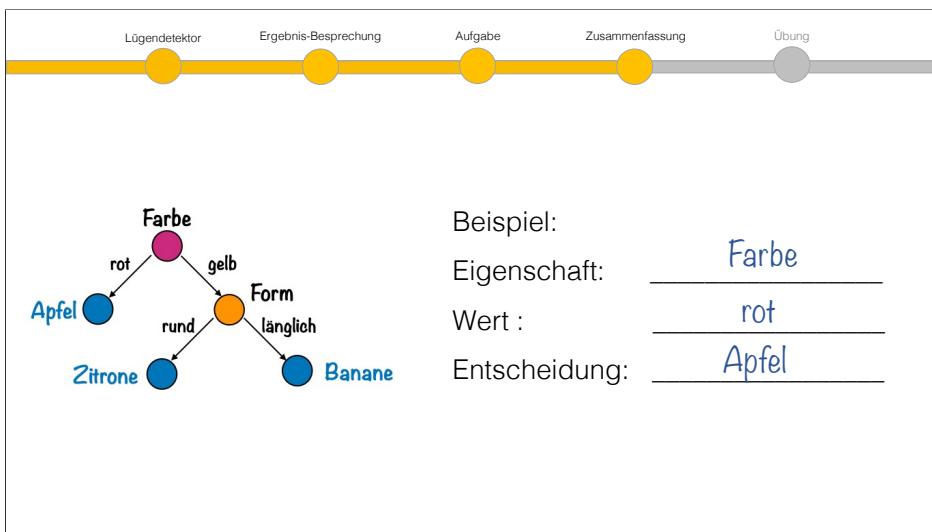
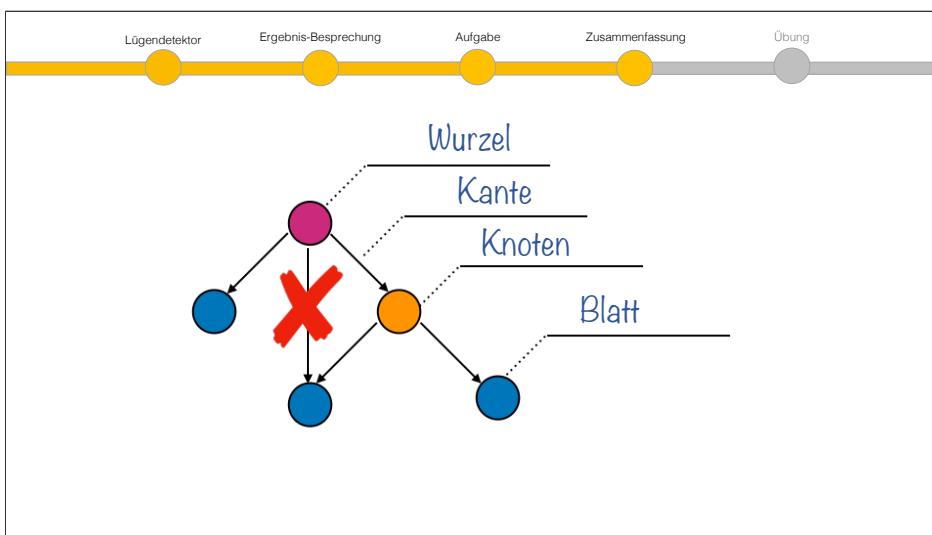
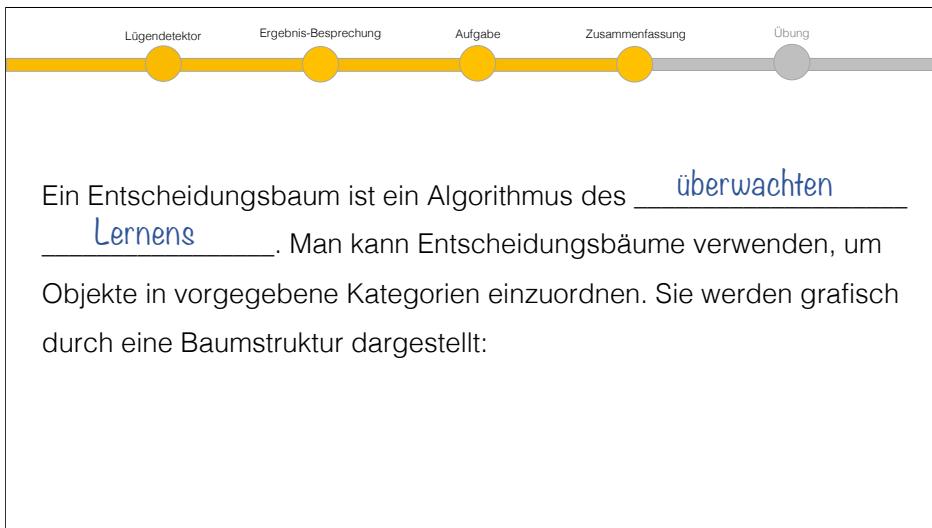
Lügendetektor Ergebnis-Besprechung Aufgabe Zusammenfassung Übung

Anwendung



Wenn der Test gut ausgefallen ist (viele richtige Ergebnisse), kann der Lernalgorithmus auf neuen Daten mit unbekannter Lösung eingesetzt werden.

Quelle Tierfotos: pixabay.com



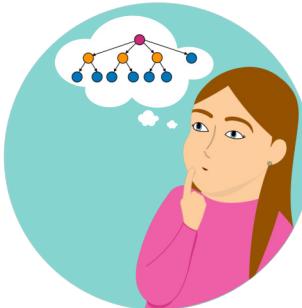
Lügendetektor	Ergebnis-Besprechung	Aufgabe	Zusammenfassung	Übung
---------------	----------------------	---------	-----------------	-------

Übung

Bearbeitet die Aufgaben auf dem Arbeitsblatt 2.

 10 Min

 Einzelarbeit



Lügendetektor	Ergebnis-Besprechung	Aufgabe	Zusammenfassung	Übung
---------------	----------------------	---------	-----------------	-------

Gut gemacht!



Literaturquellen

- E. Alpaydin, *Maschinelles Lernen*, Berlin/Boston: De Gruyter Oldenbourg, 2019.
- H. Aust, *Das Zeitalter der Daten: Was Sie über Grundlagen, Algorithmen und Anwendungen wissen sollten*, Berlin und Heidelberg: Springer, 2021.
- P. Buxmann und H. Schmidt, „Grundlagen der Künstlichen Intelligenz und des Maschinellen Lernens,“ in *Künstliche Intelligenz: Mit Algorithmen zum wirtschaftlichen Erfolg*, Berlin und Heidelberg, Springer Gabler, 2021, pp. 3-25.
- A. Géron, *Praxiseinstieg Machine Learning mit Scikit-Learn, Keras und TensorFlow: Konzepte, Tools und Techniken für intelligente Systeme*, Heidelberg: O'Reilly, 2020.

Literaturquellen

- T. Jo, Machine Learning Foundations: Supervised, Unsupervised, and Advanced Learning, Cham: Springer, 2021.
- I. Kalder, „Eine kurze Geschichte der künstlichen Intelligenz,“ in *Wie Maschinen lernen: Künstliche Intelligenz verständlich erklärt*, Wiesbaden, Springer, 2019, pp. 135-140.
- J. Kossen, F. Kuruc und M. E. Müller, „Einleitung,“ in *Wie Maschinen lernen: Künstliche Intelligenz verständlich erklärt*, Wiesbaden, Springer, 2019, pp. 3-10.
- J. Kossen, M. E. Müller und M. Ruckriegel, „Entscheidungsbäume,“ in *Wie Maschinen lernen: Künstliche Intelligenz verständlich erklärt*, Wiesbaden, Springer, 2019, pp. 111-118.

Literaturquellen

- M. Krause und E. Natterer, „Maschinelles Lernen,“ in *Wie Maschinen lernen: Künstliche Intelligenz verständlich erklärt*, Wiesbaden, Springer, 2019, pp. 21-27.
- T. M. Mitchell, Machine Learning, Singapore: McGraw Hill Education, 1997.
- S. Rosebrock, Geometrische Gruppentheorie: Ein Einstieg mit dem Computer. Basiswissen für Studium und Mathematikunterricht, Wiesbaden: Vieweg, 2004.
- SAP, „Was ist maschinelles Lernen?,“ SAP, [Online]. Available: <https://www.sap.com/germany/insights/what-is-machine-learning.html>. [Zugriff am 30 März 2022].

Literaturquellen

- S. Seegerer, T. Michaeli und R. Romeike, „So lernen Maschinen!,“ *LOG IN!*: Vol. 40, No. 1, pp. 27-31, 2020.
- L. Wuttke, Praxisleitfaden für Künstliche Intelligenz in Marketing und Vertrieb: Beispiele, Konzepte und Anwendungsfälle, Wiesbaden: Springer Gabler, 2021.
- L. Wuttke, „Training-, Validierung- und Testdatensatz,“ datasolut, 24 Juli 2020. [Online]. Available: <https://datasolut.com/wiki/trainingsdaten-und-testdaten-machine-learning/>. [Zugriff am 30 März 2022].

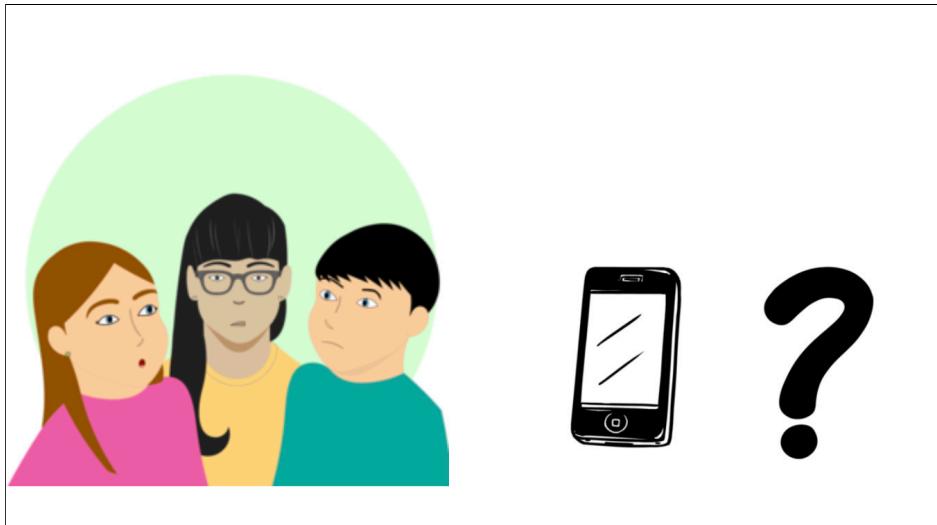
F.3 Video Problemstellung



Sprechertext: Der Lügendetektor.



Sprechertext: Das hier sind Lisa und Tim. Die beiden besuchen die Klasse 7b. Sie sind beste Freunde.



Sprechertext: Heute ist die ganze 7b in Aufregung. Tims Handy ist verschwunden! Wurde es womöglich gestohlen?



Sprechertext: Um herauszufinden was passiert ist, möchte Lisa ihre Mitschülerinnen und Mitschüler befragen. Da sie nicht davon ausgehen kann, dass alle die Wahrheit sagen, muss sie sich etwas einfallen lassen.



Lügendetektor

Sprechertext: Sie beschließt einen Lügendetektor zu entwickeln. Sie hat auch schon eine Idee, wie das funktionieren könnte.



**Was ist mit
Tims Handy
passiert?**

Sprechertext: Helft Lisa dabei, einen Lügendetektor zu entwickeln und Tims Handy wiederzufinden, indem ihr die Aufgaben auf dem Arbeitsblatt 1 löst.

Create your simpleshow now!



videomaker.simpleshow.com

F.4 Arbeitsblatt 1



Arbeitsblatt 1

Der Lügendetektor

Die Klasse 7b ist in Aufregung. Tims Handy ist verschwunden. Wurde es womöglich gestohlen? Lisa ist Tims beste Freundin. Darum möchte sie den Fall unbedingt lösen. Um herauszufinden, was geschehen ist, will sie ihre Mitschülerinnen und Mitschüler befragen. Da sie aber nicht davon ausgehen kann, dass alle die Wahrheit sagen, muss sie sich etwas einfallen lassen. Sie beschließt einen Lügendetektor-Algorithmus zu entwickeln.

- a) Scannt den QR-Code mit eurem Tablet oder Smartphone ab und seht euch das Video an:



Arbeitsauftrag:

Seht euch die Schülerinnen und Schüler der Klasse 7a in Abbildung 1 genau an. Findet Merkmale, mit denen man entscheiden kann, ob eine Person in die Kategorie „lügt“ oder „lügt nicht“ eingeteilt wird.



Gefundene Merkmale:

- b) Nehmt die Merkmale aus Aufgabenteil a) und überlegt euch einen Algorithmus, der Lisa dabei hilft zu entscheiden, ob eine Person lügt oder nicht. Haltet den Algorithmus in einer **Skizze** fest.

Platz für eure Skizze

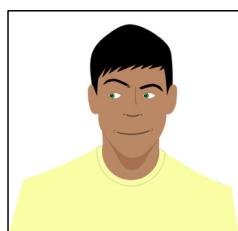
Das Icon in der Kopfzeile stammt von Flaticon.com.

- c) Scannt den QR-Code mit eurem Tablet oder Smartphone ab und seht euch das Video an:

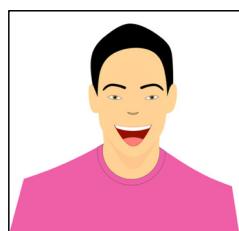


Arbeitsauftrag:

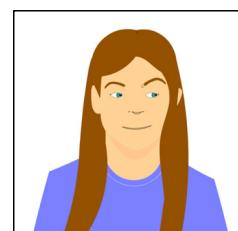
Ordnet die Schülerinnen und Schüler in Abbildung 2 mithilfe des Algorithmus in die Kategorien „lügt“ und „lügt nicht“ ein. Überprüft euer Ergebnis mit der Lösungskarte 1. Ist euer Algorithmus geeignet, einen Lügner/eine Lügnerin zu entlarven?



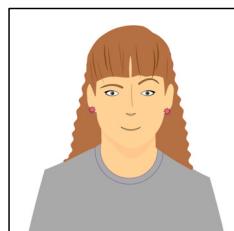
- lügt
 lügt nicht



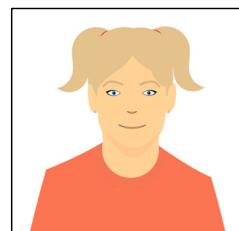
- lügt
 lügt nicht



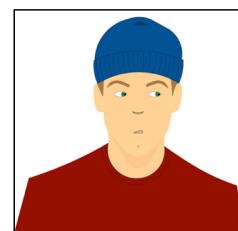
- lügt
 lügt nicht



- lügt
 lügt nicht



- lügt
 lügt nicht



- lügt
 lügt nicht

Abbildung 2

- d) Lisa befragt nun die Schülerinnen und Schüler aus der 7b, die am „Tatort“ anwesend waren.

Verwendet euren Algorithmus, um herauszufinden, welche Schülerinnen und Schüler aus der 7b die Wahrheit sagen und welche nicht (Abbildung 3). Findet heraus, was mit Tims Handy passiert ist und schreibt es in das vorgegebene Feld.

Lügen



Menschen, die Lügen erzählen, können sich durch ihren Gesichtsausdruck oder ihre Körpersprache verraten. Aber auch körperliche Stressanzeichen wie Schwitzen oder Herzklagen können einen Lügner oder eine Lügnerin überführen. 1902 erfand der Herzspezialist James Mackenzie den *Polygraphen*, einen Lügendetektor, der die körperlichen Stressanzeichen misst. Allerdings gibt es verschiedene Möglichkeiten diesen Lügendetektor zu überlisten. Heutzutage werden lernende Algorithmen in der Forschung zur Lügnerkennung eingesetzt. Noch gibt es aber kein sicheres Mittel, um einen Lügner oder eine Lügnerin zu entlarven.

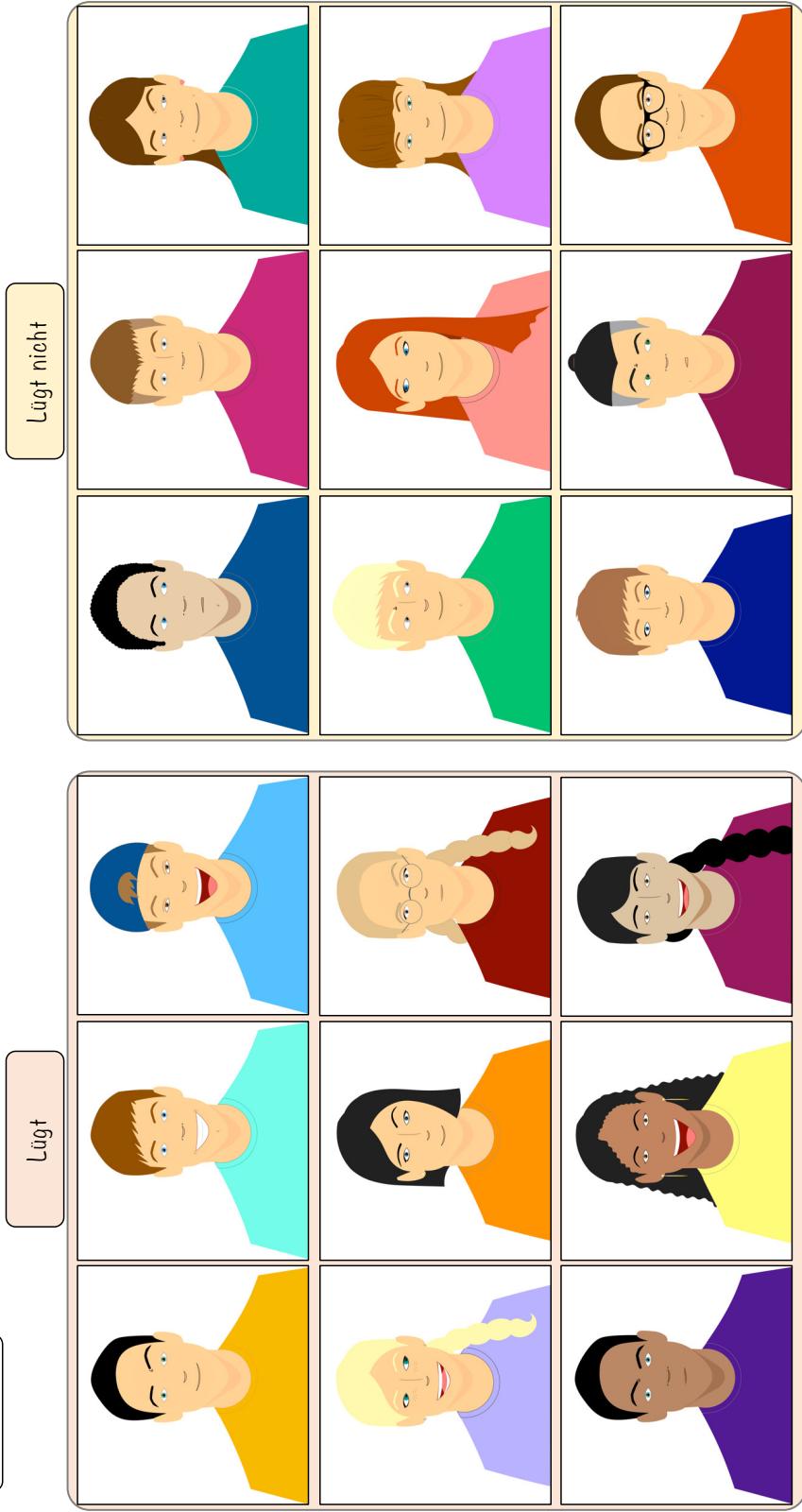
Quellen:

- <https://www.planet-wissen.de/gesellschaft/kommunikation/koerpersprache/pwieluegen100.html>
- <https://www.frontiersin.org/articles/10.3389/frobt.2019.00064/full>



Arbeitsblatt 1

Abbildung 1



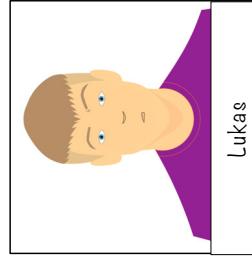
Das Icon in der Kopfzeile stammt von Flaticon.com.



Arbeitsblatt 1

Abbildung 3

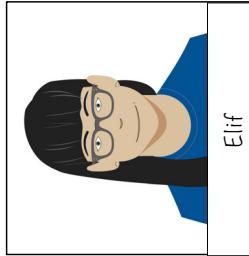
Das Icon in der Kopfzeile stammt von Flaticon.com.



Ich habe das Handy zuletzt auf der Fensterbank gesehen.

- lügt
 lügt nicht

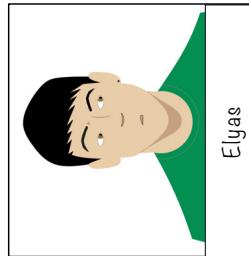
Lukas



Anna hat ein Handy auf der Fensterbank gefunden. Sie hat mir gesagt, dass sie es ins Fundbüro bringen möchte, weil sie nicht weiß, wem es gehört.

- lügt
 lügt nicht

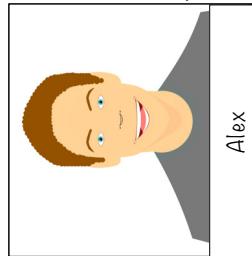
Elif



Vorhin haben ein paar Jungs auf dem Pausenhof damit angegeben das Handy versteckt zu haben, um Tim einen Streich zu spielen.

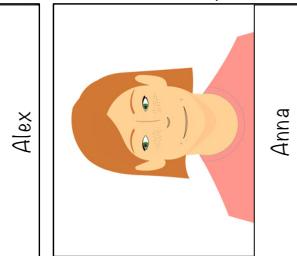
- lügt
 lügt nicht

Elias



Anna hat das Handy von der Fensterbank genommen und in ihren Rucksack gepackt.

- lügt
 lügt nicht



Ich wollte das Handy auf der Fensterbank ins Fundbüro bringen, aber dann wurde ich abgelenkt und habe es liegen lassen.

- lügt
 lügt nicht



Lukas und Alex haben Tims Handy nicht versteckt. Elias lügt.

- lügt
 lügt nicht

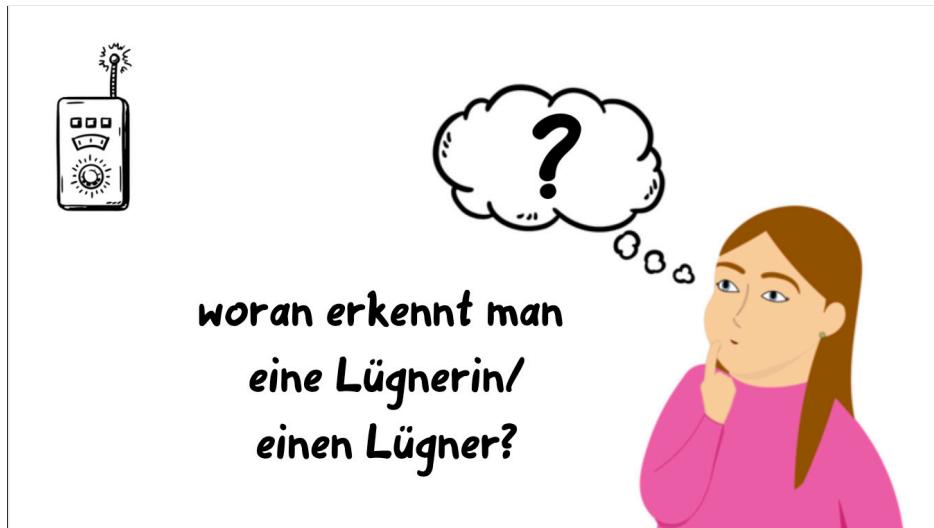
Das ist mit Tim's Handy passiert:

F.5 Verlinktes Video zu Aufgabenteil a) auf Arbeitsblatt 1



Merkmale finden

Sprechertext: Merkmale finden.



Sprechertext: Damit Lisa einen Lügendetektor entwickeln kann, muss sie ja erst einmal wissen, woran man einen Lügner oder eine Lügnerin erkennt.

Klasse 7a



Sprechertext: Um das herauszufinden, bittet sie die Klasse 7a um Hilfe. Lisa wählt zunächst 18 Schülerinnen und Schüler aus der 7a aus. Auf die restlichen sechs wird sie später noch zurückgreifen.

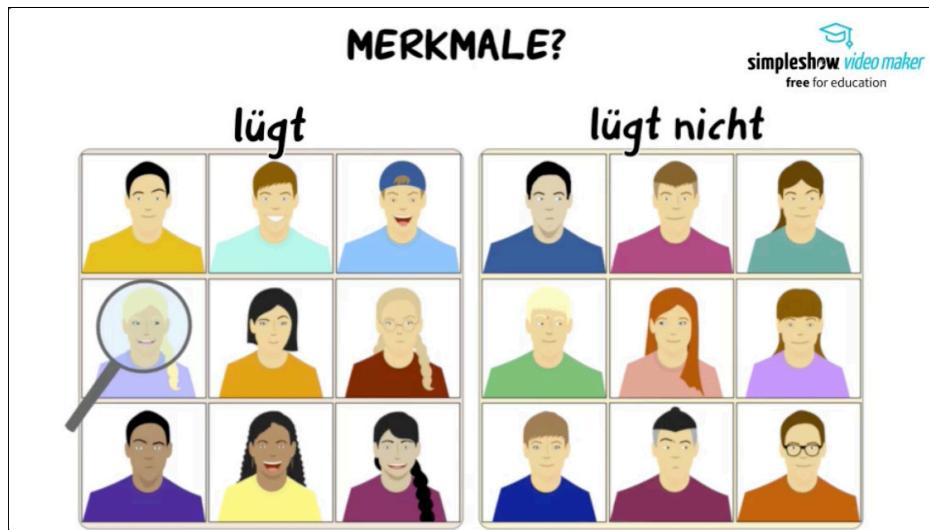
TRAININGSDATEN

simpleshow video maker
free for education

lügt lügt nicht

The diagram shows two groups of students. The left group is labeled "lügt" and the right group is labeled "lügt nicht". Each group is a 3x3 grid of student icons. The simpleshow logo and text "free for education" are in the top right corner.

Sprechertext: Die ausgewählten Schülerinnen und Schüler werden von ihr in zwei Gruppen eingeteilt: die eine Gruppe erzählt Lisa eine Lüge, die andere Gruppe eine wahre Geschichte. Wir werden diese Schülerinnen und Schüler als Trainingsdaten bezeichnen.



Sprechertext: Wichtig ist, dass für jede Person in den Trainingsdaten bekannt ist, in welche Kategorie sie gehört. So kann Lisa die Trainingsdaten auf Merkmale untersuchen, die lügende Personen von nicht lügenden unterscheiden.

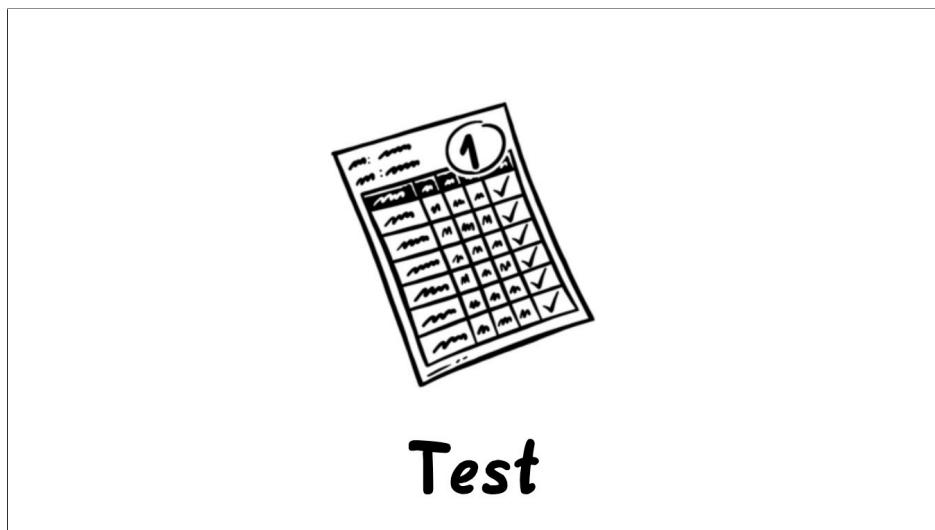
ARBEITSAUFTAG

- Schüler*innen in Abb. 1 ansehen
- Merkmale finden, die lügende und nicht lügende Personen unterscheiden

Sprechertext: Arbeitsauftrag: Seht euch die Schülerinnen und Schüler der Klasse 7a in Abbildung 1 genau an. Findet Merkmale, mit denen man entscheiden kann, ob eine Person lügt oder die Wahrheit sagt.



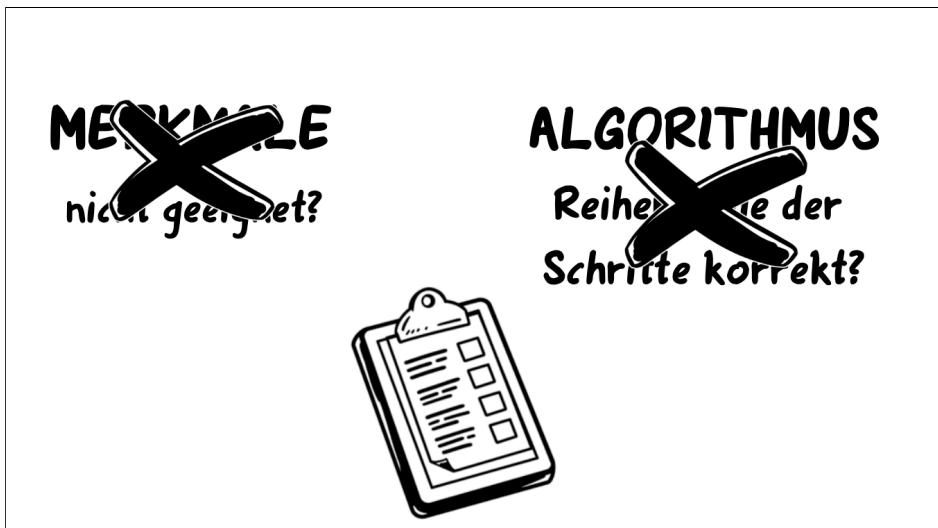
F.6 Verlinktes Video zu Aufgabenteil c) auf Arbeitsblatt 1



Sprechertext: Den Algorithmus testen.



Sprechertext: Nun da Lisa einen Algorithmus entwickelt hat, möchte sie sichergehen, dass er auch wirklich Lügnerinnen und Lügner entlarven kann.



Sprechertext: Es könnte ja sein, dass die Merkmale, die sie entdeckt hat, garnicht geeignet sind, um lügende Personen von nicht lügenden zu unterscheiden. Oder vielleicht hat sie ja einen Fehler beim Erstellen des Algorithmus gemacht? Um das auszuschließen, will sie also die Zuverlässigkeit des Algorithmus testen.



Sprechertext: Dafür zieht Lisa Beispiele heran, die dem Algorithmus noch unbekannt sind. Nun kommen die Schülerinnen und Schüler der Klasse 7a ins Spiel, die noch übrig sind. Wir werden sie als Testdaten bezeichnen.



Sprechertext: Die Schülerinnen und Schüler erzählen Lisa eine Lüge oder eine wahre Geschichte. Und Lisa versucht mithilfe des Algorithmus zu bestimmen, wer lügt und wer nicht. Nachdem der Algorithmus alle Personen eingeordnet hat, sieht sie sich die Lösung an, um die Zuverlässigkeit des Algorithmus zu bewerten.

ARBEITSAUFTAG

- Bestimmen welche Schüler*innen in Abb. 2 lügen und welche nicht
- Ergebnis mit Lösungskarte 1 vergleichen
- Zuverlässigkeit bewerten

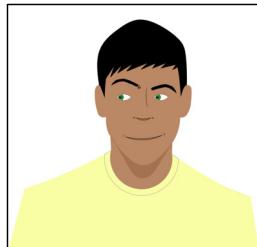
Sprechertext: Arbeitsauftrag: Bestimmt mithilfe des Algorithmus, welche Schülerinnen und Schüler in Abbildung 2 lügen und welche nicht. Überprüft eure Lösung mit Lösungskarte 1. Überlegt euch, ob euer Algorithmus zuverlässig genug arbeitet.

Create your simpleshow now!



videomaker.simpleshow.com

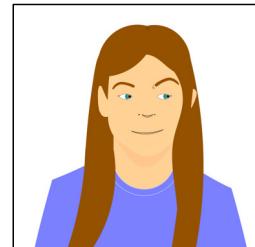
F.7 Lösungskarten zu Arbeitsblatt 1



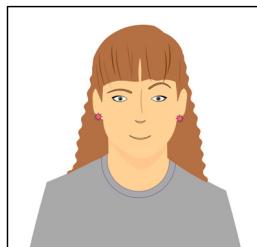
lügt
 lügt nicht



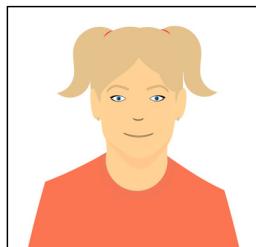
lügt
 lügt nicht



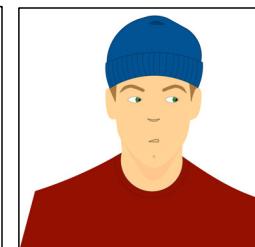
lügt
 lügt nicht



lügt
 lügt nicht



lügt
 lügt nicht



lügt
 lügt nicht

Fragen zur Lösung? Dann werft einen Blick auf Lösungskarte 2.

Merkmale:

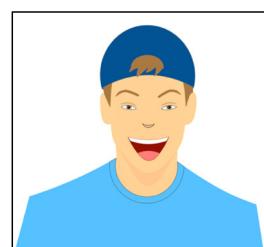
– **Mund:**

Schülerinnen und Schüler, die lachen oder lächeln und dabei ihre Zähne zeigen, findet man nur in der Kategorie „lügt“.

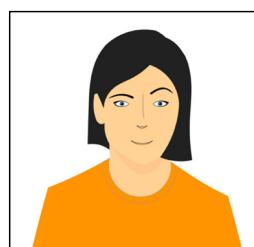
– **Blickrichtung und Augenbraue:**

Dass Schülerinnen und Schüler nach vorne blicken, ist in beiden Kategorien zu finden. Allerdings findet man die Kombination dieser Blickrichtung mit einer hochgezogenen Augenbraue nur bei den lügenden Personen.

Beispiele



Das Zeigen der Zähne beim Lachen ist ein Kriterium, um eine Person in die Klasse „lügt“ einzuteilen.



Der Blick nach vorn und die hochgezogene Augenbraue sind der Grund, warum diese Schülerin zur Klasse „lügt“ gehört.



Lösungskarte 1



Lösungskarte 2

F.8 Lösung zu Arbeitsblatt 1



Lösung zu Arbeitsblatt 1

Der Lügendetektor

Die Klasse 7b ist in Aufregung. Tims Handy ist verschwunden. Wurde es womöglich gestohlen? Lisa ist Tims beste Freundin. Darum möchte sie den Fall unbedingt lösen. Um herauszufinden, was geschehen ist, will sie ihre Mitschülerinnen und Mitschüler befragen. Da sie aber nicht davon ausgehen kann, dass alle die Wahrheit sagen, muss sie sich etwas einfallen lassen. Sie beschließt einen Lügendetektor-Algorithmus zu entwickeln.



- a) Scannt den QR-Code mit eurem Tablet oder Smartphone ab und seht euch das Video an:



Arbeitsauftrag:

Seht euch die Schülerinnen und Schüler der Klasse 7a in Abbildung 1 genau an. Mithilfe welcher Merkmale kann man entscheiden, ob eine Person in die Kategorie „lügt“ oder „lügt nicht“ eingeteilt wird?

Lösung

Die Schülerinnen und Schüler nehmen die Rolle eines Lernalgorithmus ein und begeben sich auf die Suche nach Merkmalen in den Daten, die die Klassen „lügt“ und „lügt nicht“ charakterisieren und voneinander abgrenzen. Die Daten sind hierbei durch die Bilder der 18 Schülerinnen und Schüler in Abbildung 1 gegeben. Das Erkennen von Mustern und Merkmalen in den Daten ist Bestandteil der Trainingsphase im maschinellen Lernprozess. Somit stellen die Daten in Abbildung 1 die Trainingsdaten dar.

Die zu extrahierenden Merkmale sind:

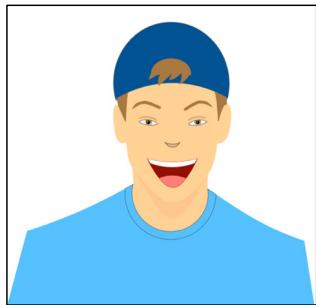
– **Mund:**

Schülerinnen und Schüler, die lachen oder lächeln und dabei ihre Zähne zeigen, findet man nur in der Klasse „lügt“.

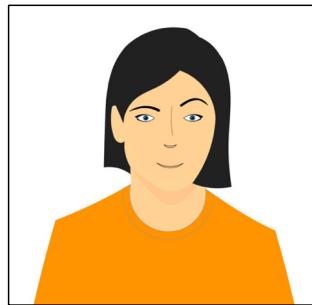
– **Blickrichtung und Augenbraue:**

Dass Schülerinnen und Schüler nach vorne blicken, ist in beiden Klassen zu finden. Allerdings findet man die Kombination dieser Blickrichtung mit einer hochgezogenen Augenbraue nur bei den lügenden Personen.

Beispiele:



Das Zeigen der Zähne beim Lachen ist ein Kriterium, um eine Person in die Klasse „lügt“ einzuteilen.



Der Blick nach vorn und die hochgezogene Augenbraue sind der Grund, warum diese Schülerin zur Klasse „lügt“ gehört.

Das Icon in der Kopfzeile wurde unter Verwendung von Ressourcen von Flaticon.com erstellt.



Lösung zu Arbeitsblatt 1

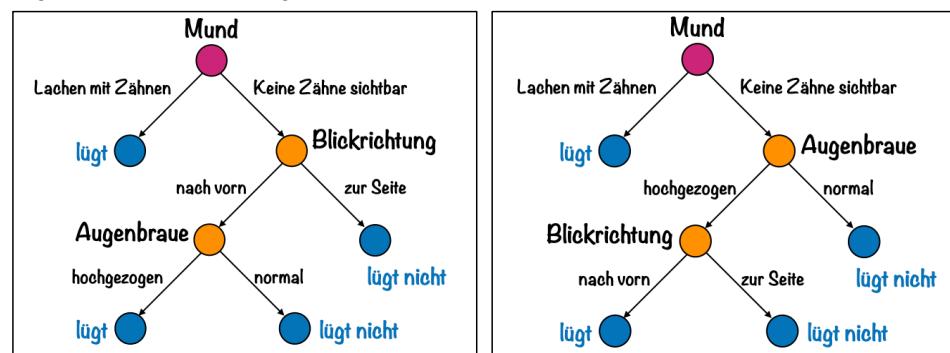
- b) Nehmt die Merkmale aus Aufgabenteil a) und überlegt euch einen Algorithmus, der Lisa dabei hilft zu entscheiden, ob eine Person lügt oder nicht. Haltet den Algorithmus in einer **Skizze** fest.

Lösung

Mit den aus Aufgabenteil a) extrahierten Merkmalen soll ein Algorithmus entwickelt werden, mit dem Personen in die Kategorien „lügt“ und „lügt nicht“ eingeteilt werden können. Der Algorithmus soll grafisch veranschaulicht werden.

Die Schülerinnen und Schüler sollen so eine eigene Darstellung entwickeln, die ggf. Elementen eines Entscheidungsbaumes nahekommen, wie etwa die Verwendung von Pfeilen. Die entwickelten Grafiken können später als Grundlage genutzt werden, um Entscheidungsbäume einzuführen.

Für den Entscheidungsbaum sind zwei Lösungen denkbar, da die Reihenfolge der Merkmale Augenbraue und Blickrichtung vertauscht werden kann:



- c) Scannet den QR-Code mit eurem Tablet oder Smartphone ab und seht euch das Video an:



Arbeitsauftrag:

Ordnet die Schülerinnen und Schüler in Abbildung 2 mithilfe des Algorithmus in die Kategorien „lügt“ und „lügt nicht“ ein. Überprüft euer Ergebnis mit der Lösungskarte 1. Ist euer Algorithmus geeignet, einen Lügner/eine Lügnerin zu entlarven?

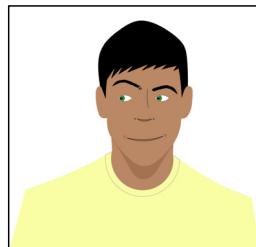
Lösung

Die Lernenden nutzen ihren Algorithmus, um zu entscheiden, ob eine Person lügt oder nicht. Dieser Aufgabenteil entspricht der Testphase im überwachten Lernprozess. Anhand von Daten (Testdaten), deren Lösung bereits bekannt ist (überwachtes Lernen), wird der Algorithmus auf seine Zuverlässigkeit geprüft. Sollten die Testdaten falsch zugeordnet werden, dann ist das ein Hinweis darauf, dass entweder die Merkmale im Algorithmus in der falschen Reihenfolge abgefragt werden oder gar nicht erst die richtigen Merkmale in a) extrahiert wurden. Die korrekte Zuordnung der Testdaten ist auf der nächsten Seite dargestellt. Die Schülerinnen und Schüler erhalten die korrekte Lösung der Testdaten über Lösungskarte 1. Sie sollen bewerten, ob ihr Algorithmus für die Problemlösung geeignet ist. Falls die Lernenden in der Testphase nicht zu der richtigen Lösung gelangen, gibt ihnen Lösungskarte 2 Informationen und Hinweise.

Das Icon in der Kopfzeile wurde unter Verwendung von Ressourcen von Flaticon.com erstellt.



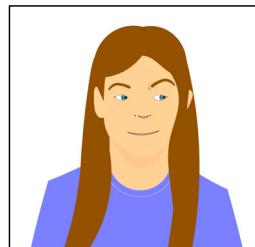
Lösung zu Arbeitsblatt 1



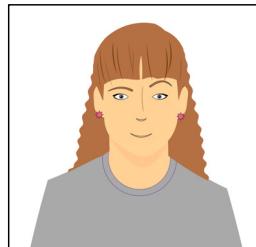
lügt
 lügt nicht



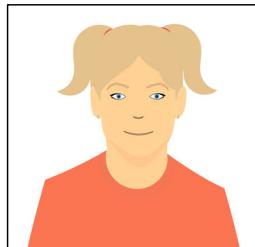
lügt
 lügt nicht



lügt
 lügt nicht



lügt
 lügt nicht



lügt
 lügt nicht



lügt
 lügt nicht

- d) Lisa befragt nun die Schülerinnen und Schüler aus der 7b, die am „Tatort“ anwesend waren.

Verwendet euren Algorithmus, um herauszufinden, welche Schülerinnen und Schüler aus der 7b die Wahrheit sagen und welche nicht (Abbildung 3). Findet heraus, was mit Tims Handy passiert ist und schreibt es in das vorgegebene Feld.

Lösung

Dieser Aufgabenteil entspricht der Anwendungsphase, die nach abgeschlossenem Training und Test erfolgt. In der Anwendungsphase wird der Algorithmus auf Daten angewandt, für die keine Lösungen vorliegen. Wie in Aufgabenteil c) verwenden die Schülerinnen und Schüler hier den Algorithmus, um die Personen zu klassifizieren. Nach der Klassifikation können sie das Rätsel um Tims Handy lösen. Es ergibt sich:

- Elif, Elyas und Anna sagen die Wahrheit.
- Lukas, Alex und Erik lügen.
- Anna hat Tims Handy auf der Fensterbank gesehen und wollte es ins Fundbüro bringen. Doch sie wurde abgelenkt. Stattdessen haben Lukas und Alex Tims Handy versteckt, um ihm einen Streich zu spielen. Elyas hat die beiden gehört als sie damit auf dem Pausenhof angegeben haben.

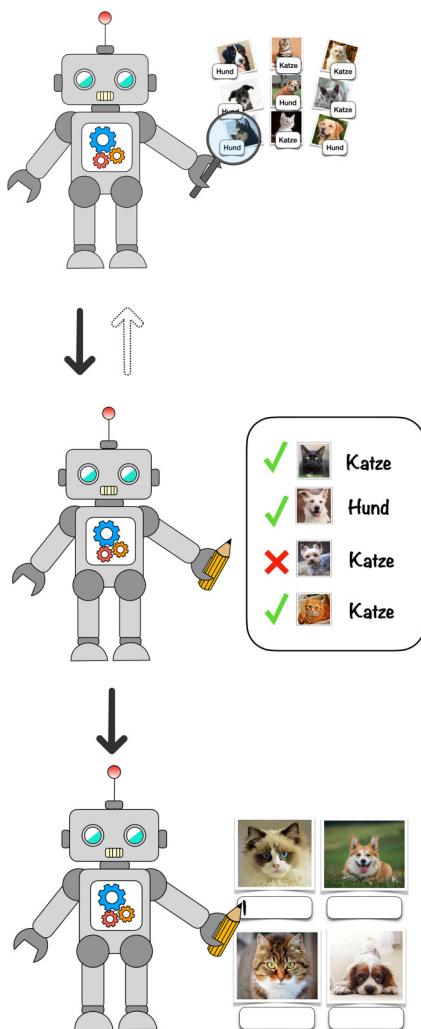
F.9 Sicherungsblatt



Unüberwachtes Lernen

Überwachtes Lernen

Beim überwachten Lernen erlernt ein Algorithmus aus Daten mit bekannter _____ eine bestimmte Verhaltensweise. Deshalb spricht man auch von _____.



In dieser Phase werden dem Lernalgorithmus _____ zur Verfügung gestellt. Der Algorithmus untersucht die Daten auf Merkmale.

Beispiel: Dem Algorithmus werden _____ Bilder von Hunden und Katzen bereitgestellt. Er untersucht die Bilder auf Merkmale, die ihm dabei helfen, Hunde und Katzen zu unterscheiden.

Mithilfe von _____ wird geprüft, wie zuverlässig der Algorithmus arbeitet. Diese Daten sind dem Algorithmus noch unbekannt. Macht der Algorithmus viele Fehler, so ist es sinnvoll zum Training zurückzukehren.

Wenn der Test gut ausgefallen ist (viele richtige Ergebnisse), kann der Lernalgorithmus auf neuen Daten mit unbekannter Lösung eingesetzt werden.

Anwendung – Lösung – Testdaten – Lernalgorithmen – beschriftete – Trainingsdaten – Test- Training

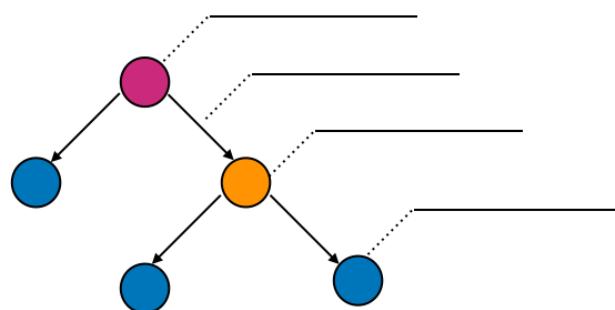
Das Icon in der Kopfzeile stammt von Flaticon.com
Quelle Tierbilder: pixabay.com



Unüberwachtes Lernen

Entscheidungsbäume

Ein Entscheidungsbau ist ein Algorithmus des _____
_____. Man kann Entscheidungsbäume verwenden, um Objekte in
vorgegebene Kategorien einzurichten. Sie werden grafisch durch eine Baumstruktur
dargestellt:



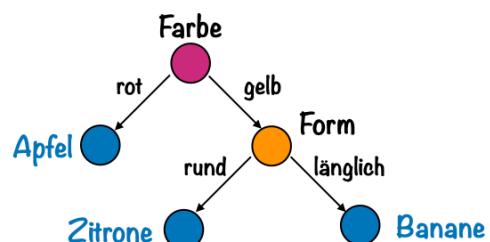
Die Blätter stellen die Entscheidungen darüber dar, in welche Kategorie ein Objekt
eingeteilt wird. Jeder andere Knoten steht für eine Eigenschaft. Die Kanten, die von
einem Knoten ausgehen, beschreiben jeweils mögliche Werte, die die zugehörige
Eigenschaft annehmen kann.

Beispiel

Eigenschaft: _____

Wert: _____

Entscheidung: _____



Will man einen Baum nutzen, um eine Entscheidung zu treffen, so beginnt man an der
Wurzel und läuft den Baum Knoten für Knoten ab.

Apfel – Blatt – überwachten Lernens – Knoten – Farbe – Wurzel – rot – Kante

F.10 Arbeitsblatt 2



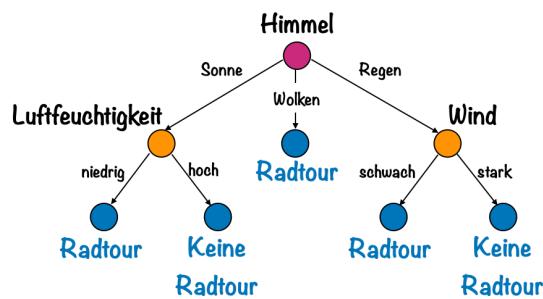
Arbeitsblatt 2

Aufgabe 1

Lisa und Tim möchten am Wochenende eine Radtour machen. Tim überprüft auf seinem Smartphone die Wettervorhersage. Welcher Tag ist für die Radtour geeignet? Markiere die Wege im Baum, die zu deiner Entscheidung geführt haben.

Lösung: Die Radtour findet am _____ statt.

	Himmel	Wind	Luftfeuchtigkeit
Samstag		stark	hoch
Sonntag		schwach	gering



Aufgabe 2

Elyas und Tim wollen sich am Freitag das Spiel ihres Lieblingsvereins ansehen und das Team anfeuern. Erstelle mithilfe der abgebildeten Trainingsdaten einen Entscheidungsbaum, mit dem man vorhersagen kann, auf welcher Tribüne ein Fan einen Platz zugewiesen bekommt.

Trainingsdaten:

Zeichnung:

Ankunft	Fan von Verein ...	Tribüne
Nach 17 Uhr	Victoria XY-Stadt	C
Vor 17 Uhr	Fortuna ABC-Hausen	A
Nach 17 Uhr	Fortuna ABC-Hausen	C
Vor 17 Uhr	Victoria XY-Stadt	B

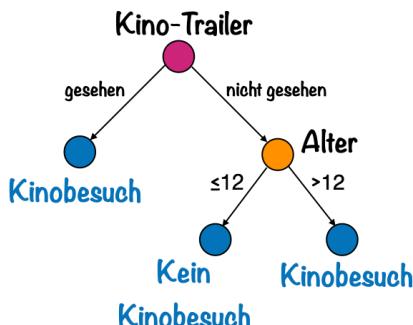
Das Icon in der Kopfzeile stammt von Flaticon.com.

Bitte umblättern ↗

Aufgabe 3

Mit folgendem Baum soll vorhergesagt werden, welche Personen den neuen Superhelden-Film im Kino ansehen werden.

- a) Überprüfe den Baum mithilfe der Testdaten. Eignet sich der Baum für die Anwendung? Begründe deine Entscheidung.



Testdaten:

Alter	Kino-Trailer	Kinobesuch	Testergebnis?
13	nicht gesehen	Kein Kinobesuch	
14	gesehen	Kinobesuch	
11	gesehen	Kein Kinobesuch	

Eignet sich der Baum für die Anwendung?

Ja Nein

Begründung: _____

- b) Die unten abgebildeten Trainingsdaten wurden für den Entscheidungsbaum verwendet. Sieh dir die Trainingsdaten an und zeichne den zugehörigen Entscheidungsbaum. Wie unterscheidet er sich vom Baum in Aufgabestellung a)?

Trainingsdaten:

Alter	Kino-Trailer	Kinobesuch
9	gesehen	Kein Kinobesuch
11	gesehen	Kein Kinobesuch
12	gesehen	Kinobesuch
11	nicht gesehen	Kein Kinobesuch
13	gesehen	Kinobesuch
10	nicht gesehen	Kein Kinobesuch
12	nicht gesehen	Kein Kinobesuch
14	nicht gesehen	Kein Kinobesuch

Zeichnung:

Wie unterscheidet er sich vom Baum in Aufgabestellung a)?

F.11 Lösung zu Arbeitsblatt 2



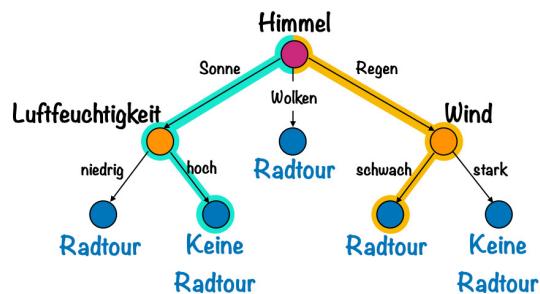
Lösung zu Arbeitsblatt 2

Aufgabe 1

Lisa und Tim möchten am Wochenende eine Radtour machen. Tim überprüft auf seinem Smartphone die Wettervorhersage. Welcher Tag ist für die Radtour geeignet? Markiere die Wege im Baum, die zu deiner Entscheidung geführt haben.

Lösung: Die Radtour findet am **Sonntag** statt.

	Himmel	Wind	Luftfeuchtigkeit
Samstag		stark	hoch
Sonntag		schwach	gering



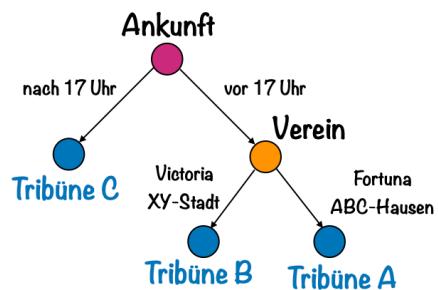
Aufgabe 2

Elias und Tim wollen sich am Freitag das Spiel ihres Lieblingsvereins ansehen und das Team anfeuern. Erstelle mithilfe der abgebildeten Trainingsdaten einen Entscheidungsbaum, mit dem man vorhersagen kann, auf welcher Tribüne ein Fan einen Platz zugewiesen bekommt.

Trainingsdaten:

Ankunft	Fan von Verein ...	Tribüne
nach 17 Uhr	Victoria XY-Stadt	C
vor 17 Uhr	Fortuna ABC-Hausen	A
nach 17 Uhr	Fortuna ABC-Hausen	C
vor 17 Uhr	Victoria XY-Stadt	B

Zeichnung:



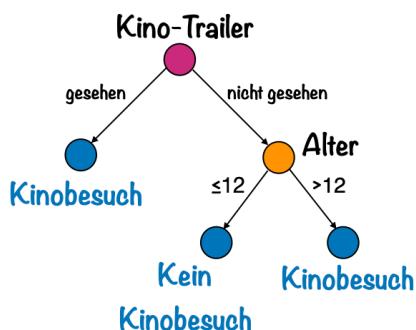


Lösung zu Arbeitsblatt 2

Aufgabe 3

Mit folgendem Baum soll vorhergesagt werden, welche Personen den neuen Superhelden-Film im Kino ansehen werden.

- a) Überprüfe den Baum mithilfe der Testdaten. Eignet sich der Baum für die Anwendung? Begründe deine Entscheidung.



Testdaten:

Alter	Kino-Trailer	Kinobesuch	Testergebnis?
13	nicht gesehen	Kein Kinobesuch	falsch
14	gesehen	Kinobesuch	richtig
11	gesehen	Kein Kinobesuch	falsch

Eignet sich der Baum für die Anwendung?

Ja Nein

Begründung: Weil 2 von 3 Testdaten falsch zugeordnet wurden, ist die Genauigkeit des Baumes nicht gut.

- b) Die unten abgebildeten Trainingsdaten wurden für den Entscheidungsbaum verwendet. Sieh dir die Trainingsdaten an und zeichne den zugehörigen Entscheidungsbaum. Wie unterscheidet er sich vom Baum in Aufgabestellung a)?

Trainingsdaten:

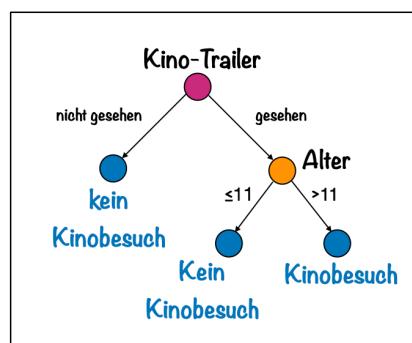
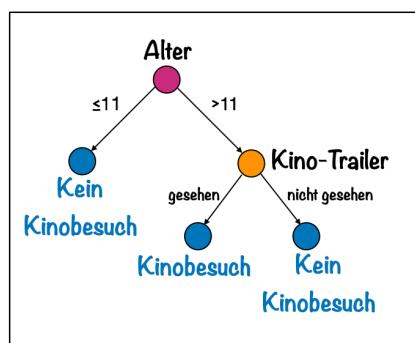
Alter	Kino-Trailer	Kinobesuch
9	gesehen	Kein Kinobesuch
11	gesehen	Kein Kinobesuch
12	gesehen	Kinobesuch
11	nicht gesehen	Kein Kinobesuch
13	gesehen	Kinobesuch
10	nicht gesehen	Kein Kinobesuch
12	nicht gesehen	Kein Kinobesuch
14	nicht gesehen	Kein Kinobesuch

Das Icon in der Kopfzeile wurde unter Verwendung von Ressourcen von Flaticon.com erstellt.



Lösung zu Arbeitsblatt 2

Zeichnung (2 mögliche Lösungen):



Wie unterscheidet er sich vom Baum in Aufgabestellung a)?

Die Altersgrenze ist falsch gesetzt. (Reihenfolge der Knoten ist vertauscht).

Bemerkung

Der Aufgabenteil b) von Aufgabe 3 ist die schwierigste Aufgabe auf dem Arbeitsblatt. Bisher mussten nur eindeutige Werte identifiziert werden, wie etwa ein Fußball-Verein, oder die Wertintervalle waren bereits vorgegeben, wie zum Beispiel in Aufgabe 2 „vor 17 Uhr“ und „nach 17 Uhr“. Bei dieser Aufgabe müssen die Lernenden die **Wertintervalle** (bis 11 Jahre und ab 12 Jahren) selbst ermitteln. Die Aufgabe muss nicht von allen Lernenden gelöst werden. Um das Erreichen der angestrebten Kompetenzen zu evaluieren, sind die Aufgaben 1 bis 3a) ausreichend. Aufgabe 3b) ist primär für leistungsstarke Lernende als Zusatzaufgabe gedacht.

F.12 Begleitbroschüre für Lehrkräfte

Wie lernen Maschinen?

Ein Workshop zur Einführung in die Konzepte
überwachten Lernens



Begleitbroschüre

Sabine Britner

Quelle Icon: Flaticon.com

Inhaltsverzeichnis

EINLEITUNG	2
HINTERGRUND	3
KÜNSTLICHE INTELLIGENZ	3
DEFINITION	3
GESCHICHTE	3
AUFGABEN	4
MASCHINELLES LERNEN	4
DEFINITION	4
LERTYPEN	5
ÜBERWACHTES LERNEN	5
ENTScheidungsbäume	7
WORKSHOP	9
ZIELE UND VORAUSSETZUNGEN	9
ZIELGRUPPE UND VORAUSSETZUNGEN	9
BEZUG ZUM BILDUNGSPLAN BADEN-WÜRTTEMBERG 2016	9
BEZUG ZU DEN RICHTLINIEN VON AI FOR K-12	9
LERNZIELE	10
MATERIALIEN	11
BEGLEITBROSCHÜRE	11
POWERPOINT-FOLIEN	11
VIDEOS	11
ARBEITSBLÄTTER	11
LÖSUNGSBLÄTTER	12
SICHERUNGSBLATT	12
VERLAUFSPLAN	13
LITERATURVERZEICHNIS	15

"We are entering a new world. The technologies of machine learning, speech recognition, and natural language understanding are reaching a nexus of capability. The end result is that we'll soon have artificially intelligent assistants to help us in every aspect of our lives." - Amy Stapleton [1]

Einleitung

Ob Bilderkennung, Sprachassistenten [2] oder Film-, Musik- und Produktempfehlungen [3] – besonders bei Jugendlichen ist heutzutage der Umgang mit künstlicher Intelligenz (Abk. KI) alltäglich. Gleichzeitig hat eine Untersuchung ergeben, dass Schülerinnen und Schüler nur wenig über die zugrunde liegenden Technologien von künstlicher Intelligenz wissen. Darüber hinaus werden diese von vielen Jugendlichen mit Gefahr oder Angst assoziiert, was auch mit dem mangelnden Wissen in Verbindung gebracht werden kann. [4]. Gerade deshalb und mit Blick auf die Entmystifikationsaufgabe des Informatikunterrichts [5] sollte die Aufklärung über die Konzepte hinter der künstlichen Intelligenz im Informatikunterricht erfolgen.

Das hier vorgestellte Workshop-Konzept richtet sich an Schülerinnen und Schüler der Sekundarstufe I (ab Klasse 7) und zielt darauf ab, sie in die grundlegenden Ideen des überwachten Lernens einzuführen. Die Jugendlichen erkunden gemeinsam mit den Hauptfiguren Lisa und Tim (siehe Abb.1) das überwachte Lernen anhand einer Unplugged-Aktivität mit Realitätsbezug.

Überwachtes Lernen ist ein Typ des maschinellen Lernens (Abk. ML) [2, 6]. Das maschinelle Lernen stellt einen Teilbereich der künstlichen Intelligenz dar [2], auf den sich die Entwicklung der KI in den letzten Jahren stark fokussiert hat [7]. Aktuell basieren die meisten KI-Systeme auf Verfahren des maschinellen Lernens [3]. Deshalb ist es naheliegend Konzepte des maschinellen Lernens im Unterricht zu thematisieren.

Auf den folgenden Seiten erfahren Sie mehr über die fachlichen Hintergründe, die Lernziele und -voraussetzungen sowie die Workshop-Struktur.



Abbildung 1: Gemeinsam mit Lisa und Tim lernen Schülerinnen und Schüler das überwachte Lernen kennen. Eigene Darstellung.

"The most important general-purpose technology of our era is artificial intelligence, particularly machine learning (ML) — that is the machine's ability to keep improving its performance without humans having to explain exactly how to accomplish all the tasks it's given." – Erik Brynjolfsson & Andrew McAfee [8]

Hintergrund

Künstliche Intelligenz

Definition

Künstliche Intelligenz (engl. *artificial intelligence*) beschreibt einen Teilbereich der Informatik [2, 7], der diejenigen Computersysteme umfasst, die Aufgaben und Probleme lösen können, für die man Intelligenz benötigt [3, 9]. Diese Definition greift den Umfang der Thematik nur grob auf, was unter anderem dadurch erklärt werden kann, dass bereits die Begriffsdefinition von Intelligenz problematisch ist [7, 9]. Darüber hinaus ist das Verständnis künstlicher Intelligenz von der zeitlichen Einordnung abhängig. Während vor einigen Jahren Navigationssysteme als künstliche Intelligenz zählten, wird die Technologie heute als selbstverständlich angesehen. [3].

Man kann zwei Arten künstlicher Intelligenz unterscheiden. Mit **starker KI** sind Methoden gemeint, die versuchen Abläufe im menschlichen Gehirn nachzuahmen. Hierzu zählen Merkmale wie Empathie oder Bewusstsein. Die erfolgreiche Entwicklung einer starken KI ist aktuell nicht bekannt. Eine **schwache KI** hingegen wird gezielt für die Lösung klar definierter und abgegrenzter Problemstellungen eingesetzt. Schwache KI-Systeme sind inzwischen technisch umsetzbar. [7].

Geschichte

Bereits im Jahre 1956 wurde der Begriff der künstlichen Intelligenz eingeführt. In diesem Jahr fand die sechswöchige Konferenz „Summer Research Project on Artificial Intelligence“ am Dartmouth College in den Vereinigten Staaten statt, die als die Geburtsstunde der künstlichen Intelligenz angesehen wird. [7, 10].

Von da an ging es mit der KI-Forschung bergauf. Schnellere und günstigere Computer sowie eine Zunahme der Speicherkapazitäten von Daten ermöglichten Fortschritte auf dem Gebiet der künstlichen Intelligenz. Ein Beispiel für einen Erfolg aus dieser Zeit ist das von Joseph Weizenbaum entwickelte Programm ELIZA, das eine(n) Gesprächspartner(in) simulieren konnte. Die Erfolge resultierten jedoch in einer zu optimistischen Erwartungshaltung an die KI-Entwicklung. Das Ausbleiben der erwarteten Erfolge führte zum sogenannten (ersten) KI-Winter. [7]. In dieser Phase ging das Interesse an der KI-Forschung und ihre finanzielle Förderung zurück [9, 10].

In den 1980er Jahren fand der erste KI-Winter ein Ende und man fokussierte sich verstärkt auf die Entwicklung sogenannter Expertensysteme. Hierunter versteht man Computerprogramme, die Menschen beim Lösen spezieller Probleme unterstützen können. Allerdings stellte sich heraus, dass Expertensysteme nur in besonderen Fällen von wirtschaftlichem Nutzen waren. [9]. Und so kam es erneut zu einem KI-Winter, welcher aber nicht lange andauerte [10].

Im Jahre 1997 gelang es dem Schachcomputer Deep Blue von IBM den damaligen Schachweltmeister Garri Kasparow zu besiegen [7, 9, 10]. Dieses Ereignis erregte große öffentliche Aufmerksamkeit, besonders da Schach jahrzehntelang als ultimativer Test für künstliche Intelligenz angesehen wurde [11].

In den letzten Jahren wurde der Fokus in der KI-Forschung auf das maschinelle Lernen gelegt [7]. So bilden heutzutage Verfahren des maschinellen Lernens zumeist die Grundlage für KI [3]. Die Wissenschaftler Erik Brynjolfsson und Andrew McAfee heben das maschinelle Lernen als wichtigste Universaldtechnologie unserer Zeit hervor [8].

Aufgaben

Künstliche Intelligenz wird für unterschiedliche Aufgaben eingesetzt [2]. Dazu gehören beispielsweise Sprachverständnis, Bilderkennung, Hilfestellungen in der Medizin und autonomes Fahren [2] sowie Produkt- und Filmempfehlungen [3].

Maschinelles Lernen

Definition

Bei maschinellem Lernen (engl. *machine learning*) handelt es sich um ein Teilgebiet der künstlichen Intelligenz [2]. Die Beziehung der Bereiche Informatik, künstliche Intelligenz und maschinelles Lernen ist in Abb.2 veranschaulicht. In der Literatur werden häufig die folgenden beiden Definitionen angeführt:

- Definition nach Arthur Samuel:

Durch maschinelles Lernen können Algorithmen aus der Erfahrung bestimmte Verhaltensweisen erlernen, ohne dass eine explizite Programmierung dieser erfolgen muss [12].

- Definition nach Tom Mitchell:

Ein Computerprogramm lernt aus der Erfahrung eine bestimmte Aufgabe zu lösen, wenn sich seine Leistung in Bezug auf diese Aufgabe mit der Erfahrung verbessert [13].

Algorithmen werden beim maschinellen Lernen als **Lernalgorithmen** bezeichnet [6]. Ein Algorithmus kann beispielsweise erlernen, Hunde und Katzen auf Bildern zu erkennen, ohne dass typische Eigenschaften der Tiere beim Programmieren im Code hinterlegt werden müssen. Stattdessen lernt der Algorithmus das Erkennen der Tiere eigenständig. [7]. Damit dieses Verhalten funktionieren kann, benötigt ein Algorithmus Erfahrung.

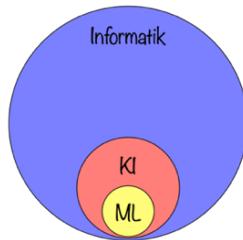


Abbildung 2: Abgrenzung und Verortung von KI und ML. Abbildung in Anlehnung an [2], S.56.

Lerntypen

Wie können Algorithmen aus der Erfahrung lernen? Beim maschinellen Lernen kann das auf zwei Arten erfolgen:

- Lernen aus Daten:

Algorithmen können aus Daten wie Bildern, Texten, Tönen oder Zahlen lernen. In diese Kategorie gehören das **überwachte** und das **unüberwachte Lernen** (engl. *supervised* und *unsupervised learning*) [9].

- Lernen durch Bestrafung und Belohnung:

Lernen Algorithmen über Belohnung und Bestrafung eine Verhaltensstrategie zu optimieren, so spricht man von **verstärkendem** oder **bestärkendem Lernen** (engl. reinforcement learning) [9, 13, 14].

Überwachtes Lernen

Das überwachte Lernen wird im Wesentlichen für **Klassifikations-** und **Regressionsaufgaben** angewandt [2, 15, 16].

- Klassifikation:

Unter einer Klassifikation versteht man das Einordnen von Objekten in vorgegebene Kategorien, die **Klassen** genannt werden [15]. Ein Beispiel für eine Klassifikationsaufgabe ist in Abb.3 zu sehen. Die Murmeln werden in die Klassen „rot“ und „blau“ eingeteilt. Beispiele für Lernverfahren, die für die Klassifikation eingesetzt werden, sind Entscheidungsbäume, logistische Regression und Naive Bayes [17].

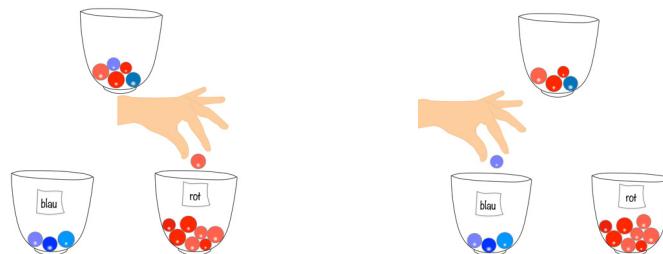


Abbildung 3: Die Murmeln werden in die beiden vorgegebenen Behälter mit der Aufschrift „rot“ bzw. „blau“ eingeordnet. Eigene Darstellung.

- Regression:

Die Regression beschäftigt sich mit der Vorhersage kontinuierlicher Zahlenwerte [2, 15]. Die vorherzusagende Größe wird als Zielgröße bezeichnet [18]. Für seine Vorhersage werden Größen aus den Eingabedaten unter der Annahme gewählt, dass sie die Zielgröße beeinflussen. Durch die Verwendung von Eingabedaten mit bekannter Lösung (Werte für Zielgröße sind bekannt), kann der Lernalgorithmus den Zusammenhang zwischen Eingabegröße und Zielgröße durch eine Funktion annähern. [16]. So kann für jeden beliebigen Wert die Zielgröße bestimmt werden [18]. Ein Beispiel ist in Abb.4 zu sehen. Mithilfe des Durchmessers einer Glas-Murmel soll das Gewicht der Murmel bestimmt werden. Das bekannteste Regressionsverfahren ist die lineare Regression [17].

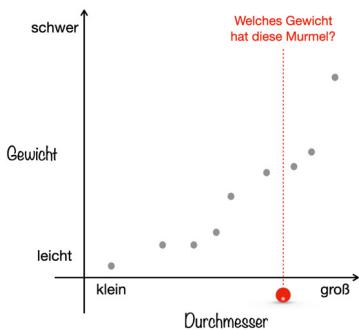


Abbildung 4a: Bekannte Murmel-Gewichte werden über dem Murmel-Durchmesser aufgetragen. In Anlehnung an [18], S. 41.

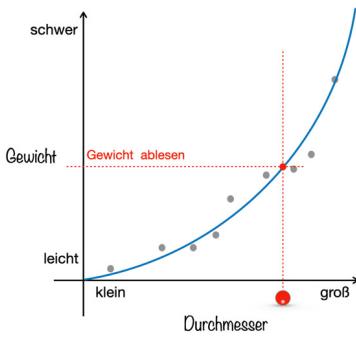


Abbildung 4b: Schnittpunkt mit Regressionskurve liefert Vorhersage für das Murmel-Gewicht. In Anlehnung an [31], S. 65.

Wie bereits geschildert, lernen Algorithmen beim überwachten Lernen aus Daten. Das überwachte Lernen zeichnet sich dadurch aus, dass die Lösung der Daten bereits bekannt ist [19, 6]. Ein Beispiel: Soll ein Algorithmus erlernen, Hunde und Katzen auf Bildern zu erkennen, so muss man dem Algorithmus Bilder zur Verfügung stellen, die mit „Hund“ oder „Katze“ beschriftet sind [7, 14]. Die Beschriftung der Bilder ist hier die bekannte Lösung. Die Beschriftungen der Daten werden als **Labels** bezeichnet [19].

Darüber hinaus werden die Daten beim überwachten Lernen in **Trainings-** und **Testdaten** unterteilt. Der Grund für diese Unterteilung wird in den folgenden Abschnitten erläutert. In der Praxis liegt das Verhältnis von Trainingsdaten zu Testdaten bei etwa 70:30. Abb.5 zeigt ein Beispiel für eine solche Aufteilung. [6].



Abbildung 5: Einteilung beschrifteter Tierbilder in Trainings- und Testdaten im Verhältnis von ca. 70:30 (Quelle Tierfotos: pixabay.com).

Wie kann man die Daten nutzen, damit ein Algorithmus ein bestimmtes Verhalten erlernt? In der ersten Phase durchläuft ein Algorithmus das sogenannte **Training**. In der Trainingsphase werden die Trainingsdaten zur Verfügung gestellt. [7]. Der Algorithmus untersucht die Daten und begibt sich auf die Suche nach Merkmalen und Mustern [6, 20]. Auf unser Beispiel bezogen: Dem Algorithmus werden Trainingsbilder von Hunden

und Katzen bereitgestellt [14]. Er untersucht die Bilder auf Merkmale, die die Klassen „Katze“ bzw. „Hund“ auszeichnen und voneinander unterscheiden. Dieser Prozess wird als das Lernen im Lernalgorithmus verstanden [7, 6].

Anschließend wird der trainierte Algorithmus einem **Test** unterzogen [7]. Anhand von Testdaten wird die Fehlerquote [6] bzw. die Treffergenauigkeit [9] des Algorithmus bestimmt. Die Treffergenauigkeit ist dabei der prozentuale Anteil der korrekt zugeordneten Testbeispiele an der Gesamtzahl der Testdaten [9]. Die Testdaten sind dem Lernalgorithmus noch unbekannt [20]. Für unser Beispiel bedeutet dies, dass der Algorithmus versucht die Tierbilder, die für den Test zurückgelegt wurden, jeweils in die richtige Klasse einzuordnen. Ob eine Zuordnung richtig ist, wird über die zugehörigen Labels geprüft. Fällt der Test gut aus, kann der Algorithmus auf neuen Daten mit unbekannter Lösung angewendet werden [6, 20]. Ist man mit dem Testergebnis nicht zufrieden, so kann man zum Training zurückkehren [21]. In Abb.6 ist der Prozess vom Training zur Anwendung dargestellt. Bei der Bilderkennung wird eine Treffergenauigkeit von mindestens 90% [14] und damit eine Fehlerquote von maximal 10% angestrebt.

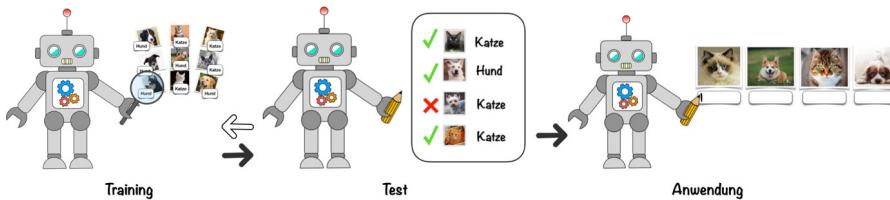


Abbildung 6: Vom Training zur Anwendung eines Lernalgorithmus (Quelle Tierfotos: pixabay.com).

Eine hohe Fehlerquote kann unterschiedliche Gründe haben. Beispielsweise kann eine zu geringe Datenmenge ursächlich für ein schlechtes Testergebnis sein. Grundsätzlich gilt, dass die Fehlerquote mit Zunahme der Trainingsdatenmenge abnimmt. In der Praxis liegt die Größenordnung im Bereich mehrerer Tausend Trainingsbeispiele. [6]. Eine weitere häufige Ursache ist die Überanpassung (engl. *overfitting*). Darunter versteht man, dass die vom Lernalgorithmus erlernten Merkmale und Muster zu stark an den Trainingsdaten orientiert und der Algorithmus somit nicht verallgemeinert werden kann. Auch muss man bei der Wahl der Trainingsdaten darauf achten, dass diese repräsentativ sind. [19]. So ist es beispielsweise nicht zielführend einem Algorithmus nur Bilder einer Hunderasse vorzuführen, wenn dieser erlernen soll Hunde auf Bildern zu erkennen.

Entscheidungsbäume

Im Workshop setzen sich Schülerinnen und Schüler mit dem überwachten Lernen auseinander. Entscheidungsbäume können schrittweise [22] und intuitiv [9] nachvollzogen werden, weshalb sie für die erste Auseinandersetzung im Workshop besonders geeignet sind.

Entscheidungsbäume werden grafisch durch eine Baumstruktur aus **Knoten** und Verzweigungen (**Kanten**) dargestellt. Der oberste Knoten in einem Baum wird als Wurzel bezeichnet. Ein Knoten führt über eine Kante zu einem weiteren Knoten. [2]. Die Baumstruktur ist in Abb.7a veranschaulicht. Im Allgemeinen erlauben es Bäume nicht,

dass mehrere Kanten zum selben Knoten führen (siehe Abb.7b) [23]. Knoten, von denen keine Kanten mehr ausgehen, werden als **Blätter** bezeichnet [2, 9].

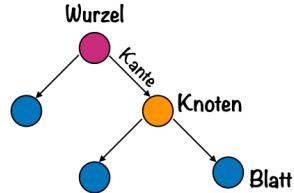


Abbildung 7a: Bestandteile einer Baumstruktur.
Eigene Darstellung.

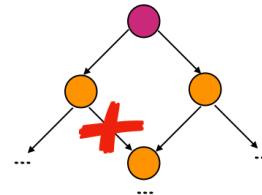


Abbildung 6b: Ein Knoten besitzt nur eine hinführende Kante. Eigene Darstellung.

Bei Klassifizierungsaufgaben liefern Entscheidungsbäume Vorhersagen darüber, in welche Klasse ein Objekt eingeteilt wird. Jedes Blatt eines Baumes repräsentiert eine Klasse. Die restlichen Knoten stehen jeweils für ein **Attribut** (eine Eigenschaft [22]). Die von einem Knoten ausgehenden Kanten sind mit **Werten** oder Wertintervallen beschriftet, die ein Attribut annehmen kann. [15]. Ein Beispielbaum mit Attributen und zugehörigen Werten ist in Abb.8 dargestellt. Will man einen Baum für eine Vorhersage nutzen, beginnt man an der Wurzel und läuft den Baum Knoten für Knoten ab, bis man bei einem Blatt landet. An jedem Knoten, der kein Blatt ist, wird das zu klassifizierende Objekt auf das zugehörige Attribut geprüft. Über die vom Knoten ausgehenden Kanten wählt man den zutreffenden Wert aus. So gelangt man über die ausgewählte Kante zum nächsten Knoten. [13].

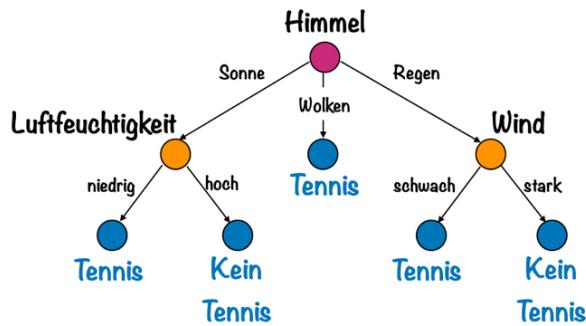


Abbildung 8: Entscheidungsbaum zur Vorhersage, ob das Wetter zum Tennisspielen geeignet ist. Attribute sind „Himmel“, „Luftfeuchtigkeit“ und „Wind“. Die Kanten sind mit den zugehörigen Werten beschriftet. Die Blätter repräsentieren die vorgegebenen Klassen „Tennis“ und „kein Tennis“. Abbildung in Anlehnung an [13], S.53.

Wie oben erwähnt, müssen beim überwachten Lernen vor der Anwendung eines Lernverfahrens Training und Test absolviert werden. Die Trainingsdaten werden dazu genutzt, Entscheidungsbäume zu erstellen. Dies ist der Lernprozess im Entscheidungsbau. [15]. Dabei spielt die Reihenfolge, in der die Merkmale der Trainingsdaten im Baum angeordnet werden, eine wichtige Rolle. Die Wahl der Position im Baum, an der auf ein bestimmtes Attribut geprüft wird, wirkt sich auf den Erfolg des Baumes hinsichtlich der Problemstellung aus. [22].

*"Artificial Intelligence, deep learning,
machine learning—whatever you're doing
if you don't understand it—learn it.
Because otherwise you're going to be a
dinosaur within 3 years." – Marc Cuban
[24]*

Workshop

Ziele und Voraussetzungen

Zielgruppe und Voraussetzungen

Der hier vorgestellte Workshop wurde für Schülerinnen und Schüler der Sekundarstufe I ab Klasse 7 konzipiert.

Um sich mit maschinellem Lernen auseinandersetzen zu können, wird ein Begriffsverständnis von **Algorithmen** vorausgesetzt. Das vorangegangene Kapitel zeigt auf, dass (Lern-)Algorithmen ein grundlegender Bestandteil maschinellen Lernens sind.

Bezug zum Bildungsplan Baden-Württemberg 2016 [25, 26]

Inhaltsbezogene Kompetenzen

- Daten und Codierung:
Im Rahmen des Workshops setzen sich die Schülerinnen und Schüler mit der Datenstruktur Baum am Beispiel von Entscheidungsbäumen auseinander.
- Algorithmen:
In der Erarbeitungsphase entwickeln Lernende einen Algorithmus, um die vorgegebene Problemstellung zu lösen.

Prozessbezogene Kompetenzen

- Modellieren und Implementieren:
Die Lernenden üben im Workshop aus einem gegebenen tabellarischen Datensatz einen Entscheidungsbaum zu modellieren.
- Kommunizieren und Kooperieren:
Im Workshop gibt es Aktivitäten, die in Zweiergruppen durchgeführt werden. Dabei lösen die Lernenden die Aufgabenstellungen gemeinsam, indem sie Lösungsideen sowie Ergebnisse diskutieren.

Bezug zu den Richtlinien von AI for K-12

AI for K-12 ist eine von der Association for the Advancement of Artificial Intelligence und Computer Science Teachers Association ins Leben gerufene Arbeitsgruppe, deren Aufgabe in der Definition von Richtlinien für die Vermittlung von künstlicher Intelligenz vom Kindergarten bis zur Klassenstufe 12 besteht. Diese Richtlinien werden durch fünf Ideen strukturiert, die der Arbeitsgruppe zufolge jeder Schülerin und jedem Schüler bekannt sein sollte, wobei sich hinter Idee Nummer drei („Computers can learn from data“) die Konzepte des maschinellen Lernens verbergen. [27].

Auf der Webseite der Arbeitsgruppe können die Richtlinien heruntergeladen werden (siehe [28]). Die Richtlinien für Idee Nummer drei sind nach Klassenstufen und Konzepten aufgeteilt. Unter dem Konzept „Nature of Learning“ ist für die Klassenstufen 6 bis 8 das Lernziel „Train and evaluate a classification or prediction model using machine learning on a tabular dataset.“ ([28], S. 3) aufgeführt. Die dem Workshop zugrunde liegenden Lernziele sind daran angelehnt.

Lernziele

Übergeordnetes Lernziel (Kernkompetenz)

Im Workshop lernen die Schülerinnen und Schüler das überwachte Lernen anhand von Entscheidungsbäumen kennen. Nach der Teilnahme können die Schülerinnen und Schüler Entscheidungsbäume erstellen, testen und anwenden.

Feinlernziele (Teilkompetenzen)

Die Schülerinnen und Schüler können ...

- ... aus tabellarischen Datensätzen einen Entscheidungsbaum erstellen, indem sie Attribute aus der Tabelle als Knoten in der richtigen Reihenfolge anordnen und über Kanten, beschriftet mit Werten, in geeigneter Weise verbinden (TK1).
- ... einen Entscheidungsbaum auf tabellarischen Datensätzen korrekt anwenden (TK2).
- ... einen Entscheidungsbaum mithilfe tabellarischer Testdatensätze bewerten, indem sie den Baum korrekt auf den Testdaten anwenden und die Ergebnisse des Baumes mit der Lösung der Testdaten vergleichen (TK3).

Komponenten der Lernziele

In Anlehnung an [29] und [30] wurden die Lernziele komponentenweise aufgebaut. Jede formulierte Teilkompetenz besteht aus den Komponenten:

- **Wer:**
Welche Personengruppe soll das Lernziel erfüllen?
- **Handlung¹:**
Was soll die Personengruppe in der Lage sein zu tun?
- **Bedingung:**
Unter welchen Rahmenbedingungen soll die Handlung ausgeführt werden?
- **Inhalt:**
Was ist der zu erlernende Inhalt?
- **Kriterium:**
Woran kann man festmachen, dass das Lernziel erfüllt wurde?

Als Beispiel wird die Teilkompetenz TK3 (siehe oben) betrachtet. Dabei beschreiben die „Schülerinnen und Schüler“ die Komponente Wer, die Handlung ist „bewerten“ und soll unter der Bedingung „mithilfe tabellarischer Testdatensätze“ erfolgen, der „Entscheidungsbaum“ ist der zu erlernende Inhalt und schließlich ist das Lernziel erfüllt,

¹ vgl. Bloom Taxonomie oder Operatoren im Bildungsplan (<https://bildungsplaene-bw.de/Lde/LS/BP2016BW/ALLG/SEK1/INF7/OP>)

wenn die Lernenden „den Baum korrekt auf den Testdaten anwenden und die Ergebnisse des Baumes mit der Lösung der Testdaten vergleichen“ können.

Evaluation der Lernzielerreichung

Im Rahmen des Workshops bearbeiten die Lernenden das Arbeitsblatt 2, auf dem alle Teilkompetenzen (TK1, TK2, TK3) adressiert werden. Dadurch entstehen Produkte, die hinsichtlich der Kompetenzerreichung analysiert werden können. Die für jedes Lernziel formulierte Komponente „Kriterium“ gibt Orientierung, um die Erfüllung des jeweiligen Lernziels zu evaluieren. Als Hilfestellung dient die in den Materialien bereitgestellte Musterlösung zu Arbeitsblatt 2 (Lösungsblatt 2).

Beispielsweise adressiert Aufgabe 1 auf dem Arbeitsblatt 2 die Teilkompetenz TK2. Die Lernenden sollen bei dieser Aufgabe einen vorgegebenen Entscheidungsbaum auf tabellarischen Datensätzen anwenden und dabei die von ihnen gewählten Wege durch den Baum einzeichnen. Mithilfe der Musterlösung kann überprüft werden, ob die Lernenden das richtige Ergebnis ermittelt haben. Entlang der Wege, die die Lernenden im Baum markieren, kann evaluiert werden, ob eine korrekte Anwendung des Entscheidungsbaums (z.B. Beginn an der Wurzel oder Auswahl zutreffender Kanten) zum Ergebnis geführt hat.

Materialien

Begleitbroschüre

- Begleitbroschüre.pdf

Die hier vorliegende Begleitbroschüre enthält Informationen über das Workshop-Konzept. Neben einem Kapitel, das Lehrpersonen in die relevanten fachsystematischen Hintergründe künstlicher Intelligenz und maschinellen Lernens einführt, werden hier die Lernziele und -voraussetzungen sowie wichtige Informationen zur Workshop-Durchführung präsentiert.

PowerPoint-Folien

- Folien.pptx

Zur Workshop-Durchführung werden animierte PowerPoint-Folien zur Verfügung gestellt. Die Folien sind zur besseren Orientierung mit Präsentationsnotizen versehen.

Videos

- Problemstellung.mp4
- Verlinkte Videos

Mithilfe des Videos Problemstellung.mp4 sollen die Lernenden in die Problemstellung für die Erarbeitungsphase eingeführt werden. Die Videos, die auf dem Arbeitsblatt 1 über einen QR-Code verlinkt sind, stellen den Lernenden Informationen bereit, die sie zur Bearbeitung der Teilaufgaben benötigen.

Arbeitsblätter

- AB1.pdf
- AB2.pdf

Das Arbeitsblatt AB1.pdf wurde für die Erarbeitungsphase konzipiert. Das zweite Arbeitsblatt ist für die Übungsphase gedacht und wird zum Zwecke der Evaluation nach dem Workshop einer Arbeitsblatt-Analyse unterzogen.

Lösungsblätter

- LB1.pdf
- LB2.pdf

Die Lösungsblätter liefern die Lösungen zu den in den Arbeitsblättern beschriebenen Arbeitsaufträgen. Sie sind ausschließlich für die Lehrpersonen konzipiert. In der Erarbeitungsphase erhalten die Schülerinnen und Schüler Lösungskarten (Lösungskarten.pdf), um eine Zwischenlösung zu kontrollieren.

Sicherungsblatt

- Sicherungsblatt.pdf

Um die erarbeiteten Inhalte des Workshops schriftlich festzuhalten, gibt es ein Blatt für die Lernenden mit Lücken, das in der Sicherungsphase gemeinsam mit der Lehrperson ausgefüllt wird.

Verlaufsplan

Zeit	Lernziel	Inhalt / Thema	Methode	Transfer	Sonstiges
5 Min Begrüßung		Die Lehrperson begrüßt die Lernenden und stellt Workshopthema sowie Agenda vor.	Plenum, Folie 1-2		Transparenz schaffen und Orientierung geben.
5 Min Problemstellung		Die Lehrperson stellt die Problemstellung (Video) vor, gibt den Arbeitsauftrag, nimmt die Gruppeneinteilung vor und teilt das Arbeitsblatt aus.	Lehrvortrag, Video Arbeitsblatt 1, Folie 3-4		Motivation der Lernenden
20 Min Erarbeitung	Kommunikation und Kooperation der Lernenden	Die Lernenden bearbeiten das Arbeitsblatt. Während der Bearbeitung steht die Lehrperson den Schülerinnen und Schülern bei Fragen und Problemen zur Seite.	Partnerarbeit, Arbeitsblatt 1, Tablet/ Smartphone, Kopfhörer, Lösungskarten, (Bunt-Stifte, (Lineal)		Ziel: Unterrichtsthematik erkunden und Problemstellung aus dem Video lösen.
25 Min Besprechung Ergebnisse		Die Lehrperson ruft Lernende auf, die ihre Teillösungen vorstellen. Die Lernenden können vorgestellte Ergebnisse ergänzen oder Fragen dazu stellen. Präsentierte Algorithmus-Skizzen sowie der Sinn des Tests werden diskutiert. Die Lehrperson führt die Begriffe „Entscheidungsbaum“, „überwachtes Lernen“ und „Lernalgorithmus“ sowie den Prozess vom Training zur Anwendung eines überwachten Lernalgorithmus ein.	Präsentation der Lernenden, Lehrvortrag, Folie 5-16		Auf Grundlage der von den Lernenden entwickelten Skizzen, wird der Entscheidungsbaum eingeführt. Bei Durchführung der Aktivität durchlaufen die Lernenden bereits die Phasen Training, Test und Anwendung. Diese sollen nun für das überwachte Lernen verallgemeinert werden.

5 Min <i>Transferaufgabe</i>	Modellierung Entscheidungsbaum: TK1 wird adressiert	Die Lehrperson stellt den Lernenden eine Transferaufgabe. Die Lernenden bearbeiten die Aufgabe.	Partherarbeit (Sitznachbarin/Sitznachbar), Papier, (Bunt-)Stifte, Folie 17 Transfer von Bild- auf tabellarische Datensätze. Werden Attribute aus der Tabelle als Knoten identifiziert und in korrekter Reihenfolge angeordnet? Werden die Knoten über Kanten verbunden, die mit den korrekten Werten beschriftet sind?
5 Min <i>Besprechung Transferaufgabe</i>		Die Lehrperson ruft eine Lernende/einen Lernenden auf, die/der einen Lösungsvorschlag präsentiert. Die anderen Lernenden können ergänzen oder Fragen stellen. Die Lehrperson ergänzt ggf. die vorgestellte Lösung.	Präsentation Lernende/ Lernender, (Lehrevortrag), Folie 18-19
10 Min <i>Sicherung</i>		Die Sicherung wird von der Lehrperson ammoderiert. Die Lernenden erhalten von der Lehrperson den Lückentext. Sie besprechen und füllen den Lückentext gemeinsam mit der Lehrperson aus.	Klassengespräch, Sicherungsblatt, Stifte, Folie 20-27 Zusammenfassung und schriftliches Festhalten der neuen Erkenntnisse.
10 Min <i>Übung/ Evaluation</i>	TK1, TK2 und TK3 werden adressiert	Die Lehrperson moderiert die Phase an. Sie teilt das Arbeitsblatt aus. Die Lernenden bearbeiten das Blatt. Die Lehrperson sammelt das Blatt am Ende der Phase ein. Die Lehrperson verabschiedet sich.	Einzelarbeit, Arbeitsblatt 2, (Bunt-)Stifte, Folie 28 Evaluation: Die eingesammelten Arbeitsblätter werden einer Arbeitsblatt-Analyse ² unterzogen.
5 Min <i>Verabschiedung</i>			Plenum, Folie 29

² siehe oben

Literaturverzeichnis

- [1] A. Stapleton, „Virtual Personal Assistants: Future Gatekeeper to Your Attention?“, Opus Research, 17 November 2015. [Online]. Available: <https://opusresearch.net/wordpress/2015/11/17/virtual-personal-assistants-future-gatekeeper-to-your-attention/>. [Zugriff am 24 März 2022].
- [2] L. Wuttke, Praxisleitfaden für Künstliche Intelligenz in Marketing und Vertrieb: Beispiele, Konzepte und Anwendungsfälle, Wiesbaden: Springer Gabler, 2021.
- [3] J. Kossen, F. Kuruc und M. E. Müller, „Einleitung,“ in *Wie Maschinen lernen: Künstliche Intelligenz verständlich erklärt*, Wiesbaden, Springer, 2019, pp. 3-10.
- [4] A. Lindner, M. Berges und M. Lechner, „KI im Toaster? Schüler:innenvorstellungen zu künstlicher Intelligenz,“ in *INFOS 2021 – 19. GI-Fachtagung Informatik und Schule*, Bonn, Gesellschaft für Informatik, 2021, pp. 133-142.
- [5] U. Lautebach, „Informatik für alle! Ein Plädoyer.,“ Gesellschaft für Informatik, 1 Februar 2018. [Online]. Available: <https://gi.de/meldung/informatik-fuer-alle-ein-plaedyoer>. [Zugriff am 24 März 2022].
- [6] M. Krause und E. Natterer, „Maschinelles Lernen,“ in *Wie Maschinen lernen: Künstliche Intelligenz verständlich erklärt*, Wiesbaden, Springer, 2019, pp. 21-27.
- [7] P. Buxmann und H. Schmidt, „Grundlagen der Künstlichen Intelligenz und des Maschinellen Lernens,“ in *Künstliche Intelligenz: Mit Algorithmen zum wirtschaftlichen Erfolg*, Berlin und Heidelberg, Springer Gabler, 2021, pp. 3-25.
- [8] E. Brynjolfsson und A. McAfee, „The Business of Artificial Intelligence—What It Can and Cannot Do for Your Organization,“ Machine Learning Times, 9 Oktober 2017. [Online]. Available: <https://www.predictiveanalyticsworld.com/machinelarningtimes/business-artificial-intelligence-can-not-organization/8990/>. [Zugriff am 5 Mai 2022].
- [9] H. Aust, Das Zeitalter der Daten: Was Sie über Grundlagen, Algorithmen und Anwendungen wissen sollten, Berlin und Heidelberg: Springer, 2021.
- [10] I. Kalder, „Eine kurze Geschichte der künstlichen Intelligenz,“ in *Wie Maschinen lernen: Künstliche Intelligenz verständlich erklärt*, Wiesbaden, Springer, 2019, pp. 135-140.
- [11] A. Rogers, „What Deep Blue And AlphaGo Can Teach Us About Explainable AI,“ Forbes, 9 Mai 2019. [Online]. Available: <https://www.forbes.com/sites/forbestechcouncil/2019/05/09/what-deep-blue-and-alphago-can-teach-us-about-explainable-ai/?sh=20c1750552fd>. [Zugriff am 28 März 2022].
- [12] A. L. Samuel, „Some Studies in Machine Learning Using the Game of Checkers,“ *IBM Journal of Research and Development* 44(1/2), pp. 210-229, 1959.
- [13] T. M. Mitchell, Machine Learning, Singapore: McGraw Hill Education, 1997.
- [14] S. Seegerer, T. Michaeli und R. Romeike, „So lernen Maschinen!,“ *LOG IN: Vol. 40, No. 1*, pp. 27-31, 2020.

- [15] T. Jo, Machine Learning Foundations: Supervised, Unsupervised, and Advanced Learning, Cham: Springer, 2021.
- [16] E. Alpaydin, Maschinelles Lernen, Berlin/Boston: De Gruyter Oldenbourg, 2019.
- [17] N. Castle, „Regression vs. Classification Algorithms,“ Oracle, 8 März 2018. [Online]. Available: <https://blogs.oracle.com/ai-and-datasience/post/regression-vs-classification-algorithms>. [Zugriff am 29 März 2022].
- [18] J. Kossen und M. E. Müller, „Regression,“ in *Wie Maschinen lernen: Künstliche Intelligenz verständlich erklärt*, Wiesbaden, Springer, 2019, pp. 39-43.
- [19] A. Géron, Praxiseinstieg Machine Learning mit Scikit-Learn, Keras und TensorFlow: Konzepte, Tools und Techniken für intelligente Systeme, Heidelberg: O'Reilly, 2020.
- [20] L. Wuttke, „Training-, Validierung- und Testdatensatz,“ datasolut, 24 Juli 2020. [Online]. Available: <https://datasolut.com/wiki/trainingsdaten-und-testdaten-machine-learning/>. [Zugriff am 30 März 2022].
- [21] SAP, „Was ist maschinelles Lernen?,“ SAP, [Online]. Available: <https://www.sap.com/germany/insights/what-is-machine-learning.html>. [Zugriff am 30 März 2022].
- [22] J. Kossen, M. E. Müller und M. Ruckriegel, „Entscheidungsbäume,“ in *Wie Maschinen lernen: Künstliche Intelligenz verständlich erklärt*, Wiesbaden, Springer, 2019, pp. 111-118.
- [23] S. Rosebrock, Geometrische Gruppentheorie: Ein Einstieg mit dem Computer. Basiswissen für Studium und Mathematikunterricht, Wiesbaden: Vieweg, 2004.
- [24] Upfront Ventures, „Mark Cuban Interview by Jason Hirschhorn | Upfront Summit 2017,“ 08 Februar 2022, [Online]. Available: https://www.youtube.com/watch?v=RtZ_H_aSTCI&list=PLavjJa4-6tbplWvLGw-LphcRA_2coZuRx. [Zugriff am 04 April 2022].
- [25] Ministerium für Kultus, Jugend und Sport Baden-Württemberg, „GEMEINSAMER BILDUNGSPLAN DER SEKUNDARSTUFE I. Aufbaukurs Informatik,“ 31 Mai 2018. [Online]. Available: http://bildungsplaene-bw.de/site/bildungsplan/get/documents/lbw/export-pdf/depot-pdf/ALLG/BP2016BW_ALLG_SEK1_INF7.pdf. [Zugriff am 21 April 2022].
- [26] Ministerium für Kultus, Jugend und Sport Baden-Württemberg, „GEMEINSAMER BILDUNGSPLAN DER SEKUNDARSTUFE I. Wahlfach Informatik an der Hauptschule, Werkrealschule und Realschule,“ 28 Juli 2018. [Online]. Available: http://bildungsplaene-bw.de/site/bildungsplan/get/documents/lbw/export-pdf/depot-pdf/ALLG/BP2016BW_ALLG_SEK1_INFWF.pdf. [Zugriff am 21 April 2022].
- [27] D. Touretzky, C. Gardner-McCune, F. Martin und D. Seehorn, „Envisioning AI for K-12: What Should Every Child Know about AI?,“ *Proceedings of the AAAI Conference on Artificial Intelligence*. vol. 33, p. 9795–9799, 2019.
- [28] AI for K-12, „Draft Big Idea 3 - Progression Chart,“ 2019 November 2020. [Online]. Available: <https://ai4k12.org/wp-content/uploads/2021/01/AI4K12-Big-Idea-3-Progression-Chart-Working-Draft-of-Big-Idea-3-v.11.19.2020.pdf>. [Zugriff am 16 Mai 2022].

- [29] R. F. Mager, *Preparing Instructional Objectives*, Belmont (Kalifornien): David S. Lake, 1984.
- [30] K. Ndihokubwayo, C. Byukusenge, E. Byusa, H. T. Habiyaremye, A. Mbonyiryivuze und J. Mukagihana, „Lesson plan analysis protocol (LPAP): A useful tool for researchers and educational evaluators,“ *Heliyon*, p. e08730, 2022.
- [31] J. Kossen und M. E. Müller, „Lineare Regression,“ in *Wie Maschinen lernen. Künstliche Intelligenz verständlich erklärt*, Wiesbaden, Springer, 2019, pp. 61-67.

Eidesstattliche Erklärung

Ich erkläre hiermit, dass ich die vorliegende Arbeit eigenständig und ausschließlich mit den angegebenen Quellen und Hilfsmitteln verfasst habe. Alle Textpassagen, die wörtlich oder sinngemäß auf Quellen anderer Autorinnen und Autoren beruhen, sind als solche kenntlich gemacht. Diese Arbeit wurde bisher keiner anderen Prüfungsbehörde vorgelegt und ebenfalls noch nicht veröffentlicht.

Bühl, 20.10.2022


(Sabine Britner)