

Apfel oder Popcorn? Eine enaktive Einführung in KI, maschinelles Lernen und Entscheidungsbäume mit Datenkarten (ca. 9 Ustd.)

Inhaltsfeld Inhaltliche Schwerpunkte	Übergeordnete Kompetenzerwartungen Die Schülerinnen und Schüler ...
<ul style="list-style-type: none"> – Automaten und künstliche Intelligenz <ul style="list-style-type: none"> ○ Maschinelles Lernen (mL) mit Entscheidungsbäumen – Information und Daten <ul style="list-style-type: none"> ○ Daten erläutern und repräsentieren 	<ul style="list-style-type: none"> – Argumentieren (A) <ul style="list-style-type: none"> ○ formulieren Fragen zu einfachen informatischen Sachverhalten ○ äußern Vermutungen zu informatischen Sachverhalten auf der Basis von Alltagsvorstellungen oder Vorwissen – Modellieren und Implementieren (MI) <ul style="list-style-type: none"> ○ erstellen informatische Modelle zu gegebenen Sachverhalten ○ überprüfen Modelle und Implementierungen – Darstellen und Interpretieren (DI) <ul style="list-style-type: none"> ○ beschreiben einfache Darstellungen von informatischen Sachverhalten ○ stellen informatische Sachverhalte in geeigneter Form dar ○ interpretieren informatische Darstellungen – Kommunizieren und Kooperieren (KK) <ul style="list-style-type: none"> ○ beschreiben einfache informatische Sachverhalte unter Verwendung von Fachbegriffen sachgerecht ○ anstelle der vorherigen KE: erläutern informatische Sachverhalte unter Verwendung von Fachbegriffen sachgerecht ○ strukturieren gemeinsam eine Lösung für ein informatisches Problem
<p>weitere Vereinbarungen</p> <p>... zur Vernetzung: UV 6.1: Detektivarbeit – Auseinandersetzung mit verschiedenen Verfahren zur Codierung und Verschlüsselung sowie deren Anwendung</p> <p>... zu Synergien: ggf. zum Fach Biologie (Ernährung)</p>	

Vorhabenbezogene Konkretisierung

In dieser Unterrichtsreihe geht es darum Schülerinnen und Schülern (SuS) eine Vorstellung von maschinellem Lernen und Künstlicher Intelligenz zu vermitteln. Dies wird anhand von datenbasierten Entscheidungsbäumen erarbeitet. Die Umsetzung in dieser Reihe basiert hauptsächlich auf unplugged Materialien, die das handlungsorientierte Lernen auf enaktiver Ebene ermöglichen. Dies wird ergänzt durch eine digitale Lernumgebung, die zum Ende der Reihe flexibel einsetzbar ist. Der ausgewählte Kontext „Lebensmittel“ ist für alle SuS relevant und insbesondere auch für jüngere SuS geeignet.

Lebensmittel kann man anhand von Nährwertangaben als „eher empfehlenswert“ oder „eher nicht empfehlenswert“ klassifizieren. Dabei müssen mehrere Merkmale wie Fettgehalt, Zuckergehalt und Kalorien berücksichtigt werden. Ein mehrstufiges Regelsystem, mit dem solche Klassifikationen durchgeführt werden können, sind sogenannte Entscheidungsbäume (engl. decision trees). Solche Entscheidungsbäume kann man basierend auf Daten erstellen. Mit Daten ist hier gemeint: Man geht von einer Menge von Lebensmitteln aus, zu dem Nährwertangaben bekannt sind, und zu denen man weiß, ob sie eher empfehlenswert oder nicht empfehlenswert sind. Darauf aufbauend kann man „manuell“ schrittweise Entscheidungsbäume erstellen, die die Lebensmittel zunehmend fehlerfreier klassifizieren. Dieser Erstellungsprozess kann auch automatisiert werden, um nach bestimmten Kriterien optimale Entscheidungsregeln zu finden. Die Automatisierung erfordert, jedes Lebensmittel als „Datenkarte“ - das ist eine Liste von Zahlenwerten zu den verschiedenen Nährwertmerkmalen - digital zu repräsentieren. Ein maschinelles Lernverfahren entwickelt zu diesen Daten einen passenden (daten-basierten) Entscheidungsbaum. In der Praxis sind neben Entscheidungsbäumen auch andere Typen von Klassifikatoren - z.B. neuronale Netze - im Gebrauch, mit darauf angepassten maschinellen Lernverfahren.

Entscheidungsbäume haben den Vorteil, dass sie als Regelsystem von SuS verstanden werden können, ebenso können die Erstellungsverfahren eines Baumes zunächst manuell erarbeitet und dann am Computer automatisiert werden. Im Unterricht werden Lebensmittel zunächst als reale Datenkarten modellhaft repräsentiert und die SuS können Karten sortieren und klassifizieren, um sich auf einer enaktiven Ebene Verfahren anzueignen. Der Anspruch ist, einen Einblick „in den Maschinenraum“ des maschinellen Lernens zu gewinnen und nicht nur vorgegebene Systeme, die eine völlige Black-Box bleiben, als Klassifikatoren mit Daten zu trainieren.

In dieser Unterrichtsreihe wird in ca. 9 Unterrichtsstunden in datenbasierte Entscheidungsbäume eingeführt. Dabei steht im Vordergrund, wie ein Entscheidungsbaum aufgebaut ist und wie die passenden Entscheidungsregeln datenbasiert hergeleitet werden. Dieser systematische, datenbasierte Erstellungsprozess kann dann als eine Methode des maschinellen Lernens automatisiert erfolgen und ein resultierender Entscheidungsbaum kann als eine Form künstlicher Intelligenz bezeichnet werden. Dazu erstellen SuS manuell mit Hilfe von Datenkarten (siehe Abb. 1) eigene Entscheidungsbäume, um zu verstehen, erstens wie ein Entscheidungsbaum als Regelsystem aufgebaut ist, und zweitens wie man systematisch bei der Konstruktion vorgehen kann, um Entscheidungsbäume mit möglichst geringer Fehlklassifikationsanzahl zu erhalten. Ergänzend gibt es eine vorbereitete digitale Lernumgebung, in der SuS Entscheidungsbäume automatisiert erstellen können. Dabei lernen sie etwas über Künstliche Intelligenz und maschinelles Lernen. Sie lernen Entscheidungsbäume als gewinnbringende

Repräsentation von Daten kennen, mit deren Hilfe Erkenntnisse gewonnen und Vorhersagen getroffen werden können, bei deren Anwendung aber auch Fehler passieren können.



Apfel		Popcorn	
			
Nährwerte pro 100g		Nährwerte pro 100g	
Energie	52 kcal	Energie	499 kcal
Fett	0,2 g	Fett	23,0 g
davon gesättigte Fettsäuren	0,0 g	davon gesättigte Fettsäuren	13,8 g
Kohlenhydrate	13,8 g	Kohlenhydrate	57,0 g
davon Zucker	11,0 g	davon Zucker	3,8 g
Eiweiß	0,3 g	Eiweiß	10,7 g
Salz	0,0 g	Salz	1,8 g

Abbildung 1: Beispiel Lebensmittelkarten

Auf fachlicher Basis der deutschen Gesellschaft für Ernährung (DGE) wird das Thema Ernährung aufgegriffen, welches in der Sekundarstufe I behandelt werden sollte, aber aktuell in den Lehrplänen der verschiedenen Fächer unterrepräsentiert ist. Auf diese Weise wird das Thema maschinelles Lernen mit einem bildungsrelevanten Sachthema verknüpft.

Sequenzierung

Die konkrete Sequenzierung des Unterrichts mit detaillierten Erläuterungen der einzelnen Unterrichtsstunden und allen zugehörigen, frei zugänglichen Arbeitsmaterialien finden sich hier:

<https://go.upb.de/Entscheidungsbaeume-KI5u6>

Nachfolgend findet sich eine Kurzübersicht über die Inhalte der Unterrichtsstunden:

Phase	Thema	Inhalt
1 1 U.-Std.	Einführung in den Kontext KI und Formulierung der Leitfrage	In dieser Phase wird als Einstieg ein fertiges KI-System (Google QuickDraw) exploriert. Dabei werden Vorkenntnisse aktiviert und erste Erkenntnisse über KI-Systeme und maschinelles Lernen gewonnen. Ziel ist es Klassifizieren von Objekten als eine Anwendung von KI einzuführen und zu erarbeiten, dass dies durch (überwachtes) maschinelles Lernen möglich ist, das auf Daten basiert. Abschließend wird als eine Leitfrage der Unterrichtsreihe die Frage aufgeworfen, wie maschinelles Lernen basierend auf Daten funktionieren kann.
2 1 U.-Std.	Exkurs: Lebensmittel (optional)	In dieser Phase erfolgt ein Einstieg in den Datenbegriff und das Thema Lebensmittel. Dazu wird die Bedeutung verschiedener Nährstoffe thematisiert. Verschiedene Lebensmittel können durch Datenkarte repräsentiert werden. Nährwertangaben als Ausprägungen von Merkmalen von Lebensmitteln aufgefasst und somit als als Daten interpretiert die jeweils ein Lebensmittel modellieren.
3 1 U.-Std.	Vorbereiten der Daten: Datenkarten mit Label versehen	Die zu entwickelnde künstliche Intelligenz soll später basierend auf Nährwertangaben vorschlagen, ob ein Lebensmittel eher empfehlenswert oder eher nicht empfehlenswert ist. Wenn ein KI-System wie in diesem Fall mit überwachtem maschinellem Lernen erstellt wird, dann werden dafür Beispiele für eher empfehlenswerte oder eher nicht empfehlenswerte Lebensmittel als Trainingsdaten benötigt. In dieser Stunde werden solche Beispiele in Form von Daten mit Hilfe der Datenkarten hergestellt.
4 1 U.-Std.	Einführung in das Aufstellen datenbasierter Entscheidungsregeln	Ziel innerhalb der Unterrichtsreihe ist es, ein mehrstufiges Regelsystem zur Klassifikation von Lebensmitteln zu erstellen. Dafür lernen die SuS in dieser Stunde zunächst Entscheidungsregeln (einstufige Entscheidungsbäume) aus den Daten abzuleiten. Dies wird mit dem Konzept des Datensplits umgesetzt, das in dieser Stunde eingeführt wird. Dabei werden die Datenkarten anhand eines Merkmals und eines sogenannten Schwellenwerts in zwei Teilgruppen aufgeteilt (z.B. Lebensmittel mit bis zu 10 g Fett und über 10 g Fett). Die Entscheidungsregeln werden in dieser Stunde exemplarisch mit Hilfe des Merkmals Energie erstellt.

5 2 U.-Std.	Erstellen einstufiger Entscheidungsbäume	<p>Die SuS wissen nun, wie man eine Entscheidungsregel aufstellen kann. Bisher wurden aber nur ein Merkmal und drei Schwellenwerte ausprobiert. Um den resultierenden Entscheidungsbaum zu verbessern und noch mehr Lebensmittelkarten richtig zu klassifizieren, wird in dieser Stunde erarbeitet, wie man systematisch nach guten Entscheidungsregeln suchen kann. Dabei werden verschiedene Merkmale einbezogen und ein strategisches Vorgehen beim Suchen des Schwellenwertes erarbeitet.</p>
6 1 U.-Std.	Erstellen mehrstufiger Entscheidungsbäume	<p>Nachdem die SuS systematisch nach guten Entscheidungsregeln gesucht haben, ist zu erkennen, dass man mit einer Entscheidungsregel nicht alle Lebensmittel korrekt klassifizieren kann. Es wird motiviert, dass man ein mehrstufiges Regelsystem benötigt. Deshalb werden in dieser Stunde basierend auf der ersten Regel weitere Merkmale einbezogen, um Entscheidungsregeln in der zweiten Stufe des Baums zu erstellen. Je nachdem wie schnell die SuS arbeiten, können sie zwei- oder mehrstufige Entscheidungsbäume erstellen.</p>
7 1 U.-Std.	Testen von Entscheidungsbäumen mit neuen Daten	<p>Nachdem verschiedene Gruppen von SuS unterschiedliche Entscheidungsbäume erstellt haben, werden diese Bäume auf neue Lebensmittel angewendet, die die SuS auf Blankokarten selbst mitgebracht haben. Die selbst mitgebrachten Lebensmittel werden mit allen Bäumen klassifiziert. Dadurch wird offensichtlich, dass in den Entscheidungsbäumen auch Unsicherheiten stecken, da einige Lebensmittel falsch klassifiziert werden. Um die Unsicherheiten in Entscheidungsbäumen systematisch weiter zu untersuchen, testet jede Gruppe ihren Entscheidungsbaum mit den 15 Testkarten, die im Katenspiel als gelbe Karten gekennzeichnet sind. So wird es möglich, die Performance der Entscheidungsbäume miteinander zu vergleichen.</p>
8 1 U.-Std.	Automatisiertes Erstellen von Entscheidungsbäumen mit dem Computer	<p>Nachdem die SuS den Erstellungsprozess eines Entscheidungsbaums kennengelernt und manuell durchlaufen haben, wird in dieser Stunde im eigentlichen Sinne des maschinellen Lernens ein Entscheidungsbaum automatisch mit Hilfe eines Computers erstellt. In einer vorbereiteten menübasierten digitalen Lernumgebung (basierend auf Jupyter Notebook), die per Link erreichbar ist, können die SuS selbst mit Entscheidungsbäumen arbeiten oder die Lehrkraft kann den Prozess demonstrieren. Die Lernumgebung ist weitgehend selbsterklärend und erfordert keine Programmierkenntnisse. Es ist möglich per Knopfdruck Einfluss auf die Daten zu nehmen und einen Entscheidungsbaum basierend auf den Daten erstellen lassen. Dies ermöglicht auch die Daten zu verändern und die Auswirkungen auf den Entscheidungsbaum zu untersuchen.</p>
9 1 U.-Std.	Reflexion	<p>Abschließend wird reflektiert, wie mit Hilfe des Computers aus Daten Entscheidungsbäume erstellt werden und dass das ein Beispiel (unter vielen) für maschinelles Lernen ist. Es wird darauf eingegangen welche Vor- und Nachteile solche Entscheidungsmodelle haben und wo die SuS in ihrem Alltag ähnliche Entscheidungsmodelle wiederfinden.</p>