Unterrichtsverlauf:

|  |  |  |
| --- | --- | --- |
| **Phase** | **Inhalt** | **Material** |
| **1**  **45 min** | **Thema: Einführung in den Kontext KI und Problematisierung**  **Zwecke der Phase**   * SuS werden motiviert, herausfinden zu wollen, wie eine KI funktioniert. * SuS gewinnen erste Erkenntnisse über KI-Systeme und maschinelles Lernen:   + Ein Beispiel für KI-Systeme sind Klassifikatoren: Sie können Objekte einer Klasse zuordnen.   + Für maschinelles Lernen benötigt man Trainingsdaten: Das sind Beispiele mit dem passenden Label einer Klasse.   **Kernaktivität der Phase**   * Die Lehrkraft demonstriert kurz Google Quick Draw (<https://quickdraw.withgoogle.com/>), ein fertiges KI-System, das Bilder klassifizieren kann.   + Arbeitsauftrag: Alle SuS malen sechs Objekte, die die KI zu erkennen versucht. Die Webseite ist nach Aufrufen des Links selbsterklärend. (20 min) * Wichtige Aspekte für die anschließende Besprechung im Plenum mit Präsentation 1:   + Folie 2: Funktionsweise der KI besprechen: Die KI klassifiziert die Objekte korrekt, wenn sie passend gezeichnet werden.   + Folien 3-4: Die KI nutzt Trainingsdaten: Nach dem Klassifizieren wurde automatisch eine Übersicht angezeigt, was erkannt bzw. nicht erkannt wurde. Außerdem werden einige Beispiele gezeigt, die andere Personen zum gleichen Objekt gemalt haben, um die Datenbasis zu veranschaulichen, auf deren Grundlage die gemalten Objekte klassifiziert wurden. Hier wird der Fachbegriff Label eingeführt.   + Folie 5: Motivieren der Leitfragen der Unterrichtsreihe: Wie können KI-Systeme Objekte klassifizieren? Welche Rolle spielen Daten und maschinelles Lernen dabei? * Für weitere Hintergrundinformationen siehe Dokument Informationen 1 zu dieser Phase.   Wortspeicher: **KI, Maschinelles Lernen, Objekt, Klasse, Label, Merkmal** | Präsentation 1  Information 1 |
| **2** | **Thema: Exkurs Lebensmittel und empfehlenswerte Ernährung**  Falls man einen vertieften Einstieg in das Thema Lebensmittel machen möchte, findet man in den Dokumenten Information 2 und Nutri-Score.pdf dazu Anregungen. Dies ist optional. Es kann sich außerdem eine Kooperation mit dem Biologieunterricht anbieten. | Information 2  Nutri-Score.pdf |
| **3**  **45 min** | **Thema: Vorbereiten der Daten: Datenkarten mit Labeln versehen**  **Zwecke der Phase**   * in Klassifikationsproblem für Lebensmitteldaten wird aufgeworfen. Es soll unterschieden werden zwischen „eher empfehlenswerten” und „eher nicht empfehlenswerten” Lebensmitteln. * Der Trainingsdatensatz für den weiteren Gebrauch in der Unterrichtsreihe wird hergestellt und im Plenum konsentiert. * Eine Verknüpfung zur Leitfrage der Unterrichtsreihe wird hergestellt. Thema dieser Unterrichtsstunden ist, wie ein KI-System als Klassifikator basierend auf Daten erstellt werden kann. Dies wird am Beispiel von Lebensmitteldaten und dem aufgeworfenen Klassifikationsproblem erarbeitet.   **Kernaktivitäten der Phase**   * Die Lehrkraft leitet die Problemstellung „Lebensmittel klassifizieren” ein und stellt einen Bezug zum vorher erkundeten Klassifikationsproblem her. Das Ziel des Ganzen ist, dass aus den Nährwertangaben ein Regelsystem von den SuS erstellt wird, mit dem entschieden werden kann, ob ein Lebensmittel „eher empfehlenswert“ oder „eher nicht empfehlenswert“ ist. Folgende Punkte sind dabei wichtig:   + Genau wie wir es bei Quickdraw gesehen haben, benötigen wir eine Menge an Objekten, die schon passende Label. haben. Im Beispiel der Lebensmittel sind es keine gemalten Bilder, sondern Lebensmittel mit Daten und mit einem Label „eher empfehlenswert“ oder „eher nicht empfehlenswert“. Wir nennen die Sammlung von Lebensmitteln mit Label: Trainingsdaten.   + Basierend auf den Trainingsdaten wollen wir dann selbst ein System herstellen, das diese Lebensmittel möglichst fehlerfrei klassifiziert.   + Später diskutieren wir, wie Maschinen solche Systeme herstellen können. * Die Lehrkraft formuliert die Aufgabe für diese Phase: Wir stellen solche Beispiele mit Label (=Trainingsdaten) gemeinsam her, um damit später unser KI-System zu trainieren.   + Die SuS bearbeiten in Partnerarbeit Arbeitsblatt 1, auf dem Bilder der Lebensmittel abgebildet sind. Sie vergeben Label basierend auf ihren eigenen Erfahrungen und entscheiden, welche Lebensmittel sie für eher empfehlenswert halten und welche sie für eher nicht empfehlenswert halten.   + Wenn alle SuS AB 1 fertiggestellt haben, werden die Kartenspiele in der Klasse verteilt. Ein Kartenspiel sowie je 50 grüne und rote Büroklammern pro Paar werden ausgegeben.   + In einer anschließenden Plenumsbesprechung wird ein allgemeiner Konsens hergestellt, welche Nahrungsmittel welche Label erhalten. Dabei werden nur die blauen Karten im Deck verwendet, da diese für die Trainingsdaten vorgesehenen sind. Die gelben Karten werden erst später als Testdaten relevant.   + Jedes SuS-Paar labelt seine Karten entsprechend mit den farbigen Büroklammern.   **Didaktische Hinweise**  Für das Konsentieren des Kartensatzes in Daten mit Label hat sich ein stringentes Vorgehen als sinnvoll in mehreren Erprobungen herausgestellt. Dazu ein Vorschlag:  Alle Karten werden einzeln nacheinander durchgegangen (Lehrkraft hält Karte hoch, alle SuS nehmen die gleiche Karte), um gemeinsam die Label (Büroklammern) an die Karten zu heften. Ablauf:   1. Jedes SuS-Paar bekommt ein Kartenspiel. Die Karten sollten auf jeden Fall in der Reihenfolge bleiben, wie sie in der Schachtel sind, damit alle SuS die Karten in der gleichen Reihenfolge vorliegen haben. 2. Im Plenum wird besprochen und festgelegt, welches Label die erste blaue Datenkarte (Haselnussschnitte) bekommen soll. Zunächst kann einfach abgestimmt werden. Wenn keine klare Mehrheit zustande kommt, wird diskutiert. 3. Falls Einigkeit erzielt wird, stecken alle SuS eine passend farbige Büroklammer an die Datenkarte. Falls auch nach der Diskussion keine Einigkeit erzielt wird, kann die Karte beiseitegelegt werden (diese wird in Phase 6 als neues Objekt verwendet). 4. 2. und 3. werden für alle blauen Karten wiederholt.   Leitung der Diskussion: Die Lehrkraft kann die Diskussionen im Plenum moderieren und ggf. bei groben Fehleinschätzungen (z. B. Gurke als „eher nicht empfehlenswert”) eingreifen. Eine Orientierung liefert die Datei Rumpfdatensatz\_28.csv, in der die besonders eindeutig zuzuordnenden Lebensmittel zu finden sind. Einzelne Abweichungen von diesem Vorschlag sind aber nicht weiter schlimm, d. h. die Diskussion muss nicht strikt geleitet werden. In vielen Unterrichtserprobungen hat sich gezeigt, dass per Mehrheitsvotum der Großteil der Lebensmittel wie im Vorschlag von SuS zugeordnet wird. Zur Weiterarbeit empfehlen sich ca. 30 Karten mit Label zu versehen (etwa 10 Karten können beiseitegelegt werden). Die Karten aus Rumpfdatensatz\_28.csv sollten größtenteils enthalten sein, um in der Weiterarbeit gute Ergebnisse zu erzielen.  AB 1 kann im späteren Unterrichtsverlauf als Erinnerung dienen, welche Karte in welcher Farbe gelabelt wurde, falls zwischen den Unterrichtsstunden Label von den Karten abgefallen sind.  Für weitere Hintergrundinformationen siehe Dokument „Informationen2“ zu dieser Phase.  Zum Aufbewahren der Karten eignet sich pro Gruppe ein großer Briefumschlag, eine Klarsichthülle oder ein Schuhkarton. Es lohnt eine Beschriftung mit den Namen der SuS.  Wenn die Büroklammern von den Karten beim nächsten Herausholen der Karten abgefallen sind, kann AB1 als Referenz genutzt werden, um wieder die korrekte Büroklammer an eine Karte zu heften. Deshalb sollten überzählige Büroklammern mit aufbewahrt werden.  Wortspeicher: **Daten, Beispiel, Objekt, Label, Trainingsdaten** | Arbeitsblatt 1 (in Farbe ausdrucken)  Datenkarten + 50 rote & 50 grüne  Büroklammern  Rumpfdatensatz  Klarsichthüllen |
| **4**  **45 min** | **Thema: Einführen des Datenbegriffs und datenbasierter Entscheidungsregeln**  **Zwecke der Phase**   * Ein Vokabular zum Sprechen über Daten am Beispiel Lebensmittel wird eingeführt. Dazu begreifen wir Nährwertangaben als Ausprägung von Merkmalen von Lebensmitteln. Somit kann ein Lebensmittel durch eine Liste von Daten (Zahlen als Nährwertangaben) repräsentiert werden. * Das Aufstellen einer Entscheidungsregel anhand von Daten wird eingeführt, um Lebensmittel zu klassifizieren. Dafür werden folgende wichtige Konzepte eingeführt: Datensplit, Mehrheitsentscheidung, Fehlklassifikation, Darstellung als Entscheidungsbaum(-diagramm). (für Hintergrundinformationen zu diesen Konzepten siehe [hier](https://www.prodabi.de/materialien/entscheidungsbaeume/))   **Kernaktivität der Phase**   * Die Lehrkraft leitet ein, dass für das gemeinsame Arbeiten ein gemeinsames Vokabular benötigt wird und erläutert an einer Datenkarte die Begriffe Objekt, Merkmal, Ausprägung und Label. * SuS bearbeiten zu den Begrifflichkeiten das Arbeitsblatt 2. * Besprechen von Arbeitsblatt 2 * Anschließend wird das Aufstellen von Entscheidungsregeln eingeführt anhand von Präsentation 2.   + Die Lehrkraft gibt folgenden Impuls: "Vielleicht kann man schon mit einem Merkmal ganz gut vorhersagen, ob das Lebensmittel empfehlenswert ist oder nicht. Wir probieren einen kleinen Datensatz aus und schauen nach dem Merkmal Energie. Um einen Überblick zu bekommen, ordnen wir die Karten nach Energie.“   + Um dies zu zeigen, nutzt die Lehrkraft Präsentation 2 . In Präsentation 2 wird ein Minibeispiel mit 6 Lebensmitteln gezeigt. Im gezeigten Beispiel wird zuerst der Optimalfall dargestellt, dass in beiden Teildatensätzen alle Label die gleiche Farbe haben.   + Entdeckung: Ab einem gewissen "Schwellenwert" sind alle Karten über dem Schwellenwert nicht empfehlenswert. Anhand des Schwellenwerts wird ein "Datensplit" durchgeführt (Einteilung in zwei Gruppen) und die naheliegende Entscheidungsregel wird formuliert.   + Die Lehrkraft hält fest, dass diese perfekte Trennung in empfehlenswert und nicht empfehlenswert die *Zielstellung* für einen Datensplit ist, die aber nicht durch jeden Datensplit erreicht wird.   + Anschließend wird mit zwei zusätzlichen Lebensmitteln ein Beispiel gezeigt, bei dem man eine Entscheidungsregel nur so aufstellen kann, dass auch manche Lebensmittel falsch klassifiziert werden. Dies ist der Normalfall. An diesem Beispiel werden die Mehrheitsentscheidung und der Begriff Fehlklassifikation eingeführt. Ziel ist es, mit einer Entscheidungsregel möglichst wenig Fehlklassifikationen zu erreichen. * Das vorher in der Präsentation gesehene wird nun durch eine Aktivität vertieft. Es wird im Plenum eine *Lebendige Statistik* durchgeführt, um die Entscheidungsregel aus Präsentation 2 mit noch mehr Daten zu überprüfen. Dabei werden die Konzepte *Datensplit*, *Schwellenwert, Mehrheitsentscheidung in den beiden Teildatensätzen* und *Fehlklassifikation* vertieft und gefestigt. Alle SuS repräsentieren je eine Lebensmittelkarte, indem sie sie festhalten (d. h. es werden so viele Karten genutzt, wie SuS in der Klasse sind). Es wird der Datensplit aus der vorherigen Präsentation betrachtet (Energie, Schwellenwert 260 kcal). Für die genaue Durchführung der Lebendigen Statistik siehe **Informationen 3**. Die Lehrkraft notiert den resultierenden einstufigen Entscheidungsbaum an der Tafel. * Die SuS gehen zurück an ihre Plätze und im Plenum wird gemeinsam anhand der Dokumentation an der Tafel für den Entscheidungsbaum die Anzahl der Fehlklassifikationen ermittelt. * Eine anschließende Besprechung dient als Zusammenfassung der bisherigen Phase und als Ausblick auf weitere Inhalte der Unterrichtsreihe. Folgende Punkte sollten in der Diskussion vorkommen: * Eine einzelne Entscheidungsregel ist ein kleiner Entscheidungsbaum mit nur einer Ebene. * Wichtige gelernte Inhalte bisher sind die Schritte zur Aufstellung einer Entscheidungsregel:   + Ein *Datensplit* basiert auf einem bestimmten Merkmal und einem Schwellenwert   + In den Teildatensätzen geschieht eine Mehrheitsentscheidung   + Auszählen von falsch klassifizierten Lebensmitteln (= Anzahl der Fehlklassifikationen) * Der Entscheidungsbaum ist schon einmal ganz gut, da er nur wenige Fehler macht. In der nächsten Phase werden nun noch mehr einstufige Entscheidungsbäume überprüft, um den besten zu finden. * Später werden noch weitere Ebenen von Entscheidungsregeln dazukommen. Die Bäume werden “wachsen”. * Hausaufgabe: * Übungen zum Nutzen von Schwellenwerten in der Entscheidungsbaumdarstellung (Arbeitsblatt 3) * optionale weitere Hausaufgabe: Mathematische Wiederholung <, >, ≤, ≥ Zeichen (z. B. https://anton.app/de/lernen/mathematik-5-klasse/thema-01-natuerliche-und-ganze-zahlen/uebungen-04-zahlen-ordnen-vergleichen/)   Wortspeicher: **Objekt, Merkmal, Merkmalsausprägung, Label, Klassifikation, Schwellenwert, Datensplit, Anzahl der Fehlklassifikationen** | Arbeitsblatt 2  Präsentation 2  Information 3  Hausaufgabe: Arbeitsblatt 3 |
| **5**  **90 min** | **Thema: Erstellen guter einstufiger Entscheidungsbäume**  **Zwecke der Phase:**   * Das Aufstellen einer Entscheidungsregel mit Hilfe von Datenkarten wird eingeübt. Dabei werden die Konzepte *Datensplit*, *Schwellenwert*, *Mehrheitsentscheidung* vertieft. * Das Vergleichen von Entscheidungsregeln wird eingeübt. Dabei wird das Konzept *Fehlklassifikation* vertieft. * SuS lernen eine Heuristik kennen, wie man bei einem vorgegebenen Merkmal einen günstigen *Schwellenwert* auswählt, um eine Entscheidungsregel aufzustellen.   **Kernaktivitäten der Phase**  Aktivität 1: Vorgehensweise für das Finden guter Schwellenwerte erarbeiten   * Als erster Schritt wird das Sortieren von Datenkarten als Hilfsmittel eingeführt. Durch das aufsteigende Sortieren können schnell verschiedene Datensplits ausprobiert werden, ohne dass man immer zwei Haufen von Datenkarten bilden muss (wie in der lebendigen Statistik). Die *Linealmethode* (siehe Information 4) kann mit sortierten Karten angewandt werden. * Die Linealmethode wird im Plenum basierend auf Präsentation 3 erarbeitet. Dabei kann man schon Strategiehinweise geben, wie man einen *Schwellenwert per Augenmaß* (siehe Dokument Information 4) wählen kann. * Anhand von Arbeitsblatt 4 wird die Linealmethode in Einzelarbeit eingeübt, um die Methode danach in Kleingruppen mit den Datenkarten anzuwenden. * Im Plenum werden die Ergebnisse von AB4 verglichen und ggf. korrigiert. Die SuS können ihre Begründungen vortragen und diskutieren. Bei der Besprechung sollte insbesondere auf folgende Aspekte geachtet werden:   + Was ist ein Kriterium für einen „besten“ Schwellenwert? (Anzahl falsch klassifizierte Karten/Fehlklassifikationen)   Aktivität 2: Beste Entscheidungsregel mit Datenkarten suchen   * In Partnerarbeit arbeiten die SuS je mit ihrem eigenen Kartensatz mit Label. Jedes Paar bekommt von der Lehrkraft ein Merkmal zugeteilt, für das eine Entscheidungsregel gesucht wird. Jedes Merkmal sollte an mindestens ein Paar vergeben werden, um im Nachhinein die Entscheidungsregeln basierend auf allen verschiedenen Merkmalen vergleichen zu können.   + Aufgabe: Jedes Paar sortiert die Karten nach dem ihm zugeteilten Merkmal, zieht mehrere Schwellenwerte (mindestens drei) in Betracht und begründet, welcher der beste Schwellenwert unter den in Betracht gezogenen ist. Es sollen mindestes drei in Betracht gezogen werden, um das Vergleichen zu üben. Es wird nicht vorausgesetzt, alle mögliche Schwellenwerte zwischen je zwei Karten zu betrachten, weil dies für einige SuS überfordernd sein kann. Optional kann Spielplan 1 als Unterstützung des Prozesses eingesetzt werden. Die Dokumentation erfolgt auf [Arbeitsblatt 5](https://unterrichtsmaterial-ddi.cs.upb.de/images/4/4a/AB5_Entscheidungsregeln_Verlgeichen.docx). * Besprechung: Alle Paare präsentieren die jeweils beste gefundene Entscheidungsregel. Dabei werden Merkmal, Schwellenwert und Anzahl der Fehlklassifikationen genannt. Somit wird im Klassenverband die insgesamt beste Entscheidungsregel aus den untersuchten Merkmalen ermittelt und an der Tafel notiert mit Angabe des Schwellenwerts und der Fehlklassifikationen   + Impuls für eine gemeinsame Besprechung: Was kann ein Computer schneller und besser machen als die SuS? An der Tafel werden Punkte gesammelt. Folgende Punkte sind mögliche Antworten: Ein Computer sucht auch die beste mögliche Entscheidungsregel und zieht dafür **alle** Merkmale in Betracht.   + Für jedes Merkmal ermittelt der Computer datenbasiert **alle** sinnvollen (d. h. zwischen zwei Kartenwerten liegenden) Schwellenwerte und vergleicht diese anhand der Anzahlen der Fehlklassifikationen.   + Ein Computer kann diese Vergleiche (und Berechnungen) typischerweise sehr **schnell** durchführen. * Was haben wir (SuS) gemachtim Vergleich zu einem Computer?   + Der Klassenverband hat die Vorgehensweise eines Computers fast vollständig simuliert, weil über alle Gruppen hinweg alle Merkmale betrachtet wurden. Es wurden vermutlich nicht alle möglichen, aber viele der sinnvollen Schwellenwerte betrachtet und verglichen.   + Der Computer kann diese Prozesse schneller als ein Mensch durchführen. Der Klassenverband hat z. B. eine halbe Stunde benötigt.   **Didaktische Hinweise:**  Der Umgang mit den Datenkarten ist zentral in dieser Phase. Um mit Fehlklassifikationen argumentieren zu können und damit die Güte verschiedener Schwellenwerte vergleichen zu können, ist es nötig, die Karten aufsteigend zu sortieren. Es kann nötig sein, dass Paare an zwei aneinander geschobenen Tischen arbeiten, um die Karten vollständig aufsteigend nach einem Merkmal sortieren zu können. Diese Sortierarbeit braucht einiges an Zeit und kann für manche SuS herausfordernd sein, bietet aber eine gute Gelegenheit für die Besprechung, was ein Computer schneller (und genauer) kann als ein Mensch.  Wortspeicher: **Datensplit, Schwellenwert, Mehrheitsentscheidung, Entscheidungsregel, Fehlklassifikation, Linealmethode** | [Präsentation 3](https://unterrichtsmaterial-ddi.cs.upb.de/images/6/63/Pr%C3%A4sentation3_Schwellenwert_suchen.pptx)  Information 4  Arbeitsblatt 4  Arbeitsblatt 5  Spielplan (nur Seite 1) (optional) |
| **6**  **90 min** | **Thema: Mehrstufige Entscheidungsbäume erstellen**  **Zwecke der Phase**   * Es wird motiviert mehrstufige Entscheidungsbäume zu betrachten, nicht nur einstufige, da alle einstufigen Entscheidungsbäume noch Fehler machen. Außerdem ist eine gute Ernährung multifaktoriell, sodass es Sinn ergibt, mehrere Merkmale heranzuziehen. * Die Vorgehensweise beim Erstellen der zweiten Stufe wird erläutert und geübt: Das Vorgehen auf der zweiten Stufe ist im Wesentlichen gleich zur ersten Stufe mit zwei Unterschieden. Auf der zweiten Stufe wird nur mit den Teildatensätze weitergearbeitet, die in der ersten Stufe entstanden sind, und es wird mit anderen Merkmalen als dem der ersten Stufe gearbeitet.   **Kernaktivität der Phase**   * Die Lehrkraft motiviert das Betrachten mehrstufiger Entscheidungsbäume wie folgt. Aus der vorherigen Phase ging hervor, dass man mit einstufigen Entscheidungsbäumen nicht alle Lebensmittel korrekt klassifizieren kann. Deshalb werden in dieser Phase basierend auf der ersten Regel weitere Merkmale einbezogen, um Entscheidungsregeln in der zweiten Stufe des Entscheidungsbaums zu erstellen, wodurch der Entscheidungsbaum im Hinblick auf die Anzahl der fehlklassifizierten Lebensmittel verbessert wird. * Als wichtige Ergänzung zum einstufigen Entscheidungsbaum wird eingeführt, dass in zweiter Stufe jeweils nur mit den Karten eines Teildatensatzes weitergearbeitet wird (entweder arbeitet man also mit den Karten im linken oder im rechten Ast weiter). Das kann entweder im Plenum mit Unterstützung von [Präsentation 4](https://unterrichtsmaterial-ddi.cs.upb.de/images/b/bd/Pr%C3%A4sentation4_Zweiter_Datensplit.pptx) besprochen werden oder optional durch eine weitere Aktivität in der lebendigen Statistik veranschaulicht werden (siehe **Information 5** zu dieser Phase). * Anhand von [Präsentation 4](https://unterrichtsmaterial-ddi.cs.upb.de/images/b/bd/Pr%C3%A4sentation4_Zweiter_Datensplit.pptx) erläutert die Lehrkraft die Dokumentation eines mehrstufigen Entscheidungsbaums gemäß [Arbeitsblatt 6](https://unterrichtsmaterial-ddi.cs.upb.de/images/5/5f/AB6_Zweiter_Datensplit.docx). * Wieder in Partnerarbeit arbeiten die SuS mit ihrem Kartensatz und erstellen weitere Stufen ihres Entscheidungsbaums. Jedes Paar darf wählen, ob es mit dem eigenen Entscheidungsbaum aus der letzten Phase weiterarbeitet oder den nimmt, der als bester einstufiger Entscheidungsbaum ausgewählt wurde. * Nächster Auftrag für die Paare: Verbessert den Baum, indem ihr weitere Stufen anbaut. Wählt dazu geeignete Merkmale aus. Dokumentation erfolgt auf Arbeitsblatt 6. Spielplan Teil 2 kann für die Organisation der Datenkarten genutzt werden, um z. B. nicht genutzte Karten eines Teildatensatzes passend abzulegen. Zur genauen Nutzung des Spielplans siehe **Information 5**.   + Zur Differenzierung für schnelle Gruppen: Weitere Verbesserungen durch Veränderung des Baumes     - Differenzierung 1: Eine dritte Stufe des Entscheidungsbaums wird angebaut. Hierfür kann Spielplan Teil 2 genutzt werden, indem „einzelne Datensplits” durch Zerschneiden von Spielplan Teil 2 angelegt werden. Auf AB6 kann eine dritte Stufe des Entscheidungsbaums selbst eingezeichnet werden.     - Differenzierung 2: Verschiedene Merkmale werden von SuS im zweiten Split ausprobiert.   + Alle SuS-Paare erstellen mindestens einen zweistufigen Entscheidungsbaum und dokumentieren ihn auf AB6. * Im Plenum werden die erstellten, mehrstufigen Bäume präsentiert mit Hinweis auf die benutzten Merkmale, Schwellenwerte und Gesamtanzahl der Fehlklassifikationen. Außerdem werden die Entscheidungsregeln verbal formuliert. * Die erstellten Bäume werden anhand des Kriteriums der Anzahl der Fehlklassifikationen miteinander verglichen. Alle Bäume werden wertgeschätzt und die besten drei Bäume (diejenigen mit der geringsten Anzahl an Fehlklassifikationen) werden besonders hervorgehoben. * Für die nächste Stunde werden die in dieser Phase erstellten Entscheidungsbäume (AB6) in der Klasse aufgehängt (bzw. durch Lehrkraft eingesammelt). Diese werden in der nächsten Stunde nochmal benötigt, um Testdaten darauf anzuwenden. * Hausaufgabe: Alle SuS füllen eine Blankokarte (AB8) für ein neues Lebensmittel zu Hause aus. Diese neuen Lebensmittel werden in der nächsten Stunde mit den erstellten Bäumen klassifiziert.   **Didaktische Hinweise**  Aus mehreren Erprobungen ist bekannt, dass SuS in dieser Phase sehr unterschiedlich schnell arbeiten. Das Ziel sollte es sein, dass alle SuS-Paare einen zweistufigen Entscheidungsbaum erstellen und mit der Anzahl der Fehlklassifikationen auf AB6 dokumentieren.  Ein typischer SuS-Fehler ist es, mit allen Karten in einem Ast weiterzuarbeiten, obwohl nur mit der Teilgruppe der Karten des entsprechenden Asts gearbeitet werden darf.  Wortspeicher: **Stufen eines Entscheidungsbaums** | Präsentation 4  Information 5  Arbeitsblatt 6  Spielplan (alle Seiten)  Hausaufgabe: Arbeitsblatt 8 |
| **7**  **45 min** | **Thema: Benutzen verschiedener Entscheidungsbäume zum Klassifizieren einer neuen Lebensmittelkarte**  **Zwecke der Phase**   * Die SuS benutzen erstmalig explizit einen mehrstufigen Entscheidungsbaum zum Klassifizieren eines Lebensmittels. Dabei erlernen sie das Benutzen des Entscheidungsbaums als Klassifikator. * Die SuS wenden unterschiedliche Entscheidungsbäume auf das gleiche Lebensmittel an, um festzustellen, dass verschiedene Entscheidungsbäume als Klassifikatoren unterschiedliche Entscheidungen für das gleiche Lebensmittel treffen können. * Es wird motiviert zu überprüfen, welcher Entscheidungsbaum besonders viele neue Lebensmittel korrekt klassifiziert. Dafür werden in der nächsten Phase Testdaten eingeführt.   **Kernaktivität der Phase**   * Vorbereitung: Die Lehrkraft legt (oder hängt) die erstellten Bäume (als ausgefüllte AB 6 aus der vorherigen Phase) als Stationen zusammen jeweils mit roten und grünen Büroklammern im Klassenraum aus. * Im Plenum wird anhand von Präsentation5\_Baum\_Anwenden eingeführt, wie man mit einem Entscheidungsbaum anhand einer Lebensmittelkarte ein Lebensmittel klassifizieren kann. Dies bereitet die SuS auf die nächste Aktion vor. * Jede:r SuS geht mit dem neuen Lebensmittel (der zu Hause ausgefüllten Blankokarte von AB 8) von Station zu Station. An jeder Station durchlaufen die SuS mit der Karte den jeweiligen Entscheidungsbaum und heften am Ende die passende, farbige Büroklammer an die Karte (je nachdem, ob der Entscheidungsbaum das Lebensmittel als eher empfehlenswert oder eher nicht empfehlenswert klassifiziert). Jedes neue Lebensmittel erhält so eine Vielzahl an Klassifizierungen. * Wichtige Aspekte für die gemeinsame Besprechung:   + Beobachtung: Ein Lebensmittel kann von verschiedenen Bäumen unterschiedlich klassifiziert worden sein, was an den verschiedenfarbigen Büroklammern an einzelnen Karten erkennbar ist.   + Schlussfolgerung: Verschiedene Entscheidungsbäume können das gleiche Lebensmittel unterschiedlich klassifizieren.   + Neues Ziel: Wir wollen herausfinden, welcher Entscheidungsbaum Lebensmittel am zuverlässigsten korrekt klassifiziert.   + Vorgehen: Dies kann durch den Einsatz von Testdaten herausgefunden werden.   Wortspeicher: **Testdaten, Klassifikator** | Aufgehängte Arbeitsblätter 6  Präsentation 5 |
| **8**  **45 min** | **Thema: Systematisches Testen von Entscheidungsbäumen anhand mehrerer Testkarten**  **Zwecke der Phase**   * .Die SuS wenden Testdaten auf ihren eigenen Entscheidungsbaum an, um herauszufinden, wie viele der 15 Test-Lebensmittel falsch klassifiziert werden. * Die SuS vergleichen verschiedene Entscheidungsbäume anhand der Anzahl der falsch klassifizierten Test-Lebensmittel. * Im Klassenverband wird anhand der Testdaten der Entscheidungsbaum mit der geringsten Anzahl falsch klassifizierter Lebensmittel ausgewählt.   **Kernaktivität der Phase**   * Die gelben Testkarten werden im Plenum eingeführt und mit Hilfe von Arbeitsblatt 9 mit Labeln versehen (wie bereits in Phase 1 für die blauen Trainingsdaten). * Jedes SuS-Paar testet mit den Testdaten den eigenen Entscheidungsbaum (Ergebnis aus der vorherigen Stunde auf AB6). Dafür wird jede Testkarte genommen und durch den Entscheidungsbaum klassifiziert. Für jede Karte wird notiert, ob die Klassifikation durch den Entscheidungsbaum mit dem im Plenum gewählten Label an der Karte übereinstimmt. * Am Ende kann auf AB6 folgender Satz ergänzt werden: „Mit dem Entscheidungsbaum unserer Gruppe wurden von den Testdaten \_\_\_ Lebensmittel richtig klassifiziert und \_\_\_\_ Lebensmittel falsch klassifiziert.” * Im Plenum wird verglichen, wie gut die Bäume jeweils mit Testdaten abschneiden. Das Vergleichskriterium ist die Anzahl falscher Klassifikationen durch den Entscheidungsbaum. * Abschließend werden im Plenum folgende Aspekte besprochen:   + Welches sind die drei besten Entscheidungsbäume gemessen an der Anzahl der falsch klassifizierten Test-Lebensmittel?   + Die Top 3 der besten Entscheidungsbäume gemessen an den Trainingsdaten werden mit den Top 3 der besten Entscheidungsbäume gemessen an den Testdaten verglichen.     - Aspekt für die Besprechung: Ein Entscheidungsbaum wird mit Trainingsdaten erstellt und für diese optimiert. Wenn man diesen Entscheidungsbaum für neue Daten benutzt, kann es sein, dass er nicht mehr das beste Ergebnis liefert. Daher ist es wichtig, basierend auf Daten erstellte Entscheidungsbäume immer zusätzlich mit Testdaten zu überprüfen.   **Didaktische Hinweise**  Für den Vergleich der Performance eines Entscheidungsbaums auf Trainings- und auf Testdaten wird normalerweise der Anteil falscher Klassifikationen in den Trainingsdaten mit dem Anteil in den Testdaten verglichen. Dies haben wir hier vereinfacht zur absoluten Anzahl an falschen Klassifikationen, um potentielle SuS-Schwierigkeiten beim Vergleich von Dezimalzahlen zu vermeiden. | Testkarten  Arbeitsblatt 9  Arbeitsblatt 6 (ausgefüllt aus vorheriger Stunde) |
| **9**  **90 min** | **Thema: Automatisches Erstellen von Entscheidungsbäumen mit dem Computer und Reflexion über Einsatzmöglichkeiten und Grenzen**  **Zwecke der Phase**   * Die SuS erarbeiten, wie ein Computer systematisch beim Auswählen von Schwellenwerten vorgeht. Dazu greifen sie auf ihre eigenen Erfahrungen beim Erstellen von Entscheidungsbäumen zurück. * Die SuS erstellen einen Entscheidungsbaum, im eigentlichen Sinne des maschinellen Lernens, automatisch mit Hilfe eines Computers und einer digitalen Lernumgebung. * SuS reflektieren über Einsatzmöglichkeiten und Grenzen ihrer eigenen Entscheidungsbäume.   **Kernaktivität der Phase**   * Im Plenum wird wieder aufgegriffen, was in einer früheren Phase (Ende Phase 4) dazu festgehalten wurde, wie ein Computer einen Entscheidungsbaum automatisiert erstellen kann. Dazu sind folgende Punkte wichtig:   + Ein Computer sucht zuerst die beste einstufige Entscheidungsregel, zieht dafür alle Merkmale in Betracht.   + Für jedes Merkmal ermittelt der Computer datenbasiert alle sinnvollen (d. h. zwischen zwei Kartenwerten liegenden) Schwellenwerte und vergleicht diese anhand der Anzahlen der Fehlklassifikationen.   + Anschließend wird das gleiche Vorgehen in der zweiten (dritten, vierten, …) Stufe wiederholt, sodass sich der Baum nach und nach aufbaut.   + Ein Computer kann diese Vergleiche (und Berechnungen) typischerweise sehr schnell durchführen, sodass wir nach Sekunden einen mehrstufigen Entscheidungsbaum angezeigt bekommen. Ein Mensch würde dafür viel länger brauchen. * Außerdem wird kurz darauf eingegangen, was ein Mensch zu diesem Prozess beitragen muss:   + Objekte müssen von einem Menschen mit einem passenden Label versehen werden, bevor der Computer damit arbeiten kann.   + Der Mensch muss dem Algorithmus ein paar „Anweisungen“ geben, wie z. B. die maximale Tiefe des Baums (= Anzahl der Stufen des Baums). * Anschließend an diese grobe Wiederholung erarbeiten die SuS noch einmal detailliert, wie ein Computer für ein Merkmal und gegebene Daten alle sinnvollen Schwellenwerte testet und darauf aufbauend entscheidet. Dies wird auf [Arbeitsblatt 10](https://unterrichtsmaterial-ddi.cs.upb.de/images/d/dd/AB10_Systematik.docx) bearbeitet und im Plenum besprochen. * Nachdem das Vorgehen eines Computer erarbeitet wurde, arbeiten die SuS in einer digitalen Lernumgebung, um Entscheidungsbäume automatisch von einem Computer erstellen zu lassen. Dabei durchlaufen sie vorbereitend verschiedene Schritte, die sie in den Phasen zuvor selbst manuell mit den Datenkarten durchgeführt hatten (Label auswählen, Sortieren, Entscheidungsregeln manuell erstellen). Im Dokument **Informationen 8.1** sind weitere Hinweise zum Umgang mit der Lernumgebung zu finden. * Im Plenum wird nacheinander über zwei wichtige Aspekte reflektiert (detailliertere Vorschläge in **Information 8.2**):   + Vergleich des Erstellungsprozesses eines Entscheidungsbaums manuell mit Datenkarten und automatisch mit dem Computer   + Reflexion über Grenzen und Einsatzmöglichkeiten der selbst erstellten Entscheidungsbäume im Hinblick auf eine gesunde Ernährung | Arbeitsblatt 10  Information 8.1  Lernumgebung  <https://go.upb.de/auto-baum>  Computer  Information 8.2 |
| **Evaluation**  (1/2 U-Std.) | **Bitte um Evaluation**  Bitte als Abschluss der Unterrichtsreihe den SuS die Evaluation (Rückmeldung) geben mit dem Hinweis, dass dies eine anonyme Rückmeldung an die „Erfinder:innen“ der Unterrichtsreihe ist. Die „Erfinder:innen“ sind sehr daran interessiert zu erfahren, wie die Reihe den SuS gefallen hat.  Eine Umfrage zur Evaluation findet sich digitalisiert auf einem Server datenschutzrechtlich unbedenklich der Universität Paderborn: <https://go.upb.de/ev-lebensmittel>  Herzlichen Dank dafür! | Online Umfrage  <https://go.upb.de/ev-lebensmittel> |