

# AI Unplugged

Wir ziehen künstlicher  
Intelligenz den Stecker

Aktivitäten und Unterrichtsmaterial  
zu künstlicher Intelligenz ohne Strom

# Vorwort

Nicht erst seit der Veröffentlichung der KI-Strategie der Bundesregierung Ende 2018 gewinnt das Thema künstliche Intelligenz (KI, engl. AI für Artificial Intelligence) zunehmend an gesellschaftlicher Bedeutung. Schon heute interagieren wir wie selbstverständlich mit KI-Systemen, etwa wenn wir Sprachassistenten wie Siri oder Alexa nutzen. Dennoch wissen laut Umfragen über 50% der Deutschen bisher nicht, was künstliche Intelligenz ist.

Damit das nicht so bleibt, finden Sie hier eine Sammlung verschiedener Unplugged Aktivitäten, die rund um das Thema KI angesiedelt sind. Unplugged Aktivitäten stellen Ansätze bereit, mit denen Lernende jeden Alters Ideen und Konzepte der Informatik enaktiv erfahren können und die explizit auf den Einsatz des Computers verzichten. Die vorliegende Broschüre umfasst fünf Aktivitäten, mit denen Sie Ideen und Konzepte, die dem Thema künstliche Intelligenz zugrunde liegen, zielgruppengerecht vermitteln können.

Im Allgemeinen wird KI heute vor allem über maschinelles Lernen realisiert, aber künstliche Intelligenz ist weit mehr als das. Das Thema KI umfasst nicht nur technische Aspekte, sondern wirft auch Fragen mit gesellschaftlicher Relevanz auf. Diese Broschüre zeigt Möglichkeiten, wie die genannten Themen mit Kindern, Jugendlichen und auch Erwachsenen diskutiert werden können.

Bei Fragen, Kommentaren oder Anmerkungen zum vorliegenden Material zögern Sie nicht, uns unter [aiunplugged@digi4all.de](mailto:aiunplugged@digi4all.de) zu kontaktieren.

Die Aktivitäten benötigen teilweise zusätzliches Material. Druckvorlagen finden sie gesammelt unter:

<https://aiunplugged.org>





# Inhaltsverzeichnis

Klassifikation mit Entscheidungsbäumen	4
#deeplearning	8
Reinforcement Learning	12
Back to the Roots	16
Der Turing-Test	20
Weitere Anregungen	23

# **Klassifikation mit Entscheidungsbäumen**

## **Das Gute-Äffchen-Böse-Äffchen-Spiel**

### **Zielgruppe**

Primarstufe, Sekundarstufe I, Sekundarstufe II

### **Darum geht's**

Wie trifft ein Computer selbstständig Entscheidungen? Wie entscheidet ein Computer, ob eine Person sportlich ist, einen Kredit erhalten sollte usw.? Solche Klassifikationsvorgänge sind eine häufige Anwendung von KI. In dieser Aktivität haben Schülerinnen und Schüler die Möglichkeit, selbst ein Klassifikationsmodell mithilfe eines Entscheidungsbaumes zu erstellen. Aus den verschiedenen Modellen wird dann das beste Modell ausgewählt.

### **Diese Ideen stecken dahinter**

- KI klassifiziert Daten auf Basis von Mustern.
- KI nutzt das Klassifikationsmodell, das am besten auf gegebene Daten passt.
- Klassifikationsmodelle sind nicht perfekt.
- Bestimmte Merkmalskombinationen lassen auf eine bestimmte Kategorie schließen.

### **Das wird benötigt**

- Äffchenkarten (Vorbereitung: Ausschneiden aus Vorlage A0, alternativ digitale Version)
- Tafel mit Magneten oder Pinnwand



### **So funktioniert's**

Die Schülerinnen und Schüler untersuchen, wie die Zugehörigkeit von Beispielelementen (*Trainingsdaten*) zu einer Kategorie zustande gekommen ist. Dazu entwickeln sie in Paaren Kriterien, mit denen neue Elemente klassifiziert werden können. Im Anschluss werden die entstandenen Modelle mit neuen Beispielen (*Testdaten*) getestet und die Genauigkeit der Vorhersage bestimmt.

## Kontext

Wir sind Tierpflegerin bzw. Tierpfleger in einem Zoo und für die Fütterung der Äffchen zuständig. Alle Äffchen sehen sehr süß aus, aber wir müssen aufpassen, denn manche Äffchen beißen. Von den Äffchen im Zoo wissen wir bereits, ob sie beißen. Allerdings werden bald neue Tiere zur Gruppe hinzukommen und wir müssen uns nun überlegen, wie wir herausfinden können, welche neuen Äffchen beißen und welche nicht – am besten ohne ihren Zähnen zu nahe zu kommen.

## Ablauf

Je nach Zielgruppe entscheiden Sie sich für die elementare Spielvariante mit 20 Bildkarten (blau) oder die fortgeschrittene Variante mit 40 Bildkarten (blau und grün). Bei diesen 20 bzw. 40 Äffchen handelt es sich um alle Tiere des Zoos, d.h. von diesen ist bereits bekannt, ob sie beißen. Sie werden in *Trainings-* und *Testdaten* aufgeteilt. Anhand der Trainingsdaten überlegen wir uns Kriterien, die bestimmen, ob die Äffchen beißen, und prüfen deren Zuverlässigkeit anhand der Testdaten. Die Trainingsdaten werden – unterteilt in die zwei Kategorien beißend und nicht beißend – gut sichtbar an die Tafel gepinnt. Die Testdaten werden zunächst zurückgelegt. Sie können sich selbst Regeln überlegen, nach denen die

### Variante 1 (blau)



**Variante 1 Trainingsdaten**  
beißend: 6, 7, 8, 15  
nicht beißend: 1, 2, 4, 9, 12, 14, 17, 18

**Variante 1 Testdaten**  
beißend: 3, 5, 11, 19  
nicht beißend: 10, 13, 16, 20

Äffchen klassifiziert werden oder einen der untenstehenden Vorschläge nutzen (hierbei ist auch die Verwendung reduzierter Subsets möglich). Die in den Beispielen geltenden Regeln sind anhand von Entscheidungsbäumen illustriert.

Machen Sie Ihren Schülerinnen und Schülern zunächst klar, auf welche Details Sie sich konzentrieren könnten, indem Sie das Vorgehen an einem Beispiel illustrieren. Stellen Sie dazu beispielsweise die Äffchenkarten 01 bis 04 und 05 bis 08 gegenüber. In diesem Beispiel ist die Mundform ein Indiz für beißende Äffchen, nicht aber die Augen (Abb. 1).

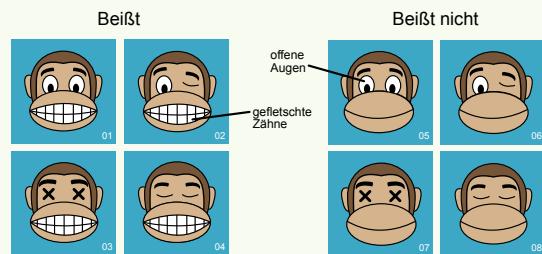
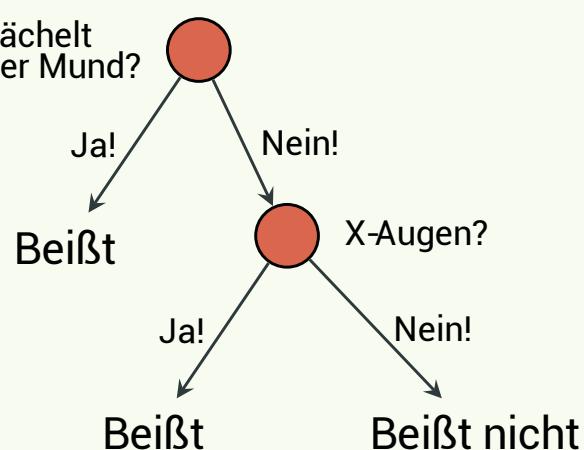
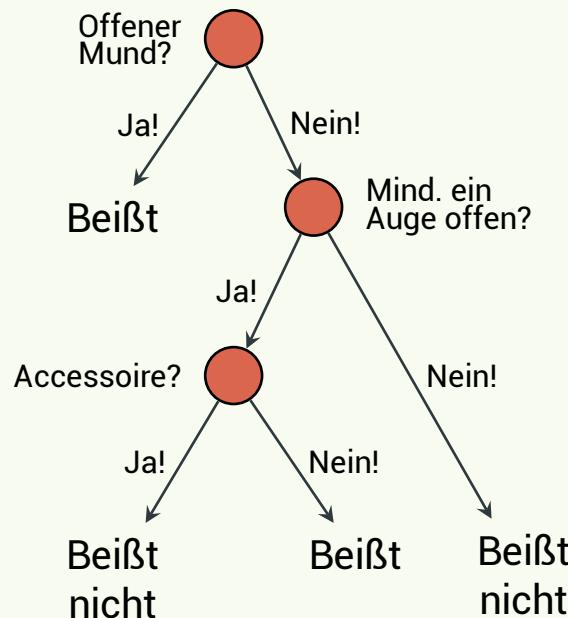


Abb. 1: In diesem einfachen Beispiel beißen alle Äffchen mit gefletschten Zähnen

Alternativ können Sie in höheren Klassenstufen zunächst auch die einfache Variante des Spiels spielen (Variante 1), um die Regeln und das erforderliche Vorgehen zu verdeutlichen.



### Variante 2 (blau & grün)



Die Schülerinnen und Schüler bilden Zweierteams und überlegen sich anhand der Trainingsdaten Kriterien, wie sie beißende von nicht beißenden Äffchen unterscheiden können. Diese müssen so eindeutig notiert werden, dass sie im Anschluss von einem anderen Team angewendet werden können. Eine Möglichkeit der Notation sind Entscheidungsbäume. Das Vorhandensein bzw. die Abwesenheit eines bestimmten Merkmals ermöglicht hier die klare Zuordnung zu einer Gruppe. Die Nutzung von Entscheidungsbäumen ist optional, alternativ ist auch das konkrete Aufschreiben von Regeln denkbar. Am Ende dieser *Trainingsphase* werden die formulierten Kriterien mit einem anderen Team getauscht.

Nun werden den Schülerinnen und Schülern nacheinander Bilder der restlichen Äffchen (Testdaten) gezeigt. Für jedes Äffchen-Bild müssen die Teams unter Verwendung des Regelschemas einer anderen Gruppe entscheiden, ob das Äffchen beißt oder

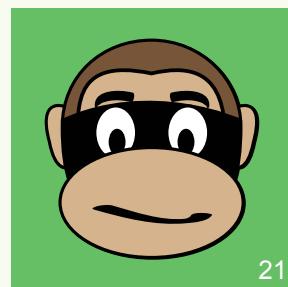
### Variante 2 Trainingsdaten

beißend: 1, 2, 5, 9, 10, 14, 15, 16, 17, 28, 33, 35, 36  
nicht beißend: 4, 7, 12, 19, 22, 23, 24, 25,  
30, 32, 37, 38, 39, 40

### Variante 2 Testdaten

beißend: 6, 13, 18, 34  
nicht beißend: 3, 8, 11, 20, 21, 26, 27, 29, 31

nicht. Jedes Team notiert die getroffene Entscheidung. Nachdem alle Äffchen gezeigt wurden, wird ausgewertet, welches Team bei den meisten Äffchen das Beißverhalten richtig bewertet hat. Dabei fällt auf, dass viele Klassifikationsmodelle die Äffchen zwar mehrheitlich richtig kategorisieren, es jedoch schwer ist, alle Tiere richtig einzuordnen. Für uns als Tierpflegerin bzw. Tierpfleger ist es also am sinnvollsten, das erfolgreichste Modell bei der Fütterung der neuen Äffchen zu nutzen, auch wenn wir dadurch nicht sicherstellen können, dass wir wirklich nie gebissen werden.



21

Abb. 2: Für Äffchen 21 kann man keine Kriterien aus den Daten ableiten

In der fortgeschrittenen Variante kann Bild Nr. 21 (siehe Abb. 2) verwendet werden, um zu thematisieren, wie problematisch es für ein KI-System ist, wenn sich die Merkmalsausprägung eines Elements deutlich von der Trainingsdaten unterscheidet. Wir haben bisher keine Erfahrung mit den Merkmalen von Bild Nr. 21, da dieses

Äffchen eine neue, unbekannte Mund-Form hat. Dementsprechend ist die Zuordnung des Äffchens nicht sinnvoll möglich. In der Praxis ist das Verhalten der KI in diesem Fall nur schwer vorhersehbar. Alternativ können Sie auch das Bild eines völlig anderen Tieres verwenden, um den Unterschied zu den bisher gezeigten Elementen noch deutlicher herauszustellen. Diese Beispiele lassen sich dann auch auf die Realität übertragen: Eine Bank vergibt an einen bestimmten Kunden unerwarteterweise keinen Kredit mehr oder das selbstfahrende Auto erkennt Blätter auf der Straße als Gefahrensituation und leitet irrtümlicherweise eine Vollbremsung ein. In derartigen Situationen kann von KI auch eine Gefahr ausgehen, wenn die Gründe für bestimmte Entscheidungen nicht mehr nachvollziehbar sind.

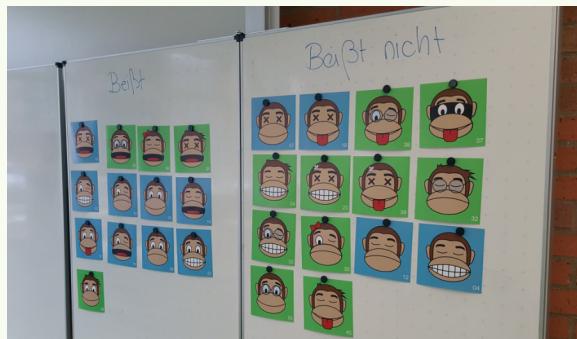


Abb. 3: Trainingsdaten unterteilt in zwei Kategorien

## Hintergrund

Kategorienbildung wird durch das Erkennen von sich wiederholenden Mustern bei einzelnen Elementen möglich. Wie aber hängen diese Aspekte mit künstlicher Intelligenz zusammen?

Beim sogenannten *überwachten Lernen* (*Supervised Learning*) beobachtet die KI eine Reihe von Ein- und Ausgabe-Paaren (Trainingsdaten) und lernt, wie diese miteinander in Beziehung stehen. Dieses Wissen wird dann verwendet, um

neue Elemente in eine der Kategorien einzuordnen. Testdaten, deren Kategorisierung uns, jedoch nicht dem Modell, bekannt ist, dienen dazu, die Güte des erlernten Klassifikationsmodells zu bestimmen.

Dasselbe Prinzip kommt auch bei neuronalen Netzwerken und anderen KI-Anwendungen zum Einsatz. Bei diesem Vorgehen kann es zu verschiedenen Problemen kommen, denn kein Modell ist perfekt. Je nach Trainingsdaten kann das Klassifikationsmodell bestimmte Eigenschaften der Trainingsdaten zu stark gewichten oder vernachlässigen, sodass keine allgemeinen Aussagen und damit keine korrekte Klassifikation unbekannter Elemente mehr möglich ist. Viele Trainingsdaten können helfen, diese Effekte zu verringern, führen allerdings nicht immer zu genaueren Ergebnissen, da zu viele Trainingsdaten auch eine Überanpassung zur Folge haben können. In diesem Fall lernt das KI-System die Trainingsdaten „auswendig“ und ist anschließend nicht mehr in der Lage, auf neue Daten zu verallgemeinern.

Es bietet sich an, diese Aspekte von maschinellem Lernen im Rahmen der Aktivität zu thematisieren. Lassen Sie die Schülerinnen und Schüler bei der Anwendung ihrer Regelwerke in der Testphase Erklärungen geben, nach welchen Merkmalen sie die Äffchen eingeordnet haben, um herauszustellen, dass unterschiedliche Regelwerke verwendet werden. Gehen Sie anschließend darauf ein, dass ein Klassifikationsmodell voraussichtlich nie zu 100 % genau ist, sondern dass das Modell gewählt wird, das die Testdaten am besten einordnet. Lassen Sie die Schülerinnen und Schüler ihren eigenen „Lernprozess“ beschreiben und dann mit dem eines Computers vergleichen.

# #deeplearning

## Erkennen von Bildern mit neuronalen Netzen

### Zielgruppe

Sekundarstufe I, Sekundarstufe II

### Darum geht's

Wie kann ein Computer eigentlich Dinge „erkennen“? Wie entscheidet ein Computer, ob auf einem Foto eine Katze abgebildet ist? Wie kann er Gebäude von Menschen unterscheiden? Gegenstände oder Objekte anhand ihrer Form bzw. ihres Äußeren zu erkennen ist für Menschen sehr leicht. Für den Computer, der beispielsweise in einem selbstfahrenden Auto die Objekte in seiner Umgebung erkennen muss, stellt dies jedoch eine komplexe Aufgabe dar. In dieser Aktivität haben Schülerinnen und Schüler die Möglichkeit, selbst nachzuvollziehen, wie Computer den Inhalt von Bildern erkennen.

### Diese Ideen stecken dahinter

- Neuronale Netze ordnen Eingaben bestimmte Ausgaben zu: Rohdaten, wie z. B. Bilder, erhalten eine Klassifizierung, indem beispielsweise den Objekten im Bild Begriffe zugeordnet werden.
- Neuronale Netze bestehen aus unterschiedlichen Abstraktionsschichten, die zunehmend komplexere Merkmale identifizieren.
- Die Klassen von Objekten, die erkannt werden sollen, müssen der KI bereits bekannt sein.

### Das wird benötigt

- Fotokarten von Häusern, Katzen und Autos für jede Gruppe



### So funktioniert's

Die Schülerinnen und Schüler stellen den Bilderkennungsprozess eines (vereinfachten) neuronalen Netzes nach. Dabei nehmen sie die Rollen der unterschiedlichen Schichten innerhalb eines solchen Netzes ein. Sie extrahieren Merkmale aus einem Foto und klassifizieren das Bild. Dabei erkennen sie die Grenzen des Systems und überlegen, welche Modifikationen am Netz notwendig sind, um bessere Ergebnisse mit ihrem Netz zu erzielen.

## Kontext

Als Menschen verlassen wir uns stark auf das, was wir sehen. Sehen wir eine Katze, wissen wir sofort, dass es eine Katze ist, sehen wir einen Hund erkennen wir auch diesen sofort. Ein Computer kann dies hingegen nicht so leicht feststellen, aber er kann es lernen – genauso wie wir das im Kleinkindalter getan haben. Dazu zeigen wir dem Computer sehr viele Bilder von Hunden, aber auch von anderen Tieren. Der Computer lernt dann selbst, anhand welcher Muster er einen Hund von einer Katze unterscheiden kann. Richtig trainiert kann der Computer dann nicht nur Bilder automatisch beschriften, sondern auch Hautkrebs erkennen oder – verbaut in einem Auto – auf Hindernisse reagieren.

Skizzen davon (je 30 Sekunden) und gibt diese an B weiter. Dabei wird darauf geachtet, dass C die Skizzen nicht sieht.

B erhält die Skizzen von A und prüft, ob viereckige Formen, dreieckige Formen oder runde Formen enthalten sind. Dann gibt B die gesammelten Informationen an C weiter.

C wertet die erhaltenen Informationen anhand der folgenden Tabelle aus und verkündet, ob es sich beim ursprünglichen Bild um ein Haus, ein Auto oder eine Katze handelt.

	Viereckige Form?	Dreieckige Form?	Runde Form?
Haus	Ja	Ja	Nein
Auto	Ja	Nein	Ja
Katze	Nein	Ja	Ja

Zuletzt stellt A fest, ob die Lösung stimmt.

Lassen Sie die Schülerinnen und Schüler das Spiel in wechselnden Rollen ausprobieren. Es besteht auch die Möglichkeit, eine Rolle von zwei Schülerinnen bzw. Schülern ausüben zu lassen bzw. jeder Rolle explizit drei Schülerinnen und Schüler zuzuordnen, welche dann jeweils ein *Neuron* (d.h. einen Knoten der Schicht) und keine ganze Schicht mehr darstellen.

Nach kurzem Ausprobieren geben Sie den Gruppen weitere Bilder, die nicht zu den Kategorien passen, welche das Netz erkennen kann, oder Merkmale aufweisen, die eine eindeutige Klassifizierung nicht zulassen. Ein Bild von einem Hund wird beispielsweise vom Netz nicht korrekt erkannt – ganz einfach, weil das Netz die Kategorie

## Ablauf

Thematisieren Sie zu Beginn, wie ein Computer wohl den Inhalt eines Bildes erkennt. Antworten werden sich häufig auf definierte Regeln oder den Vergleich mit einer Bilddatenbank beziehen, heutzutage gehen Computer dabei jedoch anders vor. Teilen Sie die Schülerinnen und Schüler in Dreiergruppen auf, jede Gruppe erhält einen Stapel von Fotokarten. In jeder Gruppe gibt es drei Rollen, die jeweils eine Schicht eines neuronalen Netzes repräsentieren (siehe Abb. 4). Die Aufgaben der Rollen sind im Einzelnen:

A zieht ein Bild aus dem Stapel von Fotokarten (B und C sollen das Bild nicht sehen!), erstellt zwei unterschiedliche

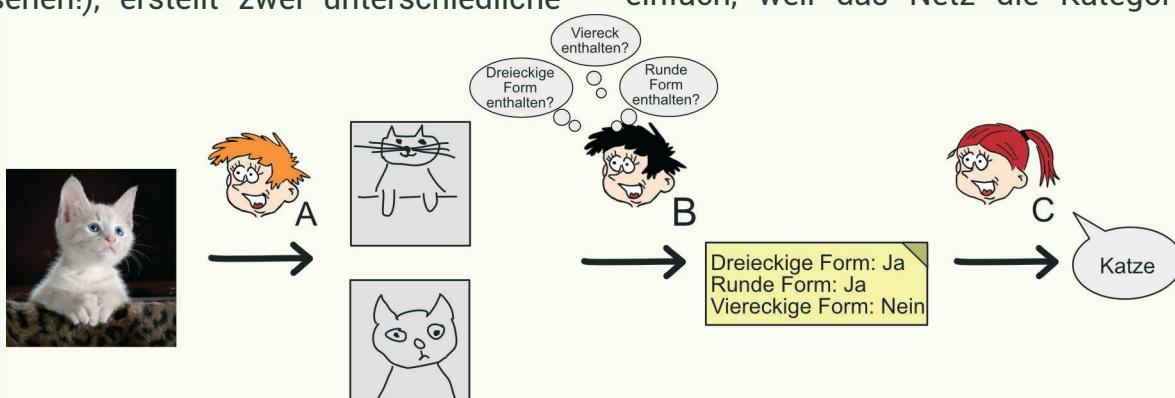


Abb. 4: Die einzelnen Rollen für Schülerinnen und Schüler

„Hund“ nicht kennt. Basierend auf dieser Feststellung überlegen die Schülerinnen und Schüler nun, wie das Netz verändert und erweitert werden kann, um in Zukunft auch Hunde oder weitere andere Objekte erkennen zu können. Zunächst muss dafür eine neue Ausgabekategorie eingeführt werden. Gleichzeitig reicht allerdings auch die Anzahl der vom Netz identifizierten Merkmale nicht mehr aus, sodass entweder weitere Merkmale gefunden werden müssen, die die Unterscheidung der Kategorien erlauben, oder mehrere Merkmale zu einem komplexen Muster zusammengefasst werden müssen. Dieses Zusammenfassen entspricht letztlich dem Hinzufügen weiterer Schichten im neuronalen Netz. Als Hilfestellung eignet sich etwa Abb. 6, anhand der Sie den Schülerinnen und Schülern das Kombinieren von einfachen Merkmalen zu komplexeren Mustern verdeutlichen können.

## Hintergrund

In Situationen, in denen ein zu lösendes Problem nur schwer in logische Regeln übersetzt werden kann, werden häufig künstliche neuronale Netze eingesetzt. Zu diesen Problemen gehören das Verständnis von Texten oder das Erkennen von Objekten in Bildern. Die Entwurfsidee für künstliche neuronale Netze entstammt der Neurobiologie und orientiert sich am Aufbau des menschlichen Gehirns. Analog zu einer menschlichen Nervenzelle, die verschiedene Reize verarbeitet und einen Impuls weitergibt, verarbeitet auch ein künstliches Neuron verschiedene Eingaben und kann ein Signal weitergeben. Die Eingabekanten sind mit einem Gewicht versehen, d.h. sie haben einen unterschiedlich starken Einfluss auf die Ausgabe des Neurons. Ein künstliches Neuron ist damit zwar entfernt an die Funktionsweise eines menschlichen Neurons angelehnt, funktioniert aber eher wie ein einfacher Taschenrechner: Es multipliziert die Kantengewichte und Eingabewerte, addiert diese und gibt ein Ergebnis weiter.

Genau wie im menschlichen Nervensystem, hängen viele Neuronen zusammen und bilden auf diese Weise ein Netz. Die Neuronen werden dabei in Schichten angeordnet. Je nach Komplexität eines Problems kann ein Netz aus zwei, aber auch aus mehr Schichten bestehen. In der Ausgangssituation dieser Aktivität sind es drei. Hat ein Netz weitere Schichten zwischen Ein- und Ausgabeschicht so spricht man von *Deep Learning*.

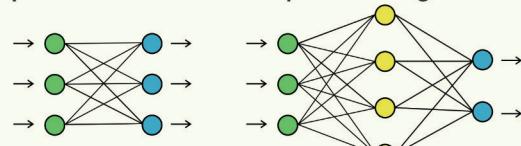


Abb. 5: Bsp. für künstliche neuronale Netze - links: einfaches Netz, rechts: „tieferes“ Netz

In der Praxis funktioniert die Bilderkennung bzw. -klassifikation meist über sogenannte *Convolutional Neural Networks*, welche darauf spezialisiert sind, selbstständig Muster zu erkennen und sich daher sehr gut eignen, um Bilder zu klassifizieren. Diese Art von neuronalen Netzen zeichnet sich dadurch aus, dass sie sogenannte *Convolutions* (Faltungen) nutzt, um Merkmale und Muster aus Eingabedaten zu extrahieren. Mittlerweile können diese Netze Bilder schneller klassifizieren als Menschen.

Wie funktioniert das genau? Digitale Fotos sind aus kleinen Farbelementen – den Pixeln – zusammengesetzt, die in einem Raster angeordnet sind. Jedes Pixel hat einen gewissen Farbwert. Für den Computer sind Fotos anders als für uns Menschen also zunächst nur Zahlenwerte. Netze zur Bilderkennung versuchen anfangs einfache Merkmale – sogenannte *Features* – zu erkennen. Dazu werden Filter über das Bild gelegt. Das ähnelt dem, was wir in Fotobearbeitungsprogrammen machen, wenn wir dort beispielsweise den Hochpassfilter anwenden (siehe Abb. 7). Letztlich handelt es sich hierbei um eine mathematische Berechnung, welche mehrere Bildpunkte erfasst und mit einem Filter zu einem neuen Bildpunkt verrechnet. Je nach Filter kann so beispielsweise in einem ersten Schritt erkannt werden, dass sich Pixel mit



Abb. 6: Neuronale Netze identifizieren immer komplexere Merkmale über mehrere Ebenen

ähnlichen Helligkeitswerten zu Kanten verbinden lassen. Auf einer weiteren Ebene werden dann Merkmale wie horizontale und vertikale Linien, Kreise oder Ecken extrahiert. Gängige Bildbearbeitungsprogramme wie Gimp erlauben das Eingeben solcher Filter in Form einer Matrix, eine gute Möglichkeit deren Wirkung selbst auszuprobieren.



Abb. 7: Anwendung eines Filters

Im Spiel fungieren die Zeichnungen, die von den Schülerinnen und Schülern angefertigt werden, ebenfalls als Filter, da hier die zentralen Elemente des auf dem Foto abgebildeten Objekts extrahiert werden. Anhand der Skizzen werden anschließend geometrische Formen in den Bildern identifiziert. Die Nutzung von nur drei Merkmalen stellt hierbei eine starke Vereinfachung gegenüber einem realen neuronalen Netz dar, welches über mehrere Millionen Neuronen in einer Vielzahl von Schichten verfügt.

Auf den ersten Ebenen eines neuronalen Netzes findet sich eine Vielzahl derartiger einfacher und eher geometrischer Filter. Die hier erkannten Muster werden (wieder durch die Anwendung von Filtern) anschließend zu komplexeren Mustern zusammengesetzt. Auf „tieferen Ebenen“ können damit nicht mehr nur Ecken und Kanten,

sondern auch Teile von Objekten, wie Augen oder Fellpartien, und schließlich auch vollständige Objekte, wie etwa Hunde oder Katzen, erkannt werden. Während des Verarbeitungsprozesses werden immer wieder (überflüssige) Informationen verworfen, da beispielsweise die exakte Position einer diagonalen Linie auf einem Bild für die Erkennung eines Objekts in vielen Fällen von untergeordnetem Interesse ist. Am Ende steht eine Wahrscheinlichkeit, mit der ein Bild einer bestimmten Kategorie zugeordnet werden kann.

Ein neuronales Netz kann jedoch nicht ohne Weiteres den Inhalt jedes Bildes erkennen. Vielmehr ist der Anwendungsrahmen sehr begrenzt: Das neuronale Netz muss mit einer sehr großen Menge von Bildern (mehrere Tausend) zunächst „trainiert“ werden. Dabei lernt es, welche Merkmale für Bilder, die zu einer bestimmten Kategorie gehören, entscheidend sind. Somit kann das neuronale Netz auch nur Bilder korrekt klassifizieren, deren Kategorie ihm bekannt ist. Z. B. kann ein Netz, welches Hunde und Katzen unterscheiden soll, keine anderen Tiere erkennen, sondern sortiert sie vielmehr in eine der beiden bekannten Kategorien ein. Ein trainiertes neuronales Netz kann seine Aufgabe jedoch deutlich schneller erfüllen, als Menschen das jemals könnten. Daher werden Bilderkennungsverfahren z. B. in selbstfahrenden Autos zur Erkennung verschiedener Objekte im Straßenverkehr (Gegenverkehr, Fußgänger etc.) oder in der Hautkrebserkennung bereits eingesetzt.

# Reinforcement Learning

## Schlag das Krokodil!

### Zielgruppe

Sekundarstufe I, Sekundarstufe II

### Darum geht's

Computer, die Schach spielen und Menschen schlagen können, kennen wir mittlerweile. Das chinesische Brettspiel Go hingegen galt lange Zeit als so komplex, dass nur Menschen es meistern können – so lange, bis Google mit *AlphaGo* menschliche Profi-Spielerinnen und Spieler das Fürchten lehrte. In dieser Aktivität sehen wir, wie Computer selbst Strategien für Spiele lernen, obwohl sie nur die Spielregeln kennen.

### Diese Ideen stecken dahinter

- Computer können durch „Belohnung“ und „Bestrafung“ lernen.
- Computer bewerten den Nutzen zufälliger Aktionen basierend auf Belohnung und Bestrafung.
- Computer erlernen Strategien bzw. Aktionsfolgen, indem sie nach maximaler Belohnung streben.

### Das wird benötigt

- Pro Schülerpaar: 1 „Mini-Schach“-Feld, 3 Äffchen- und 3 Krokodil-Karten, 1 Übersicht über mögliche Züge
- bunte Süßigkeiten (z. B. Schokolinsen) oder Papiermarken zum Bewerten der Spielzüge in 4 verschiedenen Farben (gelb, rot, orange, blau; ca. 20 pro Farbe)



### So funktioniert's

Je zwei Schülerinnen bzw. Schüler spielen eine Partie „Mini-Schach“ gegeneinander. Dabei übernimmt eine Schülerin bzw. ein Schüler die Rolle eines „Papier“-Computers. Der Computer wählt seine Züge zunächst zufällig, lernt aber nach und nach mit einem Süßigkeiten-Token-System, welche Züge ihm zum Sieg verhelfen und welche in einer Niederlage enden. Mit Hilfe der Strategie, die sich auf diese Weise entwickelt, wird der Computer mit der Zeit immer besser.

*Diese Aktivität basiert auf einer Idee von CS4Fun.*

(<http://www.cs4fn.org/machinelearning/sweetlearningcomputer.php>)

## Kontext

Wie lernen wir als Mensch ein Brett- oder Videospiel zu spielen? Vielleicht beobachten wir andere beim Spielen, vielleicht probieren wir aus, wie bestimmte Aktionen oder Spielzüge das Spiel beeinflussen. Je öfter wir gewinnen, desto besser werden wir in einem Spiel. Wir entwickeln Strategien, mit welchen Zügen wir in bestimmten Spielsituationen am erfolgreichsten sind. Auf dieselbe Weise lernt auch ein Computer Spiele zu spielen.

## Ablauf

Das Spiel folgt einfachen Schachregeln: Jede Spielfigur bewegt sich wie ein Bauer, d. h. sie kann nur vorwärts gehen und gegnerische Figuren nur diagonal schlagen. Eine Schülerin bzw. ein Schüler übernimmt die Äffchen und agiert als menschliche Spielerin bzw. menschlicher Spieler. Eine andere Schülerin bzw. ein anderer Schüler übernimmt die Rolle des Computers in Form der Krokodile.

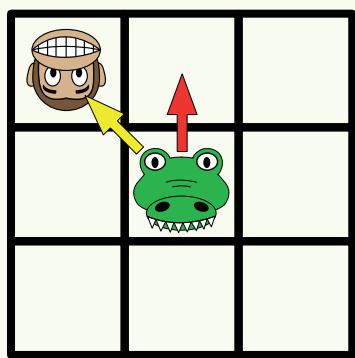


Abb. 8: Mögliche Bewegungen einer Spielfigur

Eine Seite hat gewonnen, wenn sie es schafft

- eine eigene Spielfigur an das andere Ende des Spielfeldes zu führen.
- alle gegnerischen Figuren zu schlagen.
- dafür zu sorgen, dass der Gegner in der nächsten Runde keinen Spielzug mehr ausführen kann.

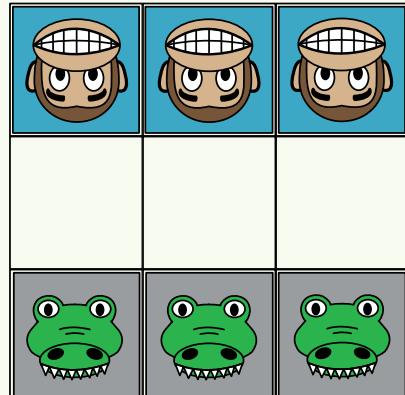


Abb. 9: Spielfeld vor Spielbeginn

Zur Vorbereitung werden Ausdrücke der Zugmöglichkeiten des Computers vor der Spielerin bzw. dem Spieler, der die Krokodile übernimmt, ausgebreitet. Dann werden Schokolinsen (oder andere farbige Tokens) auf diese Zugmöglichkeiten verteilt. Dabei legen Sie für jeden farbigen Pfeil die entsprechende Farbe als Token in den Bereich rechts der gezeigten Spielsituation (siehe Abb. 11).

Die menschliche Spielerin bzw. der menschliche Spieler beginnt. Sie bzw. er kann sich frei gemäß der Spielregeln bewegen. Danach sind die Krokodile an der Reihe. Die Spielerin oder der Spieler vergleicht das aktuelle Spielfeld mit den Zugmöglichkeiten und wählt die passende Spielsituation aus den gegebenen Möglichkeiten aus. Zur schnelleren Orientierung ist jeweils angegeben, um welchen Spielzug es sich handelt. In der ersten Runde müssen nur die beiden Möglichkeiten für Zug 1, in der zweiten Runde die 10 Spielzüge für Zug 2 und in Runde 3 die 7 Spielzüge für Zug 3 betrachtet werden. Dabei sind symmetrische Spielsituationen nicht doppelt aufgeführt. Im Anschluss schließt sie bzw. er die Augen und zieht zufällig eines der neben der jeweiligen Spielsituation ausgelegten Tokens und verschiebt es auf die Zugvorgaben. Die Farbe des Tokens bestimmt, welcher Zug ausgeführt wird und die Spielerin bzw. der Spieler

bewegt die Spielfigur entsprechend des gleichfarbigen Pfeils. Wird beispielsweise ein rotes Token gezogen, wird das Krokodil dem roten Pfeil folgend bewegt.

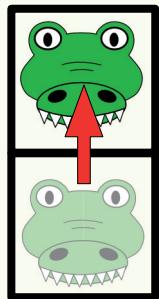


Abb. 10: Krokodil wird entlang des roten Pfeils bewegt

Dieses Vorgehen wird solange wiederholt, bis die Siegerin bzw. der Sieger der Spielrunde feststeht. Bevor eine neue Runde gespielt wird, passt der Computer nun seine Strategie wie folgt an:

- Krokodile haben gewonnen: Ein Token in der Farbe des letzten, siegbringenden Spielzugs wird zusätzlich auf das Feld dieses Spielzugs gelegt.
- Äffchen haben gewonnen: Die Schokolinse, die den letzten Zug der Krokodil-Spielerin bzw. des Krokodil-Spielers bestimmt hat, wird entfernt. Die Spielerin bzw. der Spieler der Äffchen darf dieses ggf. essen.

Zudem werden alle Tokens wieder rechts vom jeweiligen Spielfeld platziert.

*Optional: Um die Regeln zu vereinfachen, können Sie die Verstärkung durch ein zusätzliches Token im Falle eines Siegs der Krokodile auch vernachlässigen.*

## Hintergrund

Zunächst wird der Computer kaum eine Chance haben zu gewinnen, da er seine Bewegungen zufällig wählt (indem mit geschlossenen Augen ein Token gezogen wird). Je mehr Spiele der Computer beendet, desto besser wird er: Er „lernt“, welche Züge ihm zum Sieg verhelfen und welche er vermeiden sollte, da sie in der Vergangenheit in einer Niederlage endeten. So wird die Strategie des Computers schrittweise verfeinert.

Da der Computer für das Verlieren bestraft und für das Gewinnen belohnt wird, sprechen wir auch von *bestärkendem Lernen (Reinforcement Learning)* – ein Lernen durch Belohnung und Bestrafung:

- Bestrafung = Wegnehmen einer Süßigkeit bei einem Spielzug, der zur Niederlage führte
- Verstärkung = Hinzufügen einer Süßigkeit bei einem Spielzug, der zum Sieg führte

Durch dieses Vorgehen werden bei den jeweiligen Spielzügen diejenigen Züge „aussortiert“, die Niederlagen zur Folge hatten, sodass irgendwann nur noch „gute“ Züge übrig bleiben. In der Praxis würde man Strategien, die nicht zum Erfolg führen, nicht sofort eliminieren, sondern nur die Wahrscheinlichkeit ihres Auftretens verringern. So lernt die KI nach und nach, welche Strategie in welcher Situation wohl am besten geeignet ist, schließt einzelne Strategien, die nicht in jedem Fall zum Erfolg geführt haben, aber nicht sofort vollständig aus. Obwohl dieses Vorgehen im Spiel vereinfacht wird, indem Züge, die zur unmittelbaren Niederlage geführt haben, sofort entfernt werden, kann nie der Fall eintreten, dass für eine Spielsituation

alle möglichen Züge eliminiert werden, da es für jede Situation mindestens eine Handlungsoption für den Computer gibt, die nicht zur sofortigen Niederlage führt.

Ein Computer kann auf diese Weise allein durch die Kenntnis der Spielregeln oder möglicher Eingaben lernen, ein Spiel zu gewinnen. Lernt ein Computer etwa das Videospiel *Super Mario* zu spielen, wird er zunächst nur wahllos Eingaben tätigen. Das könnte dazu führen, dass er minutenlang nur stehen bleibt oder mehrfach in denselben Gegner hineinläuft. Er analysiert dabei die Objekte bzw. Pixel im Bild und reagiert mit Eingaben. Sein Ziel ist es, die erreichte Punktzahl im Spiel zu maximieren, die Punkte fungieren hier als Belohnung. Je weiter der Computer nach rechts vorrücken kann, desto größer fällt die positive Verstärkung aus. Mit der Zeit wird er so beispielsweise lernen, dass Springen seine Belohnung erhöht, wenn sich ein Gegner unmittelbar rechts von ihm befindet, da er durch das Überspringen des Gegners weiter im Level vorankommt. Auf diese Weise verbessert sich das Vorgehen eines KI-Systems in einem Spiel Stück

für Stück, wobei das System immer versucht, seine Belohnung (oder genauer: eine bestimmte Funktion) zu maximieren.

Lassen Sie die Schülerinnen und Schüler im Rahmen der Dekontextualisierung analysieren, wie sich das Verhalten des Computers entwickelt. Es soll deutlich werden, dass der Computer durch Verstärkung von rein zufälligem Handeln zu einer effizienten Spielstrategie kommt. Im Anschluss kann beispielsweise mit Hilfe eines Videos zum Spiel *Super Mario* (siehe Website) gezeigt werden, wie das verstärkende Lernen in einem neuronalen Netz vonstatten geht. Lassen Sie die Schülerinnen und Schüler hierbei insbesondere über die Grenzen der erlernten Strategien reflektieren. Sie können diese Aktivität sehr gut mit der Aktivität *Back to the Roots: Krokodil-Schach und klassische KI* kombinieren, um den Kontrast zwischen lernenden Systemen und klassischen KI-Anwendungen, wie regelbasierten Systemen, herauszustellen.

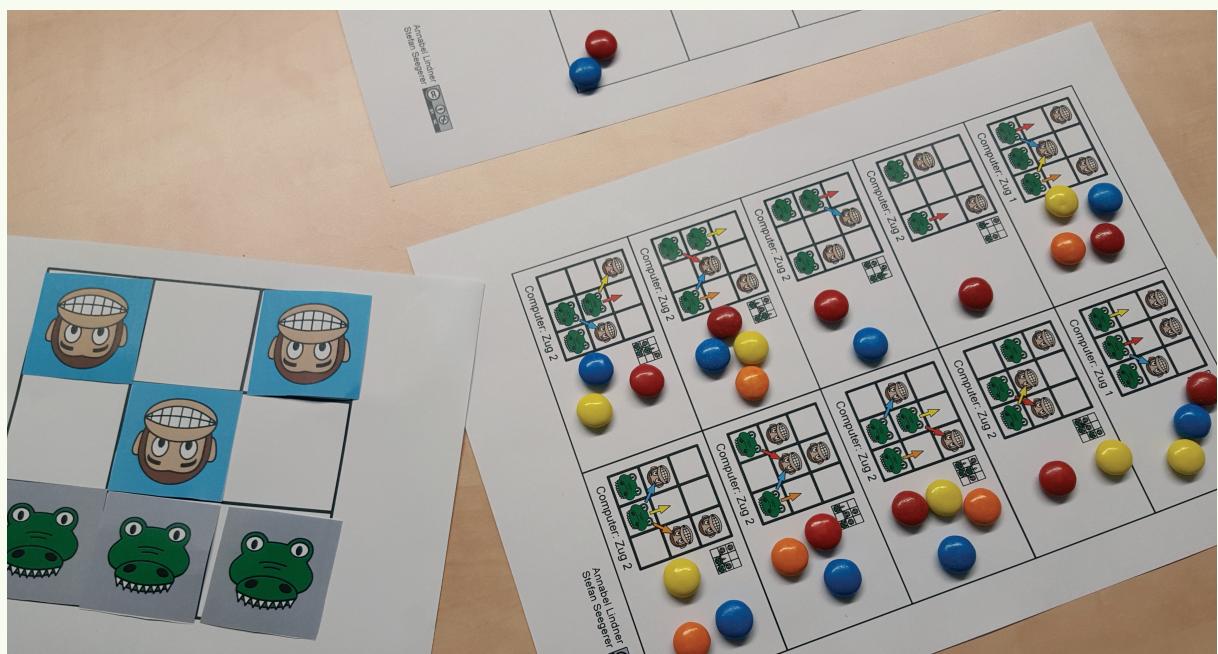


Abb. 11: Spieldurchführung: Die Verteilung der Schokolinsen stellt die gelernten Strategien dar

# *Back to the Roots*

## Krokodil-Schach und klassische KI

### Zielgruppe

Sekundarstufe I, Sekundarstufe II

### Darum geht's

Die bisherigen Übungen beschäftigen sich intensiv mit lernenden KI-Systemen. Das ist jedoch nicht alles, was die KI zu bieten hat: Die Ursprünge der KI liegen in der Logik und der Idee, Wissen mathematisch zu formalisieren und es auf diese Weise für Maschinen zugänglich zu machen. Welche Unterschiede es zwischen lernender KI und klassischen Ansätzen gibt und wo die Grenzen der Systeme liegen, wird in dieser Aktivität aufgezeigt. Dazu wird die vorangehende *Reinforcement Learning* Aktivität mit einem Expertensystem umgesetzt und so die stark unterschiedlichen Herangehensweisen verdeutlicht.

### Diese Ideen stecken dahinter

- Wissen muss formal darstellbar sein, um maschinell verarbeitet werden zu können.
- Expertensysteme können Regeln und Fakten kombinieren, um neues Wissen zu generieren.
- Solche KI-Systeme treffen keine selbständigen Entscheidungen, sondern arbeiten nach den Regeln der Logik.
- KI-Systeme verfügen über Verarbeitungsmechanismen zum automatischen Schließen aus vorhandenem Wissen.

### Das wird benötigt

- Pro Schülerpaar: 1 „Mini-Schach“-Feld, 3 Äffchen- und 3 Krokodil-Karten, 1 Übersicht über Regeln für den nächsten Zug

### So funktioniert's

Wie in der Aktivität *Reinforcement Learning* spielen je zwei Schülerinnen bzw. Schüler eine Partie „Mini-Schach“ gegeneinander. Dabei übernimmt eine Schülerin bzw. ein Schüler die Rolle eines „Papier“-Computers. Wird die Aktivität, wie von uns empfohlen, direkt mit der vorangehenden *Reinforcement Learning*-Version kombiniert, bietet es sich an, dass ein Rollentausch stattfindet. Anstatt seine Züge zufällig zu wählen, arbeitet der Computer nun jedoch nach vorgegebenen Regeln, die als Kopie zur Verfügung gestellt werden.

## Kontext

Wie kann man einen Computer so programmieren, dass er Brett- oder Videospiele spielen kann? Computer können die Regeln eines Spiels nur „verstehen“ und entsprechend handeln, wenn diese so dargestellt sind, dass ein Computer sie auch verarbeiten kann. Wissen muss daher für die maschinelle Verarbeitung formal repräsentiert werden (z. B. durch mathematische Terme). Dann kann der Computer es mit Hilfe von Logik auswerten und seine Handlungen daraus ableiten. KI-Systeme sind also gar nicht wirklich intelligent, sondern nutzen geschickt verschiedene Möglichkeiten, aus vorhandenem Wissen ihr Verhalten abzuleiten.

## Ablauf

Das Spiel folgt einfachen Schachregeln und hat dieselben Grundbedingungen wie in der Aktivität *Reinforcement Learning*: Jede Spielfigur bewegt sich wie ein Bauer, d. h. sie kann nur vorwärts gehen und gegnerische Figuren nur diagonal schlagen. Eine Schülerin bzw. ein Schüler übernimmt die Äffchen und agiert als menschliche Spielerin bzw. menschlicher Spieler. Eine andere Schülerin bzw. ein anderer Schüler übernimmt die Rolle des Computers in Form der Krokodile.

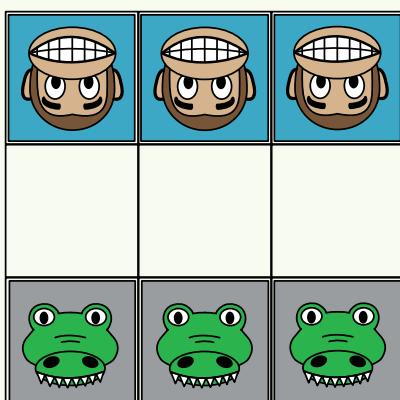


Abb. 12: Spielfeld vor Spielbeginn

Eine Seite hat gewonnen, wenn sie es schafft

- eine eigene Spielfigur ans andere Ende des Spielfeldes zu führen.
- alle gegnerischen Figuren zu schlagen.
- dafür zu sorgen, dass der Gegner in der nächsten Runde keinen Spielzug mehr ausführen kann.

Zur Vorbereitung wird der Spielerin bzw. dem Spieler, der die Krokodile übernimmt, ein Ausdruck der Regeln für die Züge des Computers ausgehändigt, diese ersetzen die Zugmöglichkeiten und Tokens aus der Aktivität *Reinforcement Learning*. Die menschliche Spielerin bzw. der menschliche Spieler beginnt. Sie bzw. er kann sich frei gemäß der Spielregeln bewegen. Danach sind die Krokodile an der Reihe. Die Spielerin oder der Spieler vergleicht das aktuelle Spielfeld mit der Regelstabelle und wählt die passende Spielsituation aus den 10 Möglichkeiten aus, dabei sind symmetrische Spielsituationen nicht doppelt aufgeführt. Dann führt sie oder er den Zug aus, welchen die Regel vorgibt.

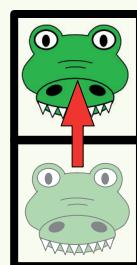


Abb. 13: Krokodil wird entlang des roten Pfeils bewegt

Dieses Vorgehen wird solange wiederholt, bis eine Siegerin bzw. ein Sieger feststeht. Es können mehrere Runden gespielt werden, um zu überprüfen, ob der Computer mit Hilfe seiner Regeln immer in der Lage ist zu gewinnen.

Präsentieren Sie den Schülerinnen und Schülern im Anschluss die erweiterte Version des Krokodil-Schachspiels mit 4x4 Feldern (siehe Website). Thematisieren Sie, dass die Regeln, über die der Computer verfügt, nun nicht mehr ausreichen. Lassen Sie die Schülerinnen und Schüler einen

Vergleich zur Variante des Mini-Schachs aus der Aktivität *Reinforcement Learning* ziehen, in welcher der Computer lernt: Auch hier reicht das „Wissen“ des Computers nicht mehr aus. Somit kommen beide Systeme hier an ihre Grenzen.

Die Schülerinnen und Schüler überlegen nun, welche Vorgänge notwendig sind, um das „Wissen“ des Computers an die 4x4-Version des Spiels anzupassen. Sie erkennen, dass der regelbasierte Computer vom Menschen durch das Hinzufügen neuer Regeln für den optimalen Zug manuell angepasst werden muss, während das lernende System auf dieselbe Weise wie in der vorangehenden Aktivität, also durch die Bewertung von zufälligem Verhalten, auch die optimalen Verhaltensweisen für das 4x4-Spielfeld erlernen kann, sobald alle neuen Zugmöglichkeiten ergänzt wurden. Das lernende System benötigt somit eine neue Trainingsphase, in der das Wissen für das erweiterte Spiel implizit erworben wird, während die neuen Regeln im regelbasierten Expertensystem explizit hinzugefügt werden müssen. Hierbei hat der Mensch die Aufgabe, zunächst selbst den idealen Spielzug für jede Spielsituation aus allen möglichen Zügen zu ermitteln und anschließend alle Regeln umfassend und vollständig zu formalisieren. Dies entfällt beim lernenden System. Diese explizite formale Darstellung von Regeln, die das Vorgehen des Systems steuern, ist jedoch bei komplexen und mehrschrittigen Problemen kaum oder gar nicht mehr umfassend möglich. Lernende KI-Systeme bieten hier gegenüber Expertensystemen den großen Vorteil, dass sie derartige Vorgehensweisen „selbstständig“ ermitteln können und sogar Zusammenhänge zwischen Daten ermitteln können, die für den Menschen nicht sichtbar sind, z. B. weil dazu die Betrachtung von vielen tausend Datensätzen notwendig ist.

## Hintergrund

Logik und Wissensverarbeitung spielen in vielen Bereichen der Informatik eine Rolle und sind darüber hinaus Kernthemen der künstlichen Intelligenz. Da die natürliche Sprache mehrdeutig und zu vielfältig ist, um Wissen für die Verarbeitung durch Maschinen zugänglich zu machen, kommt der Frage nach der für Maschinen bestmöglichen Repräsentation von Wissen seit den Anfängen der KI eine entscheidende Bedeutung zu.

Ansätze dieser „klassischen“ KI setzen hier auf symbolische Wissensrepräsentation, also die explizite Darstellung von Wissen in Computersystemen beispielsweise mit Hilfe von Logik. Diese ermöglicht die eindeutige, einheitliche und präzise Darstellung von Wissen, welche für die Verarbeitung durch einen Computer notwendig ist. Entsprechende Darstellungsweisen kommen z. B. in regelbasierten Expertensystemen zum Einsatz, die in kommerziellen Anwendungen heutzutage immer noch eine Rolle spielen. In solchen Systemen werden aus logischen Aussagen, die Fakten und Regelwissen repräsentieren, automatisch Schlüsse gezogen, wie der Computer handeln muss. Im Falle unserer Aktivität entsprechen die Fakten der aktuellen Spielsituation und das Regelwissen den Anweisungen, welcher Zug durchzuführen ist.

Die Faktenbasis stellt gültige Aussagen dar. Eine Menge von Regeln, die in „wenn ... dann“-Form repräsentiert sind, bildet die Regelbasis. Formal kann diese „wenn ... dann“-Form beispielsweise durch Aussagenlogik dargestellt werden. Ein Kontrollsystem wählt basierend auf der Faktenlage geeignete Regeln aus, bewertet sie und handelt entsprechend. In komplexen Experten-

systemen können die aus Regeln gezogenen Schlussfolgerungen auch wieder als Eingabefakten für weitere Regeln dienen und tragen so zur Erweiterung der Faktenbasis bei. Die Aufgabe des Kontrollsystems wird in dieser Aktivität von der Schülerin bzw. dem Schüler übernommen, der die Rolle des Computers inne hat. Mit Hilfe der Spielzug-Vorgaben, welche die Regelbasis darstellen, muss sie oder er aus der aktuellen Spielsituation den nächsten Zug schließen.

In einem regelbasierten System wird diese Vorgehensweise als datengetrieben bzw. als Vorwärtsverkettung bezeichnet, da auf Grundlage von Fakten versucht wird, ein noch unbekanntes Ziel zu erreichen. Im Gegensatz dazu steht die Rückwärtsverkettung, die versucht, eine getroffene Hypothese zu beweisen. In diesem Fall würden wir das Spiel quasi „rückwärts“ spielen und versuchen, von einer Spielsituation aus, in der der Computer gewonnen hat, die für diesen Gewinn notwendigen Züge abzuleiten.

Auch das maschinelle Lernen, das heutzutage die dominante Methode im Bereich der künstlichen Intelligenz darstellt und in vielen Bereichen Expertensysteme und andere klassische KI-Anwendungen bereits ersetzt hat, kommt nicht ohne Wissensrepräsentation aus. In Technologien wie den neuronalen Netzen erfolgt diese jedoch implizit, sodass hier auch von sub-symbolischen Systemen gesprochen wird: Es wird ein systematisches Verhalten antrainiert und somit eine Art implizites Wissen über zugrunde liegende Zusammenhänge erworben. Einblicke in den konkreten Lösungsweg sind in diesen Netzen allerdings nur schwer möglich, da allgemeine, den Daten zugrunde liegende Regeln im neuronalen Netz auch nur indirekt, beispielsweise durch Kantengewichte und Aktivierungsschwellenwerte der Neuronen, repräsentiert werden. Mehr zu neuronalen Netzen finden Sie auch in der Aktivität #deeplearning. CS4FN bietet zudem eine weitere Aktivität, die ein Expertensystem für Tic Tac Toe unplugged modelliert.

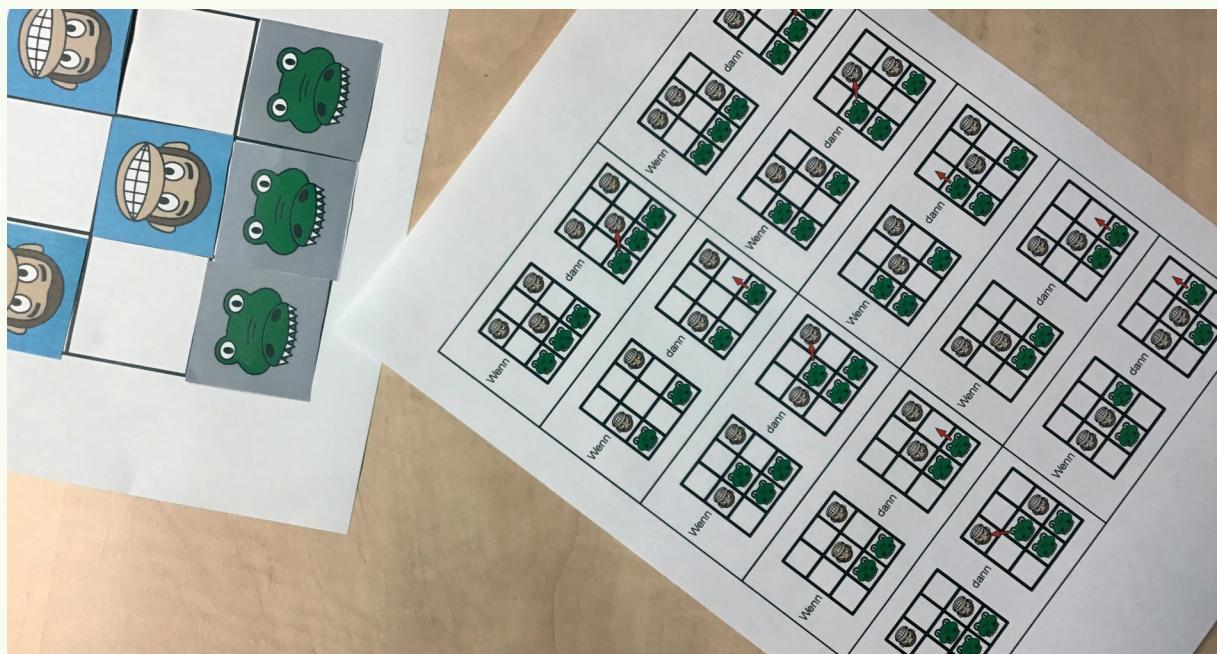


Abb. 14: Spielaufbau: Die Regeln für die Person, die den Computer übernimmt, sind klar definiert

# Der Turing-Test

„Ach wie gut, dass niemand weiß,  
dass ich ein Computer bin!“

## Zielgruppe

Sekundarstufe I, Sekundarstufe II

## Darum geht's

Wie muss sich eine Maschine verhalten, um als intelligent zu gelten? Was bedeutet künstliche Intelligenz genau? Diese Fragen beschäftigen Forscher seit den Anfängen der künstlichen Intelligenz. Alan Turing hat 1950 mit dem Turing-Test eine Idee entwickelt, wie man feststellen könnte, ob eine Maschine intelligent ist. Diese Aktivität stellt den Turing-Test mit Schülerinnen und Schülern nach und soll zur Diskussion anregen, ob Computer tatsächlich so etwas wie menschliche Intelligenz zeigen können. Sie deckt außerdem auf, wie leicht man durch sorgfältig gewählte Beispiele von „Intelligenz“ von einer Maschine in die Irre geführt werden kann.

## Diese Ideen stecken dahinter

- Intelligente Systeme nutzen bestimmte Strategien, um menschliches Verhalten zu imitieren.
- Es braucht spezielle Verfahren, um die Intelligenz von Maschinen zu bewerten.
- Die Definition von (künstlicher) Intelligenz ist nicht eindeutig.

## Das wird benötigt

- Arbeitsblätter/Folie mit vorgegebenen Turing-Test Fragen für die ganze Klasse
- eine Kopie der Antworten auf die Turing-Test-Fragen
- 4 freiwillige Schülerinnen oder Schüler in den Rollen Computer (1x), Mensch (1x) und Läufer (2x)

## So funktioniert's

Die Schülerinnen und Schüler spielen in dieser Aktivität ein Frage-Antwort-Spiel, bei dem sie allein durch das Stellen von Fragen und die Analyse der zugehörigen Antworten versuchen müssen, einen Computer von einem Menschen zu unterscheiden. Dazu nehmen je eine Schülerin bzw. ein Schüler die Rolle eines Computers und eine menschliche Rolle ein. Sie werden von ihren Mitschülerinnen und Mitschülern befragt und anhand der Antworten muss die Klasse bestimmen, wer welche Rolle repräsentiert.

Diese Aktivität entstammt den ursprünglichen CS Unplugged Materialien. Diese sind unter Creative Commons CC-BY-SA durch Bell, Witten und Fellows lizenziert. Das ursprüngliche Material wurde in der vorliegenden Beschreibung frei übersetzt und adaptiert.

## Kontext

Seit Jahrhunderten streiten Philosophen, ob eine Maschine über menschliche Intelligenz verfügen kann bzw., ob das menschliche Gehirn nicht vielleicht auch nur eine sehr gute Maschine ist. Einige halten künstliche Intelligenz für eine absurde Idee, andere glauben, dass wir irgendwann Maschinen entwickeln werden, die genauso intelligent sind wie wir. Künstliche Intelligenz hat viel Potenzial, andererseits schürt die Vorstellung von intelligenten Maschinen auch Ängste.

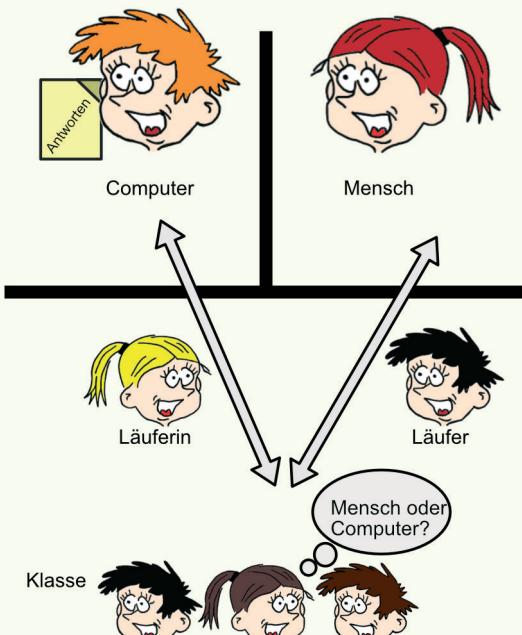


Abb. 15: Aufbau des Turing-Tests

## Ablauf

Thematisieren Sie mit den Schülerinnen und Schülern vor dem Spiel, ob sie Computer für intelligent halten oder annehmen, dass Computer jemals intelligent sein werden. Fragen Sie nach, wie man entscheiden könnte, ob ein Computer intelligent ist und stellen Sie kurz den Turing-Test vor, welcher mit der Aktivität nachgestellt wird.

Vor Beginn der eigentlichen Aktivität werden vier freiwillige Schülerinnen bzw. Schüler bestimmt, sie übernehmen die Rollen des Computers und des Menschen (siehe Abb. 15). Außerdem

gibt es zwei Läufer, die den fairen Verlauf des Spiels absichern und mit Zettel und Stift ausgestattet werden, um die Antworten zu notieren. Die Rollen von Mensch und Computer werden von der Lehrkraft im Geheimen zugeteilt, bevor diese beiden Schülerinnen bzw. Schüler das Klassenzimmer in zwei getrennte Räume verlassen (alternativ können Sie auch eine Trennwand nutzen, es muss jedoch sichergestellt werden, dass sich die Schülerinnen und Schüler nicht gegenseitig sehen). Die Schülerin bzw. der Schüler in der Rolle des Computers erhält dort eine Kopie der Antworten auf die Fragen des Turing-Tests. Jeder der Läufer ist für eine Rolle zuständig, für welche wird ebenfalls geheim gehalten.

Die Klasse muss nun herausfinden, welche Schülerin bzw. welcher Schüler die Rolle des Computers eingenommen hat. Dazu wählt sie vom ausgeteilten Arbeitsblatt pro Runde eine Frage, die an Computer und Mensch gestellt werden soll. Nachdem eine Frage gewählt wurde, sollen die Schülerinnen und Schüler zunächst erklären, warum sie diese Frage für geeignet halten, um den Computer vom Menschen zu unterscheiden. Diese Argumentation ist das zentrale Element dieser Aufgabe, da die Klasse auf diese Weise darüber nachdenkt, wie sich die Antworten einer Person und eines „intelligenten“ Computers unterscheiden.

Dann wird die Frage von den Läufern an ihre Mitschülerinnen bzw. Mitschüler in den anderen Räumen übermittelt, ebenso werden die Antworten von dort von den Läufern zurück in die Klasse gebracht. Dabei muss der Mensch kurz und ehrlich auf die gestellte Frage antworten – also eine menschliche Antwort geben. Der Computer wählt hingegen die entsprechende Antwort vom Arbeitsblatt aus. Wenn die Anweisungen kursiv geschrieben sind, muss der Computer selbst eine Antwort ausarbeiten (z. B. die aktuelle Uhrzeit

angeben). Bei der Übermittlung der gegebenen Antworten müssen die Läufer besonders darauf achten, nicht zu offenbaren, mit wem sie interagieren.

Die Klasse diskutiert nun, welche Antwort wahrscheinlich von einem Computer stammt. Wiederholen Sie den Vorgang mit einigen weiteren Fragen, möglichst so lange, bis die Klasse eine klare Entscheidung treffen kann, wer der Computer ist. Kann die Klasse Mensch und Computer nicht verlässlich unterscheiden, hat der Computer den Turing-Test bestanden.

### Hintergrund

Obwohl noch kein aktuelles Computerprogramm über so etwas wie allgemeine Intelligenz verfügt, ist die Frage, ob Computer dazu grundsätzlich in der Lage sind, noch unbeantwortet. Dies hängt maßgeblich damit zusammen, dass bereits die Definition von Intelligenz kontrovers diskutiert wird.

Vor diesem Hintergrund schlug der britische Mathematiker Alan Turing 1950 ein Verfahren vor, mit dem die Intelligenz einer Maschine bestimmt werden kann ohne eine genaue Definition von Intelligenz zu benötigen. Dieser sogenannte Turing-Test lässt den Computer seine „Intelligenz“ demonstrieren. Das Szenario des Tests ähnelt der oben beschriebenen Aktivität: Eine Fragestellerin bzw. ein Fragesteller interagiert sowohl mit einer Person als auch mit einem Computer via Chat. Kann sie oder er die beiden nicht zuverlässig voneinander unterscheiden, hat der Computer den Turing-Test bestanden. Da die Kommunikation über Chat erfolgt, kann sich der Computer nicht durch physische Eigenschaften, etwa die Stimmlage, verraten. Ein bekanntes Beispiel für ein solches Interaktionssystem ist der Chatbot *Eliza*.

Die Antworten, die eine Schülerin bzw. ein Schüler in der Rolle des Computers gibt, sind denen, die ein „intelligentes“ Computerprogramm erzeugt, nicht unähnlich. Einige der Antworten werden den Computer sehr schnell enttarnen: Ein Mensch wird die Wurzel von 2 kaum auf 20 Stellen genau angeben können. Andere Fragen, bei deren Beantwortung der Computer immer ein bestimmtes Antwortmuster verwendet, entlarven ihn erst nach einiger Zeit. Beispielsweise sind Antworten auf „Magst du XY?“-Fragen, unabhängig voneinander betrachtet, nicht auffällig. Wenn man allerdings mehrere Fragen dieses Typs kombiniert, wird deutlich, dass der Computer formelhaft vorgeht, um Antworten aus den Fragen zu generieren. Die Antworten können auch zeigen, dass der Computer eine Frage falsch interpretiert hat, obwohl dies natürlich auch einem Menschen passieren könnte. Viele Antworten sind vage gehalten und eine Nachfrage würde deutlich machen, dass der Computer den Inhalt der Frage nicht wirklich verstanden hat. Zudem ist es für den Computer oft sicherer mit „Ich weiß es nicht“ zu antworten (z. B. auf die Frage nach der Wurzel von 2). Dies verleiht menschliche Züge, kann aber auch zur Enttarnung führen, wenn diese Taktik zu oft oder bei zu einfachen Fragen angewendet wird. Auch verzögerte und fehlerhafte Antworten, beispielsweise auf arithmetische Probleme, können den Fragesteller länger in die Irre führen. Computer sind somit zwar z. B. durch formelhafte Antworten, Spiegeln der Aussagen des Gesprächspartners, Reaktionen auf Schlüsselwörter, die Verwendung von Redensarten und das Wiederaufnehmen von Themen in der Lage, Gesprächsfähigkeit zu suggerieren, dies ist jedoch nur eine Fassade, die leicht zu durchschauen ist.



# Weitere Anregungen

Links und Details zu diesen Aktivitäten finden Sie auf unserer Website.

- Gesichtserkennung** Unsere Haustüre kann uns vom Postboten unterscheiden, unsere Fotoverwaltungssoftware markiert automatisch unsere Freunde: Gesichtserkennung ist ein bekanntes Anwendungsfeld der KI. Dabei sollte die Technologie so robust sein, dass sie uns auch im Winter trotz Mütze und im Sommer mit Sonnenbrille erkennt. Diese Aktivität vermittelt dieses Prinzip anhand von Cartoon-Charakteren.
- Äffchen, Sherlock Äffchen** Wie kann Wissen so dargestellt werden, dass ein Computer es „verstehen“ und logische Schlüsse daraus ziehen kann? Logik und formale Wissensrepräsentation sind hier von großer Bedeutung! KI-Systeme sind also gar nicht wirklich „intelligent“, sondern nutzen nur sehr geschickt verschiedene Möglichkeiten, um Wissen zu repräsentieren. Diese Art der Wissensrepräsentation kann auch in Logikrätseln abgebildet werden: Entsprechende Rätsel erfordern die Kombination verschiedener Fakten nach gewissen Regeln, um dann zur Lösung zu kommen.
- Brain-in-a-Bag** In dieser Aktivität stellen Schülerinnen und Schüler die Funktionsweise eines neuronalen Netzes selbst mit Schnüren und Klopapierrollen nach. Das fertige Netz ist dann in der Lage, ein Spiel zu spielen.
- Unsupervised Learning** Neben *Supervised* und *Reinforcement Learning* gibt es auch noch sogenannte *Unsupervised Learning* Verfahren: Hier lernen Computer ohne vorher bekannte Zielwerte und ohne Belohnungen. Alleine aus einer Menge von Datenpunkten sollen so z. B. Kategorien (kaufkräftige Kunden in Webshops) oder Anomalitäten (verdächtige Aktivitäten auf Webservern) erkannt werden. Zeichnen Sie mit Kreide ein Koordinatengitter (z. B. auf dem Schulhof) und fordern Sie Ihre Schülerinnen und Schüler auf, sich anhand der beiden Achsen passend im Gitter zu positionieren. Je nach gewählten Achsen können so nicht nur Cluster, sondern auch Ausreißer bzw. Anomalitäten identifiziert werden.

## *Impressum*

**Herausgeber:**

Professur für Didaktik der Informatik  
Friedrich-Alexander-Universität Erlangen-Nürnberg  
Martensstraße 3  
91058 Erlangen  
<https://aiunplugged.org>

**Redaktion und Gestaltung:**

Annabel Lindner, Stefan Seegerer

**Auflage**

2. Auflage, 1000 Stück

Alle Texte und Grafiken dieser Broschüre (außer das FAU Logo) sind, sofern nicht anders angegeben, unter CC BY NC 3.0 lizenziert, das heißt insbesondere, dass Sie das Material in jedwedem Format oder Medium bearbeiten, vervielfältigen und weiterverbreiten dürfen, allerdings nicht kommerziell. Dabei müssen Sie lediglich den Urheber nennen.