

Arbeitsblatt 1 Reinforcement Learning: Die Gerüchteküche

Reinforcement Learning erleben – ganz ohne Computer

Lehrerhandreichung für eine Unplugged-Aktivität im Klassenraum



Reinforcement Learning ist eines der zentralen Konzepte der Künstlichen Intelligenz – aber wie vermittelt man es an junge Schülerinnen und Schüler, die noch keinen Zugang zu den dahinterliegenden mathematischen Modellen haben?

Diese Handreichung nutzt ein einfaches Maus-Käse-Labyrinth-Szenario und übersetzt das abstrakte Prinzip der Wertpropagation in eine körperliche Erfahrung im Klassenraum: Schülerinnen und Schüler sitzen in einem Stuhlrastrer und bilden ein lebendiges Labyrinth.

Ein Schüler hält den „Käse“ (20 Punkte), leere Plätze sind Hindernisse. Durch die einfache Regel „Bester Nachbar minus 1“ breiten sich die Werte Runde für Runde wellenförmig über das Spielfeld aus – genau wie beim echten Reinforcement Learning. Am Ende navigiert eine „Maus“ blind durch das Labyrinth, indem sie nur den Zahlen folgt.

Das Dokument enthält eine vollständige Aufbauanleitung, den ausdruckbaren Drei-Takt-Regelzettel für Schüler, einen kommentierten Beispielablauf Runde für Runde, einen optionalen Wettbewerbsmodus für zwei Teams sowie Transfertabellen und Erweiterungsvarianten.

Überblick

Dauer: ca. 25–35 Minuten | **Schüler:** mindestens 10 (8 Sitzende + 1 Maus + 1 Beobachter) |

Material: Mini-Whiteboards oder Karteikarten, Stifte, 1 Käse-Emoji-Schild

Idee: Schüler sitzen in einem Stuhlrastrer und bilden ein lebendiges Labyrinth. Durch eine einfache Regel – „Bester Nachbar minus 1“ – breitet sich das Wissen über den Käse wellenförmig aus, genau wie beim Reinforcement Learning. Am Ende navigiert eine „Maus“ durch das Labyrinth, indem sie nur die Zahlen der Schüler abliest.

Phase 1: Aufbau des Labyrinths

Bauen Sie mit den Schülern ein Stuhlraster auf – zum Beispiel 3 Reihen à 4 Plätze. Einige Plätze bleiben bewusst **leer** (Stuhl wird weggenommen oder ein Rucksack draufgelegt). Diese leeren Plätze sind die **Hindernisse**.

Beispiel-Aufbau: 3 Reihen × 4 Plätze

(Blick von der Tafel aus auf die Klasse – Reihe 1 ist vorne)

Anna		Chris	Dana 🧀
Emil	Fiona	Greta	
Hakan	Ida		Jan

Varianten: 2×4 (8 Plätze, einfacher), 3×4 (12 Plätze, Standard), 4×4 (16 Plätze, anspruchsvoll). Je nach Klassengröße und gewünschtem Schwierigkeitsgrad anpassen. **Wichtig:** Der Käse sollte in einer Ecke liegen, damit die Wellen sich über das gesamte Feld ausbreiten müssen.

✅ Checkliste Aufbau

Jeder sitzende Schüler hat ein Mini-Whiteboard (oder große Karteikarte) und einen Stift.
 Der Käse-Schüler bekommt das Käse-Schild und schreibt eine große „20“ auf sein Board.
 Alle anderen schreiben eine große „0“ auf ihr Board.
 Jeder hält sein Board so, dass die Nachbarn (links, rechts, vorne, hinten) die Zahl sehen können.
 Ein Schüler (die spätere „Maus“) steht zunächst seitlich und beobachtet.

Phase 2: Die Gerüchte-Regel

Dies ist das Herzstück der Aktivität. Die Regel wird an die Tafel geschrieben und als Regelzettel ausgeteilt. Beides ist wichtig, damit Schüler jederzeit nachschauen können.

REGELZETTEL FÜR SCHÜLER

Die Gerüchteküche: So breitet sich das Wissen aus!

Die Lehrkraft ruft jede Runde drei Kommandos. Du machst immer genau das:

1 „UMSCHAUEN!“

Schau dir die Zahlen deiner **direkten Nachbarn** an: links, rechts, vorne, hinten.
Leere Plätze (Hindernisse) zählen nicht – überspringe sie!
Merke dir die **höchste Zahl**, die du siehst.

2 „RECHNEN!“

Ist die höchste Nachbar-Zahl **größer** als deine eigene Zahl?
→ **Ja: Deine neue Zahl = höchster Nachbar – 1**
→ **Nein:** Du behältst deine Zahl.

3 „ZEIGEN!“

Halte dein Board hoch, damit alle deine **neue Zahl** sehen können.

Wiederhole, bis die Lehrkraft „Stopp!“ ruft.

Warum funktioniert das? (Hintergrund für die Lehrkraft)

Das Drei-Takt-Protokoll („Umschauen → Rechnen → Zeigen“) löst ein zentrales Problem: Wenn alle gleichzeitig updaten, könnten Schüler neue statt alte Werte ihrer Nachbarn ablesen. Durch die Trennung in drei Phasen ist sichergestellt, dass alle Schüler in einer Runde dieselbe Datenbasis nutzen.

Mathematisch entspricht jede Runde einer **synchronen Bellman-Iteration**. Die Vereinfachung „bester Nachbar – 1“ (statt „ $\times \gamma$ “) produziert ganzzahlige Werte und hält das Kernprinzip intakt: Werte sinken mit der Entfernung, und die optimale Policy folgt dem Gradienten.

Phase 3: Durchführung – Runde für Runde

Die Lehrkraft moderiert die Runden. Hier der typische Ablauf am Beispiel des 3×4-Aufbaus:

Startzustand (vor Runde 1)

Dana hält „20“ hoch (Käse). Alle anderen halten „0“ hoch. Die Lehrkraft beschreibt: „Im Moment weiß nur Dana, wo der Käse ist. Alle anderen wissen noch nichts.“

Runde 1

Lehrkraft: „Umschauen!“ – Alle schauen auf ihre Nachbarn.

Lehrkraft: „Rechnen!“ – Chris sitzt neben Dana (20). Sein bester Nachbar ist 20, das ist größer als seine 0. Neue Zahl: $20 - 1 = 19$. Alle anderen sehen nur Nullen → keine Änderung.

Lehrkraft: „Zeigen!“ – Chris hält jetzt „19“ hoch. Alle anderen weiterhin „0“.

Lehrkraft kommentiert: „Das Gerücht über den Käse hat sich einen Platz weit ausgebreitet!“

Runde 2

„Umschauen! – Rechnen! – Zeigen!“ Jetzt sehen Greta und Anna die 19 von Chris. Neue Werte: $19 - 1 = 18$. Fiona sieht nur Nullen → bleibt bei 0.

Lehrkraft: „Das Gerücht breitet sich weiter aus – aber Achtung: Fiona hat noch nichts gehört, weil Hindernisse den Weg versperren!“

Weitere Runden (Welle für Welle)

So geht es weiter, bis **alle Schüler** einen Wert größer als 0 haben. In jeder Runde breitet sich das Wissen um genau eine Stuhlposition weiter aus. Die Lehrkraft kann eine Runde überspringen lassen und fragen: „Wer hat sich verändert? Wer wartet noch?“

Moderationstipp: Der „Aha!“-Moment

Nachdem ca. 3–4 Runden gelaufen sind, halten Sie inne und fragen:

„Was fällt euch auf? Wo sind die höchsten Zahlen? Wo die niedrigsten?“

Die Schüler werden erkennen: Hohe Zahlen sind nahe am Käse, niedrige weit weg. Die Zahlen bilden eine Art „Hügel“ mit dem Käse als Gipfel – das ist die Q-Map!

Typische Schülerfehler und wie Sie damit umgehen

Diagonale Nachbarn zählen: „Nur links-rechts-vorne-hinten! Stell dir vor, die Maus kann nicht schräg laufen.“

Über Hindernisse hinweg schauen: „Der leere Stuhl ist eine Wand. Da kannst du nicht durchschauen – auch nicht drüber rufen!“

Vergessen, die höchste Zahl zu nehmen: Kommt selten vor, da die meisten Nachbarn am Anfang bei 0 stehen. Wird später bei Aufgabe 2 (4×4) relevant.

Wert wird kleiner statt größer: Die Regel „nur updaten, wenn der Nachbar GRÖßER ist“ verhindert das. Betonen Sie: „Ein Gerücht wird nur besser, nie schlechter!“

Phase 4: Die Maus navigiert!

Jetzt kommt der dramatische Moment: Ein Schüler spielt die Maus und testet die soeben erstellte Q-Map.



Anleitung für die Maus

1. Stelle dich hinter den Schüler, der dir als Startposition zugewiesen wurde.
2. Frage diesen Schüler nach seiner Zahl. Schau dann auf die Zahlen der direkten Nachbarn (links, rechts, vorne, hinten).
3. Gehe zum Nachbarn mit der **höchsten Zahl**.
4. Wiederhole, bis du beim Käse angekommen bist.
5. Zähle deine Schritte!

Wichtig: Die Maus darf **nur die Zahlen ablesen** – sie weiß nicht, wo der Käse tatsächlich ist! Trotzdem findet sie den optimalen Weg. Dieser Moment ist der zentrale didaktische Beweis, dass die Q-Map funktioniert: **Reines Zahlenfolgen führt zum Ziel.**

Reflexionsfragen nach der Navigation:

- „Hat die Maus den kürzesten Weg genommen?“
- „Was wäre passiert, wenn die Maus nicht dem höchsten Wert gefolgt wäre?“
- „Was müsste sich ändern, wenn wir den Käse jetzt woanders hinlegen?“

Phase 5 (optional): Der Wettbewerb – Zwei Teams, ein Labyrinth

Für höhere Motivation kann die Aktivität als Wettbewerb zwischen zwei Teams durchgeführt werden.

Spielmodus: „Wer berechnet schneller & fehlerfrei?“

1. Beide Teams bauen **denselben** Stuhlaufbau nebeneinander auf (identische Position von Käse und Hindernissen).
2. Auf „Start!“ beginnen beide Teams gleichzeitig mit den Runden („Umschauen – Rechnen – Zeigen“).
3. Ein Team signalisiert „Fertig!“, sobald alle Schüler einen Wert > 0 haben.
4. Die Lehrkraft prüft stichprobenartig 2–3 Werte. Stimmen sie? → **Die Maus darf starten!** Ist ein Fehler drin? → **Strafzeit** (z.B. eine Extra-Runde rechnen).
5. Die Maus startet am selben Feld. **Erste Maus beim Käse gewinnt.**

Beurteilung: Pro und Contra

✓ Vorteile

- Hohe Motivation und Energie durch Wettkampfcharakter
- Tempo erzwingt sichere Regelanwendung statt Abwarten
- Fehler werden sofort bestraft (Strafzeit) – das betont Genauigkeit
- Die „Maus-Phase“ wird zum dramatischen Finale mit Zuschauern
- Beobachter-Schüler können beide Teams vergleichen und Fehler suchen

⚠ Herausforderungen

- Braucht doppelten Platz und min. 18 Schüler (2×8 + 2 Mäuse)
- Zeitdruck kann dazu verleiten, Werte zu raten statt zu rechnen
- Lehrkraft muss zwei Gruppen gleichzeitig moderieren (Lösung: je ein Schiedsrichter)
- Kompetitive Schüler könnten Frustration bei Verlieren zeigen
- Lerngespräch wird durch Wettkampf-Aufregung schwieriger

💡 Empfehlung

Erst verstehen, dann wetteifern! Führen Sie die Aktivität zuerst einmal **gemeinsam und langsam** im Plenum durch (Phase 1–4). Erst wenn die Regel sitzt, starten Sie den Wettbewerb als Wiederholung oder in der nächsten Stunde. So verbinden Sie tiefes Verständnis mit Spielfreude.

Alternative ohne Platzbedarf: Statt zwei physischer Aufbauten können Sie ein Team „live“ und ein Team „auf Papier“ (mit dem Q-Map-Arbeitsblatt) arbeiten lassen. Beide Lösungen werden verglichen – „Stimmt ihr überein?“ Damit werden Unplugged und Arbeitsblatt elegant verknüpft.

Phase 6: Transfer – Von der Gerüchteküche zur KI

Zum Abschluss stellt die Lehrkraft die Verbindung zum echten Reinforcement Learning her. Hier eine Gegenüberstellung, die Sie an die Tafel schreiben oder als Folie zeigen können:

Gerüchteküche	App (MouseMaze)	Echtes RL
Schüler auf Stühlen	Felder im Gitter	Zustände (States)
Leere Plätze / Rucksäcke	Schwarze Blöcke	Hindernisse
Käse-Schüler (20)	Zielfeld	Reward
Zahlen auf den Boards	Farbige Heatmap	Q-Werte / Value-Map
„Bester Nachbar – 1“	Bellman-Update	$V(s) = \max[R + \gamma \cdot V(s')]$
Eine Runde (U–R–Z)	Eine Iteration	Trainingsschritt
Pfeil zum besten Nachbarn	Policy-Pfeile	Optimale Strategie
Maus-Schüler folgt Zahlen	Agent folgt Policy	Exploitation

Verknüpfung mit der App und dem Arbeitsblatt

Idealer Unterrichtsfluss:

1. **Gerüchteküche (Unplugged)** → Intuition aufbauen, Regel erleben
2. **Q-Map-Arbeitsblatt** → Regel auf Papier anwenden, Werte und Pfeile selbst berechnen
3. **MouseMaze-App** → Gleiche Prinzipien in der Simulation wiedererkennen, mit Parametern experimentieren
4. **Quiz** → Gelerntes überprüfen und transferieren

Erweiterungen und Varianten

„**Käse wandert**“: Nach dem ersten Durchlauf wird der Käse an eine neue Position verschoben. Alle Werte müssen auf 0 zurückgesetzt und neu berechnet werden. Die Schüler erleben, warum eine KI neu trainieren muss, wenn sich die Umgebung ändert.

„**Hindernis verschwindet**“: Mitten im Spiel wird ein Hindernis entfernt (ein Schüler setzt sich auf den leeren Platz). Die Q-Map muss angepasst werden. Wessen Wert ändert sich? Eine hervorragende Transferfrage!

„**Zwei Käse**“ (**Fortgeschritten**): Zwei Schüler halten Käse (beide = 20). Die Wellen treffen sich in der Mitte. Welchen Käse findet die Maus? → Den nächsten! Das demonstriert, dass Q-Learning immer den optimalen Weg findet.

„**Blinde Maus**“: Die Maus trägt eine Augenbinde und kennt die Aufstellung nicht. Sie fragt bei jedem Schritt nur den aktuellen Schüler und dessen Nachbarn nach ihren Zahlen – und findet trotzdem den Weg. **Das verstärkt die Erkenntnis: Die Q-Map enthält alles, was man wissen muss.**