3 Theorie zu maschinellem Lernen: Kann Quanteninformation bei der Schaffung künstlicher Intelligenz helfen?

V. Dunjko et al.: Quantum-Enhanced Machine Learning
Phys. Rev. Lett. 117, 130501 (2016)

3.1 Zu maschinellem Lernen in klassischen und Quantensystemen

Sehr abstrakte Behandlung im Artikel → Nur die Konzepte sind wichtig, keine konkrete Implementierung / kein Beispiel

3.1.1 Arten des maschinellen Lernens

a) Supervised – überwacht

Idee: Maschine lernt anhand gegebener korrekter Antworten

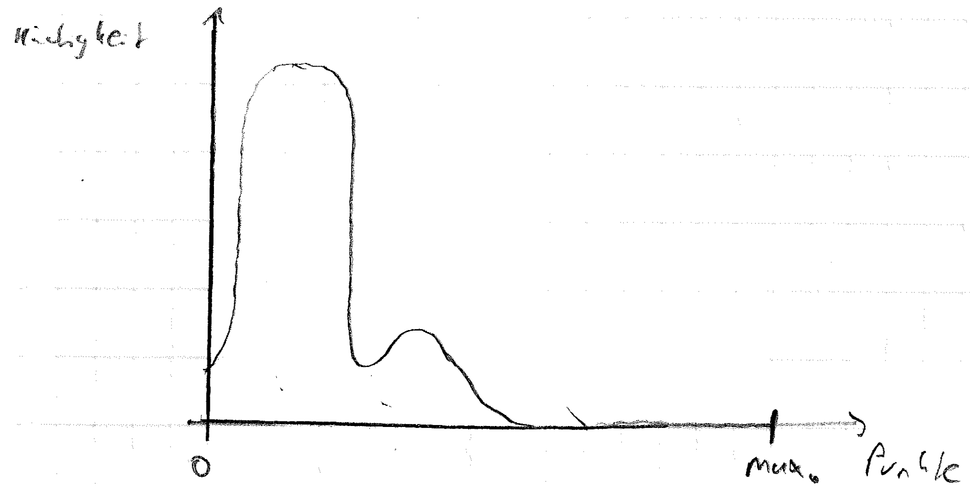
Maschine Daten x → Umgebung
← Label y

→ wird mit Paaren (x, y) trainiert → Maschine erkennt daraus einen Zusammenhang
→ primitives Beispiel: Interpolation von $y = f(x)$

b) unsupervised – unüberwacht

Idee: Man gibt eine Menge an Datenpunkten ein $f(x)$:
→ Die Maschine sucht eigenständig nach Gemeinsamkeiten / Mustern und gruppiert die Datenpunkte danach.

Beispiel: MMP-Übungsklausur 2017



↑ ↑

Maschine erkennt 2 Gruppen.

c) <u>reinforcement – bestärkend</u>

Idee: Maschine führt eine Reihe von Aktionen aus und lernt anhand der Rückmeldung erfolgreich (Belohnung) oder unerfolgreich (Strafe), ob das Vorgehen sinnvoll ist.

Beispiel: Maschine lernt ein Spiel, führt Züge aus:
- -> nähert sich Gewinnen: positive Bewertung der Züge
- -> nähert sich Verlieren: negative Bewertung der Züge

3.1.2 <u>Einzug der Quantenmechanik im Sinn des Artikels</u>

- Ausnutzung von Quantenalgorithmen um
  - die Komplexität der Rechnung zu reduzieren (typisch für Quanteninformationsverarbeitung)
  - den notwendigen Satz an Datenpunkten, der zum

lernen notwendig ist, reduzieren.

3.1.3 Genutzter Quantenalgorithmus

· Grover-Algorithmus zur inversen Suche

Problemstellung: Telefonbuch nach Namen sortiert, es wird aber nach dem Namen gesucht wenn man die Nummer kennt.

· Grundlegende Idee: Man besitzt eine Orakel-Funktion, die zurück gibt

$P(\text{Name}) = 1$, falls hinter dem Namen die gesuchte Nummer liegt.

$P(\text{Name}) = 0$, falls hinter dem Namen eine andere Nummer liegt.

· Das kann in einigen Schritten ausgenutzt werden, um die Amplitude eines Zustands, der die richtige Nummer enthält, zu verstärken

Anfang: $|\psi\rangle = \frac{1}{\sqrt{3}}\left(|a\rangle + |b\rangle + |c\rangle\right)$

Ende: $|\psi'\rangle = 10^{-7}|a\rangle + |b\rangle + 10^{-8}|c\rangle$

→ Messung in der Namen-Basis liefert sicher den richtigen Wert.

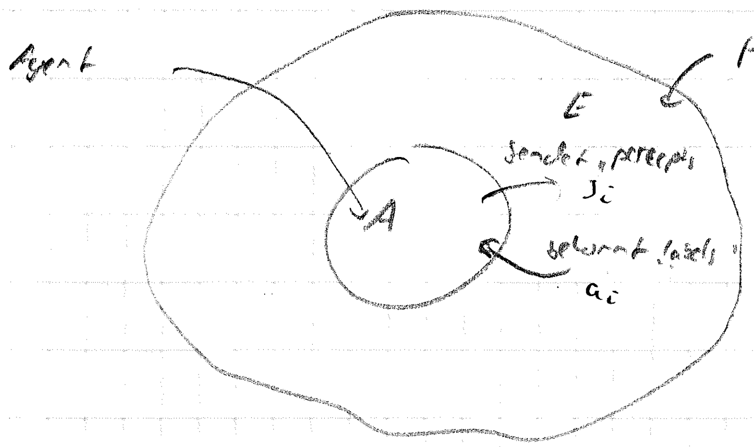
→ Benötigt $O(\sqrt{N})$ statt $O(N)$ Schritte.

3.2 Herangehen im Artikel

- Konstruktion eines Beispiels zur Anwendung des Grover-Algorithmus im „reinforcement learning". → sehr restriktive Bedingungen, dabei bleibt das Problem aber vollkommen abstrakt.

3.2.1 Grundlegender Aufbau

a) Umgebung und Maschine



$(\mathcal{H}_A, \mathcal{H}_S)$ enthält die Information, über die A lernen soll

Reinforcement learning → zusätzlich zu den a_i wird zurückgeliefert ob die Paare (s_i, a_i) (nachdem mehrere ausgetauscht wurden) wertvoll waren (z.B. gute Kombination von Spielzügen) → reward status.

In der QM: Hilberträume $\mathcal{H}_A$, $\mathcal{H}_S$

Werte werden zu Zuständen $|a_i\rangle$ $|s_i\rangle$

b) <u>Register</u>

· In der QM: Speichern der Zustände in Registern,
gemeinsames Rechen-Register R_C
→ s. Abb. 1

Operatoren wirken in der Form

$$\begin{pmatrix} (M_A) & \\ & \mathbb{1} \end{pmatrix} \begin{pmatrix} \\ \\ \\ \end{pmatrix} \begin{matrix} \}R_A \\ \}R_C \\ \}R_S \end{matrix} \qquad \begin{pmatrix} \mathbb{1} & \\ & (M_E) \end{pmatrix} \begin{pmatrix} \\ \\ \\ \end{pmatrix} \begin{matrix} \}R_A \\ \}R_C \\ R_S \end{matrix}$$

· Im wesentlichen werden kontrollierte Gatter angenommen, um die Historie auszulesen

$$U_t^T \left(|x\rangle_{R_C} \otimes |\psi\rangle_{R_T} \right) = |x\rangle_{R_C} \otimes U_t^x \, |\psi\rangle_{R_T}$$

Warum kein Widerspruch zum no cloning-Theorem?
→ Es wird immer dieselbe Basis verwendet, und Zustände in dieser angenommen! → „classical Basis"
→ no cloning-Theorem gilt natürlich, man muss so nicht den tatsächlichen Zustand von R_T oder R_C erhalten. Der gemessene Zustand von R_T wird als Schritt definiert!

c) <u>Betrachtet wird</u>

A und E sind Quantensysteme → QQ-Variante
→ Konsequenz: nur Komplexität der Rechnung kann verbessert werden

3.2.2 Konzept

a) Ausgangspunkt und Eingrenzung

Das Problem

- lässt sich auf ein Quantensystem abbilden
  $E \to E^q$, es gibt ein E^q!
- Die Frage lässt sich auf ein Orakel zurückführen, das dem „reward status' entspricht. $\to |a_1, \dots, a_n\rangle$
  $\to (-1)^{R(a_i)} |a_1, \dots, a_n\rangle$
- E ist deterministisch
- E ist epochal (kehrt nach einer Reihe von Zeitschritten zum ursprünglichen Zustand zurück.)
- muss so gestaltet sein, dass Erfolg in der Vergangenheit, beim Weiterverfolgen desselben Schemas (also Anwendung genau dieses Wissens) diesen auch in der Zukunft garantiert → mehr Erfolg in der Vergangenheit = mehr Erfolg in der Zukunft
  → luck favoring

b) Vorgehen

- Suche die Sequenz, die die Maschine bestmöglich trainiert und trainiere sie danach
  → „reward status" muss danach konstruiert sein.
  → Suche kann mit Grover-Algorithmus optimiert werden!

→ Diskussion anhand Abb. 2