Deep Learning im gesellschaftlichen Kontext

Jerom Schult
jerom.schult@haw-hamburg.de
Ausarbeitung Grundseminar an der HAW Hamburg
Berliner Tor 7, 20099 Hamburg, Deutschland

28. Februar 2019

1 Motivation

Die Relevanz der verschiedenen Machine Learning Verfahren, darunter Deep Learning, wächst kontinuierlich. Sowohl in wissenschaftlichen Publikationen, als auch in den Medien ist das Thema mittlerweile stark vertreten. Dies liegt zum großen Teil an der, mit der Zeit rasant steigenden, Rechengeschwindigkeit moderner Computer.

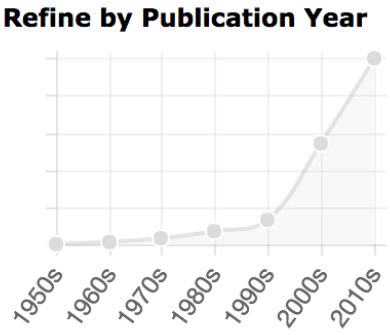

Abbildung 1: Suchhäufigkeit nach „Machine Learning“ in der ACM Digital Library (vergl. ACM Digital Library (2018))

Dabei reicht die Spanne der Anwendungsfelder von Experimenten zu selbstfahrenden Autos (vergl. Bojarski u. a. (2016)) über Anomaliedetektion in Krankendaten hin zu sehr guten automatischen Spracherkennungsverfahren mit Reccurent Neural Networks (vergl. Hannun u. a. (2014)). Bei allen diesen Beispielen wird eine Klassifikation oder Entscheidungsfindung durch Algorithmen vorgenommen.
Als weiteres Beispiel dient hierfür die Lösung von Videospielen mittels Machine Learning Verfahren wie z.B. von DeepMind (vergl. DeepMind (2018)) gezeigt.
Mnih u. a. (2015) lösen mithilfe von Künstlichen Neuronalen Netzen und einer angepassten Version des Q Learning Verfahrens (vergl. Kapitel 2) unterschiedliche, klassische Videospiele mit einem generalistischen Ansatz. Daran zeigt sich die Fähigkeit von Deep Learning Verfahren Probleme ohne die Verwendung von Domänenwissen zu lösen. Diese Fähigkeit beherbergt viele spannende Möglichkeiten und Anwendungsfelder, aber auch einige Risiken.
Dabei ist das Bewusstsein über die Risiken

einer solchen algorithmischen Entscheidungsfindung, die vor allem außerhalb von Simulationen erheblich sein können, sehr schwach ausgeprägt. Forschende und IT-Professionals haben oft vor allem die Möglichkeiten und die Profitabilität solcher Verfahren im Blick. Beim Rest der Gesellschaft geht das Bewusstsein über mögliche Gefahren gegen Null (Meinung des Autors).

2 Reinforcement Learning mit KNN

Um ein erstes Verständnis für Deep Learning Verfahren und ihre Möglichkeiten zu bekommen eignen sich besonders Spiele, als Simulationen der Realität.

2.1 Agent & Umgebung

Die Grundlage für Machine Learning Verfahren auf spielen wurde bereits bei Sutton und Barto (1998) gelegt, wo Reinforcement Learning Probleme durch zwei Komponenten dargestellt werden. Die Komponente die Entscheidungen trifft und versucht aus ihnen zu lernen ist der „Agent“, die Komponente, die alles außerhalb des Agenten darstellt ist die „Umwelt“. Nach jeder Aktion a_t aus den aktuell Möglichen Aktionen aus dem gesamten Aktionsraum, die der Agent an die Umgebung mitteilt, reicht diese ihm einen neuen Zustand $s_t + 1$ und eine Belohnung $r_t + 1$ zurück. Von diesen Parametern abhängig erstellt der Agent eine Policy π mit dem Ziel, diese an eine optimale Policy anzunähern.

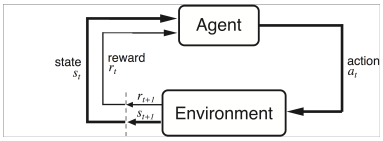

Abbildung 2: Darstellung der Interaktion von Agent und Umgebung nach Sutton und Barto (1998)

2.2 Q Learning mit KNN

Mnih u. a. (2015) bedienten sich in ihrem Paper einer Form des Q Learning um die unterschiedlichen Atari-Spiele zu lösen. Klassisch basiert Q-Learning auf einer Funktion, die Belohnungswerte für Zustände approximiert.

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha\left[r_{t+1} + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t)\right] \tag{1}$$

one-step Q Learning nach Sutton und Barto (1998)

Dabei bekommt die Approximationsfunktion Q in jedem Schritt einen Zustand und eine dazugehörige Belohnung r_{t+1} für die zuletzt ausgeführte Aktion und verrechnet diese mit der bestmöglichen, erwarteten Belohnung des nächsten Schrittes $\max_a Q(s_{t+1}, a)$. Über den Discount-Faktor γ kann gesteuert werden wir stark die erwartete, zukünftige Belohnung in die Bewertung eines Zustand-Aktion Paares einfließen soll. Die Lernrate α steuert die Geschwindigkeit in der die Approximationsfunktion Q in die gewählte Richtung läuft. Die vorgestellten γ und α sind nur zwei von vielen verschiedenen, möglichen Hyperparameter beim Tuning von Machine Learning Verfahren.

Initialize replay memory D to capacity N
Initialize action-value function Q with random weights θ
Initialize target action-value function $\hat{Q}$ with weights $\theta^{-} = \theta$
For episode = 1, M **do**
  Initialize sequence $s_1 = \{x_1\}$ and preprocessed sequence $\phi_1 = \phi(s_1)$
  For t = 1,T **do**
    With probability ε select a random action a_t
    otherwise select $a_t = \text{argmax}_a Q(\phi(s_t), a; \theta)$
    Execute action a_t in emulator and observe reward r_t and image x_{t+1}
    Set $s_{t+1} = s_t, a_t, x_{t+1}$ and preprocess $\phi_{t+1} = \phi(s_{t+1})$
    Store transition $(\phi_t, a_t, r_t, \phi_{t+1})$ in D
    Sample random minibatch of transitions $(\phi_j, a_j, r_j, \phi_{j+1})$ from D
    Set $y_j = \begin{cases} r_j & \text{if episode terminates at step j+1} \\ r_j + \gamma \max_{a'} \hat{Q}\left(\phi_{j+1}, a'; \theta^{-}\right) & \text{otherwise} \end{cases}$
    Perform a gradient descent step on $\left(y_j - Q\left(\phi_j, a_j; \theta\right)\right)^2$ with respect to the network parameters θ
    Every C steps reset $\hat{Q} = Q$
  End For
End For

Abbildung 3: Kompletter Algorithmus zum Q Learning mit Künstlichen Neuronalen Netzen mit Double Q Learning und Memory Replay nach Mnih u. a. (2015)

Das Team von DeepMind erweiterte das vorgestellte Q Learning um Memory Replay (vergl. Liu und Zou (2017a)) und Double Q Learning (vergl. Van Hasselt u. a. (2016)) und erreichte damit sehr gute Ergebnisse in einigen Spielen ohne dem Algorithmus Domänenwissen beizuführen. Allerdings sind die Ergebnisse stark von der Komplexität der zu lösenden Spiele abhängig (vergl. Mnih u. a. (2015)).

Der vorgestellte Algorithmus initiert eine Replay Memory D mit N Plätzen, in der gemachte Erfahrungen gespeichert werden (vergl. Liu und Zou (2017b)). Diese mappen Ausgangs- und Folgezustände mit ihren jeweiligen Belohnungen. Es wird ein Neuronales Netz mit den Gewichtsbelegungen Θ zum Training initiiert, und ein zweites Netz Θ_i^{-} das für das von Van Hasselt u. a. (2016) vorgestellte Double Q Learning Verfahren benötigt wird. Dieses führt zu einem verlässlicheren Modell und hat in Mnih u. a. (2015) großen Einfluss auf den Lernerfolg der Modelle.

Das Lernverfahren wird für M Episoden durchlaufen, wobei eine Episode einer Spielpartie entspricht. In einer Schleife wird jeweils bis zum Spielende der folgende Algorithmus ausgeführt. Über eine ϵ-Greedy Policy wird die nächste Aktion vom Neuronalen Netz Q ausgewählt. Die Explorationsrate ϵ steuert dabei die Exploration des Lernverfahrens (vergl. Kaelbling u. a. (1996)) nach einer ϵ-Greedy Policy. Dabei ergibt es mit fortgeschrittener Lerndauer Sinn auf bereits bewertete Zustände zuzugreifen, um den Agenten zum Ziel zu führen. Die gewählte Aktion wird nun vom Agenten an die Umwelt kommuniziert, die den Folgezustand und eine Belohnung zurückgibt. Die so herbeigeführte Erfahrung wird in der Replay-Memory D gespeichert, aus der später ein Batch an Erfahrungen für ein Gradientenabstiegsverfahren des Künstlichen Neuronalen Netztes benutzt wird.

2.3 Probleme

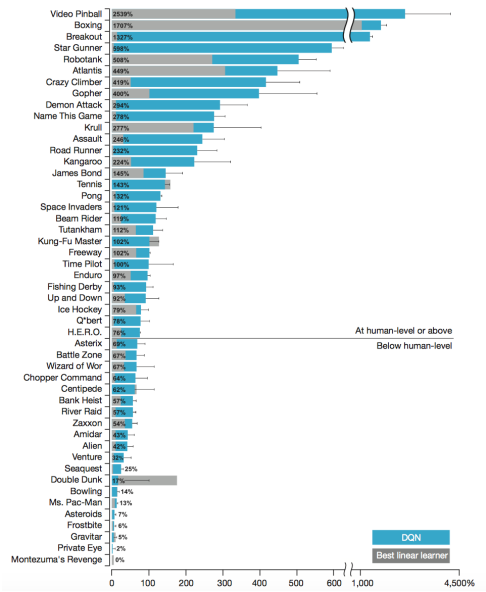

Abbildung 4: Vergleich des Netzes von DeepMind mit den damals besten Vergleichswerten aus der Literatur und menschlichen Vergleichswerten für eine Reihe klassischer Atari-Spiele Mnih u. a. (2015)

Das Team von DeepMind zeigt, dass Machine Learning Verfahren in der Lage sind Simulationen der Realität befriedigend zu lösen, aber bei realen Anwendungsfällen oft noch Probleme mit der schieren Komplexität der zu lösenden Aufgaben haben (vergl. Abbildung 4). Weiterhin ist es bisher meist nur möglich eine optimale Belegung aller möglichen Hyperparameter experimentell zu bestimmen. Domhan u. a. (2015) beschreiben z.B. einen automatisierten Ansatz für das Hyperparameter-Tuning, der jedoch weiterhin auf randomisierter Suche beruht. Mit solchen Verfahren kann für die Qualität der Lernverfahren nur eine Annahme getroffen werden, da nicht alle möglichen Eingaben überprüft werden und keine beweisbar optimale Parameterbelegung gefunden werden kann.

3 Deep Learning in der Gesellschaft

Machine Learning Verfahren, darunter Deep Learning, erhalten langsam auch in der Gesellschaft einen immer größeren Fokus. 2018 veröffentlichte die deutsche Bundesregierung ein Eckpunktepapier und eine Strategie zum Thema Künstliche Intelligenz. Darin wird unter anderem die Absicht beschrieben einen starken Fokus auf die „ethischen und rechtlichen“ (vergl. deutsche Bundesregierung (2018)) Grenzen bei der Anwendung von Machine Learning Verfahren zu legen und stark auf die Sicherheit der genannten Verfahren zu achten.
Ein Beispiel für eine möglicherweise fatale (Meinung des Autors) Anwendung von algorithmischer Entscheidungsfindung ist der in China geplante „Social Score“ (vergl. Lee (2018)), durch den Bürger von einer Software in unterschiedliche Kategorien eingeteilt werden, abhängig von über sie verfügbaren Daten. Ein großes Problem ist dabei, neben der möglichen Diskriminierung aufgrund fehlerhafter oder nicht-neutraler Trainingsdaten, z.B. Diskriminierung nach Hautfarbe, Geschlecht (vergl. Sam Corbett-Davies (2017)), die Nicht-Nachvollziehbarkeit der Entscheidungen. Der gemeinnützige Verein „Algorithm Watch“ zeigte am Beispiel der „Schufa“, dass ein Nachvollziehen der Entscheidungen durch algorithmische Verfahren (und in diesem Sinne können die Bewertungskriterien der Schufa als solches gesehen werden) im Nachhinein eine große Herausforderung darstellt (vergl. Algorithmwatch (2018)). Für ein Reverse-Engineering des Entscheidungs-Modells müssten im besten Fall die gesamten und gleichen Trainingsdaten vorliegen. Selbst dann wäre es durch die oftmali-

ge, randomisierte Vorbelegung der einzelnen Gewichts- und Hyperparameter bei Machine Learning Verfahren sehr wahrscheinlich, dass ein stark unterschiedliches Modell hergeleitet wird.

4 Aktuelle Forschung

Das Problem der Nicht-Nachvollziehbarkeit ist in der Forschung angekommen und wird unter dem Begriff „Explainable AI“ zusammengefasst. Eine Explainable AI ist in diesem Fall ein Modell der künstlichen Intelligenz, das z.B. die für eine Entscheidung auschllaggebenden Attribute des Input nachvollziehbar macht.

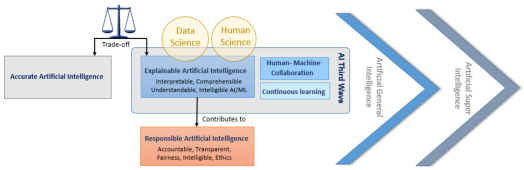

Abbildung 5: Darstellung der Wirkungsbereiche von XAI nach Adadi und Berrada (2018)

Dazu gibt es laut der DARPA (vergl. Gunning (2017)) unterschiedliche Ansätze: Bei der „Deep Explanation“ wird direkt beim Training des Modells auf den Trainingsdaten ein eigens dafür genutzter, beobachtender Lernvorgang z.B. mit Recurrent Neural Networks genutzt, der zu jedem Output des beobachteten Modells lernt welche Features des Input für die Entscheidung wichtig waren, also welche Attribute des Input die Entscheidung definieren. So Kann zum Beispiel für jede Kategorisierung eines Bildes durch ein Modell nachvollzogen werden, welche Regionen des Bildes die Entscheidung beeinflussten. Beim Bild eines Schweins, könnten dies die Bereiche des Bildes sein, die Nase und Schwanz abbilden.
Bei „Interpretable Models“ wird direkt ein strukturiertes, interpretierbares, kausales Model für die Entscheidungsfindung gewählt. Ein einfaches Beispiel dafür sind Entscheidungsbäume, die vom Grundsatz viel besser nachvollzogen werden können, als z.B. Ansätze mit Künstlichen Neuronalen Netzen.

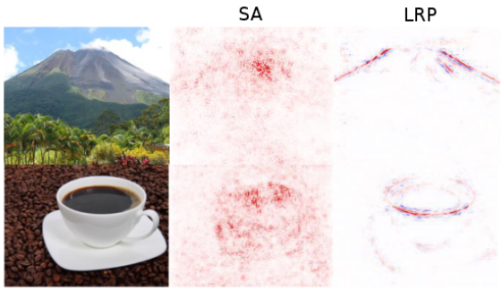

Abbildung 6: Darstellung einer Model Induction nach Samek u. a. (2017)

Bei der „Model Induction“ wird das Model als Black-Box betrachtet und durch taktische Eingabe von Input-Werten und Beobachtung der Outputs des beobachteten Modells hergeleitet welche Input-Regionen für die Entscheidung ausschlaggebend waren. So könnte bei einem Modell, dass erkennen soll ob auf dem Input-Bild ein Zebra abgebildet ist oder nicht überprüft werden, ob das Modell auch ein Bild einer Jacke mit Zebrastreifen-Aufdruck als positiv erkennt. In dem Fall könnte man darauf schließen, dass das Modell das Zebra allein aufgrund der Streifen erkennen will und überprüfen ob dieses Vorgehen für den gewählten Anwendungsfall ausreicht.

Einen weiteren Forschungsbereich stellen die interdisziplinären „Science and Technology

Studies“ dar, die sich mit den Auswirkungen der Nutzung digitaler Errungenschaften in der Gesellschaft befassen, darunter Machine Learning Verfahren (vergl Jasanoff u. a. (2001)). Aus der Zusammenarbeit von Soziologen, MINT-Forschern und weiteren Disziplinen sollen Möglichkeiten und Gefahren von Anwendungen Künstlicher Intelligenz in der Gesellschaft erkannt und moralisch, ethische Leitfäden und Grenzen entwickelt werden. Bryson und Winfield (2017) liefert einen ersten Vorschlag, wie sich Standards für die Entwicklung von Verfahren Künstlicher Intelligenz einsetzen lassen könnten.

5 Fazit

Machine Learning Verfahren wie Deep Learning können in sehr unterschiedlichen, komplexen Anwendungsfällen eingesetzt werden. Dabei könnten sie dabei helfen Aufgaben zu lösen, die als für andere Verfahren zu komplex gelten. Die sich daraus ergebenden Möglichkeiten stehen aktuell stark im Fokus der öffentlichen Debatte über Künstliche Intelligenz. Dabei werden, wie in Kapitel 3 skizziert, die Gefahren und ethischen Implikationen der Technologie oft unterschätzt, bzw. vernachlässigt. Ansätze wie Explainable AI und die Science and Technology Studies sind ein guter Anfang, um die Thematik zu erforschen. Die Tatsache, dass auch die deutsche Politik sich in einem Strategie-Papier mit der Thematik auseinandersetzt macht Hoffnung. Jedoch sind dies alles nur erste Schritte. Es reicht nicht sich theoretisch mit dem Thema auseinanderzusetzen. Im Bedarfsfall wird es, wie bei anderen die Gesellschaft betreffenden Themen, darauf ankommen ob, und inwiefern die Experten des Bereichs auf die Thematik und ihre Implikationen aufmerksam machen können.

Literatur

[ACM Digital Library 2018] ACM Digital Library: *ACM Digital Library (2018).* https://dl.acm.org/. 2018. – Accessed: 2018-09-29

[Adadi und Berrada 2018] Adadi, A. ; Berrada, M.: Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI). In: *IEEE Access* 6 (2018), S. 52138–52160. – ISSN 2169-3536

[Algorithmwatch 2018] Algorithmwatch: *OpenSchufa - Warum wir diese Kampagne machen.* https://algorithmwatch.org/de/openschufa-warum-wir-diese-kampagne-machen/. 2018

[Bojarski u. a. 2016] Bojarski, Mariusz ; Testa, Davide D. ; Dworakowski, Daniel ; Firner, Bernhard ; Flepp, Beat ; Goyal, Prasoon ; Jackel, Lawrence D. ; Monfort, Mathew ; Muller, Urs ; Zhang, Jiakai ; Zhang, Xin ; Zhao, Jake ; Zieba, Karol: End to End Learning for Self-Driving Cars. In: *CoRR* abs/1604.07316 (2016). – URL http://arxiv.org/abs/1604.07316

[Bryson und Winfield 2017] Bryson, Joanna ; Winfield, Alan: Standardizing ethical design for artificial intelligence and autonomous systems. In: *Computer* 50 (2017), Nr. 5, S. 116–119

[deutsche Bundesregierung 2018] Bundesregierung deutsche: *Eckpunkte der Bundesregierung für eine Strategie Künstliche Intelligenz.* https://www.bmwi.de/Redaktion/DE/Downloads/E/eckpunktepapier-ki.pdf?__blob=publicationFile&v=10. 2018

[DeepMind 2018] DeepMind: *DeepMind*. https://deepmind.com/research/. 2018. – Accessed: 29-09-2018

[Domhan u. a. 2015] Domhan, Tobias ; Springenberg, Jost T. ; Hutter, Frank: Speeding up automatic hyperparameter optimization of deep neural networks by extrapolation of learning curves. In: *Twenty-Fourth International Joint Conference on Artificial Intelligence*, 2015

[Gunning 2017] Gunning, David: Explainable artificial intelligence (xai). In: *Defense Advanced Research Projects Agency (DARPA), nd Web* (2017)

[Hannun u. a. 2014] Hannun, Awni Y. ; Case, Carl ; Casper, Jared ; Catanzaro, Bryan ; Diamos, Greg ; Elsen, Erich ; Prenger, Ryan ; Satheesh, Sanjeev ; Sengupta, Shubho ; Coates, Adam ; Ng, Andrew Y.: Deep Speech: Scaling up end-to-end speech recognition. In: *CoRR* abs/1412.5567 (2014). – URL http://arxiv.org/abs/1412.5567

[Jasanoff u. a. 2001] Jasanoff, Sheila ; Markle, Gerald E. ; Peterson, James C. ; Pinch, Trevor: *Handbook of science and technology studies*. Sage publications, 2001

[Kaelbling u. a. 1996] Kaelbling, Leslie P. ; Littman, Michael L. ; Moore, Andrew P.: Reinforcement Learning: A Survey. In: *Journal of Artificial Intelligence Research* 4 (1996), S. 237–285. – URL http://people.csail.mit.edu/lpk/papers/rl-survey.ps

[Lee 2018] Lee, Felix: *Die AAA-Bürger*. https://www.zeit.de/digital/datenschutz/2017-11/china-social-credit-system-buergerbewertung. 2018

[Liu und Zou 2017a] Liu, Ruishan ; Zou, James: The Effects of Memory Replay in Reinforcement Learning. In: *CoRR* abs/1710.06574 (2017). – URL http://arxiv.org/abs/1710.06574

[Liu und Zou 2017b] Liu, Ruishan ; Zou, James: The Effects of Memory Replay in Reinforcement Learning. In: *CoRR* abs/1710.06574 (2017). – URL http://arxiv.org/abs/1710.06574

[Mnih u. a. 2015] Mnih, Volodymyr ; Kavukcuoglu, Koray ; Silver, David ; Rusu, Andrei A. ; Veness, Joel ; Bellemare, Marc G. ; Graves, Alex ; Riedmiller, Martin ; Fidjeland, Andreas K. ; Ostrovski, Georg ; Petersen, Stig ; Beattie, Charles ; Sadik, Amir ; Antonoglou, Ioannis ; King, Helen ; Kumaran, Dharshan ; Wierstra, Daan ; Legg, Shane ; Hassabis, Demis: Human-level control through deep reinforcement learning. In: *Nature* 518 (2015), 02, S. 529 EP –. – URL http://dx.doi.org/10.1038/nature14236

[Sam Corbett-Davies 2017] Sam Corbett-Davies, Avi Feller Sharad Goel Aziz H.: *Algorithmic decision making and the cost of fairness*. https://arxiv.org/abs/1701.08230. 2017

[Samek u. a. 2017] Samek, Wojciech ; Wiegand, Thomas ; Müller, Klaus-Robert: Explainable Artificial Intelligence: Understanding, Visualizing and Interpreting Deep Learning Models. In: *CoRR* abs/1708.08296 (2017). – URL http://arxiv.org/abs/1708.08296

[Sutton und Barto 1998] Sutton, Richard S. ; Barto, Andrew G.: *Reinforcement Learning:*

An Introduction. MIT Press, 1998. – URL http://www.cs.ualberta.ca/~sutton/book/the-book.html

[Van Hasselt u. a. 2016] Van Hasselt, Hado ; Guez, Arthur ; Silver, David: Deep Reinforcement Learning with Double Q-Learning. In: *AAAI* Bd. 2 Phoenix, AZ (Veranst.), 2016, S. 5