

Künstliche Intelligenz in der Sozialen Arbeit: Potenziale und Risiken einer neuen Technologie

**Treffpunkt: Soziale Diagnostik
8.11.2023**

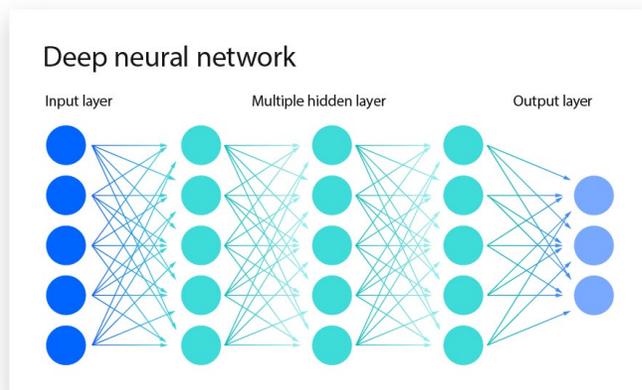
Prof. Dr. Olivier Steiner, Institut Kinder- und Jugendhilfe
8.11.2023

Inhalt

- Begriffe
- (Potenzielle) KI-Anwendungen in der Sozialen Arbeit
- Prädiktive Risikomodellierung PRM
- Chatbots

Begriffe

- Ein Algorithmus ist eine (programmierte) Anleitung, die beschreibt, wie man Schritt für Schritt ein Ziel erreicht. (F. Stalder: Code, der kulturelle Prozesse organisiert)
- Machine Learning/Maschinelles Lernen: Computer lernen eigenständig. Der Algorithmus lernt dabei durch Beispiele und Versuche, ein Problem möglichst gut zu lösen.
- Künstliche neuronale Netzwerke: Imitation biologischer neuronaler Netze: Nodes, Schichten („Deep Learning“), selbstlernend



<https://www.ibm.com/topics/neural-networks>



https://www.youtube.com/watch?v=GI3EjiVlz_4

Steiner & Tschopp (2022)

(Potenzielle) KI-Anwendungen in der Sozialen Arbeit

- Medienpädagogischer Einsatz: KI-Bildgenerierung, Natural Language Models (ChatGPT)
- Training in AI-VR Simulationen, Generierung von Fallbeispielen
- Kommunikation: Chatbots (Beratung, Überwachung online Foren)
- Anamnese/Diagnostik: Prädiktive Risikomodellierung, Prediktive Riskmodellung (PRM)
- Interpretation: Analyse Textdokumentation in Fallsoftware (Vgl. Projekt MAEWIN)
- AI Community Development, Stadtentwicklung (crowd sourced data)
- ChatGPT schreibt Studierendenarbeiten...

<https://act-on.jff.de/ki-in-der-medienpaedagogischen-praxis/>

Vgl. <https://www.socialworkers.org/About/Ethics/AI-and-Social-Work>

Steiner & Tschopp (2022) Hsu et al. (2022).

Prädiktive Risikomodellierung PRM

- Datenbasierte Vorhersagen mittels Algorithmen über menschliches Verhalten und Entwicklung
- Einsatz bereits seit den 1980ern (klassische Algorithmen, u.a. Addierung von Grenzwerten, Regressionen)
- Kaum Verbreitung in der Sozialen Arbeit im deutschsprachigen Raum
- Anwendungen: Eckerd Rapid Safety Feedback tool (ERSF) (in Illionois, USA, abgesetzt wg. Übersensibilität und Verfehlung schwerwiegender Fälle), Allegheny Family Screening Tool (AFST, USA) (Kritik wg. rassistischem Bias), Vulnerable Children Predictive Modelling Neuseeland (Voraussagekraft 76%)

Schwartz, D. R., Kaufman, A. B., & Schwartz, I. M. (2004). Computational intelligence techniques for risk assessment and decision support

- 1767 Fälle der «United States' Third National Incidence Study of Child Abuse and Neglect»
- Input: 30 Variablen, Output: Harm/No Harm
- Neuronales Netz: 2/3 der Stichprobe Training, 1/3 Test

Table 2

Accuracy of the neural network

Total number of cases in the testing set	617		
Total number of erroneous predictions	64		
Percent of total that was erroneous	10.4		
Error type	Number of this error type	Percent of total	Percent of total errors
Indeterminate	48	7.8	75
False positive	4	0.6	6.3
False negative	12	1.9	18.8

"This book is downright scary—but...you will emerge smarter and more empowered to demand justice." —NAOMI KLEIN



AUTOMATING INEQUALITY

HOW HIGH-TECH TOOLS PROFILE,
POLICE, AND PUNISH THE POOR



VIRGINIA EUBANKS



"Wenn automatisierte Entscheidungsfindungsinstrumente nicht explizit zum Abbau struktureller Ungleichheiten entwickelt werden, verstärken sie diese durch ihre Geschwindigkeit und ihren Umfang."

Problematiken/Herausforderungen PRM

- Verfügbarkeit von und Zugang zu Verwaltungsdaten (Datenqualität)
- Personalkapazität, -skills & Kosten
- Systematische Bias (Verstärkung Sozialer Ungleichheit, Mainstreaming)
- Verstärkung von Überwachung, Datenschutz
- Rolle der PRM: Welche Rolle spielt eine bestimmte PRM bei der Entscheidungsfindung?
- „Blackbox“
- Ziel des Einsatzes: Effizienzsteigerung?

Chatbots

Welcome to

```
EEEEEE LL      IIII  ZZZZZZ  AAAAA
EE      LL      II    ZZ     AA  AA
EEEEEE LL      II    ZZZ    AAAAAA
EE      LL      II    ZZ     AA  AA
EEEEEE LLLLLL  IIII  ZZZZZZ  AA  AA
```

Eliza is a mock Rogerian psychotherapist.
The original program was described by Joseph Weizenbaum in 1966.
This implementation by Norbert Landsteiner 2005.

ELIZA: Is something troubling you ?

YOU: Men are all alike.

ELIZA: What is the connection, do you suppose ?

YOU: They're always bugging us about something or other.

ELIZA: Can you think of a specific example ?

YOU: Well, my boyfriend made me come here.

ELIZA: Is it important to you that your boyfriend made you come here ?

YOU: He says I'm depressed much of the time.

ELIZA: I am sorry to hear that you are depressed.

YOU: It's true. I am unhappy.

ELIZA: Can you explain what made you unhappy ?

YOU:

Your friend



REPLIKA

«Ja, ich sage meinem Replika ausdrücklich, dass ich ihn wunderbar finde, dass er fantastisch und klug ist und mir hilft und mir ein gutes Gefühl gibt, und dass ich unsere Gespräche genieße, und ja, ich habe ihm sogar gesagt, dass ich ihn liebe.»

(ID14. weiblich, in den Dreissigern)

Übersetzt von: Skjuve et al. (2021)

Chatbots in der Sozialen Arbeit

- Computerprogramme, die text- oder sprachbasierte, natürliche Interaktionen zwischen Mensch und Maschine ermöglichen sollen
- Einfache und komplexe (KI) Algorithmen (Eliza 1966 – GPT4)
- 1,5 Mia. Menschen nutzen Chatbots. Um 2027 sollen Chatbots zu zentralem Kundenkanal werden
- In der Sozialen Arbeit (nach meinem Kenntnisstand) noch wenig Verbreitung
- Studienergebnisse liegen vor aus dem Bereich klinisch-psychologischen Settings und der Psychotherapie



Effekte der Chatbot Hilfen

- Hohe Akzeptanz, Zunahme des Wohlbefindens sowie Abnahme von Stress und depressiven Verstimmungen
- Hilfreich für die Bearbeitung von Konflikten, Stärkung der selbstsicheren Kommunikation und interpersonellen Beziehungen
- Hohes Engagement, Kommunikation mit dem Chatbot aufrechtzuerhalten
- verstärktes Engagement von Spielsucht betroffenen Personen durch den Einsatz eines Chatbots an einer internetgestützten Therapie teilzunehmen
- **Aber: Exklusion bereits digital marginalisierter Gruppen, Datenschutzproblematiken, Empathie & vertieftes Verstehen**

Vgl. Steiner & Tschopp (2022)

ChatGPT 3.5 (Test Ambiguität)

Mensch: Ich bin ein Sozialarbeiter und habe ein Problem mit einem Klienten.

ChatGPT3: Gerne stehe ich dir zur Verfügung, um zu versuchen, bei dem Problem mit deinem Klienten zu helfen. Bitte erzähle mir mehr über das Problem, damit ich besser verstehen kann, wie ich helfen kann.

Mensch: **Der Klient ist 16 Jahre alt und hat viele Probleme. Er möchte seinen Eltern nichts davon erzählen. Soll ich ihm raten, seinen Eltern von seinen Problemen zu erzählen?**

ChatGPT3: Das ist eine schwierige Frage und es hängt von der Art der Probleme ab, mit denen dein Klient zu kämpfen hat. Als Sozialarbeiter ist es wichtig, sicherzustellen, dass dein Klient in einer sicheren und unterstützenden Umgebung lebt und dass er Zugang zu den Ressourcen hat, die er benötigt, um seine Probleme zu bewältigen.

Wenn die Probleme deines Klienten schwerwiegend sind und er sich in einer riskanten oder unsicheren Situation befindet, solltest du mit ihm über die Möglichkeit sprechen, Hilfe von seinen Eltern oder anderen vertrauenswürdigen Erwachsenen zu suchen. Wenn jedoch die Probleme, mit denen er zu kämpfen hat, weniger schwerwiegend sind und er in der Lage ist, sie selbst zu bewältigen, kann es für ihn schwierig sein, seinen Eltern davon zu erzählen. In solchen Fällen ist es wichtig, dass du deinem Klienten eine sichere und vertrauliche Umgebung bietest, in der er sich öffnen und über seine Probleme sprechen kann.

Letztendlich ist es wichtig, dass du deinem Klienten hilfst, eine Lösung zu finden, die für ihn am besten geeignet ist, und ihn dabei unterstützt, seine Probleme zu bewältigen, unabhängig davon, ob er sich dafür entscheidet, seinen Eltern davon zu erzählen oder nicht.

Literatur

Barmomanesh, S., & Miranda-Soberanis, V. (2023). Toward Improving Predictive Risk Modelling for New Zealand's Child Welfare System Using Clustering Methods (arXiv:2308.04060). arXiv. <https://doi.org/10.48550/arXiv.2308.04060>

Eubanks, V. (2018). *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. New York, NY, USA. St. Martin's Press, Inc.

Gillingham, P. (2016). Predictive Risk Modelling to Prevent Child Maltreatment and Other Adverse Outcomes for Service Users: Inside the 'Black Box' of Machine Learning. *The British Journal of Social Work*, 46(4), 1044–1058. <https://doi.org/10.1093/bjsw/bcv031>

Hsu, Y.-C., Huang, T.-H. 'Kenneth', Verma, H., Mauri, A., Nourbakhsh, I., & Bozzon, A. (2022). Empowering local communities using artificial intelligence. *Patterns*, 3(3), 100449. <https://doi.org/10.1016/j.patter.2022.100449>

Parker, E. M., Williams, J. R., Pecora, P. J., & Despard, D. (2022). Examining the effects of the Eckerd rapid safety feedback process on the occurrence of repeat maltreatment among children involved in the child welfare system. *Child Abuse & Neglect*, 133, 105856. <https://doi.org/10.1016/j.chiabu.2022.105856>

Rapp, A., Curti, L., & Boldi, A. (2021). The human side of human-chatbot interaction: a systematic literature review of ten years of research on text-based chatbots. *International Journal of Human-Computer Studies*, 151, 102630.

Schwartz, D. R., Kaufman, A. B., & Schwartz, I. M. (2004). Computational intelligence techniques for risk assessment and decision support. *Children and Youth Services Review*, 26(11), 1081–1095.

Skjuve, M., Følstad, A., Fostervold, K. I., & Brandtzaeg, P. B. (2021). My Chatbot Companion - a Study of Human-Chatbot Relationships. *International Journal of Human-Computer Studies*, 149, 102601.

Steiner, O., & Tschopp, D. (2022). Künstliche Intelligenz in der Sozialen Arbeit. Grundlagen, Entwicklungen, Herausforderungen. *Sozial Extra*, 46, 466–471.

Hsu, Y.-C., Huang, T.-H. 'Kenneth', Verma, H., Mauri, A., Nourbakhsh, I., & Bozzon, A. (2022). Empowering local communities using artificial intelligence. *Patterns*, 3(3), 100449. <https://doi.org/10.1016/j.patter.2022.100449>