

Künstliche Intelligenz insbesondere Sprachmodelle, im sozialen Bereich

Kurzbeschreibung des Moduls

Das Modul „Künstliche Intelligenz, insbesondere Sprachmodelle, im sozialen Bereich“, vermittelt einen fundierten Einstieg in das Thema „Künstliche Intelligenz“ mit Schwerpunkt auf KI-basierte Sprachmodelle und deren Einsatz und Auswirkungen im sozialen Bereich. Im Fokus stehen sowohl das grundlegende technische Verständnis als auch die gesellschaftlichen, ethischen und nachhaltigkeitsbezogenen Implikationen dieser Technologie. Das Modul hat einen Umfang von ca. 120-180 Minuten und kann außerdem flexibel als Unterrichtsmodul für die Sekundarstufe II eingesetzt werden.

Zielgruppe: Studierende, SchülerInnen der Oberstufe

Voraussetzungen und Vorwissen

Für die Teilnahme an dem Modul sind keine Vorkenntnisse im Bereich Künstliche Intelligenz erforderlich. Vorausgesetzt werden lediglich grundlegende digitale Kompetenzen, wie sie im schulischen oder beruflichen Alltag üblich sind. Hilfreich ist ein allgemeines Interesse an sozialen, gesellschaftlichen oder pädagogischen Fragestellungen sowie die Bereitschaft, sich kritisch mit neuen Technologien auseinanderzusetzen. Vorwissen aus den Bereichen soziale Arbeit, Bildung, Psychologie, Verwaltung oder Nachhaltigkeit kann den Einstieg erleichtern, ist jedoch keine Voraussetzung. Das Modul ist bewusst für Einsteiger konzipiert und führt schrittweise in die technischen, sozialen und ethischen Aspekte von KI-Sprachmodellen ein.

Zielsetzung des Moduls

Ziel des Moduls ist es, den Teilnehmenden ein reflektiertes Grundverständnis von Funktionsweise, Einsatzmöglichkeiten und Grenzen von Sprachmodellen zu vermitteln und sie für einen verantwortungsvollen Umgang mit KI in sensiblen sozialen Kontexten zu sensibilisieren. Die Teilnehmenden sollen befähigt werden, Chancen des KI-Einsatzes zur Verbesserung von Zugänglichkeit, Teilhabe und Effizienz im sozialen Bereich zu erkennen, gleichzeitig jedoch potenzielle Risiken und ethische Herausforderungen kritisch zu bewerten. Darüber hinaus wird ein Bezug zur Bildung für nachhaltige Entwicklung (BNE) und zur Agenda 2030 hergestellt.

Nach Abschluss des Moduls können die Teilnehmenden

- grundlegende Funktionsprinzipien von Sprachmodellen knapp und verständlich erläutern
- Die Bedeutung der Trainingsdaten für den Erfolg eines Sprachmodells beschreiben,
- typische Anwendungsfelder von Sprachmodellen im sozialen Bereich benennen,
- Chancen und Risiken des KI-Einsatzes in sozialen Handlungsfeldern abwägen,
- exemplarische Bezüge zwischen KI-Anwendungen und ausgewählten Zielen der Agenda 2030 herstellen,

- eine eigene, reflektierte Haltung zum Einsatz von KI im sozialen Kontext formulieren.

Zielkonflikte:

- **Effizienzsteigerung und menschlicher Nähe:**

Während KI Prozesse beschleunigen, Fachkräfte entlasten und den Zugang zu Hilfe erleichtern kann, besteht gleichzeitig die Gefahr, dass persönliche Beziehung, Empathie und individuelle Wahrnehmung in der sozialen Arbeit an Bedeutung verlieren.

- **Datennutzung und Datenschutz**

Für leistungsfähige KI-Systeme werden große Datenmengen benötigt, gleichzeitig erfordert der soziale Bereich einen besonders sensiblen Umgang mit personenbezogenen und oft hochsensiblen Informationen.

- **Automatisierung und Verantwortung**

KI kann Entscheidungen vorbereiten oder unterstützen, die letztendliche Verantwortung für Beratung, Bewertung und Intervention muss jedoch bei Menschen verbleiben.

Inhalte und Ablauf

Zu Beginn des Moduls erfolgt eine kurze Einführung in das Thema Künstliche Intelligenz und Sprachmodelle. Dabei wird geklärt, was unter Sprachmodellen zu verstehen ist, wo sie im Alltag bereits eingesetzt werden und warum ihr Einsatz im sozialen Bereich besondere Anforderungen mit sich bringt. Dazu wird ein kurzes Übersichtsvideo zum Thema Sprachmodelle geschaut, in denen die Grundbegriffe sowie die allgemeine Funktionsweise behandelt wird. Dazu werden in einem Kreuzworträtsel die wichtigsten Begriffe und Anhaltspunkte des Videos aufgegriffen und abgefragt. Die Teilnehmenden erhalten ein vereinfachtes Verständnis darüber, wie Sprachmodelle mit großen Textmengen trainiert werden, wie sie auf statistischen Wahrscheinlichkeiten von Wortfolgen basieren und warum sie besonders gut Weltwissen aus Texten wiedergeben können, zugleich jedoch bei komplexen Bewertungen oder moralischen Fragen an Grenzen stoßen. Konzepte wie Prompt, Tokenisierung sowie die vektorielle Darstellung von Wörtern (Word2Vec) werden anschaulich und an Beispielen erklärt. Weiterhin sind tiefergehende Texte und Videos für die technische Funktionsweise von Sprachmodellen verlinkt, in denen sich die Teilnehmenden selber informieren können.

Darauf aufbauend werden ausgewählte Anwendungsfelder im sozialen Bereich vorgestellt. Dazu zählen insbesondere die psychosoziale Beratung, die öffentliche Verwaltung sowie Gewaltprävention und Kinderschutz. Dazu werden je 3 Texte bereitgestellt die die Themen sowie den Einsatz von KI beschreiben.

Psychosoziale Beratung:

Das Projekt KIA – KI-gestützte Assistenz für digitale psychosoziale Beratung ist ein vom

Bundesministerium für Familie, Senioren, Frauen und Jugend gefördertes Gemeinschaftsprojekt der Technischen Hochschule Nürnberg und der Bundeskonferenz für Erziehungsberatung (bke). Ziel des dreijährigen Projekts ist die Entwicklung eines KI-basierten Assistenzsystems, das Fachkräfte in der schriftbasierten psychosozialen Onlineberatung unterstützt, ohne die menschliche Beratung zu ersetzen. Im Mittelpunkt steht die Unterstützung von Beratenden bei der Reflexion und Strukturierung umfangreicher Beratungstexte. Die KI soll unter anderem Dialoge zusammenfassen, methodische und themenspezifische Hinweise geben, Fallverläufe übersichtlich aufbereiten und Beziehungen zwischen beteiligten Personen visualisieren. Die Entscheidungshoheit verbleibt dabei stets bei den menschlichen Fachkräften. Großer Wert wird auf Datenschutz, Anonymisierung und ethische Verantwortung gelegt. Authentische Beratungstexte werden nur mit freiwilliger, jederzeit widerrufbarer Einwilligung der Ratsuchenden und in vollständig anonymisierter Form genutzt. Fachkräfte der bke-Onlineberatung sind aktiv in Entwicklung, Testung und Reflexion des Systems eingebunden.

Verwaltung:

Der Blogbeitrag zum Thema „Verwaltung“ beschreibt, wie KI-basierte Sprachmodelle das Potenzial haben, die öffentliche Verwaltung grundlegend zu verändern. Durch ihre Fähigkeit, große Mengen an Textdaten zu verstehen, zu analysieren und zu erzeugen, können sie Verwaltungsprozesse effizienter gestalten, Mitarbeitende bei Routineaufgaben entlasten und die Kommunikation mit Bürgerinnen und Bürgern verbessern. Einsatzmöglichkeiten reichen von Chatbots für Bürgeranfragen über die automatische Klassifizierung und Weiterleitung von Anträgen bis hin zur Unterstützung bei der Erstellung und Analyse von Dokumenten. Gleichzeitig weist der Text auf zentrale Herausforderungen hin, etwa Datenschutz- und Datensicherheitsfragen, mögliche Verzerrungen durch Trainingsdaten (Bias), fehlende Transparenz sowie die Notwendigkeit menschlicher Kontrolle. Insgesamt wird betont, dass KI-Sprachmodelle große Chancen für eine bürgernähere und effizientere Verwaltung bieten, ihr Einsatz jedoch klaren rechtlichen, ethischen und organisatorischen Rahmenbedingungen folgen muss.

Gewaltprävention und Kinderschutz

Der Text der Bundeszentrale für Kinder- und Jugendmedienschutz beschreibt den Einsatz von Künstlicher Intelligenz als wichtiges Instrument zum Schutz von Kindern und Jugendlichen im digitalen Raum. KI kann große Mengen an Texten, Bildern und Videos effizient analysieren und dabei helfen, jugendgefährdende Inhalte wie Gewalt, Missbrauchsdarstellungen oder problematische Kommunikationsmuster frühzeitig zu erkennen. Besonders im Bereich sozialer Netzwerke und nutzergenerierter Inhalte bietet KI gegenüber klassischen, regelbasierten Filtersystemen deutliche Vorteile, da sie komplexe Muster und Zusammenhänge besser erfassen kann. Gleichzeitig macht der Text deutlich, dass der Einsatz von KI mit erheblichen Herausforderungen verbunden ist. Dazu zählen Fehlalarme, begrenztes Kontextverständnis, Risiken für Datenschutz und Privatsphäre sowie das Missbrauchspotenzial von KI durch Täterinnen und Täter, etwa bei der Erstellung manipulativer Inhalte oder beim Cybergrooming. Insgesamt wird KI als unterstützendes, aber keinesfalls fehlerfreies Werkzeug verstanden, dessen Einsatz stets von menschlicher Kontrolle, klaren ethischen Leitlinien und dem Schutz der Rechte von Kindern und Jugendlichen begleitet werden muss.

Anhand dieser konkreter Beispiele wird verdeutlicht, wie KI-Sprachmodelle Fachkräfte unterstützen können, etwa durch Textauswertung, Strukturierung von Informationen oder mehrsprachige Kommunikation, ohne dabei menschliche Entscheidungshoheit zu ersetzen. Die Texte können als Gruppenpuzzle ausgegeben werden, sodass verschiedene Teilnehmende sich nach dem Lesen über die Texte austauschen können und die verschiedenen Multiple Choice Fragen dazu lösen. So erhält jeder einen detaillierten Einblick in die Umsetzung von KI im sozialen Bereich, und kann die in den anderen Bereichen einen allgemeinen Einblick erhalten, ohne selber zu viel Lesen zu müssen.

Ein weiterer zentraler Bestandteil des Moduls ist die Gegenüberstellung von Chancen und Risiken des KI-Einsatzes. Chancen liegen insbesondere im niedrigschwelligen Zugang zu Hilfeleistungen, in der Effizienzsteigerung sozialer Arbeit, in der Förderung von Inklusion sowie in der Möglichkeit früher Krisenerkennung. Gleichzeitig werden Risiken wie der mögliche Verlust menschlicher Nähe, Datenschutzprobleme, Verzerrungen durch Trainingsdaten (Bias), digitale Ausgrenzung und ungelöste Verantwortungsfragen thematisiert. Dazu sollen die Teilnehmenden zuerst sich verschiedene Risiken und Chancen anschauen. In dem dazu gezeigten Bild haben sich ein paar Fehler eingeschlichen. Die Teilnehmenden sollen über die Chancen und Risiken diskutieren und dabei mögliche falsche Zuordnungen finden. Anschließend werden nochmals alle richtig zugeordnet gezeigt und beschrieben.

Abschließend wird der Einsatz von Sprachmodellen in den Kontext der Bildung für nachhaltige Entwicklung gestellt. Die Teilnehmenden reflektieren, inwiefern KI-gestützte Anwendungen zur Erreichung ausgewählter Ziele der Agenda 2030 beitragen können, beispielsweise in den Bereichen Bildung, Gesundheit, soziale Gerechtigkeit, Geschlechtergleichheit und Abbau von Ungleichheiten. Dabei wird betont, dass technologische Innovation allein nicht ausreicht, sondern stets von ethischen Leitlinien, Transparenz und menschlicher Kontrolle begleitet werden muss. Dazu sollen die Teilnehmenden zum Schluss ein Essay zu einer der folgenden Leitfragen formulieren:

- Wie können Sprachmodelle soziale Arbeit unterstützen und den Zugang zu Hilfeleistungen verbessern – und wo liegen dabei mögliche Risiken oder Grenzen?
- Inwiefern tragen Sprachmodelle zur Erreichung ausgewählter Ziele der Agenda 2030 im sozialen Bereich bei (z. B. Bildung, Gesundheit, Gleichheit)?
- Welche ethischen, rechtlichen und gesellschaftlichen Herausforderungen entstehen beim Einsatz von Sprachmodellen in sensiblen sozialen Kontexten?

Die Leitfragen beziehen sich auf die zuvor gelernten Inhalte des Moduls und gibt den Teilnehmenden die Möglichkeit selber Stellung zu beziehen, eigene Erfahrungen und Gedanken zu teilen und das Gelernte zusammenfassend widerzugeben.

Lehr- und Lernformen

Das Modul kombiniert kurze inhaltliche Inputs mit interaktiven Elementen wie Videoausschnitten, kurzen Quiz- oder Diskussionsformaten, Gruppen- oder Partnerarbeit sowie reflexiven Gesprächsphasen. Es bietet sich an, aufgrund des Gruppenpuzzels im 2 Arbeitsschritt, das Modul in Gruppen von drei Personen zu bearbeiten. Es kann jedoch auch

alleine oder zu zweit bearbeitet werden, dabei entsteht nur mehr Lesearbeit, da für die Multiple Choice Fragen alle Texte benötigt werden.

Zeitlicher Rahmen

- Einführung und Einstieg: ca. 5 min
- Grundlagen von Sprachmodellen (Video und Kreuzworträtsel): ca. 15-20 min
- Gruppenpuzzle und Multiple Choice: ca. 30-45 min
- Anwendungsfelder & Chancen/Risiken: ca. 15-20 Minuten
- BNE-Bezug und Reflexion: (je nach Umfang des Essays): ca. 30-60 Minuten

Quellen:

Erklärvideo: https://www.youtube.com/watch?v=8Rumn1O-RkY&ab_channel=KernAI

<https://katzberger.ai/2019/06/25/word2vec-verstaendlich-erklaert>

<https://datascientest.com/de/word2vec>

<https://www.bke-beratung.de/sites/default/files/dateien/bericht-der-bke-onlineberatung-2023.pdf> (S.32-35)

<https://www.cgi.com/de/de/blog/oeffentliche-verwaltung/wie-ki-sprachmodelle-revolutionieren>

<https://www.bzj.de/resource/blob/233866/eb855aab99192e7a8b58fe824b68e2db/20234-ki-als-instrument-des-kinder-und-jugendmedienschutzes-data.pdf>