



Der folgende Text wurde mittels eines sprachbasierten KI-Tools vom Englischen ins Deutsche übersetzt. Für die Demografieberatung Digi+ wurde er nur minimal angepasst und zeigt somit gleich beispielhaft die Stärken und Schwächen der Übersetzung mittels KI.

Bewerten Sie selbst Verständlichkeit, Flüssigkeit, Stil und Leserlichkeit des Textes.

Im Original stammt der Text von Sara Walker und wurde als Blogeintrag für das KI-Entwicklungsunternehmen „saifr“ verfasst.

(Siehe auch: [Garbage in, garbage out: Why data quality is critical to AI | Saifr](#), Stand: 5.1.2023)



Müll rein, Müll raus: Was bestimmt die Qualität von KI-Modellen?

Da KI-Modelle darauf ausgelegt sind, Daten zu verarbeiten und Entscheidungen auf der Grundlage dieser Daten zu treffen, ist die Datenqualität entscheidend für die Genauigkeit und Zuverlässigkeit des Modells.

Daten bilden das Fundament eines jeden KI-Systems, unabhängig von dessen Komplexität oder Leistungsfähigkeit. Da KI-Modelle darauf ausgelegt sind, Daten aufzunehmen, zu verarbeiten und dann Entscheidungen oder Vorhersagen auf der Grundlage dieser Daten zu treffen, ist die Qualität der Daten entscheidend für die Genauigkeit und Zuverlässigkeit des Modells. Schlechte Datenqualität kann zu falschen Ergebnissen oder negativen Ergebnissen führen, während hochwertige Daten handlungsorientierte Erkenntnisse und aussagekräftige Vorhersagen liefern können. Aber was bestimmt die Datenqualität? Es gibt mehrere Faktoren, die die Qualität von Daten beeinflussen können, von der Menge bis zur Sicherheit.

Menge

Wenn es um KI geht, ist die Menge der Daten von größter Bedeutung. Je mehr Daten ein Modell zugreifen kann, desto besser kann es lernen und arbeiten. Oft ist es ein reines Zahlenspiel, und größere Modelle übertreffen kleinere. Modernste KI-Modelle werden jetzt auf Hunderte von Milliarden Datensätzen trainiert.

Allerdings gibt es zusehends Techniken, die versuchen, die Anforderungen an die Datenmenge zu umgehen. Zum Beispiel können Zero-Shot- und Few-Shot-Learning ein Modell mit begrenzten Daten trainieren. Zusätzlich hat das Aufkommen von Grundlagenmodellen eine neue Perspektive auf die Rolle von Daten in KI-Systemen gebracht. Grundlagenmodelle wie GPT-3 erfordern eine enorme Datenmenge, da sie auf Billionen von Token vorab trainiert sind. Sobald diese Modelle jedoch erstellt sind, werden sie für nachgelagerte Aufgaben recht dateneffizient, da sie wesentlich weniger Daten benötigen, um sich an spezifische Anwendungen anzupassen. Dieser Ansatz ermöglicht es KI-Systemen, die



Leistung von Grundlagenmodellen zu nutzen, während der Bedarf an umfangreicher Datensammlung für nachgelagerte Aufgaben minimiert wird.

Genauigkeit

Obwohl im Allgemeinen die Menge über die Qualität für KI-Modelle steht, ist die Genauigkeit der Daten selbst immer noch entscheidend, da ungenaue Daten zu falschen Ergebnissen führen können. Viele KI-Modelle werden mit Internetdaten trainiert, und wie wir alle wissen, kann das Internet ein Ort für Fehlinformationen sein. Es ist entscheidend, dass Daten aus glaubwürdigen Quellen stammen.

Voreingenommenheit

Im Zusammenhang mit der Genauigkeit tritt Voreingenommenheit auf, wenn die Daten nicht repräsentativ für die Zielbevölkerung sind. Dies kann durch verschiedene Faktoren verursacht werden, einschließlich der Daten selbst, der Algorithmen, die zur Verarbeitung der Daten verwendet werden, und der Datenvorverarbeitungstechniken, die zur Vorbereitung der Daten für das Modell verwendet werden. Voreingenommenheit kann auch durch den Datensammelungsprozess verursacht werden, der möglicherweise auf bestimmte Arten von Daten oder bestimmte Arten von Benutzer*innen voreingenommen ist.

Beim Sammeln von Daten und der Entwicklung von KI-Modellen ist es wichtig, sich potenzieller Quellen von Voreingenommenheit bewusst zu sein und Maßnahmen zu ergreifen, um sie zu mindern. Voreingenommenheit kann zu unfairen Behandlung bestimmter Gruppen von Menschen führen, was ernsthafte ethische Implikationen und greifbare Ergebnisse haben kann. Zum Beispiel wird ein KI-System, das zur Diagnose von Krankheiten verwendet wird und nicht auf pädiatrischen Daten trainiert wurde, voreingenommen gegenüber Erwachsenen sein und wahrscheinlich Kinder nicht genau diagnostizieren.

Dazu zwei reale Beispiele:

Die Personalverantwortlichen einer großen Firma sortierten mithilfe eines Algorithmus die eingegangenen Bewerbungen mit Blick auf die Passfähigkeit auf die ausgeschriebene Stelle vor. Was sie nicht wussten: Bewerbungen von Frauen bewertete der Algorithmus systematisch schlechter als jene von Männern. Denn die Datenbasis, mit der der Algorithmus trainiert wurde, beinhaltete die Daten der in den letzten zehn Jahren erfolgreich eingestellten Beschäftigten. Da diese überwiegend männlich waren, hatte der Algorithmus gelernt, die Eigenschaft „männlich“ als positiv zu bewerten.

Bekannt sind etwa auch Beispiele bei denen Bildgenerierende KI-Modelle Stereotype reproduzieren. Im folgenden Prompt lautete die Aufgabe an Dall.e 2: „friendly doctor explaining illness to person, looking concerned“. Trotz der genderneutralen Formulierung von *Doctor* im Englischen, wurde stets ein Mann im weißen Kittel dargestellt und keine Frau.

Vielfalt

Es ist auch wichtig zu berücksichtigen, wie effektiv die Daten ein Problem oder Anwendungsszenario lösen können. Ein vielfältiger Datensatz, der verschiedene Aspekte des Problemfelds repräsentiert,



kann KI-Modelle besser generalisieren lassen und robustere Vorhersagen ermöglichen. Die Gewährleistung, dass die Daten eine Vielzahl von Szenarien, Randfällen und Variationen abdecken, kann die Fähigkeit des KI-Systems verbessern, sich an neue oder nicht gesehene Daten anzupassen.

Kuration und Vorverarbeitung

Der Prozess des Reinigens, Organisierens und Transformierens von Daten vor der Verwendung für das Training von KI-Modellen ist entscheidend für die Verbesserung der Datenqualität. Eine ordnungsgemäße „Daten-Kuration“ und Vorverarbeitung können Probleme wie Inkonsistenzen, fehlende Werte und Rauschen in den Daten angehen und letztendlich zu einer besseren Leistung der KI führen.

Aktualität

Es ist wichtig zu verstehen, wie aktuell die Daten sind, da dies sich auf die Ausgabe eines Modells auswirken kann. Ältere Daten sind möglicherweise nicht relevant für die aktuelle Situation oder wurden durch neuere Daten ersetzt. Das bedeutet nicht, dass ein älteres Modell nicht genau sein würde, aber ein aktuelleres Modell könnte genauer sein.

Privatsphäre und Sicherheit

Da KI-Systeme stark auf Daten angewiesen sind, ist es entscheidend, Datenschutz- und Sicherheitsbedenken im Zusammenhang mit der Datensammlung, -speicherung und -verarbeitung zu berücksichtigen. Sensible Informationen sollten ordnungsgemäß geschützt und anonymisiert werden, um das Vertrauen der Benutzer*innen zu wahren und die Datenschutzbestimmungen einzuhalten.

KI ist nur so gut wie ihre Daten

All diese Faktoren können die Qualität der Daten beeinträchtigen und somit die Faktizität und Zuverlässigkeit der Ausgabe des KI-Systems beeinflussen. Denken Sie daran, dass KI tatsächlich nichts weiß - sie verwendet einfach Mathematik, um Ergebnisse vorherzusagen. Es kann nicht beurteilen, ob die Daten, die seine Algorithmen speisen, wahr oder falsch sind, daher wird es weiterhin falsche Aussagen produzieren, wenn seine Daten ungenau sind.

Schlechte Datenqualität kann ernste Konsequenzen aufgrund ungenauer Vorhersagen oder Entscheidungen haben. Damit ein KI-System genauere Ergebnisse produziert, ist es unerlässlich, dass es auf qualitativ hochwertigen Daten trainiert wird.

Es gibt einen Spruch in der KI: Müll rein, Müll raus. Wenn Sie Ihr Modell mit schlechten Daten trainieren, wird es schlechte Ergebnisse produzieren. Wenn Sie es mit guten Daten trainieren, besteht eine höhere Wahrscheinlichkeit, dass es gute Ergebnisse produziert. Es ist so einfach: Ein Modell wird immer nur so gut sein wie seine Daten.

Die hier bereitgestellten Informationen zu KI-Tools dienen ausschließlich zu Informationszwecken und sollen keine Empfehlung, Entwicklung oder Sicherheitsbewertung jeglicher Art darstellen.



Faire Nutzung von LLMs

Ökologische Auswirkungen

Die Nutzung von Large Language Models (LLMs) bringt erhebliche ökologische Herausforderungen mit sich. Insbesondere der Energieverbrauch für das Training und die Ausführung solcher Modelle ist enorm. Eine einzelne Anfrage an ein großes Sprachmodell kann mehrere Wattstunden Strom verbrauchen, abhängig von der Komplexität und dem Modell. Laut einer Studie von Strubell et al.¹ kann das Training eines großen NLP-Modells mehrere hunderttausend Kilogramm CO₂-Emissionen verursachen. Dies verdeutlicht die Notwendigkeit, KI-Systeme energieeffizienter zu gestalten, nachhaltige Rechenzentren zu nutzen und die Technologie verantwortungsbewusst einzusetzen.

KI zur Förderung von Fairness

KI kann aktiv zur Förderung der Gleichstellung der Geschlechter und zur Bekämpfung von Diskriminierung eingesetzt werden. Durch gezielte Datenanalyse lassen sich diskriminierende Muster erkennen und beheben. Beispielsweise können KI-Systeme Bewerbungsprozesse auf geschlechtsneutrale Sprache prüfen oder sicherstellen, dass Bildgenerierungsmodelle diverse Personengruppen repräsentieren. Zudem können LLMs zur Sensibilisierung und Schulung in Unternehmen beitragen, indem sie Inhalte zu Diversity und Inklusion bereitstellen.

Unternehmen können KI aber auch gezielt einsetzen, um ökologische Nachhaltigkeit zu fördern. Beispielsweise lassen sich durch KI-basierte Analysen Energieverbrauch und Emissionen optimieren. LLMs können bei der Erstellung von Nachhaltigkeitsberichten unterstützen oder Vorschläge für umweltfreundliche Maßnahmen generieren. Auch die Optimierung von Lieferketten im Hinblick auf CO₂-Emissionen ist ein Anwendungsfeld für KI.

Beispielprompts zur Förderung von Nachhaltigkeit und Gleichstellung

Hier sind einige Beispielprompts, die gezielt für nachhaltige und diskriminierungsfreie Anwendungen von LLMs genutzt werden können:

- Erstelle einen genderneutralen Bewerbungstext für eine Führungsposition.
- Analysiere den Energieverbrauch von XY und schlage Optimierungen vor.
- Generiere eine Liste von Maßnahmen zur Förderung von Diversität am Arbeitsplatz.
- Erstelle einen Nachhaltigkeitsbericht für ein mittelständisches Unternehmen.
- Identifiziere diskriminierende Formulierungen in einem Stellenangebot.
- Schlage umweltfreundliche Alternativen für Verpackungsmaterialien vor.
- Erkläre die Vorteile von Diversity-Trainings für Unternehmen.
- Simuliere die Auswirkungen einer CO₂-Reduktionsstrategie auf die Lieferkette.
- Erstelle ein Schulungskonzept zur Sensibilisierung für Gender Bias in KI-Systemen.
- Generiere Ideen für nachhaltige Produktentwicklung in der Konsumgüterindustrie.

¹ Strubell, E., Ganesh, A., & McCallum, A. (2019). Energy and Policy Considerations for Deep Learning in NLP. <https://arxiv.org/abs/1906.02243>