

SoekiaGPT - ein didaktisches Sprachmodell

Hielscher, M.
Pädagogische Hochschule Schwyz

DOI: 10.18420/ibis-01-01-04

Zusammenfassung

Statistische Sprachmodelle, insbesondere Textgeneratoren wie ChatGPT, haben in letzter Zeit große Aufmerksamkeit erregt. Künstliche Intelligenz ist auch ein Thema im Unterricht und die kritische Auseinandersetzung sollte ein Verständnis der grundlegenden Funktionsweisen umfassen. Sprachmodelle wie ChatGPT oder Bard sind jedoch komplexe Informatiksysteme und basieren auf riesigen Dokumentensammlungen und massiver Rechenleistung. Die Lernumgebung SoekiaGPT nimmt eine Reihe didaktischer Vereinfachungen vor und ermöglicht Schülerinnen und Schülern so einen Blick unter die Motorhaube von statistischen Sprachmodellen zu werfen und mit einem selbst gebauten Modell zu experimentieren.

Einleitung

Seit vielen Jahren werden Verfahren der künstlichen Intelligenz zum Beispiel bei Übersetzungen, Bilderkennung, Empfehlungssystemen, bei Social Media oder Predictive Maintenance eingesetzt. Im Herbst 2022 hat das Sprachmodell ChatGPT einen regelrechten Hype auch in der breiten Bevölkerung ausgelöst. Generative KI-Systeme, die aus einer Eingabe Text, Bild, Ton, Video, 3D-Modell usw. generieren, werden in immer mehr Bereichen eingesetzt. Große Sprachmodelle sind auch zu einem aktuellen Thema für die Schule geworden.

Nach der Empfehlung des Dagstuhl-Dreiecks zur Bildung in einer digital geprägten Welt (vgl. Gesellschaft für Informatik 2016) lassen sich digitale Alltagsphänomene im Informatikunterricht aus drei Perspektiven betrachtet und erklären. Die Anwendungsperspektive (Wie nutze ich das?), die gesellschaftlich-kulturelle Perspektive (Wie wirkt das?) und die technologische Perspektive (Wie funktioniert das?). Inzwischen gibt es eine schon fast unübersehbare Anzahl von Handreichungen für Lehrkräfte und Unterrichtsmaterialien für den Einsatz auf verschiedenen Schulstufen. Diese Materialien beziehen sich meist auf die konkrete Anwendung von Sprachmodellen oder auf die Auseinandersetzung mit Fragen rund um die Auswirkungen auf Schule und Gesellschaft. Zur grundlegenden Funktionsweise von großen Sprachmodellen (engl. LLM, Large Language Model) gibt es hingegen kaum Materialien. Hier setzt SoekiaGPT an.

SoekiaGPT ist eine kostenlos verfügbare, webbasierte Lernumgebung ab der Sekundarstufe I, die vollständig im eigenen Browser arbeitet und ohne persönliches Konto genutzt werden kann. Mit Hilfe von vorgegebenen oder eigenen, sehr kleinen Dokumentensammlungen (von wenigen Sätzen bis einigen hundert Seiten) kann unter die Motorhaube von Textgeneratoren

und statistischen Sprachmodellen geschaut werden. Im Gegensatz zu reinen Visualisierungen können die Schülerinnen und Schüler mit dem Werkzeug vielfältig interagieren und experimentieren.

SoekiaGPT fordert mit einem Eingabefeld zu einer ersten Eingabe (Prompt) auf. Voreingestellt ist eine Kollektion mit verschiedenen Märchen. Der Textgenerator übernimmt die Aufforderung, ein neues Märchen mit "Es war einmal" zu schreiben, und beginnt einen Märchentext. Der Aufbau ähnelt damit den bekannten Chatbots (siehe Abb. 1). Die "Schau hinein" Funktion legt die dahinterliegenden Verarbeitungsschritte und Trainingsdaten offen und ist der zentrale Teil der Lernumgebung.



Abbildung 1: SoekiaGPT generiert ein "Märchen"

SoekiaGPT verzichtet bewusst auf die Berücksichtigung neuronaler Netzwerke und maschinellen Lernverfahren und nutzt ein klassisches statistisches Sprachmodell, welches sich einfach erklären und nachvollziehen lässt. Große Sprachmodelle wie etwa ChatGPT oder Bard nutzen hingegen neuronale Netze mit Milliarden von Parametern und basieren auf gigantischen Datenmengen und Rechenleistung. Im Unterricht ist es daher weder ressourcentechnisch möglich noch aufgrund der Komplexität didaktisch sinnvoll, so große Sprachmodellen zu verwenden, um daran die grundlegende Funktionsweise von statistischen Sprachmodellen zu vermitteln. Neuronale Netzwerke lassen sich zwar ebenfalls im Unterricht erklären, sind aber ein eigenes komplexes Thema, welches für das Verständnis von statistischen Sprachmodellen nicht zwingend notwendig ist. Unabhängig vom eingesetzten Verfahren geht es bei statistischen Sprachmodellen immer um die Aufbereitung und Analyse einer Dokumentenkollektion und anschließend um eine Sprachsynthese basierend auf den zuvor gewonnenen statistischen Daten.

Was sind statistische Sprachmodelle?

Im Unterschied zu wissensbasierten Sprachmodellen, bei denen Expertinnen und Experten den Aufbau von Sprache in präzisen Regeln beschreiben, verwenden statistische Sprachmodelle große Datenmengen zum Extrahieren von Zusammenhängen. Typischerweise werden riesige Dokumentensammlungen dafür verwendet. Sie enthalten Texte aus einer Vielzahl von Quellen, z.B. aus der Wikipedia, aus Büchern und Zeitschriften und aus News-Artikeln und Social Media Plattformen. Diese Texte dienen als Trainingsdaten für das Sprachmodell. Aufgrund einer mathematisch-statistischen Analyse dieser Texte lernt das Sprachmodell zum Beispiel, welches Zeichen oder welches Wort in einem Text mit grosser Wahrscheinlichkeit als nächstes folgt. Das Sprachmodell kann anschließend für einen beliebigen Textanfang für jedes mögliche Folgewort eine Wahrscheinlichkeit berechnen und durch eine Auswahl den Text fortlaufend ergänzen. Sind mehrere Wörter ähnlich wahrscheinlich, wird in der Regel mit einer gewissen Zufälligkeit ein Wort ausgewählt (siehe Abb. 2), weshalb bei wiederholter Durchführung unterschiedliche Texte entstehen können.

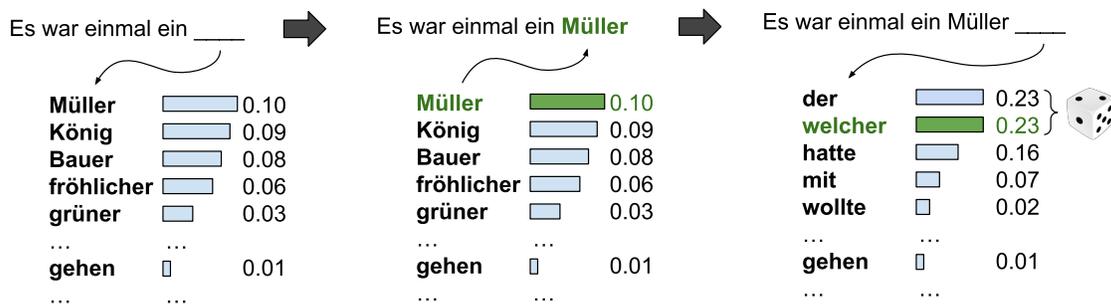


Abbildung 2: Vereinfachtes Funktionsprinzip statistischer Sprachmodelle

Bei der Analyse werden die Trainingsdaten zunächst in sogenannte Tokens zerlegt. Token sind in SoekiaGPT ganze Wörter und Satzzeichen. In anderen Sprachmodellen können das aber auch einzelnen Buchstaben oder Silben sein. Der Analyseprozess führt bei SoekiaGPT zu einer Liste von sogenannten N-Grammen. N-Gramme sind Wortfolgen (oder Tokenfolgen) der Länge N, welche in den Dokumenten vorkommen (siehe Beispiel in Tab. 1). Für jedes N-Gramm wird die Häufigkeit notiert, wie oft es über alle Dokumente hinweg angetroffen wurde. In Tabelle 1 würde damit in der ersten Spalte der Punkt nur einmal notiert und mit einem Zähler 2 versehen. Der Aufwand für den Aufbau eines Sprachmodells steigt mit der Anzahl von Trainingsdokumenten.

Bei Sprachmodellen mit neuronalen Netzwerken werden ebenfalls Textfragmente verschiedener Länge gebildet und als Eingabe für das Netzwerk genutzt, um es zu trainieren. Das Training wird wiederholt, bis das Netzwerk für möglichst alle Textfragmente das aus der Dokumentensammlung bereits bekannte nächste Wort hoch bewertet und damit vorhersagen kann (überwachtes maschinelles Lernen). Die Abkürzung GPT für Generative Pre-Trained Transformer benennt diesen vorbereitende Trainingsphase.

1-Gramme	2-Gramme	3-Gramme	4-Gramme
Der	Der Müller	Der Müller hatte	Der Müller hatte eine
Müller	Müller hatte	Müller hatte eine	Müller hatte eine Tochter
hatte	hatte eine	hatte eine Tochter	hatte eine Tochter .
eine	eine Tochter	eine Tochter .	eine Tochter . Sie
Tochter	Tochter .	Tochter . Sie	Tochter . Sie war
.	. Sie	. Sie war	. Sie war sehr
Sie	Sie war	Sie war sehr	Sie war sehr schön
war	war sehr	war sehr schön	war sehr schön .
sehr	sehr schön	sehr schön .	
schön	schön .		
.			

Tabelle 1: N-Gramme mit $N \leq 4$ für "Der Müller hatte eine Tochter. Sie war sehr schön."

Bei der Texterzeugung werden wiederholt die zuvor gesammelten statistischen Daten genutzt, um einen fortlaufenden Text zu generieren. In SoekiaGPT werden dazu die letzten Wörter des bisherigen Texts mit den Anfängen aller N-Gramme verglichen. Passt ein N-Gramm zur Eingabe, wird es in die Liste aller Kandidaten für das nächste Wort aufgenommen. Für die Eingabe "Ein Bauer hatte eine" würde gemäß Tabelle das 3-Gramm "hatte eine Tochter" als Kandidat mit der längsten Übereinstimmung gewählt und die Ausgabe mit "Tochter" fortgesetzt (siehe Tab. 1). Je häufiger ein N-Gramm in den Dokumenten vorkommt, desto eher wird es für die Fortsetzung ausgewählt. Passt ein N-Gramm mit einem höheren N, wird dieses bevorzugt, da mehr "Kontext" vorhanden ist. Notfalls wird ein 1-Gramm ausgewählt (Backing-off). Abbildung 3 fasst die Arbeitsweise von SoekiaGPT zusammen.

Bereits mit wenigen Dokumenten als Datenbasis lassen sich mit einem N-Gramm-Modell wie in SoekiaGPT kleine Texte generieren. Die Qualität großer neuronaler Sprachmodelle lässt sich damit aber nicht erreichen. Für detaillierte Ausführungen zu neuronalen und statistischen Sprachmodellen sei auf die Fachliteratur verwiesen (z.B. NLP-Kurs von Elena Voita 2023).

Ansätze zur Nutzung von SoekiaGPT im Unterricht

SoekiaGPT kann auf verschiedenen Schulstufen und mit unterschiedlicher Tiefe eingesetzt werden. Die nachfolgenden Ausführungen zeigen einige Aspekte von statistischen Sprachmodellen, die mit Hilfe der Lernumgebung bearbeitet oder erforscht werden können. In einem ersten Schritt können die Daten und die Datenanalyse als Ausgangsbasis von statistischen

Sprachmodellen genauer betrachtet werden. Im zweiten Schritt wird die Texterzeugung mit Parametern und deren Grenzen thematisiert. Ausführliche Hinweise zu möglichen Aufgabenstellungen im Unterricht finden sich in der Online-Handreichung.

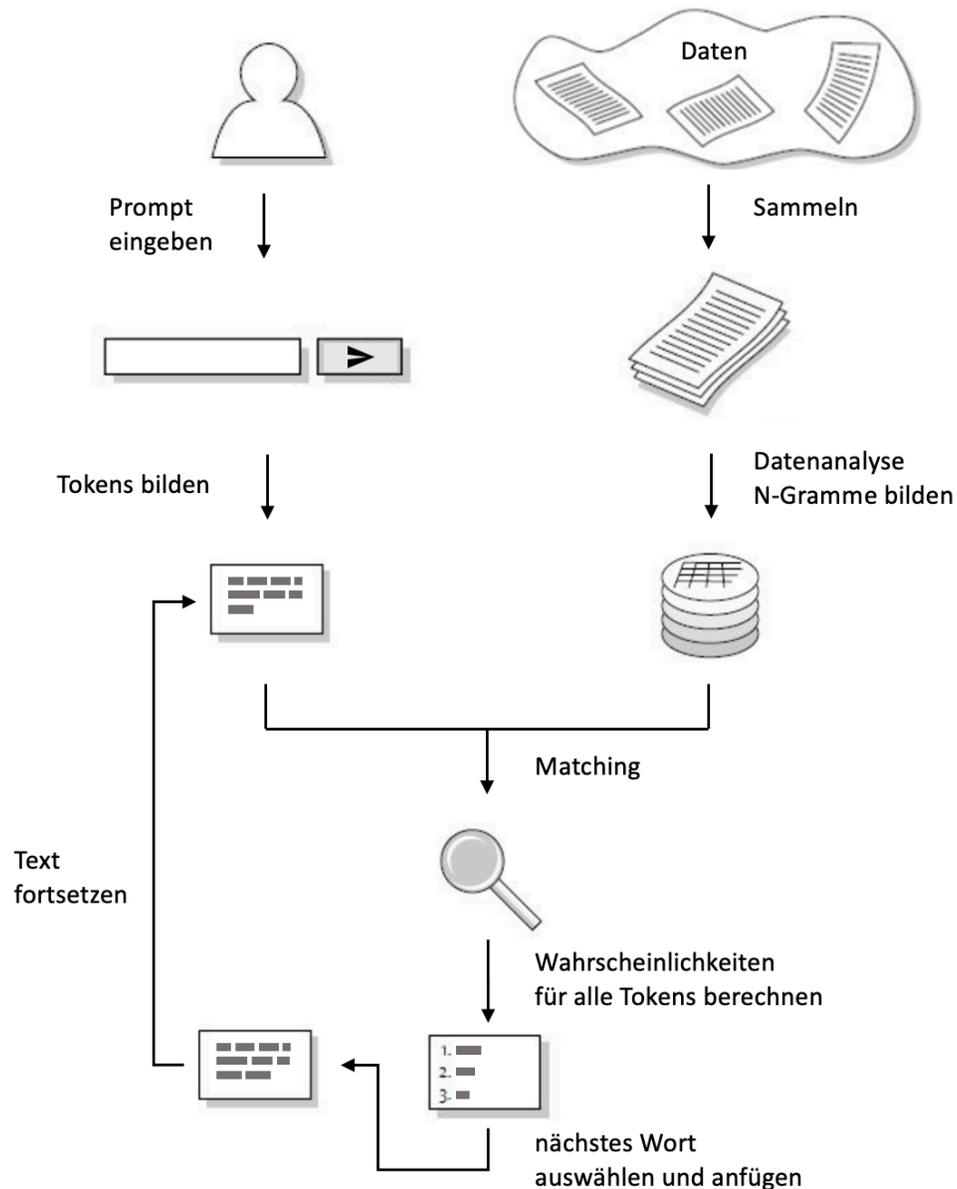


Abbildung 3: Funktionsweise von SoekiaGPT

Daten und Datenanalyse

Statistische Sprachmodelle benötigen Daten. Modelle wie GPT4 werden mit Millionen von Dokumenten, Büchern und Internetseiten gefüttert. SoekiaGPT begrenzt sich auf 25 Dokumente mit je maximal 20'000 Zeichen. Zum Experimentieren werden bereits einige Dokumentensammlungen (z.B. Märchen oder Wettervorhersagen) im Werkzeug zur Auswahl angeboten. Die

Schülerinnen und Schüler können aber auch selbst eine Kollektion anlegen, Dokumente hinzufügen oder verändern. Ebenso kann die Lehrperson eine eigene Kollektion vorbereiten und als Datei den Lernenden zur Verfügung stellen.

In der voreingestellten Kollektion "Märchen" kommen rund 3000 verschiedene Wörter vor. Es werden entsprechend viele 1-Gramme gebildet und ihre Häufigkeit gezählt. Zum Beispiel kommt das Wort "und" über 900-mal vor und ist damit das am häufigsten verwendete 1-Gramm in der Kollektion. 3-Gramme gibt es bereits rund 18'000 verschiedene, wobei das häufigste "in den Wald" noch 28-mal gezählt wurde. Solche statistischen Beobachtungen können mit SoekiaGPT sehr einfach für verschiedene Textsorten durchgeführt werden. Die Schülerinnen und Schüler können in den N-Gramm-Listen die relative und absolute Häufigkeit von Wörtern und Wortfolgen ablesen und sich so einen ersten statistischen Überblick über die Dokumentenkollektion verschaffen (siehe Abb. 4).

The screenshot displays the SoekiaGPT interface for N-gram analysis. The top bar shows 'N-Gramme' with navigation buttons (-1, +1, ≡) and 'Dokumente: 12' with document management icons (+, 📄, 🗑️, 🗑️). Below the top bar, there are tabs for 1er, 2er, 3er, 4er, and 5er N-grams, with '3er' selected. A list of N-grams is shown on the left, including 'Es war einmal', 'Hänsel und Gretel', 'auf einen Streich', 'siebene auf einer', 'und als sie', '. Der König', 'auf den Weg', 'hinaus in den', 'rucke di guck', 'um den Hals', 'und als der', '. Als der', '. Rotkäppchen aber', and 'Stroh zu Gold'. A tooltip for 'Es war einmal' indicates it was found 7 times across documents A, B, D, G, I, J, and L, with an option to 'in Eingabe einfügen'. A red vertical button labeled 'N-Gramme erstellen' is positioned between the N-gram list and the document collection. The document collection on the right consists of four cards: 'Dokument E' (Die goldene Gans), 'Dokument F' (Hänsel und Gretel), 'Dokument G' (Die sieben Raben), and 'Dokument H' (Aschenbuttel). Each card contains a snippet of text from the respective document.

Abbildung 4: Dokumentenkollektion zu "Märchen" und generierte N-Gramm-Listen

Eine Diskussion über Sprache und die allgemeine Häufigkeit von Wörtern bietet sich an und kann mit Experimenten zur Lesbarkeit von “beschädigten” Texten und dem Informationsgehalt (Entropie) ergänzt werden.

Im Unterschied zu grossen Sprachmodellen kann SoekiaGPT bei seiner Ausgabe die Herkunft von Textfragmenten aus der Dokumentensammlung visualisieren (siehe Abb. 5). Was in einem kleinen Sprachmodell mit wenigen Quellen noch funktioniert, ist in grossen Sprachmodellen nicht mehr praktikabel. Jedes Textfragment hätte dort viele tausende oder gar Millionen möglicher Quellen, weshalb man etwa bei ChatGPT nicht alle Quellen auflisten kann, die zur Antwort beigetragen haben. Mit der Frage, wie sich die Genese von Antworten solcher Systeme trotzdem erklären oder aufzeigen lässt, beschäftigt sich das Forschungsfeld Explainable AI, kurz XAI.

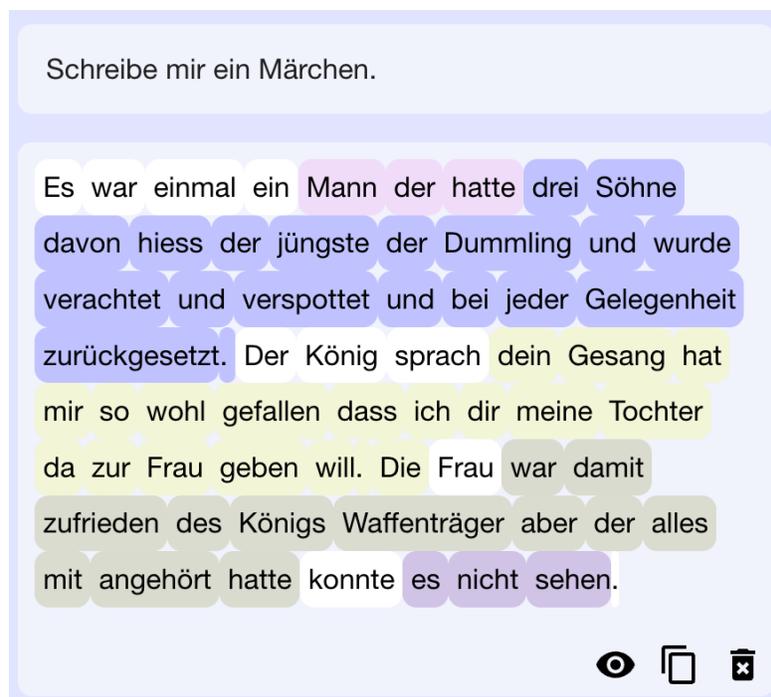


Abbildung 5: Farbliche Hervorhebung von Textfragmenten, wenn eine eindeutige Quelle benannt werden kann.

Mit SoekiaGPT sehen die Schülerinnen und Schüler an konkreten, überschaubaren Beispielen, wie Textgeneratoren neue Texte durch Kombination aus Fragmenten verschiedener Dokumente erstellen. Sprachmodelle nutzen auch den Zufall, um immer eine leicht andere Ausgabe zu erzeugen. So können im Unterricht erstellte Texte untereinander verglichen werden. Welches generierte Märchen klingt am besten? Welche typischen syntaktischen und semantischen Fehler generiert das Sprachmodell?

Durch Manipulationen der Dokumente können die Schülerinnen und Schüler die Ausgabe des Sprachmodells für einen gewählten Prompt gezielt verändern. Es wird schnell klar, dass die Ausgaben stark von den Trainingsdaten abhängen. Wurde etwa immer nur ein “König” in den

Dokumenten verwendet, wird das Sprachmodell in SoekiaGPT auch keinen Text mit “Königin” erzeugen können. Dies eröffnet zum Beispiel den Einstieg in eine Diskussion über “Verzerrte Wahrnehmungen und Vorurteile in den Daten” (Bias) und die Auswirkungen auf KI-Systeme.

Texterzeugung

Sprachmodelle nutzen die Häufigkeit von Wortfolgen, um das nächste Wort in einem bestehenden Text zu bestimmen. Dazu stehen in der Regel in jedem Schritt mehrere Kandidaten zur Auswahl.

In SoekiaGPT sucht ein Auswahlalgorithmus passende N-Gramme heraus. Dieser kann in den Einstellungen in seinem Verhalten gesteuert werden (siehe Abb.7).

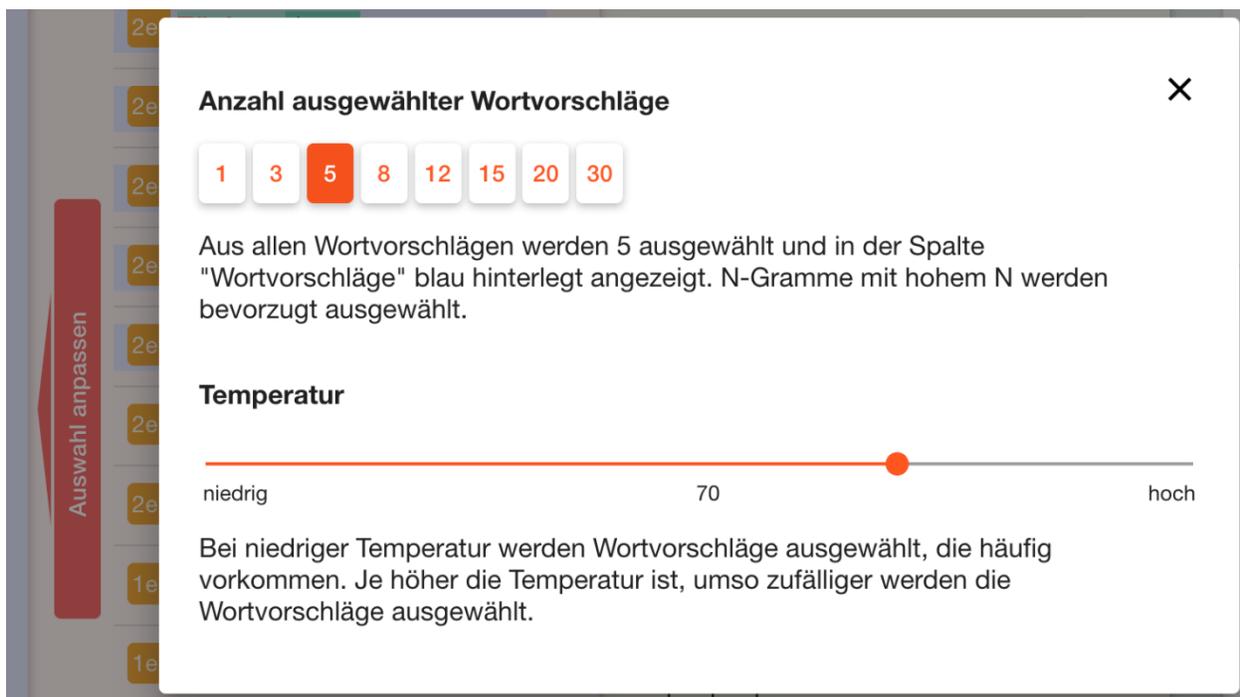


Abbildung 6: Einstellungen zur Auswahl der Wortvorschläge

Über die sogenannte “Temperatur” kann eingestellt werden, wie zufällig die Auswahl aus den statistisch wahrscheinlichen N-Grammen erfolgen soll. In der Welt der Sprachmodelle wird Temperatur häufig auch damit verbunden, wie “kreativ” das Modell sein soll. Die Schülerinnen und Schüler können mit der Temperatur experimentieren und sofort die Auswirkungen in der generierten Ausgabe beobachten. Experimentiert man mit verschiedenen Prompts und Dokumentensammlungen stellt man schnell fest, dass es keinen optimalen Wert für die Temperatur gibt, sondern dieser vielmehr situativ angepasst werden sollte. Eine Erkenntnis, die zum Beispiel auch bei der Nutzung von Bildgeneratoren und anderen KI-Werkzeugen nützlich sein kann.

Über Schalter kann das Sprachmodell angewiesen werden, auch syntaktisch ähnliche Wörter (geringe Editierdistanz) und Synonyme einzubeziehen, um die Auswahl an N-Grammen zu

vergrößern und so bessere Ergebnisse zu erzielen. SoekiaGPT nutzt dafür eine Liste gebräuchlicher deutscher Synonyme als zusätzliche Wissensbasis. Auch Systeme zur maschinellen Sprachverarbeitung nutzen in der Regel eine Vielzahl von Vor- und Nachbearbeitungsschritten, um die Qualität der Trainingsdaten und die Ausgabe zu verbessern oder unerwünschte Inhalte herauszufiltern.

Aufgrund der kleinen Dokumentenkollektion und dem eingesetzten Verfahren wird SoekiaGPT nie die Qualität und das Potenzial von ChatGPT oder Bard erreichen können. Um diese Defizite etwas auszugleichen, können wir dem Sprachgenerator etwas "helfen". Statt mit dem Algorithmus das wahrscheinlichste nächste Wort aus der Kandidatenliste wählen zu lassen (greedy), kann auf einen manuellen Modus gewechselt werden (siehe Abb. 7). Die Lernenden können so ihr eigenes Sprachwissen anwenden, um syntaktisch und semantisch sinnvolle Sätze zu generieren. Die didaktische Möglichkeit, Schritt für Schritt die Ausgabe zu generieren, unterstützt nicht nur die Vorstellung der Arbeitsweise des Sprachgenerators. Es entsteht schnell der Wunsch nach einer grösseren Auswahl an ganz bestimmten Wörtern zur Fortsetzung. Die Trainingsdaten können gezielt überarbeitet und Einstellungen so vorgenommen werden, dass die Wunschkombinationen auch zur Auswahl angeboten werden (z.B. den eigenen Wohnort in die Kollektion zu Wettervorhersagen aufnehmen).

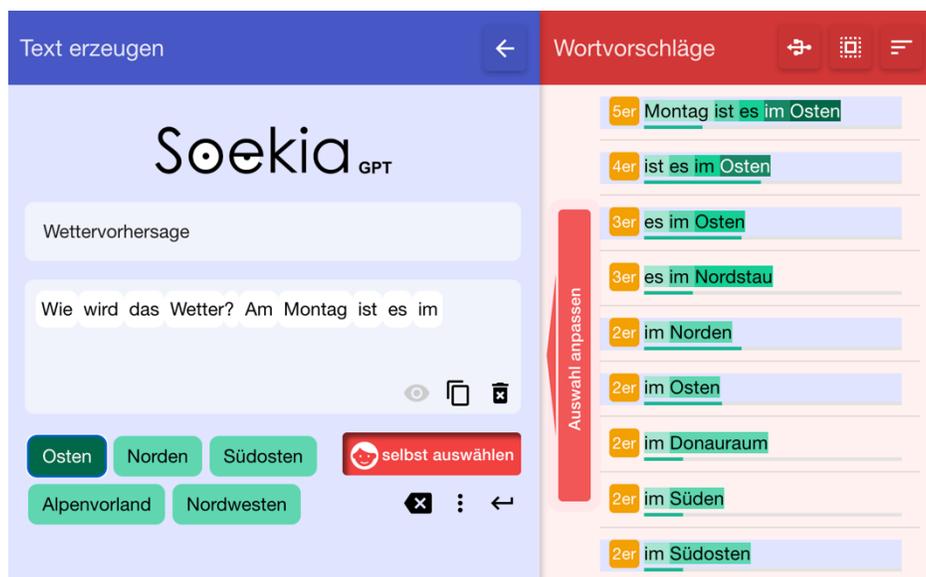


Abbildung 7: Manuelle Auswahl in SoekiaGPT

Grenzen von statistischen Sprachmodellen

Neben der Funktionsweise von Sprachmodellen ist es auch wichtig, deren Grenzen zu verstehen. Statistische Sprachmodelle erzeugen Ausgaben auf Basis von Statistik, nicht auf Basis von klar definierten Fakten oder Wissen. Dies lässt sich mit SoekiaGPT am Beispiel "Wettervorhersage" gut aufzeigen. Die Kollektion beinhaltet Formulierungen aus Wetterberichten. Generiert man zu "Wie wird heute das Wetter?" eine Ausgabe, klingt das Ergebnis durchaus plausibel, aber ist frei erfunden. Es wäre reiner Zufall, wenn sich das Wetter tatsächlich so verhalten

würde. Sprachmodelle sind gut darin, sprachlich wohlgeformte Texte zu generieren, deren Inhalt aber nicht der Wahrheit entsprechen muss. Mit den Schülerinnen und Schülern sollte über die Schwächen statistischer Sprachmodelle wie erfundene Fakten und Quellen diskutiert werden.

Von zentraler Bedeutung für die Ausgabequalität eines Sprachmodelles ist neben den Trainingsdaten der "Kontext". Bei Systemen wie ChatGPT wird dieser durch den eingegebenen Prompt und frühere Dialoge bestimmt. Gern wird im Hintergrund auch eine Verhaltensanweisung (z.B. „Converse as a WordPress expert. Be helpful, friendly, concise, avoid external URLs and commercial solutions.“) der Eingabe vorangestellt oder die Eingabe in eine solche Anweisung eingebettet. Das sogenannte "Prompt Engineering" ist inzwischen ein eigenständiges Arbeits- und Forschungsfeld.

Während ChatGPT mehrere tausend Tokens als Eingabe für die nächste Ausgabe als Kontext einbezieht, ist SoekiaGPT auf nur die letzten fünf Wörter limitiert. Die Schülerinnen und Schüler können in SoekiaGPT aber das maximale N zurückgeschauter Wörter selbst festlegen und experimentieren, welchen Einfluss dies auf die Textqualität hat.

Fazit und Ausblick

SoekiaGPT erlaubt einen Einblick in die komplexe Welt der Sprachmodelle. Damit dies möglich wird, wurden verschiedene didaktische Vereinfachungen vorgenommen. Dennoch lassen sich die grundlegenden Ideen hinter statistischer Spracherzeugung von den Schülerinnen und Schülern auf eine motivierende und spielerische Art entdecken und mit Experimenten erforschen. Selbstverständlich darf sich die Auseinandersetzung mit Sprachmodellen im Unterricht nicht auf SoekiaGPT beschränken. Es müssen auch reale große Sprachmodelle wie ChatGPT angewendet werden, Sprachmodelle als eine Anwendung im großen Kontext künstlicher Intelligenz eingeordnet und über die Chancen und Risiken von KI-Anwendungen reflektiert werden (vgl. dazu zum Beispiel die umfassenden Materialien auf ENARIS 2023).

Eine weitere Vertiefungsmöglichkeit ist der Vergleich zwischen Sprachmodellen und Suchmaschinen. In beiden Systemen wird eine große Dokumentensammlung durchsucht und analysiert, in einem effizienten Speicher aufbereitet und anschließend mit einem Algorithmus basierend auf der Eingabe eine möglichst passende Ausgabe generiert. Im Unterschied zu Suchmaschinen liefert das Sprachmodell als Ausgabe jedoch nicht eine Liste mit Suchtreffern, sondern einen einzigen Text auf Basis aller analysierten Dokumente. Die auf Soekia.ch ebenfalls angebotene Lernumgebung zu Suchmaschinen lässt sich hervorragend mit SoekiaGPT kombinieren.

Auf Soekia.ch stehen beide Werkzeuge mit Leitfäden für die Lehrperson kostenfrei zur Verfügung. Die didaktische Suchmaschine Soekia wird schon lange im Unterricht genutzt. SoekiaGPT wurde erst an wenigen Schulen erprobt und wird aufgrund der Rückmeldungen laufend überarbeitet. Wir freuen uns deshalb über Feedback und Anregungen.

Quellen

SoekiaGPT (2023): SoekiaGPT – das didaktische Sprachmodell. <https://soekia.ch/gpt.html>

Soekia 2.0 (2019): Soekia 2.0 – die didaktische Suchmaschine. <https://soekia.ch/search.html>

Gesellschaft für Informatik (2016): Dagstuhl-Erklärung: Bildung in der digitalen vernetzten Welt. <https://dagstuhl.gi.de/dagstuhl-erklaerung>

Elena Voita (2023): Natural Language Processing course. https://lena-voita.github.io/nlp_course/language_modeling.htm

ENARIS (2023): Möglichkeiten, Grenzen und innere Funktionsweisen von KI-Systemen. <https://enaris.org/material/de/>

Lizenz



Dieser Artikel steht unter der Lizenz CC BY 4.0 zur Verfügung.

Kontakt

Michael Hielscher

Pädagogische Hochschule Schwyz

E-Mail: michael.hielscher@phsz.ch