

Solutions English

1. *Why does the context always have to be considered in POS tagging?*
Since certain words are spelt the same but have two different meanings, i.e. are ambiguous, the context is the decisive factor in deciding which word type to assign in this case.
2. *How does the selection of corpora influence the results of a corpus analysis?*
The choice of corpora limits the possibilities of the research question. Newspaper corpora, for example, cannot be used to analyse the oral present, as they do not represent the oral present at all.
3. *What is sentiment analysis used for?*
Sentiment analysis: short texts are used to represent positive, neutral, or negative emotions. The sentiment analysis is often used for customer reviews for online platforms or for detecting online comments.
4. *What is the application called that creates automatically online chats?*
Chatbots
5. *What distinguishes a Large Language Model (LLM) like GPT-4 from traditional rule-based AI systems?*
LLMs like GPT-4 use machine learning to understand and generate language based on patterns learned from data, unlike traditional rule-based AI systems which follow predefined, hardcoded rules.
6. *Large Language Models (LLMs) are trained on vast datasets. What is the primary purpose of this training?*
The primary purpose of training LLMs on vast datasets is to enable them to learn language patterns, nuances, and contextual understanding, thereby improving their ability to generate human-like text and responses.
7. *What is historical bias and how can it be exemplified?*

Historical bias arises when the training data, derived from historical sources, already embodies certain prejudices or imbalances. Consequently, the model may learn and replicate stereotypes or discriminatory patterns associated with specific identity groups present in the training data.

Examples of historical bias:

In online translation, when the historical data predominantly consists of examples featuring male doctors, the model is prone to translating "the doctor" into the masculine form "der Arzt" in German.

Further examples of historical bias that could be mentioned could involve applications such as chatbots, in which it can lead to the generation of responses that perpetuate stereotypes, sexism, or racism.

8. *How did the researchers identify linguistic cues indicative of bias in the Wikipedia articles?*

The researchers identified cues by analysing edits made to remove bias. They located the words that introduced the bias, categorising them into framing and epistemological bias.

Lösungen Deutsch

1. *Weshalb ist der Kontext beim POS-Tagging immer zu berücksichtigen?*
Da gewisse Wörter zwar gleich geschrieben werden, aber zwei verschiedene Bedeutungen haben, also ambig sind, gibt der Kontext jeweils den entscheidenden Ausschlag, welche Wortart in diesem Fall zuzuordnen ist.
2. *Wie beeinflusst die Auswahl der Korpora die Ergebnisse einer korpusanalytischen Untersuchung?*
Die Auswahl der Korpora grenzt die Möglichkeiten der Fragestellung ein. So kann zum Beispiel mit Zeitungskorpora nicht die mündliche Gegenwart untersucht werden, da diese die mündliche Gegenwart überhaupt nicht abbilden.
3. *Wofür gibt es Gefühlsanalysen?*
Sentiment-Analyse: Kurze Texte werden verwendet, um positive, neutrale oder negative Emotionen darzustellen. Die Sentiment-Analyse wird häufig für Kundenbewertungen auf Online-Plattformen oder zur Erkennung von Online-Kommentaren verwendet.
4. *Wie heißt die Anwendung, die automatisch Online-Chats erstellt?*
Chatbots
5. *Was unterscheidet ein Large Language Model (LLM) wie GPT-4 von herkömmlichen regelbasierten KI-Systemen?*
LLMs wie GPT-4 nutzen maschinelles Lernen, um Sprache auf der Grundlage von aus Daten gelernten Mustern zu verstehen und zu generieren, im Gegensatz zu herkömmlichen regelbasierten KI-Systemen, die vordefinierten, fest kodierten Regeln folgen.
6. *Große Sprachmodelle (Large Language Models, LLMs) werden auf umfangreichen Datensätzen trainiert. Was ist der Hauptzweck dieses Trainings?*
Der Hauptzweck des Trainings von LLMs auf riesigen Datensätzen besteht darin, ihnen das Erlernen von Sprachmustern, Nuancen und kontextbezogenem Verständnis zu ermöglichen und dadurch ihre Fähigkeit zu verbessern, menschenähnliche Texte und Antworten zu erzeugen.
7. *Was ist historischer Bias und was ist ein Beispiel dafür?*
Historischer Bias entsteht, wenn die Trainingsdaten, die aus historischen Quellen stammen, bereits bestimmte Vorurteile oder Ungleichgewichte enthalten. Folglich kann das Modell Stereotypen oder diskriminierende Muster, die mit bestimmten Identitätsgruppen in den Trainingsdaten verbunden sind, erlernen und wiederholen.

Beispiele für historische Vorurteile:

Wenn bei der Online-Übersetzung die historischen Daten überwiegend aus Beispielen mit männlichen Ärzten bestehen, neigt das Modell dazu, "der Arzt" in die männliche Form "der Arzt" im Deutschen zu übersetzen.

Weitere Beispiele für historischen Bias, die genannt werden könnten, könnten Anwendungen wie Chatbots betreffen, bei denen er zur Generierung von Antworten führen kann, die Stereotypen, Sexismus oder Rassismus aufrechterhalten.

8. *Wie haben die Forscher sprachliche Hinweise auf Bias in den Wikipedia-Artikeln identifiziert?*

Die Forscher identifizierten die Hinweise, indem sie die Änderungen analysierten, die zur Beseitigung von Bias vorgenommen wurden. Sie ermittelten die Wörter, die den Bias verursachten, welche sie als Framing und epistemologische Bias kategorisierten.

Soluzioni Italiani

1. *Perché il contesto deve sempre essere preso in considerazione nella marcatura POS?*
Poiché alcune parole si scrivono allo stesso modo ma hanno due significati diversi, cioè sono ambigue, il contesto è il fattore decisivo per decidere quale tipo di parola assegnare in questo caso.
2. *In che modo la selezione dei corpora influenza i risultati di un'analisi dei corpora?*
La scelta dei corpora limita le possibilità della domanda di ricerca. I corpora dei giornali, ad esempio, non possono essere utilizzati per analizzare il presente orale, poiché non rappresentano affatto il presente orale.
3. *A che cosa serve l'analisi dei sentimenti?*
Analisi dei sentimenti: Testi brevi vengono utilizzati per rappresentare emozioni positive, neutre o negative. L'analisi dei sentimenti è spesso utilizzata per le recensioni de* client* su piattaforme online o per rilevare commenti online.
4. *Come si chiama l'applicazione che crea automaticamente delle chat online?*
Chatbot
5. *Cosa distingue un Large Language Model (LLM) come GPT-4 dai tradizionali sistemi di intelligenza artificiale basati su regole?*
Gli LLM come il GPT-4 utilizzano l'apprendimento automatico per comprendere e generare il linguaggio sulla base di modelli appresi dai dati, a differenza dei tradizionali sistemi di IA basati su regole che seguono regole predefinite e codificate.
6. *I modelli linguistici di grandi dimensioni (LLM) vengono addestrati su vaste serie di dati. Qual è lo scopo principale di questo addestramento?*
Lo scopo principale dell'addestramento dei LLM su vasti insiemi di dati è quello di consentire loro di apprendere modelli linguistici, sfumature e comprensione del contesto, migliorando così la loro capacità di generare testi e risposte simili a quelli umani.
7. *Che cos'è il bias storico e come può essere esemplificato?*
Il bias storico si verifica quando i dati di addestramento, derivati da fonti storiche, contengono già determinati pregiudizi (bias) o squilibri. Di conseguenza, il modello può apprendere e replicare stereotipi o modelli discriminatori associati a specifici gruppi di identità presenti nei dati di addestramento.
Esempi di bias storici:
Nella traduzione online, quando i dati di addestramento sono costituiti prevalentemente da esempi di medici maschi, il modello inclina a tradurre "il dottore" nella forma maschile "der Arzt" in tedesco.
Altri esempi di bias storici che potrebbero essere citati potrebbero riguardare applicazioni come i chatbot, che possono portare alla generazione di risposte che

perpetuano stereotipi, sessismo o razzismo.

8. *Come i ricercatori hanno identificato gli indizi linguistici indicativi di bias negli articoli di Wikipedia?*

Hanno individuato le parole che introducevano i bias, classificandole in bias di contestualizzazione e bias epistemologici.