

Annotation
mit
Sprachmodellen
zur Sprachmodellierung
mit
Annotation



Anatomie der Sprache I: Sprachliche Zeichen

Signifiant (Zeichenkörper)

Signifié (Zeicheninhalt)

(Für 2 Seiten einer Medaille vgl. Tafelbild)

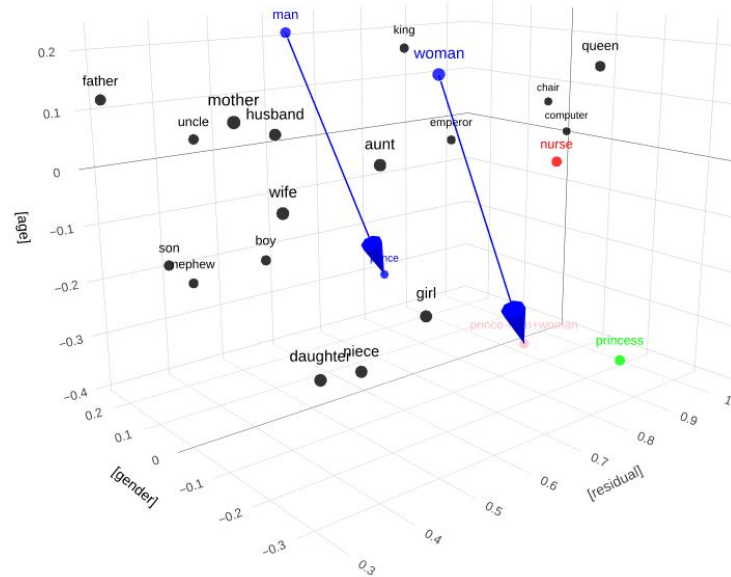
Digitale Repräsentation sprachlicher Zeichen

Zeichenkörper?



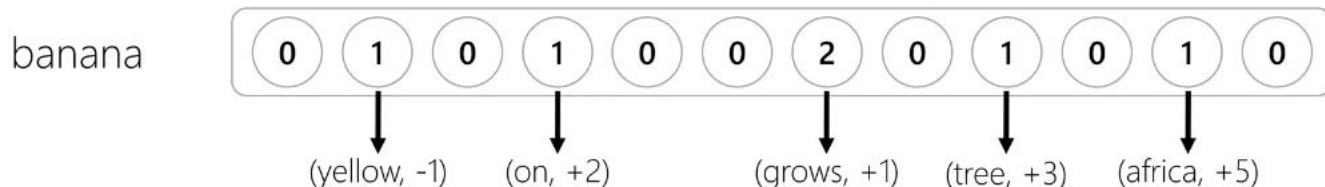
Digitale Repräsentation sprachlicher Zeichen

Zeicheninhalt?



Voraussetzungen I: Kontextvektoren

Represent an item (e.g., word) as a vector of numbers.



The vector can correspond to **neighboring word context**.

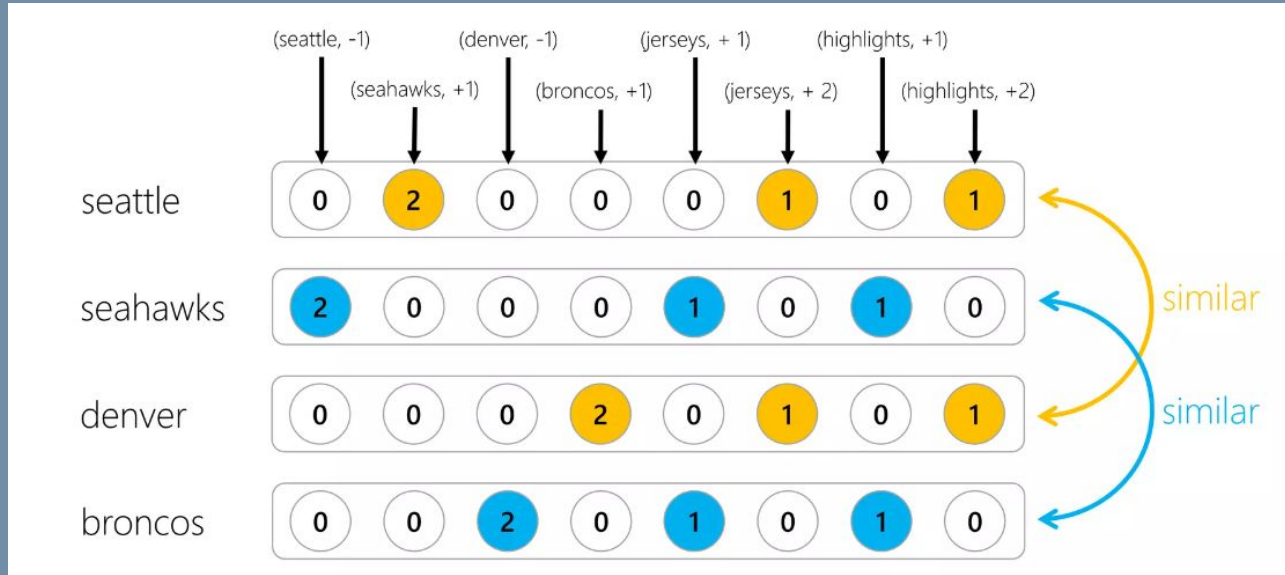
e.g., "yellow banana grows on trees in africa"

-1 0 +1 +2 +3 +4 +5

Aus (der insgesamt sehr übersichtlichen Präsentation von Bhaskar Mitra):

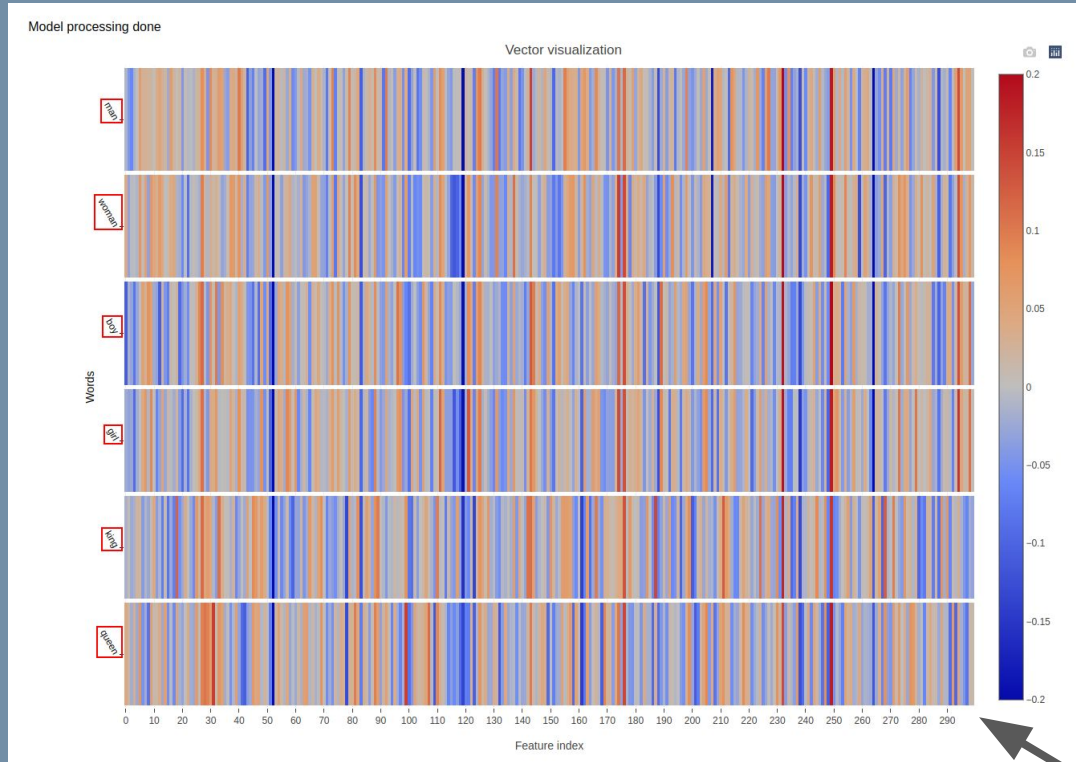
<https://www.slideshare.net/BhaskarMitra3/a-simple-introduction-to-word-embeddings>

Voraussetzungen I: Vektorrepräsentationen

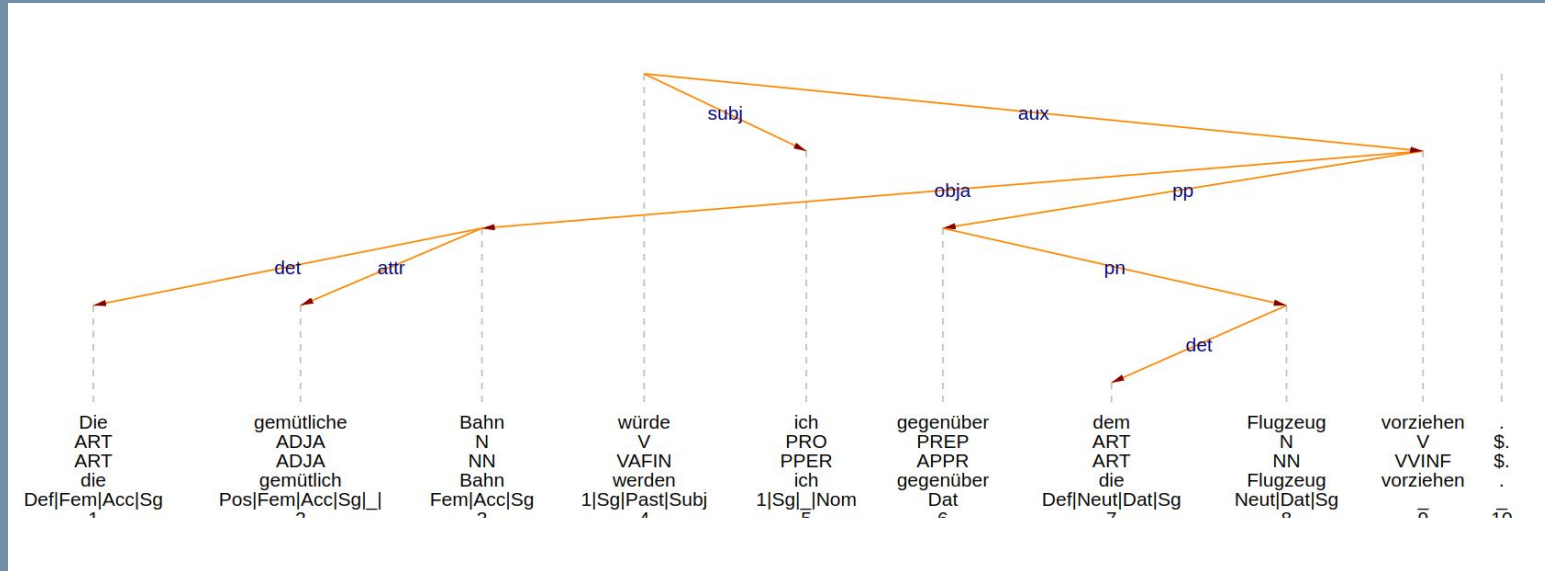


Aus (der insgesamt sehr übersichtlichen Präsentation von Bhaskar Mitra):
<https://www.slideshare.net/BhaskarMitra3/a-simple-introduction-to-word-embeddings>

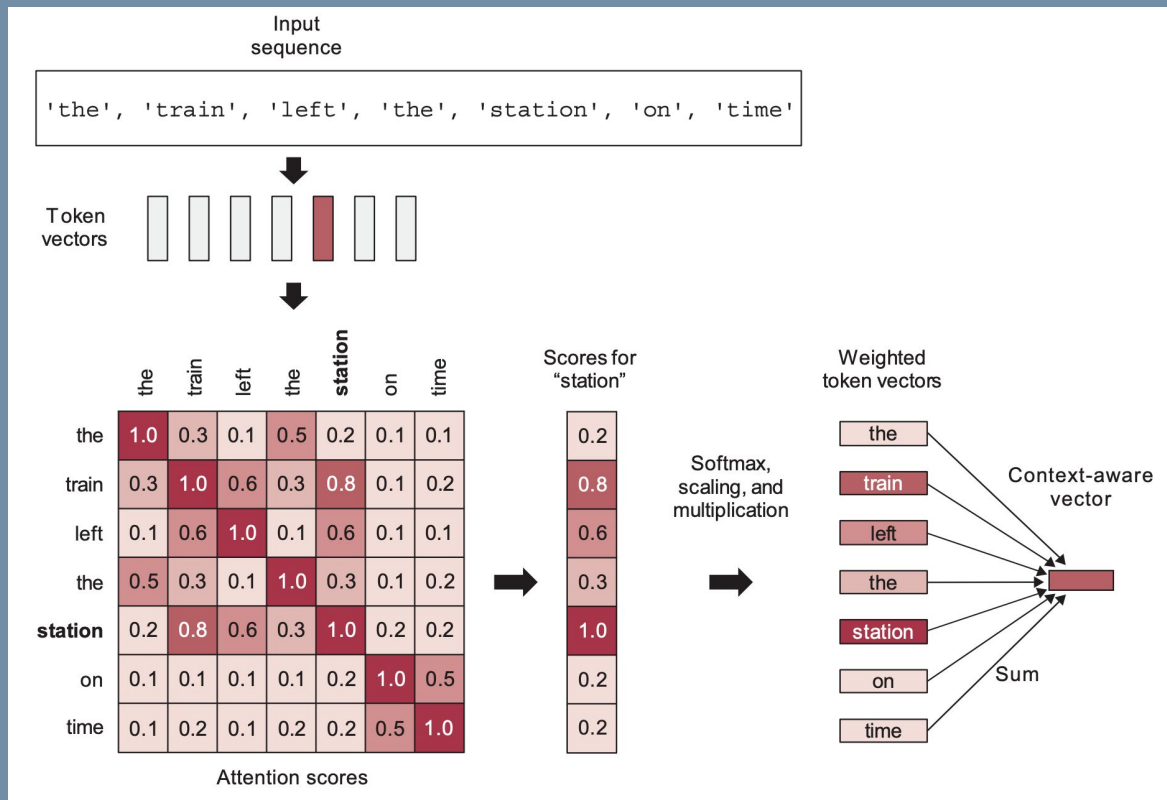
Voraussetzungen I: Dense Vectors / Embeddings



Voraussetzungen II: Attention

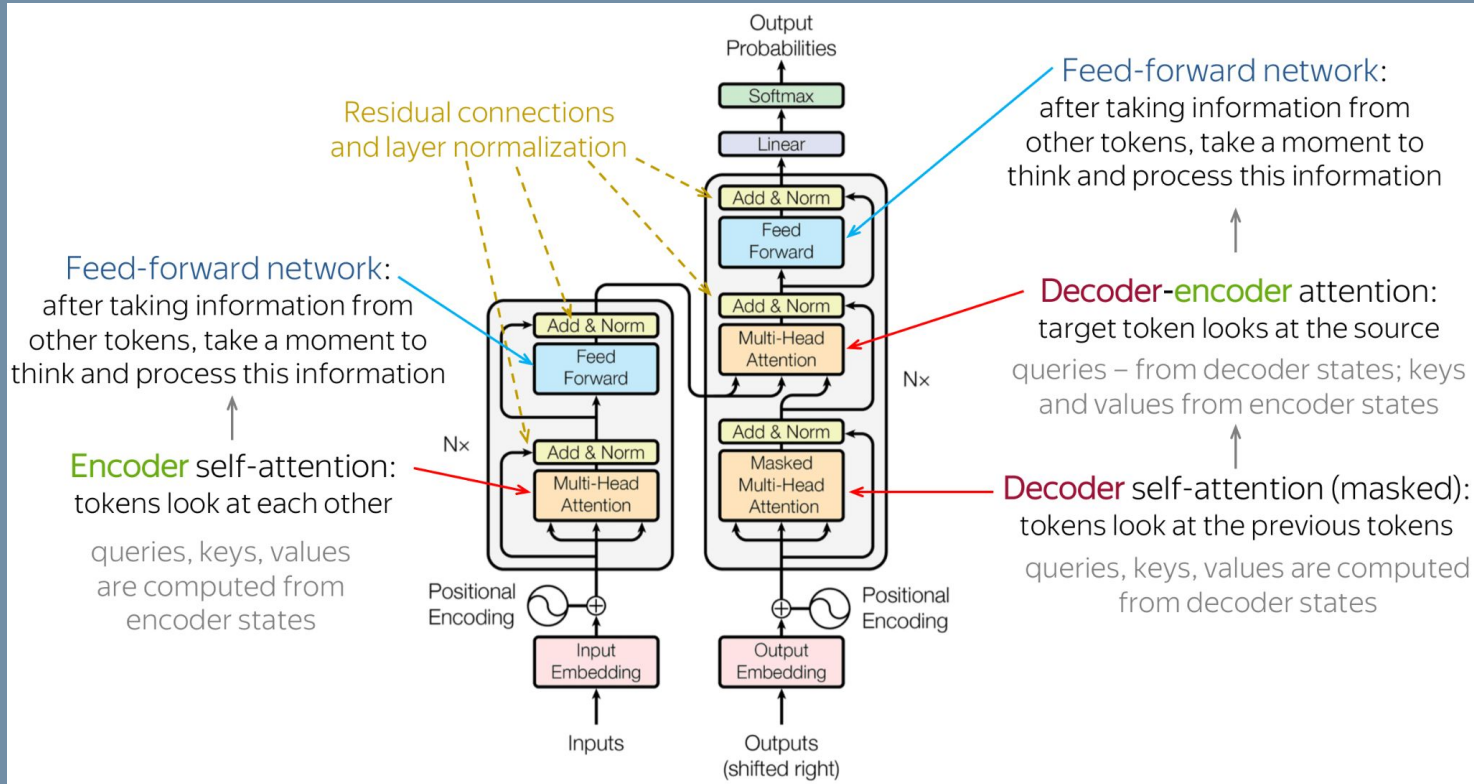


Voraussetzungen II: Attention



aus Chollet
(2023)

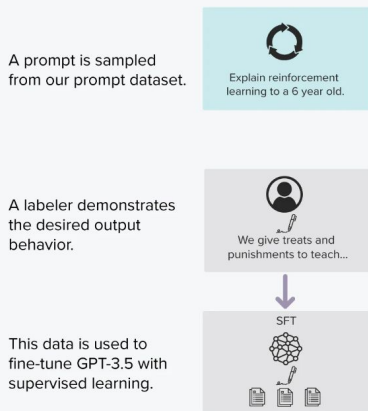
Voraussetzungen II: Attention



Finetuning: RLHF * 3

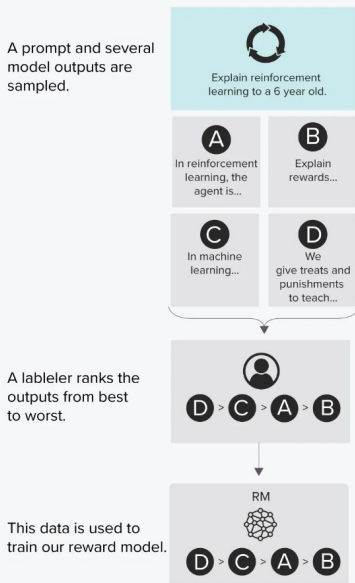
Step 1

Collect demonstration data and train a supervised policy.



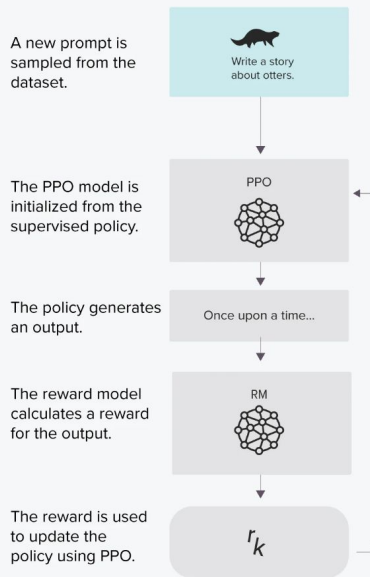
Step 2

Collect comparison data and train a reward model.



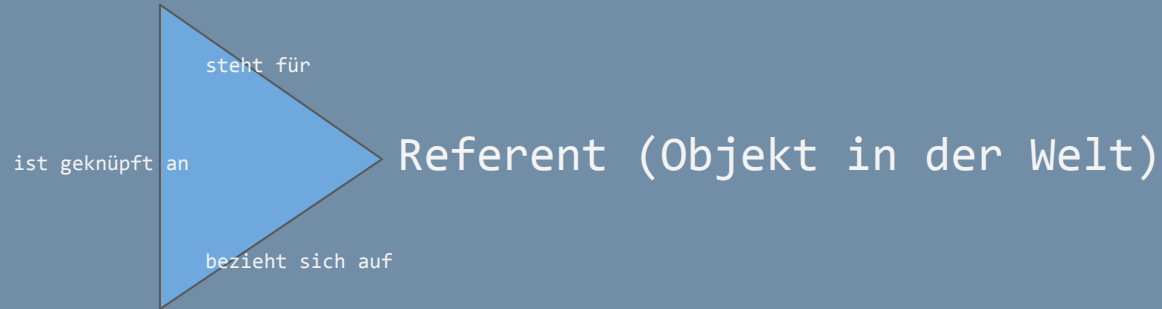
Step 3

Optimize a policy against the reward model using the PPO reinforcement learning algorithm.



Das semiotische Dreieck

Signifiant (Zeichenkörper)



Signifié (Zeicheninhalt)

Das semiotische Dreieck

Position in Codetabelle
Signifiant (Zeichenkörper)



Signifié (Zeicheninhalt)
Position im Vektorraum

Vgl. Piantadosi / Hill: Meaning without reference in large language models?
& Underwood: The empirical triumph of theory

Anatomie der Sprache II: Doppelte Gliederung

Bedeutungsunterscheidende Einheiten
(Phoneme / Grapheme)

Bedeutungstragende Einheiten
(Morpheme)