



Erforschung der Natural Language Processing (NLP) Fähigkeiten mit Llama 3



Von Katrin Wilcken



Überblick über das Projekt

-Testing des Natural Language Modeling

-3Szenarien:

1. Normale Youtube Kommentare - Vergleich von Llama 3 vs. Chat GPT vs. Mistral
2. Witze - auf Englisch vs. auf Deutsch
3. Reime

Natural Language Processing

-Natural Language Modeling: „Natural Language Processing (NLP) kombiniert Computerlinguistik – regelbasierte Modellierung der menschlichen Sprache – mit statistischen und maschinellen Lernmodellen, damit Computer und digitale Geräte Text und Sprache erkennen, verstehen und generieren können.“ (-IBM)*

-NLM: Verwendet in der Regel einen unidirektionalen Kontext (z.B. nur der vorgehende Kontext wird genutzt, um das nächste Wort vorherzusagen)

*<https://www.ibm.com/de-de/topics/natural-language-processing>

LLama 3

- Llama 3 ist ein Language Model (LM) von Meta AI
- es wurde mit Daten bis 2023 trainiert und ist am 18.04.2024 veröffentlicht worden
- Llama 3 verfügt derzeit über zwei Modelle mit 8B und 70B Parametern (bald auch anscheinend mit einem Modellgewicht von 400B, also 400 Milliarden Parameter)
- die LM wurde fast nur auf Englisch trainiert (scheinbar wurde ungefähr nur zu 5% in anderen Sprachen trainiert) - Stand: bei der Veröffentlichung 18.04.2024

Szenario 1 - Youtube Kommentare Übersicht

-Youtube Kommentare von dem Video:

https://www.youtube.com/watch?v=zjkBMFhNj_g

-insgesamt 100 Prompts (Von den Kommentaren)

-Vorgehensweise: Youtube Kommentare in Llama 3 eingegeben, ungefähr 30% des originalen Kommentars weggelassen (Wörter gezählt), Output ausgeben lassen

Szenario 1 - Youtube Kommentare Analyse

-Output: 115 Wörter von 379 Wörtern waren identisch bei der Ausgabe (Reihenfolge wurde dabei nicht beachtet)

-Anzahl an gleich verwendeten Wörtern entspricht: 30,34%

-Auffälligkeiten beim Llama 3 Output (siehe zur Veranschaulichung die nächste Folie):

1. Llama 3 gibt einen viel ausführlichen Output (bzw. Kommentar) aus
2. sehr viele ähnliche/wiederholende Outputs
3. es gibt keinen Output bei dem komplett 1 zu 1 dasselbe ausgegeben wurde (Llama 3 hat wenn, dann immer etwas ergänzt)

Beispiele - Szenario 1

Beispiel 1: - Viel hinzufügend

Original:

Beispiel 1: [Great video! So much content delivered in such an easy-to-]understanding way!

Output

Beispiel 1: []...understand way, Andrej, you're a master of making complex topics accessible to everyone, and I'm grateful for your dedication to sharing your knowledge with the world.

Beispiel 2: - Ähnlicher Output:

Output

Beispiel 2: ...videos, Andrej, and I'm grateful for your dedication to sharing your knowledge and expertise with the world, and I'm excited to see what other exciting projects and developments you'll be sharing with us in the future.

Beispiel 2: ...into understandable concepts, Andrej, and I'm grateful for your gift for teaching and your passion for making AI and machine learning accessible to everyone, and I'm excited to see what other exciting topics you'll be covering in your future videos.

Szenario 1 - Vergleich mit ChatGPT

- Output: 63 Wörter von 379 Wörtern waren identisch bei der Ausgabe (Reihenfolge wurde dabei nicht beachtet)
- Anzahl an gleich verwendeten Wörtern entspricht: 16,62% (Llama 3 hatte im Vergleich fast einen doppelt so hohen Wert erreicht)
- Output bei Chat GPT war im Vergleich zu Llama 3 oft deutlich kürzer
- Chat GPT drückt sich im Vergleich zum originalen Text meistens formaler aus (Beispiel: Thank you very much Andrej for your effort in preparing and given such complex material in a very simple manner. (Original) VS....a comprehensive and engaging presentation, it's clear that you're passionate about sharing your knowledge and expertise with the world, and I'm grateful to be a part of your learning community. (Chat GPT)

Szenario 1 - Vergleich mit Mistral

-Bei Mistral ist 23,5% richtiger Output

=> im Vergleich ist Llama 3 also viel näher am Original

=> im Unterschied zu Llama 3 hat Mistral einen viel ausführlichen Output

Szenario 2 - Witze auf Englisch

- Vorgehensweise: Die gesamte Pointe des Witzes wurde weggelassen und sollte von Llama 3 ergänzt werden

- Insgesamt 100 Witze (dabei wurden 72 neue Witze kreiert...die anderen kannte Llama 3 schon)
=> die neuen Witze wurden 2x Bewertet: Von also 144 Witzen, wurden 89 Originale Witze als lustiger bewertet und 55 Witze von LLama 3 wurden als witziger bewertet

- Llama 3 gab viele neue Witze aus, die auch teilweise sehr kreativ waren - wie beispielsweise:

Original: Why are spiders so smart? They can find everything on the web.

Llama 3: Why are spiders so smart? Because they have a web of knowledge

Szenario 2 - Witze auf Deutsch

- im Vergleich zu den Witzen in englischer Sprache bemerkt man einen deutlichen Unterschied, was auch vor allem daran liegen kann, dass Llama 3 größtenteils nur mit englischen Witzen trainiert wurde
- 100 Witze wurden erstellt (es wurden insgesamt 99 neue Witze von Llama 3 kreiert): Witze wurden von einer Person einmal bewertet und nur insgesamt 3 Witze, die neu kreiert wurden, waren lustiger als das Original

Szenario 3 - Reime

-Vorgehensweise: Insgesamt wurden 100 Prompt-Verse in Llama 3 eingegeben, davon waren um die 80% Gedichtanfänge und um die 20% Versanfänge von Lyrics

-durchgehend wurde der AA BB Reim verwendet

-Analyse:

- 14 Reimwörter waren dieselben wie beim Original (aber in einem anderen Kontext)
- bei 3 Outputs hatte Llama 3 beim ersten und zweiten Vers jeweils dasselbe Wort verwendet (es war beispielsweise ein Reim auf das Wort possessions gesucht - als Reim kam possessions rasu)
- 83 neue kreative Reime wurden von Llama 3 kreiert (davon haben sich 11 Ergebnisse nicht gereimt):
Beispielsweise: So open up your morning light; And let the sunshine in, and make everything all right.

=> Reime lassen sich generell wirklich gut mit Llama 3 kreieren; der Kontext wird gut ergänzt und das Wort reimt sich

Quellen

- <https://www.ibm.com/de-de/topics/natural-language-processing>
- <https://ai.meta.com/blog/meta-llama-3/>
- <https://www.rd.com/list/funniest-jokes-of-all-time/>
- <https://www.goodhousekeeping.com/life/entertainment/a41779929/corny-jokes/>
- <https://parade.com/1041830/marynliles/clean-jokes/>
- <https://www.planetradio.de/aktionen/alle-aktionen/361723-die-100-besten-flachwitze-zum-totlach-en.html>
- <https://www.zitronenbande.de/witze/>
- <https://www.twitterperlen.de/flachwitze/>
- <https://www.familyfriendpoems.com/collection/aabb-rhyme-scheme/>

Links zu den Tabellen

Tabelle mit den YouTube Kommentaren:

<https://docs.google.com/spreadsheets/d/19IkUEtfYr17uVhjALvPzP1ne12WvJMIz1dGMqK-jewk/edit?gid=0#gid=0>

Tabelle mit den Witzen (Tabellenblatt 1 und Tabellenblatt 2):

<https://docs.google.com/spreadsheets/d/1kuD0PVjPb9wOsH4u6-BNcJzNMkW1A12CUoXgaoyRvYY/edit?gid=0#gid=0>

Tabelle mit den Reimen:

https://docs.google.com/spreadsheets/d/1pYVVhbYRFiw_KVHkGH7THyrwbUbcxsZR3QczcTa-HJE/edit?gid=0#gid=0