

# Praxiseinstieg Large Language Models

Strategien und Best Practices für den Einsatz von  
ChatGPT und anderen LLMs

» Hier geht's  
direkt  
zum Buch

# DAS VORWORT

Obwohl die Verwendung von großen Sprachmodellen – *Large Language Models* (LLMs) – schon in den letzten fünf Jahren stetig zugenommen hat, ist das Interesse daran geradezu explodiert, als OpenAI sein Produkt ChatGPT veröffentlichte. Der KI-Chatbot hat die Leistungsfähigkeit von LLMs demonstriert und eine einfach zu bedienende Schnittstelle eingeführt, die es Menschen aus allen Gesellschaftsschichten ermöglicht, die Vorteile dieses bahnbrechenden Tools zu nutzen. Jetzt, da diese Untergruppe der Verarbeitung natürlicher Sprache – *Natural Language Processing* (NLP) – zu einem der meistdiskutierten Bereiche des maschinellen Lernens geworden ist, wollen viele Menschen sie in ihre eigenen Angebote integrieren. Diese Technologie fühlt sich tatsächlich so an, als könnte es sich um künstliche Intelligenz handeln, auch wenn es lediglich um die Vorhersage von aufeinanderfolgenden Token anhand eines probabilistischen Modells geht.

*Praxiseinstieg Large Language Models* ist ein exzellenter Überblick über das Konzept der LLMs sowie deren praktische Anwendung, und zwar für Programmierinnen und Programmierer mit und ohne Vorkenntnisse in Data Science. Die Mischung aus Erklärungen, visuellen Darstellungen und praktischen Codebeispielen macht das Buch zu einer fesselnden und leicht verständlichen Lektüre, die dazu anregt, immer wieder umzublättern. Sinan Ozdemir deckt viele Themen in einer anschaulichen Art und Weise ab und macht dieses Buch damit zu einer der besten Informationsquellen, die zur Verfügung stehen, um etwas über LLMs, ihre Fähigkeiten und den Umgang mit ihnen zu lernen und damit die besten Ergebnisse zu erzielen.

Sinan wechselt geschickt zwischen verschiedenen Aspekten von LLMs und gibt dem Leser alle Informationen, die er braucht, um LLMs effektiv zu nutzen. Beginnend mit der Diskussion, wo LLMs innerhalb von NLP angesiedelt sind, und der Erklärung von Transformern und Encodern, geht er auf Transfer Learning und Feintuning, Attention und Tokenisierung in einer verständlichen Art und Weise ein. Außerdem befasst er sich mit vielen weiteren Aspekten von LLMs, zu denen gehören: die Kompromisse zwischen Open-Source-Modellen und kommerziellen Optionen, wie man Vektordatenbanken nutzt (schon für sich genommen ein sehr beliebtes Thema), das Schreiben eigener APIs mit Fast API, das Erstellen von Embeddings

und das Überführen von LLMs in die Produktion – etwas, das sich für jede Art von Machine-Learning-Projekt als Herausforderung erweisen kann.

Ein großer Teil dieses Buchs beschäftigt sich sowohl mit visuellen Schnittstellen – wie zum Beispiel ChatGPT – als mit auch Schnittstellen für die Programmierung. Sinan stellt hilfreichen Python-Code zur Verfügung, der leicht verständlich ist und klar veranschaulicht, was im Einzelnen passiert. Im Rahmen des Prompt Engineering führt er vor, wie sich drastisch bessere Ergebnisse von LLMs erzielen lassen, und – was noch besser ist – er demonstriert, wie man diese Prompts sowohl in der visuellen GUI als auch über die Python-Bibliothek von OpenAI bereitstellen kann.

Dieses Buch hat mich so inspiriert, dass ich versucht war, dieses Vorwort mit ChatGPT zu schreiben, um all das zu demonstrieren, was ich gelernt habe. Dies zeigt, wie gut geschrieben, ansprechend und informativ das Buch ist. Auch wenn ich dazu in der Lage gewesen wäre, habe ich dieses Vorwort doch selbst geschrieben, um meine Gedanken und Erfahrungen über LLMs auf die authentischste und persönlichste Art und Weise zu formulieren, die ich kenne. Mit Ausnahme des letzten Teils des letzten Satzes, der von ChatGPT stammt, einfach weil ich es konnte.

Für jemanden, der mehr über die vielen Aspekte von LLMs lernen möchte, ist dies das richtige Buch. Es wird Ihnen helfen, die Modelle zu verstehen und sie in Ihrem täglichen Leben effektiv zu nutzen. Und was vielleicht am wichtigsten ist: Sie werden diese Reise genießen.

*– Jared Lander, Editor der Reihe bei Addison-Wesley*