

Pressemitteilung

Technische Universität Berlin

Stefanie Terp

30.01.2025

<http://idw-online.de/de/news846599>

Forschungs- / Wissenstransfer, Forschungsprojekte
Informationstechnik
überregional



TU Berlin: Expert*innen zum jüngsten Erfolg von „DeepSeek“

Expert*innen der TU Berlin zum jüngsten Erfolg des Large Language Models (LLM) „DeepSeek“ aus China, dem Unterschied zwischen Open-Source-Anwendungen wie DeepSeek und anderen LLMs sowie der Rolle Europas bei der Entwicklung von Künstlicher Intelligenz (KI)

Die Expert*innen:

Dr. Vera Schmitt (Forschungsgruppenleiterin) und Dr. Nils Feldhus (Postdoktorand) forschen in der XplaiNLP-Gruppe des Quality and Usability Lab an der TU Berlin zu Hochrisiko-KI-Anwendungen und entwickeln KI-gestützte Systeme zur intelligenten Entscheidungsunterstützung. Ihr Fokus liegt auf leistungsstarken, transparenten und erklärbaren KI-Lösungen für Anwendungsfelder wie die Erkennung von Desinformation und die Analyse medizinischer Daten. Im Bereich der natürlichen Sprachverarbeitung arbeitet die Gruppe an Schlüsselthemen wie erklärbarer KI, der Robustheit großer Sprachmodelle (LLMs), der Modellierung von Argumentationsstrukturen sowie der Mensch-Maschine-Interaktion.

Dr. Oliver Eberle ist Postdoktorand in der Machine Learning Group des Berlin Institute for the Foundations of Learning and Data (BIFOLD) an der TU Berlin. In seiner Forschung widmet er sich vor allem der erklärbaren Künstlichen Intelligenz sowie Natural Language Processing und deren Anwendungen in den Wissenschaften, wie den Digital Humanities (zum Beispiel der computergestützten Textverarbeitung) und der Kognitionswissenschaft. Er fokussiert sich dabei besonders auf die Interpretierbarkeit der Modelle und entwickelt Methoden zum besseren Verständnis der zugrundeliegenden Mechanismen von Large Language Models (LLM).

1. Wie unterscheiden sich die Konzepte von DeepSeek und ChatGPT?

Schmitt und Feldhus: DeepSeek steht für Open-Source-Transparenz und Effizienz, während ChatGPT auf massive Rechenleistung und Skalierung setzt. Ersteres ermöglicht Anpassung und niedrigere Kosten, letzteres bietet optimierte Performance, bleibt aber proprietär und ressourcenintensiv. Man muss allerdings sehen, dass DeepSeek nicht 100 Prozent Open-Source ist, denn zum Beispiel sind nicht alle Trainingsdaten bekannt, die in das Modell hineingeflossen sind. Die Verfügbarkeit der Modellparameter hingegen und die deutlich offenere Kommunikation seitens DeepSeek erlaubt es Initiativen der Open-Source-Community wie zum Beispiel „Open-R1“ die Reproduktion des Modells in Angriff zu nehmen und dabei auf viel weniger Ressourcen zurückgreifen zu müssen im Vergleich zur riesigen und teuren Infrastruktur von OpenAI, Microsoft und anderen.

Eberle: DeepSeek ist integriert in die „Huggingface Community“, eine Plattform, die bereits Hunderte von Open-Source-Modellen und Quellcode der Modelle verfügbar macht und eine wichtige Rolle in der Verfügbarkeit, Zugänglichkeit und Transparenz von LLMs sowohl in der Forschung als auch der Industrie spielt. DeepSeek verwendete in der Vergangenheit bereits andere Open-Source-Modelle (zum Beispiel das Llama-Modell von Meta) als Grundlage

(zum Beispiel bei „DeepSeek-R1-Distill-Llama-70b“). Dies spart Rechenaufwand, da die Destillation von Modellen deutlich weniger rechenintensiv ist als das Trainieren eines neuen Modells von Grund auf. DeepSeek veröffentlicht detaillierte Beschreibungen und technische Reports seiner Modelle und beschreibt hierin auch Negativ-Resultate. Dies ist ein hilfreicher Beitrag für die Open-Source-Community, weil so die Verbesserung von künftigen offenen LLM-Systemen vorangebracht wird. ChatGPT ist im Vergleich proprietär und nur das Interface ist zugänglich, die genaue Spezifikation des Modells und die trainierten Parameter sind nicht im Detail bekannt oder offen zugänglich. Der Code zum Trainieren sowie spezifische Datensätze werden meines Wissens nach weder von DeepSeek noch ChatGPT veröffentlicht.

2. Arbeiten Sie bereits mit anderen Open Source Large Language Modellen (LLMs)?

Schmitt und Feldhus: Wir arbeiten viel mit unterschiedlichen LLMs wie zum Beispiel LLaMa, Mistral, Qwen, Bloom, Vicuna und haben auch angefangen, mit DeepSeek zu experimentieren. Diese Open-Source-Modelle setzen wir gezielt in verschiedenen Anwendungsbereichen ein. Ein besonderer Fokus liegt auf der Desinformationserkennung, wo wir LLMs nutzen, um Narrative in digitalen Medien zu analysieren, Falschinformationen aufzudecken und Erklärungen für erkannte Fehlinformationen bereitzustellen. Wir setzen LLMs auch für die Anonymisierung und Verarbeitung medizinischer Daten in gemeinsamen Projekten mit der Charité ein.

Eberle: Wir arbeiten mit verschiedenen Modellen, zum Beispiel Llama, Mistral, Gemma, Qwen, Mamba und wir fokussieren uns dabei besonders auf die Interpretierbarkeit und entwickeln Methoden zum besseren Verständnis der zugrundeliegenden Mechanismen von LLMs.

3. Wie unterstützt der Open-Source-Ansatz bei Large Language Modellen konkret Ihre Forschungen? Wird Deepseek Ihre Forschung noch weiter voranbringen?

Schmitt und Feldhus: Ein Open-Source-Ansatz bei LLMs ermöglicht uns, Modelle gezielt für unsere Forschung anzupassen. Durch offenen Zugang können wir Transparenz gewährleisten und spezifische Architektur Anpassungen vornehmen. Zudem können wir so Modelle evaluieren, sie weiterentwickeln und in Mensch-KI-Prozesse effektiver integrieren. DeepSeek könnte unsere Forschung weiter voranbringen, da es effizientere Modellarchitekturen und neue Trainingsansätze bietet und diese auf Rechnern der TU Berlin reproduzierbar macht. Besonders spannend sind potenzielle Verbesserungen bei der Ressourceneffizienz, aber auch in der mehrsprachigen Verarbeitung und der Adaptierbarkeit für spezifische Domänen, die unsere bestehenden Methoden ergänzen und optimieren könnten.

Eberle: DeepSeek reiht sich ein in andere Open-Source-Modellfamilien (Llama, Mistral, Qwen und so weiter) und ermöglicht es uns, Aussagen über eine breitere Menge an LLMs zu machen. Die Struktur dieser Modelle ist weitestgehend vergleichbar und unterscheidet sich vor allem durch den Trainingsansatz und die verwendeten Datensätze. DeepSeek ermöglicht uns nun Zugang zu einem Modell mit state-of-the-art Fähigkeiten in logischem Denken (reasoning capabilities), was zu neuen Einblicken in das Lösen komplexer Aufgaben durch LLMs führen könnte.

4. Warum sind mit dem Erfolg/Misserfolg von KI auch Chiphersteller wie NVIDIA verknüpft?

Schmitt und Feldhus: Der Erfolg oder Misserfolg von KI ist eng mit Chipherstellern wie NVIDIA verknüpft, weil moderne KI-Modelle enorme Rechenleistung benötigen, die hauptsächlich durch spezialisierte GPUs (Graphics Processing Units) und KI-Beschleuniger bereitgestellt wird. NVIDIA ist führend in diesem Bereich mit leistungsstarken Chips wie der H100- und A100-Serie, die speziell dafür entwickelt wurden, künstliche Intelligenz zu trainieren und ihre Ergebnisse schnell bereitzustellen. Dazu bietet NVIDIA mit CUDA auch die passende Software an, die diese Berechnungen effizient

ermöglicht. Wenn KI-Technologien florieren, steigt natürlich die Nachfrage nach diesen Chips stark an – Unternehmen, Forschungseinrichtungen und Cloud-Anbieter investieren massiv in GPU-Cluster. Dies treibt den Umsatz und den Aktienkurs von NVIDIA in die Höhe. Umgekehrt würde ein Rückgang der KI-Nachfrage oder technologische Verschiebungen hin zu alternativen Architekturen (wie wir das jetzt mit DeepSeek R1/V3 beobachten können) die Abhängigkeit von NVIDIA verringern und deren Geschäft zum Teil negativ beeinflussen. Die doppelte Monopolstellung NVIDIAs – Hardware und Software – macht es hingegen schwer, die KI-Erfolge von dem Unternehmen zu entkoppeln. Solange auch DeepSeek GPUs von NVIDIA beziehungsweise CUDA benutzt, ist NVIDIA aus dem KI-Diskurs nicht wegzudenken. Kurz gesagt: Die Hardware-Entwicklung und der Erfolg von KI sind symbiotisch – Fortschritte in KI treiben die Chipindustrie an, während leistungsfähigere Chips neue KI-Modelle ermöglichen.

5. Wussten alle in der Community schon von der großen Durchschlagkraft des neuen, chinesischen LLM?

Schmitt und Feldhus: Ja, es war absehbar, dass China verstärkt in die Entwicklung leistungsfähiger LLMs investiert. Die Fortschritte von DeepSeek und anderen chinesischen Modellen kamen nicht aus dem Nichts – bereits in den letzten Jahren gab es enorme Investitionen und strategische Initiativen im KI-Sektor. Daher ist DeepSeek keine große Überraschung, sondern eine natürliche Weiterentwicklung, ressourceneffizientere LLMs zu erstellen. Zudem baut DeepSeek stark auf bestehende Open-Source-Modellfamilien wie LLaMA, Mistral und Qwen auf und erweitert unsere Möglichkeiten, eine breitere Palette an LLMs zu analysieren. Besonders Qwen, ebenfalls ein Produkt chinesischer Forschung, hat für uns schon deutlich gemacht, dass China hier ein entscheidender, nicht zu unterschätzender Player ist. Bemerkenswert an DeepSeek R1 ist, dass sich die Argumentationsfähigkeit deutlich verbessert hat und uns neue Einblicke in die Fähigkeit von LLMs zur Lösung komplexer Aufgaben eröffnet. Das ist vor allem für schwierigere Aufgaben mit einem höheren Komplexitätsniveau, wie die Desinformationserkennung, interessant.

Eberle: DeepSeek ist durchaus bekannt, und dessen Vorgängermodell DeepSeek-V2 war bereits recht erfolgreich, zum Beispiel in der Generierung von Code. Daher bin ich etwas überrascht über die starke Reaktion der Medien und Märkte. DeepSeek-V3 ist klar eine beeindruckende technische Leistung und kann dazu beitragen, Open-Source-Modelle auf eine Stufe mit den Fähigkeiten von proprietären Modellen wie ChatGPT zu bringen. DeepSeek ist dennoch im Zusammenhang mit der erfolgreichen Entwicklung von anderen Open-Source-LLM zu sehen.

6. Wie ist die Stellung von Europa auf diesem Gebiet?

Schmitt und Feldhus: Aktuell liegt der Fokus innerhalb der EU vor allem auf der Regulierung von KI und es werden nicht genügend Ressourcen gebündelt bereitgestellt, um auch nur entfernt ein Gegengewicht zu den USA oder China bilden zu können. Vor allem, wenn wir die Investitionspläne wie Stargate berücksichtigen, kann die EU aktuell nicht mithalten. Die EU kann derzeit nicht konkurrenzfähig bleiben, da vielversprechende KI-Startups oft von US-Unternehmen übernommen werden und/oder ihren Hauptsitz in die USA verlagern. Regulierungen und Steuern beeinflussen die Innovationskraft von NLP-Unternehmen (Natural Language Processing) erheblich innerhalb der EU. Wir sehen an der Innovationsfreudigkeit von kleinen europäischen Labs wie Mistral oder Flux (Bildgenerierung), dass die europäische Forschungsgemeinschaft trotzdem an der globalen KI-Entwicklung teilnehmen möchte, auch einen recht großen Einfluss hat und mit mehr Investitionen diese Ambitionen befeuert werden können und Europa zu einem echten KI-Player aufsteigen könnte.

Eberle: Europa und Deutschland fokussieren sich auf die Entwicklung von vertrauenswürdigen und transparenten KI Methoden. Ich habe auch den Eindruck, dass Europa sich auf spezifische Anwendungen von LLMs spezialisiert, zum Beispiel LLM-Grundlagenmodelle für Anwendungen in der Medizin (z.B. aignostics' RudolfV-Modell zur Erkennung von Pathologie-Daten), den Rechtswissenschaften (Legal LLMs wie LEGAL-BERT zur Bearbeitung und Erstellung von Rechtstexten) oder KI-Methoden für Quanten-Chemie.

7. Die Anwendung DeepSeek unterliegt ja der chinesischen Zensur. Inwiefern beeinflussen solche Einschränkungen die Leistungsfähigkeit von Large Language Modellen?

Eberle: Die Einschränkungen werden meist nach dem eigentlichen Modell-Training auferlegt, sind also wie ein Filter zu sehen, der ungewollte Ausgaben unterdrückt. Daher würde ich nicht grundsätzlich davon ausgehen, dass themenoffene Systeme generell leistungsfähiger sind. Falls jedoch größere Datenmengen bereits vor dem Training gefiltert werden, könnte das Auswirkungen auf die Generalisierungsfähigkeit dieser Modelle haben. Es ist hierbei ein wichtiger Unterschied, ob das Modell keine Daten über sensible Themen bekommt, oder ob das Modell nichts über diese sagen soll.

Weitere Informationen erteilen Ihnen gern:

Dr. Vera Schmitt und Dr. Nils Feldhus:

XplaiNLP-Gruppe

Quality and Usability Lab

TU Berlin

E-Mail: vera.schmitt@tu-berlin.de, nils.feldhus@campus.tu-berlin.de

Tel.: +49 30 314-77129 (Sekretariat XplaiNLP Group / Quality and Usability Lab)

Dr. Oliver Eberle

Machine Learning Group

Berlin Institute for the Foundations of Learning and Data (BIFOLD)

TU Berlin

E-Mail: oliver.eberle@tu-berlin.de

Tel.: +49 30 314-78621 (BIFOLD Coordination Office)