

4 Vorteile von Open Source-basierten Small Language Models

Small Language Models transformieren die Strategien für Unternehmens-KI

Proprietäre LLMs (Large Language Models) zeigen beim Einsatz für allgemeine Anwendungen überragende Leistungen. Für unternehmensgerechte KI-Lösungen (künstliche Intelligenz) sind sie jedoch nicht immer die beste Alternative. Sie sind mit beachtlichen Rechenanforderungen, undurchsichtigen Entscheidungsprozessen und hohen Lizenzkosten verbunden, was die Flexibilität einschränken und die operative Komplexität erhöhen kann. Small Language Models (SLMs) – insbesondere solche, die auf Open Source-Prinzipien basieren –, stellen eine Alternative für Unternehmen dar, die individuell angepasste KI-Lösungen entwickeln wollen.

Hier finden Sie 4 gute Gründe, warum Open Source-SLMs die richtige Wahl für Ihr nächstes KI-Projekt sein könnten.

1 Zugriff auf Community-Innovationen

Open Source-SLMs bieten Unternehmen dank ihrer Flexibilität, Zusammenarbeit und Innovation eine Basis für die Entwicklung hochgradig anpassbarer und spezialisierter KI-Anwendungen. Open Source-KI-Projekte gewähren nicht nur Zugriff auf Softwarekomponenten, sondern auch auf vortrainierte Modellgewichtungen. Dies ermöglicht Ihnen die Zusammenarbeit mit einer globalen Community von Fachkräften aus den Bereichen Entwicklung und Forschung, sodass Sie Ihre generativen KI-Technologien (gen KI) kontinuierlich weiterentwickeln und verbessern können. Die in der Community geteilten Innovationen geben Ihnen die Möglichkeit, mit modernen, fortschrittlichen Tools zu arbeiten und diese an die technischen Anforderungen Ihrer Unternehmens-KI-Lösungen anzupassen.

Bei Open Source-SLMs wie den gen KI-Modellen der **IBM Granite-Familie** können Sie mit Ihrem Wissen und Ihrer Domain-Expertise einen direkten Beitrag zu Basismodellen leisten. Statt auf Updates bei proprietären LLMs zu warten, können Sie Open Source-SLMs aktiv anpassen und so deren Relevanz und Performance für Ihre KI-Anwendungen erhöhen. Dieser interaktive Ansatz ermöglicht Ihnen schnellere Iterationen und sorgt dafür, dass Ihre Modelle mit sich ändernden geschäftlichen Anforderungen Schritt halten.

Open Source-SLMs bieten die unerlässliche Flexibilität, die Deployments in dynamischen Umgebungen mit verschiedenen Onsite-Rechenzentren und Public Cloud-Infrastrukturen erfordern. Sie bieten Ihnen vollständige Kontrolle über Ihre Modelle, damit Sie diese für verschiedenste Deployment-Szenarien optimieren können – von Umgebungen mit strengen Compliance-Anforderungen bis hin zur KI-Verarbeitung in Echtzeit. Und weil Sie bei Open Source-SLMs die Kontrolle über Ihren KI-Technologie-Stack behalten, stellen diese Modelle sicher, dass Sie Ihre innovativen KI-Lösungen auch dann noch anpassen und skalieren können, wenn sich Ihre technischen und geschäftlichen Anforderungen ändern.

2 Kontrolle über Trainingsdaten

Im Vergleich zu proprietären Alternativen bieten Open Source-SLMs mehr Transparenz. Da bewährte Anbieter die Daten offenlegen, mit denen diese Modelle vortrainiert wurden, können Sie die Modellqualität gründlich prüfen und bestätigen, dass keine schädlichen oder verzerrten Informationen enthalten sind. Diese Transparenz erlaubt es Ihnen, fundierte Entscheidungen in Bezug auf die Einführung und Bereitstellung von Modellen zu treffen. Dadurch können Sie sicherstellen, dass Ihre KI-Lösungen ethischen Standards und geschäftlichen Zielen gerecht werden, bevor Sie Ihre eigenen proprietären, vertraulichen Daten einfließen lassen.

Da Sie SLMs in Onsite-Rechenzentren und Private Cloud-Ressourcen für die gesamten IT-Umgebungen Ihres Unternehmens bereitstellen können, erhalten Sie außerdem die vollständige Kontrolle über Ihre Trainingsdaten. Diese Kontrolle ist von entscheidender Bedeutung für Unternehmen, die höchst vertrauliche oder regulierte Daten verarbeiten, da sie so sicherstellen können, dass proprietäre Informationen externen Anbietern gegenüber niemals offengelegt werden. Wenn Sie Ihre gen KI-Modelle in Ihrer eigenen Umgebung verwalten, können Sie den Zugriff kontrollieren, die Einhaltung von Richtlinien optimieren, die Datensicherheit erhöhen und bei sämtlichen KI-Lösungen für mehr Transparenz sorgen.

Außerdem bieten die Modelle der IBM Granite-Familie **Sicherheitsrichtlinien**, die Kunden von der Haftung gegenüber Ansprüchen freistellen, die auf einer behaupteten Verletzung der geistigen Eigentumsrechte von Dritten durch die bereitgestellte Open Source-Software- oder KI-Modelle beruhen. Durch die Wahl dieser Modelle und Anbieter können Sie Ihr Unternehmen in der komplexen und sich ändernden technologischen Landschaft von KI noch besser schützen.

3 Anpassbare KI-Lösungen

Open Source-SLMs ermöglichen Ihnen die schnelle und effiziente Entwicklung von KI-Lösungen, die auf Ihre spezifischen Geschäftsanforderungen zugeschnitten sind. Diese Modelle wurden gezielt für bestimmte Use Cases konzipiert und entwickelt und lassen sich daher präzise an Ihre spezifischen Herausforderungen anpassen. Gleichzeitig vermeiden sie die Komplexität und den Ressourcenbedarf, den allgemeine LLMs mit sich bringen.

Über das Fine Tuning der SLMs mit Ihren Unternehmensdaten können Sie Ihre betrieblichen Informationen und Domain-Expertise direkt in die Modellparameter einbetten. Dieser Ansatz verbessert die Relevanz der Antworten der SLMs, reduziert die Häufigkeit sowie die Kosten von Retrainings und verkürzt die Entwicklungszeit wichtiger KI-Anwendungen und -Services.

Dank ihrer kompakten Größe und reduzierter Datenanforderungen im Vergleich zu LLMs lassen sich SLMs einfacher anpassen. So können Sie genaue, effiziente Modelle entwickeln, die für bestimmte Aufgaben oder Domains optimiert wurden. In ressourcenbeschränkten Umgebungen und Edge Deployments ermöglichen SLMs zudem, dass Echtzeitanwendungen direkt auf Nutzergeräten ausgeführt werden. Dies vereinfacht die Entwicklung und lässt die Notwendigkeit einer externen Cloud-Infrastruktur entfallen.

Mit SLMs wie den IBM Granite-Modellen lässt sich auch der Übergang von Experimenten zur Produktion optimieren. Durch die vereinfachte Integration von SLMs mit diversen Hardware- und Softwareinfrastrukturen können Sie Ihre gen KI-Lösungen auf die IT-Umgebung Ihres Unternehmens zuschneiden. Diese Anpassbarkeit reduziert die operative Komplexität und ermöglicht Ihnen gleichzeitig, die Kontrolle über das Deployment und die Performance zu behalten.

4 Reduzierte KI-Modellkosten

Die Rechenanforderungen von KI zu reduzieren ist für viele Unternehmen entscheidend, um ihre Ausgaben erfolgreich verwalten zu können. Open Source-SLMs liefern die notwendige Performance für fortschrittliche gen KI-Lösungen und senken im Vergleich zu LLMs die Kosten für Training und Inferenzen sowie die erforderliche Rechenleistung.

Aufgrund der geringeren Größe – oft tausendmal kleiner als führende LLMs – erfordern SLMs viel weniger Rechenressourcen, Daten und Strom. Diese Effizienz unterstützt schnellere Trainingszeiten, vereinfacht das Fine Tuning und fördert einen nachhaltigeren Ansatz für die KI-Entwicklung.

Darüber hinaus lassen sich Open Source-SLMs effizient für mehrere verschiedene Projekte und Organisationen skalieren, ohne dass teure Hardware-Upgrades erforderlich werden. Durch das Deployment dieser Modelle in Ihrer bestehenden IT-Infrastruktur können Sie personalisierte KI-Lösungen entwickeln, ohne Abstriche bei der Performance zu machen oder Budgeteinschränkungen zu überschreiten.

Die Kostenersparnisse gehen über die Infrastruktur hinaus. Mit Open Source-SLMs beseitigen Sie auch die Lizenzgebühren, die proprietäre Modelle mit sich bringen. Dadurch bieten Sie kostengünstigen Zugriff auf moderne gen KI-Funktionen ohne anbieterabhängige Einschränkungen oder Auflagen.

Innovationen mit Open Source-SLMs von Red Hat und IBM

Die Open Source-basierten gen KI-Modelle der Granite-Familie – von IBM entwickelt und in **Red Hat® Enterprise Linux® AI** enthalten – sind auf die spezifischen Anforderungen von Unternehmens-KI-Anwendungen ausgerichtet.

Erfahren Sie mehr über Open Source-basierte gen KI-Modelle

Im E-Book [Maximierung von KI-Innovationen mit Open Source-Modellen](#) erfahren Sie mehr zu den aufgabenspezifischen SLMs und Open Source-basierten gen KI-Lösungen.



Über Red Hat

Red Hat unterstützt Kunden dabei, ihre Umgebungen zu standardisieren, cloudnative Anwendungen zu entwickeln und komplexe Umgebungen mit **vielfach ausgezeichnetem** Support, Training und Consulting Services zu integrieren, zu automatisieren, zu sichern und zu verwalten.

f facebook.com/redhatinc
X @RedHatDACH
in linkedin.com/company/red-hat

**EUROPA, NAHOST,
UND AFRIKA (EMEA)**
00800 7334 2835
de.redhat.com
europe@redhat.com

TÜRKEI
00800 448820640

ISRAEL
1 809 449548

VAE
8000-4449549