DIE GOLDENE MITTE

Welche Parameter für eine Infrastruktur angesetzt werden sollten, die für den Betrieb von Künstlicher Intelligenz und Machine Learning ausgelegt ist.

Autor: Hans Hallitzky Redaktion: Diana Künstler

➤ Künstliche Intelligenz (KI) und verwandte Technologien wie Maschinelles Lernen (ML) und Deep Learning werden in vielen Branchen immer häufiger eingesetzt. Die Technologien müssen jedoch zunächst trainiert werden, was in den meisten Fällen eine riesige Menge an Daten und auch Zeit erfordert. Das wiederum hat einen erheblichen Einfluss darauf, wie die Infrastruktur für den Betrieb dieser Technologien aussehen sollte. Während dabei viel für die Nutzung entsprechender Cloud-Angebote zu sprechen scheint, funktioniert eine Hybridlösung jedoch in den meisten Fällen besser. Unternehmen, die den Einsatz von KI oder Maschinellem Lernen planen, sollten ihre spezifischen Anforderungen daher genau prüfen.

Datenmengen

Es gibt gute Gründe für Unternehmen, Cloud-Dienste zu nutzen, um ihre KI- oder ML-Anwendungen auszuführen. In erster Linie kann der benötigte Speicherplatz und die Rechenleistung auf einer Pay-per-use-Basis eingesetzt werden und ist leichter skalierbar – solange man das Geld dafür hat. Ein anderer Grund ist, dass die großen Cloud-Anbieter ein umfangreiches Angebot an leistungsstarken KI- und ML-Anwendungen offerieren.

Dennoch kann es sich für Unternehmen lohnen, ein hybrides Modell zu wählen und die Daten im eigenen Rechenzentrum zu speichern und zu bearbeiten, bevor sie in die Cloud wandern: KI- und ML-Algorithmen können in der Regel nicht direkt auf den Rohdaten ausgeführt werden. Die Daten müssen vorverarbeitet werden, damit die Algorithmen richtig funktionieren, einschließlich Aggregation und Normalisierung. Entsprechende Standard-Speicher- und -Verarbeitungsaufgaben lassen sich jedoch in den meisten Fällen günstiger im eigenen Rechenzentrum realisieren. Beispielsweise lassen sich die Kosten für die Übertragung der Daten in die Cloud vermeiden sowie auch die Kosten für die Speicherung selbst reduzieren. Und da fortgeschrittene KI- und ML-Methoden eher mehr als weniger Daten benötigen, ist dieser Grund ein entscheidender. Das auch dann, wenn Unternehmen gegebenenfalls mit Infrastrukturanbietern zusammenarbeiten, die ähnliche Vorteile wie die Cloud-Provider bieten. Zwei Punkte, die dafür sprechen:

▸ Unternehmen können sowohl kurz- als auch langfristig Kosten reduzieren, indem sie auf Konsolidierung, Automatisierung und operative Effizienz setzen. Erster Schritt hierzu ist der Wechsel von einer teuren, Speichermedien-orientierten zu einer Software-optimierten Architektur, um so die Benutzerfreundlichkeit und die Automatisierung zu verbessern.

▸ Unternehmen sollten flexible Verbrauchsmodelle nutzen, um Vorabinvestitionen zu minimieren und nur zu bezahlen, was sie nutzen. Die Vorteile hierbei sind sofortige Skalierbarkeit ohne Vorabinvestitionen. Ein Punkt, der für diesen Ansatz spricht, besteht darin, dass Projekte auch zurückgefahren werden können, wodurch die hohen Betriebskosten aus der Anfangsphase des Projekts reduziert werden.

Einhaltung gesetzlicher Vorschriften

Ein weiterer wichtiger Grund, warum bestimmte Verarbeitungsschritte im eigenen Rechenzentrum erfolgen sollten, ist die Einhaltung von Datenschutzbestimmungen. In vielen Fällen enthalten Daten, die von KI und ML-Algorithmen verarbeitet werden sollen, persönliche Informationen, die entfernt werden müssen, bevor die Datenanalyse beginnen kann. Beispielweise in Telekommunikationsunternehmen: Diese verfügen über eine enorme Menge an wertvollen Daten, die sie nutzen können, um tiefe Einblicke in das Verhalten ihrer Kunden zu gewinnen, um ihre Dienste und ihre Netzabdeckung zu optimieren. Gemäß Datenschutzbestimmungen wie der DSGVO müssen sie jedoch zunächst alle persönlich identifizierbaren Informationen aus diesen Daten entfernen und sie anonymisieren. Dies kann am besten im eigenen Rechenzentrum erledigt werden.

Der Hauptgrund ist dabei nicht einmal, dass Cloud-Anbieter selbst keine Dienste zur Anonymisierung von Daten anbieten. Er besteht vielmehr darin, dass die korrekte Konfiguration der Dienste eine sehr komplexe Aufgabe ist und sich – da sich die Cloud-Konfiguration ständig ändert – zu jedem Zeitpunkt Fahler einschleichen können. Da die Nichteinhaltung von Datenschutzrichtlinien zu empfindlichen Strafen führen kann, ist es daher am besten, alle Daten zu anonymisieren, bevor sie in die Cloud gesendet werden.

Grundsätzlich gilt: Die großen Cloud-Anbieter sind US-Unternehmen, die unter US-Gesetze fallen. Das bedeutet, dass sie im Ernstfall die Einhaltung der europäischen Datenschutzbestimmungen nicht garantieren können, zumal der EuGH das EU-US-Privacy-Shield-Abkommen für ungültig erklärt hat. Man könnte meinen, dass ein Cloud-Anbieter aus den von ihm anonymisierten Daten keine personalisierten Informationen gewinnen kann. Doch um diese Daten zu anonymisieren, benötigt er zunächst Zugriff auf die Rohdaten. Unternehmen sollten also auf Nummer sicher gehen und ihre Daten selbst anonymisieren, bevor sie diese in die Cloud schicken.

Sowohl Kosteneffizienz als auch Compliance sprechen also für eine hybride Infrastruktur, um von Künstlicher Intelligenz und Maschinellem Lernen profitieren zu können. Der private Teil der Infrastruktur wird dabei idealerweise die Form einer Private Cloud annehmen, die in Bezug auf Flexibilität, Verbrauchsmodell und Zuverlässigkeit weitgehend die gleichen Vorteile wie auch die Dienste der Public-Cloud-Provider bietet.

Hans Hallitzky, Sales Manager DACH bei Infinidat