

Deep Learning für KI erweitert Data-Science-Horizonte

DatenwissenschaftlerInnen der Stanford University erforschen neue Anwendungsgebiete für Deep-Learning-Modelle mit extremer Parallelverarbeitung.



Kundenprofil

Führende Forschungsuniversität

Informationstechnologie | USA



„Wir sind aus einem ganz neuen Blickwinkel an unsere KI-Ansätze herangegangen und haben überlegt, wie wir die Technologien von Dell Precision und NVIDIA dafür einsetzen können.“

Jure Leskovec

Associate Professor of Computer Science,
Stanford University

Anforderungen des Unternehmens

Um Muster in komplexen, unstrukturierten Multi-Terabyte-Datenvolumen zu erkennen und Erkenntnisse daraus zu gewinnen, ist eine enorme Parallelverarbeitungsgeschwindigkeit erforderlich. Die DatenwissenschaftlerInnen der Stanford University verwendeten zunächst handelsübliche Gaming-GPUs (Graphics Processing Units). Sie benötigten aber eine höhere GPU-beschleunigte Verarbeitungsgeschwindigkeit, die von CPUs (Central Processing Units) unterstützt wird, um die GPU-Auslastung bei der Ausführung von KI-Modellen für Deep Learning zu maximieren.

Unternehmensergebnisse

- Gibt ForscherInnen die Möglichkeit, umfangreichere und größere Probleme zu lösen
- Ermöglicht zuvor unmögliche Inferenzen aus Daten
- Verkürzt die Entwicklung von KI-Algorithmen sowie ML-Trainingszeiten um bis zu 50 %
- Bietet die Leistung eines Supercomputers in einem mobilen, überall nutzbaren Laptop
- Verarbeitet Datenvolumen von bis zu 1,5 TB

Die Lösungen im Überblick

- [Dell Precision 7750 Mobile Data Science Workstation](#)
- [Dell Precision 7920 Tower Data Science Workstation](#)
- [NVIDIA Quadro RTX-basierte Workstations](#)

Data Science ist die Disziplin, die eine der am schnellsten wachsenden – und transformativen – technologischen Entwicklungen der Welt vorantreibt: die künstliche Intelligenz (KI). Die historischen Wurzeln der KI reichen bis in die 1950er Jahre zurück, aber selbst vor ein paar Jahren waren die derzeitigen Möglichkeiten der KI noch unvorstellbar: Heute setzen DatenwissenschaftlerInnen immer mehr Leistung für extreme Parallelverarbeitungsgeschwindigkeit bei ihren Datenmethoden und Lernmodellen ein.

Genau das ist auch an der Stanford University der Fall. Dort leiten Jure Leskovec, Associate Professor of Computer Science, und Rok Sosič, Senior Researcher, ein Team von 40 DatenwissenschaftlerInnen. Sowohl individuell als auch in Teams untersuchen Leskovec und Sosič KI-Anwendungen in Form von Netzwerkanalysen und Deep Learning für eine Vielzahl von Problemen. Die Netzwerkanalysen beziehen sich dabei nicht auf IT-Netzwerke, sondern meinen die Suche nach komplexen Mustern zwischen den Beziehungen von mehreren Datenentitäten.

„Unsere Aufgabe besteht darin, bahnbrechende Forschung zu betreiben, um zahlreiche ausgesprochen anspruchsvolle Forschungsprobleme zu bewältigen. Wir entwickeln neue Methoden, um aus Daten zu lernen. Dann nutzen wir diese Methoden, um neue Erkenntnisse aus diesen Daten zu gewinnen, die wir auf praktische Probleme anwenden können“, sagt Leskovec und erklärt, dass das Stanford Lab weltweit führend bei der Entwicklung von KI-Modellen für Life-Sciences-Anwendungen ist.

Deep Learning offenbart komplexe Netzwerkeffekte

Zur Veranschaulichung von KI-basierten Netzwerkanalysen vergleicht Sosič die Forschungen seines Teams mit den Nebenwirkungen der Polypharmazie und erläutert: „Wenn eine Person mehrere verordnete Medikamente einnimmt, können die Wechselwirkungen sowie deren Dynamik je nach Krankheit(en), Genetik, Geschlecht, Alter usw. variieren.“

Sosič erklärt: „Traditionell werden in sehr kostenintensiven klinischen Studien große Patientengruppen untersucht, um daraus Folgerungen und Rückschlüsse zur Wirksamkeit von Medikamenten zu ziehen. Solche Versuche sind jedoch mit PatientInnen, die mehrere Medikamente einnehmen, nicht machbar, da es zu viele mögliche Kombinationen gibt. Daher entwickeln wir eine Deep-Learning-Methodik, um die komplexen Netzwerke dieser Wechselwirkungen – also Interaktionen – zusammen mit den verschiedenen Merkmalen der Proteine zu analysieren, auf die diese Medikamente reagieren. Dieser Ansatz bietet dem Pharmabereich mehr Einblicke in die Funktionsweise ihrer Produkte und vor allem darüber, wie sie potenziell besser wirken können.“



”

„Dank der Dell Precision 7920 Tower Workstation mit den NVIDIA-GPUs können wir jetzt viel größere und komplexere Probleme in Datenvolumen mit einer Größe von bis zu 1,5 TB untersuchen.“

Rok Sosič

Senior Researcher,
Stanford University



Die Workstation ist einsatzbereit, **was das Setup vereinfacht** und spart Zeit.

Für diese Forschung, in der meist Datenvolumen im Multi-Terabyte-Bereich analysiert werden, nutzen die Lab-Teams eine ganze Reihe von Servern, die zuvor mit handelsüblichen GPUs ausgestattet wurden. Ursprünglich wurden diese GPUs für die Renderinganforderungen von Videospiele entwickelt, um die CPU im Gaming-PC zu entlasten.

„Wir verbauten so viele dieser GPU-Karten wie möglich in unseren Servern, aber wir hatten nicht genug RAM pro GPU. Somit wurde die begrenzte Verarbeitungsgeschwindigkeit zur einschränkenden Variable für den Umfang unserer Forschung“, führt Sosič aus.

Erweiterung des Forschungshorizonts durch extreme Parallelverarbeitung

Um dem Team einen Quantensprung bei der verfügbaren Parallelverarbeitungsgeschwindigkeit zu ermöglichen, hat das Stanford Data Lab eine Dell Precision 7920 Tower Data Science Workstation (DSW) erworben. Die Dell Precision 7920 Tower DSW wird in Zusammenarbeit mit NVIDIA und anderen führenden Technologieanbietern wie Canonical (das Unternehmen hinter Ubuntu, dem weltweiten De-facto-Linux-Betriebssystem für KI auf Workstations) als vollständig integriertes, einsatzbereites KI-Hardware- und -Softwarepaket ausgeliefert. Dieses Paket vereinfacht das Setup für NutzerInnen: Mit dem Bare-Metal- und Build-it-yourself-Ansatz lassen sich Tage einsparen, ganz zu schweigen von den Konfigurationen, die ansonsten erforderlich wären.

Mit der Kühlung, dem Bus-Design und dem Engineering kann die Dell Precision 7920 Tower DSW mit bis zu drei NVIDIA Quadro RTX 8000-GPUs ausgestattet werden, davon jede mit 48 GB Speicher und 4.608 Prozessor-Cores für die Parallelverarbeitung. Dank der Kooperation zwischen den Entwicklungsteams von Dell Technologies und NVIDIA konnte die CPU-Performance der DSW – verfügbar in zwei Prozessoren und mit bis zu 2 TB RAM – auf die GPUs ausgerichtet werden, um die Auslastung im gesamten Trainingsdurchlauf eines KI-Modells zu maximieren.

Das Lab-Team setzt die Dell Precision 7920 Tower DSW mit NVIDIA RTX 8000-GPUs als gemeinsame Ressource für die Entwicklung von KI-Algorithmen und das maschinelle Lernen ein. „Das war für uns ein Riesensprung nach vorne“, sagt Leskovec. „Wir sind aus einem ganz neuen Blickwinkel an unsere KI-Ansätze herangegangen und haben überlegt, wie wir die Technologien von Dell Precision und NVIDIA dafür einsetzen können.“

Sosič fügt hinzu: „Dank der Dell Precision 7920 Tower Workstation mit den NVIDIA-GPUs können wir jetzt viel größere und komplexere Probleme in Datenvolumen mit einer Größe von bis zu 1,5 TB untersuchen.“



„Wir haben einige Benchmarks für unsere Server mit den handelsüblichen GPUs durchgeführt und festgestellt, dass die Dell Precision 7750 DSW mit den NVIDIA-GPUs um 50 % schneller ist.“

Rok Sosič

Senior Researcher,
Stanford University

”

„Die Nutzung einer eigenen Dell Precision DSW mit NVIDIA RTX-GPUs ist, wie einen Ferrari zu fahren. Unsere anderen Lab-Server schneiden im Vergleich quasi wie gemeinsam genutzte LKWs ab.“

Rok Sosič

Senior Researcher,
Stanford University

Supercomputing für unterwegs passt in einen Rucksack

Das Stanford University Stanford Data Lab hat zudem eine Dell Precision 7750 Mobile DSW erworben: ein 15"-Laptop mit 8-Core-CPU, 64 GB RAM und einer NVIDIA Quadro RTX 5000-GPU mit 16 GB RAM. Ebenso wie die Dell Precision Tower DSW ist der Laptop ein vollständig integriertes, einsatzbereites KI-Gerät mit speziellem Kühlsystem, das die erzeugte Wärme ableitet, wenn CPU und GPU gemeinsam für ein KI-Modell genutzt werden.

Sosič und Leskovec betonen, dass die Dell Precision 7750 Mobile DSW im Vergleich zu ihren Servern besonders schnell ist. „Wir haben einige Benchmarks für unsere Server mit den handelsüblichen GPUs durchgeführt und festgestellt, dass die Dell Precision 7750 DSW mit den NVIDIA-GPUs um 50 % schneller ist“, sagt Sosič.

Laut Sosič ist die Dell Precision 7750 Mobile DSW für die DatenwissenschaftlerInnen des Teams gedacht, damit diese ihre KI-Modelle auch zu Hause oder an einem anderen Ort entwickeln können. „Diese Mobilität kommt unseren ForscherInnen im Homeoffice zugute. Sie können ihre Modelle mit der Dell Precision 7750 entwickeln und sie dann für das anschließende Training in unser Rechenzentrum hochladen“, erklärt er.

Laut Sosič bietet die vollständig integrierte, KI-fähige Technologie der Dell Precision DSW in Kombination mit den leistungsstarken NVIDIA RTX-GPUs dem Stanford Data Lab die benötigte Parallelverarbeitung, um die bahnbrechende Forschung fortsetzen zu können. „Wir können nun Dateninferenzen ableiten, zu denen wir bisher nicht in der Lage waren“, so Sosič. „Die Nutzung einer eigenen Dell Precision DSW mit NVIDIA RTX-GPUs ist, wie einen Ferrari zu fahren. Unsere anderen Lab-Server schneiden im Vergleich quasi wie gemeinsam genutzte LKWs ab.“

