

De impact van DeepSeek op de toekomst van AI



DeepSeek heeft een transformatieve benadering van AI-ontwikkeling geïntroduceerd. DeepSeek is niet zomaar een AI-bedrijf; het is een paradigmaverschuiving. Het team erachter bestaat uit engineers die niet gebonden zijn aan traditionele AI-methodologieën. Ze zijn wiskundigen in het hart, pragmatisten in uitvoering en niet bezorgd over hoe AI in het verleden is gedaan. Ze zagen een nieuwe weg vooruit, een die opnieuw definieert hoe we denken over AI-modellen, datamanagement en uiteindelijk de infrastructuur die het allemaal aandrijft.

De echte impact van DeepSeek zit niet alleen in de modellen die ze hebben gebouwd, maar ook in de roadmap die ze hebben onthuld voor de toekomst van AI-infrastructuur. Dit is hoe die toekomst eruit ziet en hoe het ondernemingen en

datamanagement zal beïnvloeden. Ik ben er vrij zeker van dat veel slimme mensen de aard van hun aanpak van AI zelf al hebben heroverwogen.

De opkomst van kleinere, slimmere modellen

De AI-industrie is al lang voorbereid op het idee van fundamentele modellen – massale, alom bekende netwerken die op alles zijn getraind. Maar DeepSeek heeft bewezen dat kleinere, meer gespecialiseerde modellen niet alleen levensvatbaar zijn, maar ook op vele manieren superieur. Hun aanpak? Een model met een mix van experts, waarbij kleinere, hoogopgeleide modellen samenwerken. Hoewel dit de aanpak te eenvoudig maakt, gebruiken ze een zeer slimme techniek om het juiste expertmodel te selecteren en zorgen ze er tegelijkertijd voor dat ze een efficiënt systeem creëren.

Dit verandert alles.

In plaats van één reusachtig model dat alles probeert te doen, kunnen ondernemingen een systeem van onderling verbonden modellen inzetten, elk gespecialiseerd in een specifiek domein. De implicaties zijn diepgaand:

- **Lagere kosten:** Kleinere modellen vereisen aanzienlijk minder rekenkracht, maar het echte voordeel is niet alleen lagere kosten. De aanpak heeft nu bewezen dat een meer capabele AI mogelijk is en tegelijkertijd de trainings- en gevolgcosten verlaagt.
- **Gemakkelijkere validatie:** Gerichte expertise betekent dat het eenvoudiger is om prestaties in real-world applicaties te testen en te verifiëren. Ik geloof dat dit het eenvoudig zal maken om meer gespecialiseerde modelmogelijkheden toe te voegen zonder zoveel moeite als het bouwen van een basismodel.
- **Redeneringsmodellen:** Kleine modellen krijgen veel gemakkelijker redeneringsmogelijkheden. Denk erover na: Het toevoegen van redenering is een belangrijke stap in de richting van AI-toezicht en transparantie.

Voor ondernemingen betekent dit dat AI-adoptie verder gaat dan chatbots. Het was nooit op de kaart dat een gewone onderneming een fundamenteel model zou bouwen; het is gewoon te duur. Dit biedt de blauwdruk voor het eenvoudig en kosteneffectief bouwen van zeer capabele modellen met hun eigen data. Zie het als het onboarden van een AI-personeel – de belemmering om zeer efficiënte en domeinspecifieke modellen te creëren is aanzienlijk gedaald. De tools hiervoor komen op de volgende plaats, evenals basismodellen die een efficiënte destillatie van data mogelijk maken om capabele kleinere modellen te bouwen.

De sleutel tot het ontsluiten van het volledige potentieel van AI is datamanagement

Al jaren is de AI-industrie geobsedeerd door het verzamelen en hamsteren van zoveel mogelijk data. De focus lag op het maximaliseren van het aantal tokens, het trainen op alles en het hopen dat brute kracht betere resultaten zou opleveren. Maar in de wereld van modellen met een mix van experts staat datamanagement centraal.

De uitdaging verschuift van het verbazen van data naar het selecteren, organiseren en verfijnen ervan. De prioriteiten zijn nu:

- **Het samenstellen van data van hoge kwaliteit:** AI is slechts zo goed als de data waarop het is getraind. Het filteren van lawaai en het focussen op relevantie is nu een strategisch voordeel.
- **Optimaliseren van datapipelines:** continue dataverrijking, versiebeheer en traceerbaarheid worden cruciaal.
- **Infrastructuur die AI ondersteunt:** Ondernemingen hebben een infrastructuur nodig die data dynamisch kan indexeren, filteren en ophalen op schaal.

DeepSeek heeft deze aanpak al gedemonstreerd met hun “met zorg ontworpen dataselectiepijplijn”. In plaats van willekeurig te trainen in alle beschikbare data,

filterden en verfijnden ze hun datasets, waardoor de efficiëntie werd verbeterd en tegelijkertijd de kosten werden verlaagd. AI-gestuurde intelligente dataselectie is de hoeksteen van toekomstige AI-training en zorgt voor efficiëntie en precisie in modelontwikkeling.

De infrastructuur van morgen

Naarmate AI overgaat op gespecialiseerde modellen en dataverfijning, moet de infrastructuur evolueren om deze nieuwe realiteit te ondersteunen. Toekomstige AI-systemen vereisen:

- **Multidimensionale prestaties:** Ondersteuning van duizenden kleinere modellen die in parallele en key-value stores werken die efficiënt draaien op inferentietijd.
- **Hoge connectiviteit en always-on beschikbaarheid** y: AI-gestuurde datapipelines zullen datasets voortdurend bijwerken en verfijnen.
- **Schaalbare, snelle opslag:** De mogelijkheid om in realtime aanzienlijke hoeveelheden data te indexeren, op te halen, te filteren en weer te geven.

In alle sectoren worden organisaties geconfronteerd met soortgelijke uitdagingen bij het beheer van AI-gestuurde datasets. AI-modellen vertrouwen op continu evoluerende datasets die exponentieel groeien. In plaats van alleen data op te slaan, indexeren en verfijnen toonaangevende AI-gedreven organisaties deze dynamisch. Naarmate modellen verbeteren, worden data opnieuw verwerkt om nieuwe inzichten te ontdekken en AI-systemen dienovereenkomstig opnieuw op te leiden. Deze continue cyclus van dataverfijning is de sleutel tot het echt effectief maken van AI.

Een nieuw tijdperk voor AI

DeepSeek heeft een fundamentele waarheid blootgelegd: De toekomst van AI gaat niet over grotere modellen – het gaat om het beheren en verwerken van

data. De aanpak van het schalen van AI is veranderd en we gaan nu een tijdperk in waarin elk model ongelooflijk capabel kan worden zonder de enorme kapitaalinvestering. Redenering is nu beschikbaar voor elk model. Infrastructuur, opslag en datamanagement zullen nu ook veranderen.

Het belangrijkste leerpunt? Het snijpunt van AI, datawetenschap en datamanagement is waar de echte innovatie plaatsvindt. Opslag gaat niet alleen meer over het bewaren van data; het gaat om raffinage, indexering, filteren, organiseren en hergebruik, terwijl het voortdurend wordt geoptimaliseerd voor de volgende evolutie van AI.

AI-platform van de toekomst

DeepSeek heeft aangetoond wat mogelijk is wanneer AI-ontwikkeling een nieuw paradigma omarmt. De uitdaging is nu voor ondernemingen en infrastructuuraanbieders om de kunst van het mogelijke te verkennen – het opnieuw vormgeven van systemen, workflows en architecturen die het potentieel van AI-gestuurde inzichten en besluitvorming volledig benutten.

Pure Storage is uniek gepositioneerd om deze AI-gedreven toekomst te ondersteunen. Met ons platform kunnen ondernemingen de kracht van AI benutten en tegelijkertijd zorgen voor eenvoud, betrouwbaarheid, flexibiliteit en efficiëntie in hun datamanagement. Onze uniforme infrastructuur vereenvoudigt dataopslag in alle omgevingen. Naarmate AI blijft evolueren, hebben bedrijven een platform nodig dat meegroeit – een platform dat complexiteit elimineert, de efficiëntie verbetert en moeiteloos schaalbaar is.

Laten we samen opnieuw definiëren wat er mogelijk is – maak vandaag nog verbinding met [Pure Storage](#) om te ontdekken hoe we u kunnen helpen nieuwe inzichten te ontsluiten en innovatie te stimuleren met uw data.