Recursion: Reengineering drug discovery

Founded in 2013, Recursion is an industry leader in applying AI to drug discovery and has created an innovative culture to bring together data science, engineering, and drug discovery teams.

RECURSION

"With AI at the forefront of our vision for biopharmaceuticals, Recursion has generated a massive 4 petabyte dataset of high-resolution biological images used to train deep learning models that support the discovery of new treatments. Part of this process involves feeding images into deep learning models to extract relevant information, identify high-dimensional disease signatures, and subsequently discover drugs that reverse the disease signature back to a 'healthy' control population."

Berton Earnshaw, ML Fellow

Solution

Recursion uses Determined, an open source DL training platform, to manage 100s of heterogeneous GPUs in their on-prem cluster, and to dynamically scale their Google Cloud Platform (GCP) footprint. Determined enables faster and more efficient training for deep learning models while improving team collaboration dramatically.

Results

14X
Increase Hyperparameter Search

Hyperparameter Search using the Determined Training Platform found a model configuration in 24 hours, which achieved similar performance metrics, as that of 2 weeks of hand tuning.

from 3 Days to 3 Hours
Model Training

Seamlessly scaled training to 64 NVIDIA V100 GPUs on GCP with no code changes and achieved similar training times with 120 NVIDIA Pascal GPUs in their on-prem.

70%
Decrease in GCP Cost

Easily switch between on-prem and cloud GPUs for long-running and large bursty training jobs, respectively. With built-in support for preemptible instances on GCP, teams can further realize significant cost savings.

Recursion is at the forefront of technological innovation using AI-powered drug discovery.. With over 4PB of high-resolution biological images, Recursion has built up a significant advantage to develop and deploy AI for drug discovery. For example, Recursion previously tested thousands of drugs in a cellular model of neurofibromatosis type 2 (NF2), a rare tumor syndrome, and identified a small molecule REC-2282 that rescued the high-dimensional morphology of *NF2*-deficient cells without inducing additional morphological changes. Recursion is rapidly advancing REC-2282 through clinical development through its subsidiary CereXis, Inc.

How Recursion Leverages the Determined Training Platform

Recursion uses Determined, an open-source Deep Learning training platform to manage hundreds of GPUs: a heterogeneous on-prem cluster with NVIDIA Pascal & Volta V100 GPUs, and NVIDIA Volta V100 instances on GCP. With this infrastructure, they train more models faster, enabling their drug discovery teams to collaborate and iterate quickly by trying out new ideas multiple times a day.

"The setup time for HP search took us less than an hour – a massive time saving for our team."

Recursion's data science team used Determined to execute an automated hyperparameter search. In less than 24 hours, they found a model configuration that matched the performance of their previous best-performing configuation – which took them more than 2 weeks of manual tuning to produce!

Using Determined's natively integrated distributed training, Recursion was able to scale data-parallel training to 64 V100 GPUs on GCP and reduce training time from ~3 days to 3 hours! This scaling carried over to Recursion's on-prem cluster as well, achieving the same training time with 120 NVIDIA Pascal GPUs.

"Our ML team can try out new ideas multiple times a day, instead of thinking which jobs they want to setup and waiting for days to see if a new loss function or other model change has the desired impact."

Most notably, this 24x training performance improvement required [minimal code changes](#) and no additional distributed training infrastructure — everything came native by having the Determined Training Platform manage the GPU cluster. In addition to significantly reduced training time, Recursion also reduced the time their platform engineers and data scientists spent setting up, managing, and configuring their deep learning infrastructure and experiments, respectively.

With Determined running both on-premise and on GCP, Recursion users can easily switch between the two, depending on whether they need persistent GPUs for long-running experiments on-prem, or to acquire GPUs for large bursty workloads dynamically on GCP. And with support for preemptible instances on GCP, Recursion can now look to run the same experiments at a 70% cost reduction.

Determined's Value to Recursion

By using Determined, Recursion saw value in several different ways:

1. **Easy to use Distributed Training that *just works***
    There are several distributed training libraries, but they typically require overhauling your model code and configuring additional software infrastructure. With Determined, users can benefit from state-of-the-art distributed training speeds (24x baseline in Recursion's case), with minimal additional work. Once a model runs on your Determined cluster, you can enable distributed training and a host of other functionality with a simple configuration change.

2. **A platform that lets you focus on DL, not DevOps**
    Determined abstracts away the infrastructure from the machine learning work your team performs, thus allowing you to devote more time to experimentation and less time to managing your GPU cluster. Track experiment metadata automatically, centralize & streamline workflows by natively integrating with applications like Jupyter Notebook and TensorBoard.

3. **Flexibility across cloud and on-prem clusters**
    Deep learning is an evolving space, and infrastructure decisions can quickly change. Determined gives you the flexibility to seamlessly run deep learning workloads both on-prem and in the cloud, so you can easily adapt your DL infrastructure as your needs change.

4. **World-class support from a team of experts**
    As a company, Determined's expertise resides at the intersection of machine learning and distributed systems. We strive to support our customers with the knowledge we continue to develop in this space. Our engineers work directly with customers to advise and address challenges, both theoretical and practical, in scaling deep learning training for real-world models.

About Determined AI

Determined AI (**determined.ai**) empowers deep learning engineers to be highly productive, enabling them to focus on building high-quality models in less time by simplifying the underlying infrastructure. Determined Training Platform, open-sourced under the Apache license, reduces time-to-market for AI applications with fast, collaborative, and best-in-class software infrastructure optimized GPU-powered on-prem or in the cloud. The Determined AI team is comprised of ML and distributed system experts, including key contributors to Spark MLlib, Apache Mesos, and PostgreSQL; PhDs from UC Berkeley, Carnegie Mellon University, and New York University; engineers from top companies (including AWS, Google, Cloudera, and Mesosphere); and faculty at Carnegie Mellon University.