



Accelerate your deep learning model development

Automated DL platform to unlock DL engineers' creativity

Determined AI is an **automated deep learning (DL) platform** that improves the productivity of DL engineers exponentially. With Determined AI, DL teams can rapidly build production-ready models and readily deploy them on the designated hardware, even on a device limited in resources such as an edge device.

Experiment 5
mnist_tf_adaptive

Add label Archive Fork

Configuration View

Model Definition Download

State: **COMPLETED**

Progress

Max Slots: Unlimited

Start Time: 2019/01/17, 4:57:17 PM

End Time: 2019/01/17, 4:59:42 PM

Metrics

Active Trials

ID ↓	H-Params	Start Time ↓	Steps	Latest Validation Metric ↓	Best Validation Metric ↓
287	View	2019/01/17, 4:57:17 PM	4	0.02578	Log Details Continue Training
288	View	2019/01/17, 4:57:17 PM	4	0.02816	Log Details Continue Training
289	View	2019/01/17, 4:57:17 PM	4	0.02647	Log Details Continue Training
290	View	2019/01/17, 4:57:17 PM	1	0.18690	Log Details Continue Training
291	View	2019/01/17, 4:57:21 PM	1	0.12679	Log Details Continue Training

Completed Trials

ID ↓	H-Params	Start Time ↓	End Time ↓	Steps ↓	Latest Validation Metric ↓	Best Validation Metric ↓
287	View	2019/01/17, 4:57:17 PM	2019/01/17, 4:59:40 PM	4	0.02578	Log Details Continue Training
288	View	2019/01/17, 4:57:17 PM	2019/01/17, 4:59:40 PM	4	0.02816	Log Details Continue Training
289	View	2019/01/17, 4:57:17 PM	2019/01/17, 4:59:40 PM	4	0.02647	Log Details Continue Training
290	View	2019/01/17, 4:57:17 PM	2019/01/17, 4:59:27 PM	1	0.18690	Log Details Continue Training
291	View	2019/01/17, 4:57:21 PM	2019/01/17, 4:59:27 PM	1	0.12679	Log Details Continue Training

Organizations using Determined AI have reported significant increases in both team and individual productivity, with some customers stating they save a minimum of one day per engineer per week. Additional benefits include:

Deliver quality models in less time

With Determined AI, DL engineers can readily take advantage of *distributed and/or parallel training* on their existing hardware infrastructure (e.g., a GPU cluster) with a few lines of code. Our specialized *hyperparameter search* optimizes the search space to get to quality models faster. Experiments are executed in containerized environments and all experimental meta-data is tracked. Collectively, these features enable seamless *reproducibility* and *warm starting* of future experiments to expedite subsequent results.

Full utilization of your GPU resources

Determined AI's industry-leading job scheduling algorithm enables DL engineers to maximize the number of experiments that can be run simultaneously. Our technology has the ability to resize jobs dynamically as new jobs come in or hardware goes on/offline. DL engineers can also share their compute resources anywhere according to their jobs' priorities. Your GPU resources will be utilized in an optimal way to accommodate all of your DL experiments.

Deep learning specialized automation

The Determined AI team has intimate knowledge about the deep learning problem space gained through research at **UC Berkeley** and **Carnegie Mellon University**. Our DL optimization technologies, including *hyperparameter search*, *DL job mapping and scheduling*, and *automatic architecture design that satisfies deployment SLAs*, are much more effective than generic solutions.

Seamless integration with your existing software and hardware ecosystems

Determined AI can be run on premises and/or in the cloud. It is compatible with:

Deep learning frameworks:

TensorFlow, Keras, PyTorch

Data storage and management systems:

NFS, Amazon S3, Hadoop, GlusterFS, and more

Container Orchestration Frameworks:

Kubernetes, DC/OS, bare metal

Key features of Determined AI for deep learning

Determined AI takes a pragmatic approach to the deep learning problem space to dramatically increase the productivity of deep learning engineers. Its key features include:

AutoML at scale

Deliver quality models in less time

- ✓ State-of-the-art hyperparameter and neural architecture search
- ✓ Push-button transfer learning and retraining
- ✓ Automatic data-parallel distributed training
- ✓ Data access accelerator
- ✓ TensorFlow, Keras, or PyTorch support

Experiment tracking

Enable reproducibility and collaboration

- ✓ Integrated metrics capture and model version management
- ✓ Reproducible, containerized training
- ✓ Label and share experiments
- ✓ Multi-tenancy in a shared infrastructure
- ✓ Real-time experiment metrics visualization

Seamless infrastructure

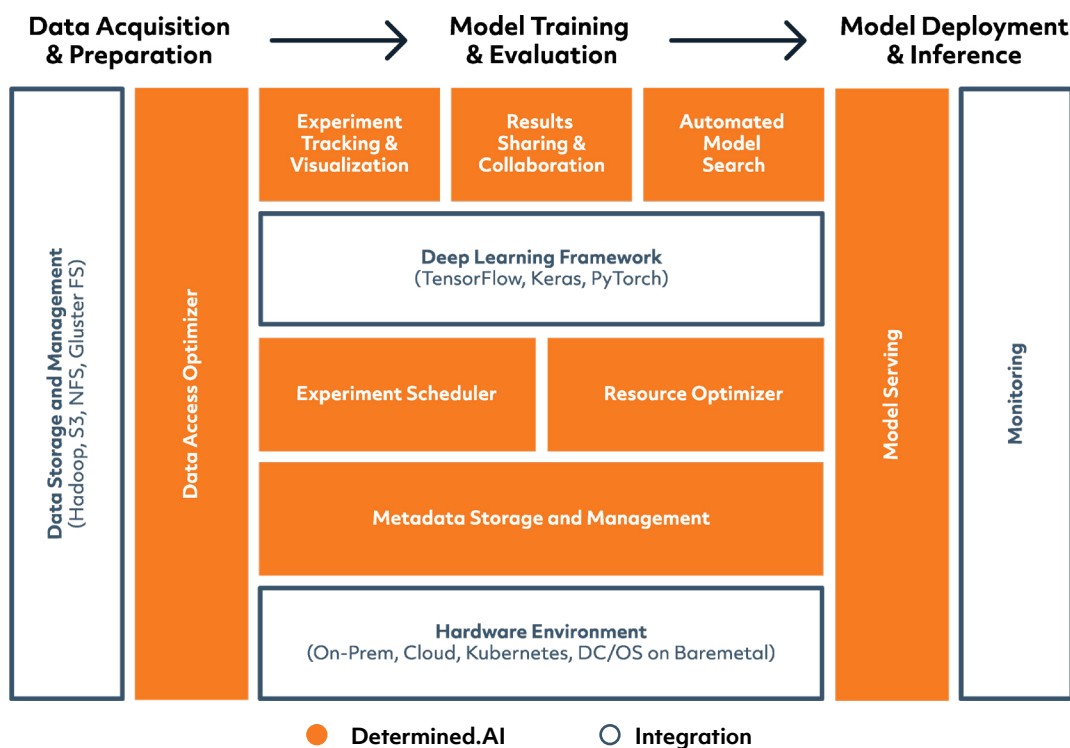
Optimize resource utilization and ensure experiment reliability

- ✓ Fault-tolerant training
- ✓ Efficient scheduling and execution of experiments
- ✓ Share GPU resources on premises, in the cloud, or both.
- ✓ Run on bare metal, and in Kubernetes

Model deployment optimization

Ensure model deployability and reduce time-to-market

- ✓ Explore models with deployment constraints
- ✓ Simulate inference speed on deployment hardware
- ✓ Automated architecture search for the edge, cloud, and mobile
- ✓ Containerized deployment in a single click



About Determined AI

Determined AI (www.determined.ai) empowers deep learning developers to produce quality models in less time using state-of-the-art optimization technologies. Our AutoML platform reduces time-to-market by increasing developer productivity, improving resource (GPU) utilization, and reducing risk. The Determined AI engineering team is comprised of machine learning and distributed systems experts, including key contributors to Spark MLlib, Apache Mesos, and PostgreSQL; PhDs from UC Berkeley, Carnegie Mellon University, New York University, University of Chicago, and UT Austin; top engineers from leading tech companies (Google, Mesosphere, Pinterest); and faculty at Carnegie Mellon University.