

Numenta Brings Brain-Based Principles to AI with SigOpt

SigOpt, the leading intelligent experimentation platform, empowers Numenta to build more robust and generalizable neural networks much faster and better than previously possible.

Translating lessons from neuroscience into more robust machine intelligence

Numenta is using their deep theoretical neuroscience research to advance the state of artificial intelligence (AI) and machine learning (ML). By applying key components of their cortical theory, they are enabling the development of a more robust and universally relevant machine intelligence architecture. Given there are limits to infrastructure scalability and efficiency that can no longer be solved solely by adding more power and data, Numenta designed an experiment to develop a systems architecture with an AI training data set where the network would synthesize data and make it generalizable with total accuracy at scale by employing sparsity, a neuroscientific concept. This approach requires significant experimentation, which was enabled by the SigOpt Intelligent Experimentation Platform. SigOpt gave Numenta the ability to design experiments by asking the right questions as well as to explore their modeling problem with significant depth and optimize their model to develop a novel sparse architecture.



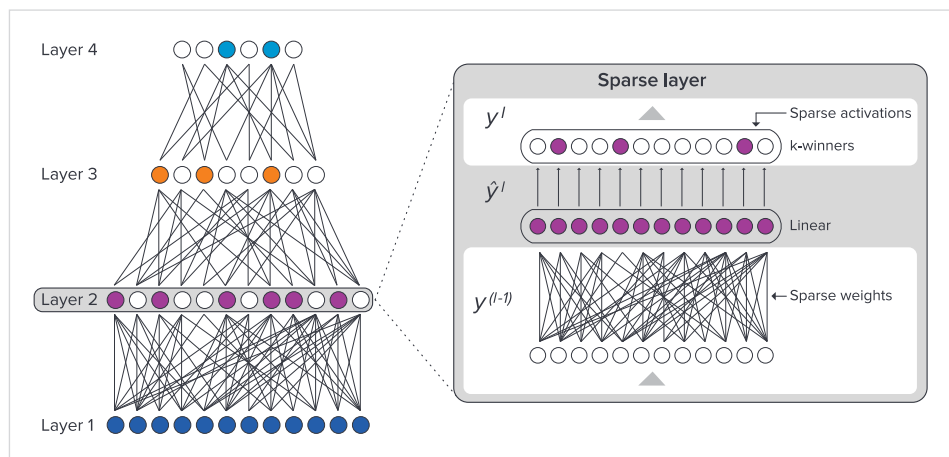
Building and training new and unconventional architecture to run at scale using existing infrastructure for results that meet industry standards

Sparsity describes the way the human brain efficiently represents information, enabling it to store and process information quickly. At any given time, only about 2 to 3 percent of neurons are actively sending signals. Additionally, the connectivity between neurons is sparse. Through the following experiment, Numenta sought to emulate the brain’s qualities of activation sparsity (limiting the number of simultaneously active neurons) and weight sparsity (limiting the interconnectedness of neurons) as well as demonstrate how sparsity enables efficiency in resource utilization, generalization, and robustness.

Applying intelligent experimentation to develop and optimize a novel neural network architecture

Developing and understanding how a model works requires experimentation, which is a nonlinear scientific process that can be messy and tough to get right. Experimentation is time and resource heavy without any guarantees for success. However, when best practice experimentation techniques are combined with best-in-class tools, this process can become more manageable, scalable, and productive.

Numenta faced this experimentation challenge when developing a novel version of the ResNet-50 architecture. In particular, they wanted to develop a sparse version of ResNet-50 that also maintained accuracy against the benchmark ImageNet data set. A good result would demonstrate the potential for sparsity to boost efficiency and capacity without deteriorating in-sample performance. The goal of this project was to arrive at a trained ImageNet using ResNet-50 with at least 75 percent sparsity and with sufficiently high accuracy to be acceptable for ML performance industry standards.



Standard dense layer vs. sparse layer. Numenta’s sparse layer contains both sparse weights and sparse activations.

Numenta developed this novel architecture with PyTorch for their modeling framework; Ray, for compute orchestration, and SigOpt, for experimentation. Numenta took ResNet-50, pretrained on ImageNet, and attempted to create sparse versions of the neural network. Throughout the process, the team encountered a variety of challenges, including:

- 1 Long training runs:** When developing a model with a data set that takes a long time to train, it is important to shorten run cycles to learn faster and conserve financial resources. The Numenta team needed to have control over the training and tuning process for sample efficiency and compute time in approaching model development.
- 2 Model interpretability and understanding:** Neural networks are black boxes. The Numenta team needed tools and techniques to understand the drivers of model performance.
- 3 Automation:** Time is at a premium when running experimentation. Accordingly, the Numenta team needed a way to effectively delegate replicable tasks for faster modeling and to focus on the priorities related to being a domain expert.
- 4 Optimization:** Training a neural network on a data set is the easy part. The challenge comes in conducting a scientific experiment that yields hyperparameters that respond in a performant manner for both in- and out-of-sample data. The Numenta team needed a simple way to implement a variety of experiments to understand model architecture behavior with sensitivity to parameters and performance across a variety of metrics.
- 5 Implementation:** Tools need to fit into existing workflows and save much more time than they cost to implement. The Numenta team wanted the right modeling tool to cost-effectively run more iterations for exploring the modeling space, which, in turn, could render more insights on the modeling problem.

For Numenta to be successful, the team needed to implement a scalable, efficient, and effective tool for experimentation and hyperparameter optimization. Altogether, the Numenta team tried a variety of solutions – including existing combinations from existing papers on dense networks – to address these challenges. None offered the right combination of resources nor delivered the desired results. They then tried leveraging the sophistication of their engineering team by testing and customizing different open source Bayesian optimization projects and prominent tools for scheduling optimization jobs. Ultimately, the Numenta team learned the time they spent implementing and optimizing the systems was greater than the value these options provided.

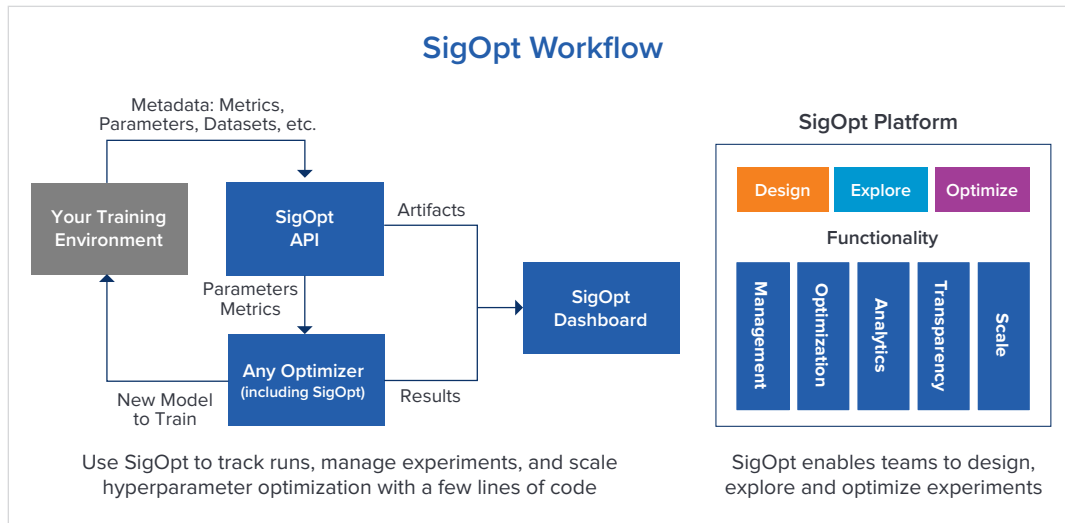
The Numenta team also considered building their own full-fledged internal solution. However, they knew they needed a best-in-class approach to hyperparameter optimization that included Bayesian optimization rather than random or grid search. A custom approach would add significant upfront engineering work and ongoing maintenance. Upon discovering the SigOpt Intelligent Experimentation Platform, the team rigorously evaluated SigOpt from a variety of perspectives, including performance, user experience, and experimentation capabilities. Their most important evaluation criteria were ease of use and reliability in scaling their jobs without breaking their workflow.

Until finding the SigOpt Intelligent Experimentation Platform, the Numenta team had spent a lot of time troubleshooting open source Bayesian optimization packages and distributed schedulers for these packages, which was an inefficient and ineffective use of time for the modeling projects. By systematically testing SigOpt, the Numenta team found they didn't have to think about the experimentation workflow anymore and could simply use an API to enable tracking and hyperparameter optimization. SigOpt worked seamlessly with higher throughput, better results, and less time wasted in the process. This efficiency made it possible for the Numenta team to tackle more modeling projects faster and to achieve better outcomes. As a result, the SigOpt Intelligent Experimentation Platform became the obvious choice for removing the headache of implementing their own experimentation tools and accelerating architecture model development.

“The SigOpt Intelligent Experimentation Platform is easy to implement as a system of record for all of your experiments—across model type, task, or package. But what sets it apart is its capacity to guide experimentation so you can uncover insights on model behavior and develop configurations of models that fit your specific needs.”

Subutai Ahmad, VP of Research, Numenta

Using SigOpt to enhance AI experimentation



The SigOpt Intelligent Experimentation Platform is a model development platform that makes it easy to track runs, visualize training, and scale hyperparameter optimization for any type of model built with any library on any infrastructure. Numenta implemented the SigOpt Intelligent Experimentation Platform to design experiments, track artifacts, explore the model space, and optimize model hyperparameters. This made the entire process much more cost-per-day resource efficient; even more, advanced features of the SigOpt Intelligent Experimentation Platform led the team to realize novel insights about their modeling problem.

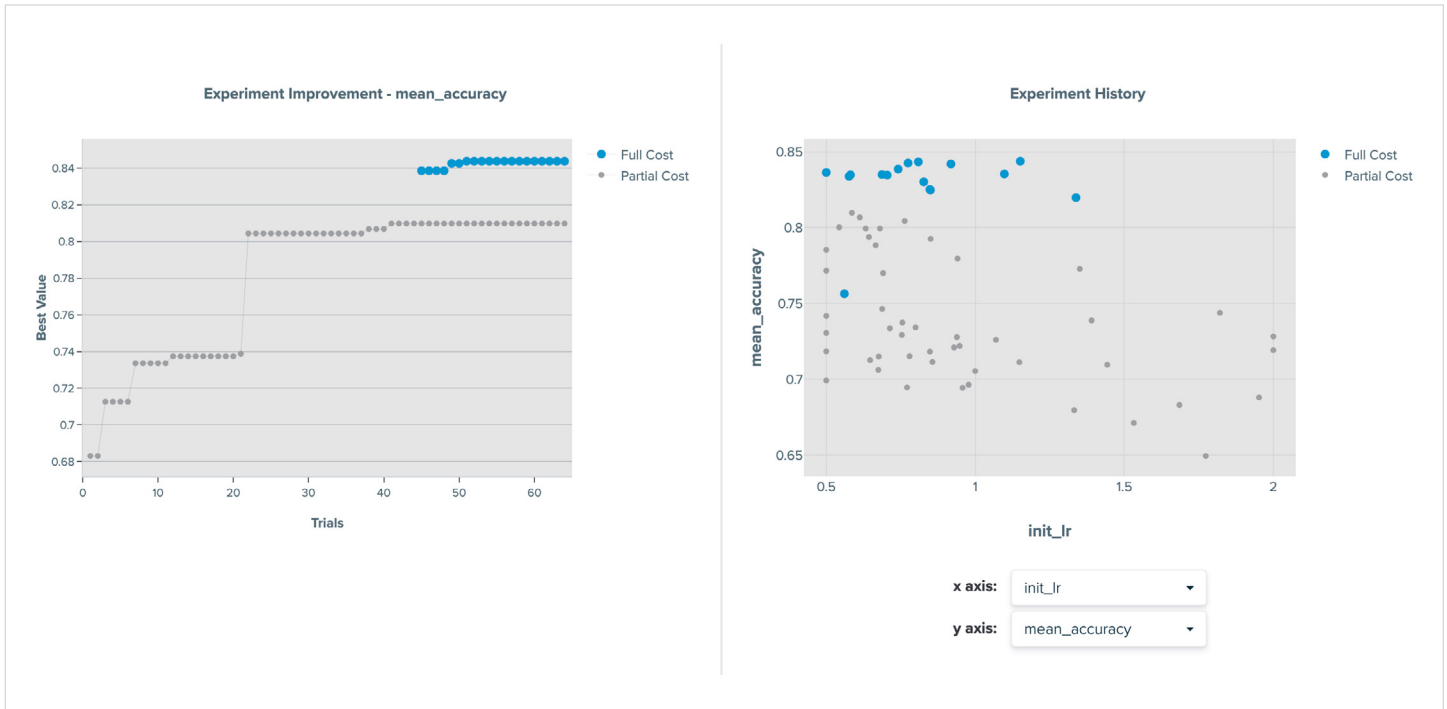
The Numenta team had the basic SigOpt Intelligent Experimentation Platform parameters—categorical and floating point parameters with different ranges for each, along with a modest sampling budget, divided into high and low fidelity tasks. A cost of running each task was assigned, dependent on the number of epochs the team wanted to run, called a multitask experiment. A total of four separate tasks were set up ranging from a 10-epoch version that was one-eighth the cost of a full iteration to the 80-epoch task that represented one full task.

The best way the Numenta team found previously to optimize the money spend on Amazon Web Services was to train smaller versions of the network; however, these parameters did not transfer well to the full network. To address this, the team structured an experiment that ran

fewer epochs per training run early in the hyperparameter optimization cycle and with more epochs later in this cycle, reducing the overall training time for the hyperparameter optimization cycle.

This particular experiment took a few weeks because each run took time to execute. Shorter runs for a single hyperparameter took four to eight hours to complete, and longer runs processed over several days. Additionally, the team hit various technical infrastructure bugs throughout the project, which took unexpected time away from experiment activities and delayed results.

One of the benefits of the SigOpt Intelligent Experimentation Platform is that it stores metadata in the cloud. The benefit of this approach is that all models and data stay private as SigOpt does not access this information; however, SigOpt does access non-sensitive information that makes it easy for modelers to run continuously. For example, when there is a system crash, the run can simply be restarted using the same experiment ID. This means that the SigOpt platform would pick up where the experiment left off and simply indicate, “This suggestion is no longer valid.” In this case, when the Numenta team experienced a system crash, they could disregard that point knowing SigOpt would not process the crash as a still-pending experiment and instead would simply disregard that particular run.



The SigOpt dashboard automatically populates a full history of your runs and experiments, including visualizations of your training runs, plots comparing metrics, and analytics like parameter importance to help you understand model behavior.

There were three specific feature sets that the Numenta team used to address challenges and achieve cost-effective experimentation results.



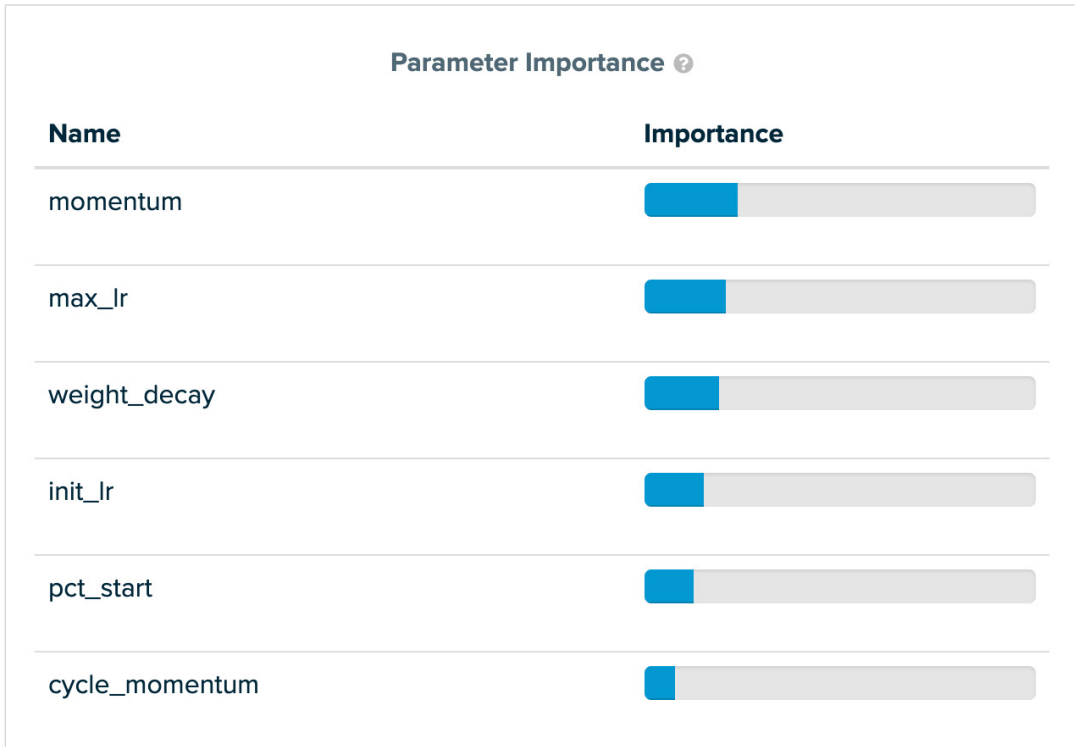
1. Managing model training wall-clock time

Challenge: Models like ResNet-50 take a long time to train on ImageNet, so it is time consuming and expensive to run experimentation on them to learn about their behavior.

SigOpt solution: Researchers can apply a variety of industry-standard techniques to reduce wall-clock time, including running in parallel across machines, applying Bayesian optimization instead of random or grid search to cut down on the number of runs required for convergence, and allowing the tracking of runs to monitor convergence if necessary. The SigOpt Intelligent Experimentation Platform enables all this functionality

while also providing a unique fourth technique: multitask optimization. This technique made it easy for Numenta to set up sample-efficient hyperparameter optimization jobs to train with fewer epochs early on and increase the number of epochs during later runs in the same tuning cycle. This reduced total training time for a hyperparameter optimization job.

Business results: As a result of implementing all these techniques, including multitask optimization, the Numenta team was able to reduce the wall-clock time for tuning ResNet-50 on ImageNet by 80 percent.¹ This made it possible to do more experimentation on the model, a requirement for developing a novel architecture that was more sparse and performant at the same time. In short, it enabled the Numenta team to do innovative experimentation that led to entirely new neural network architectures.



The SigOpt web dashboard allows you to derive the information you need to feel confident in your modeling. Every created run is an opportunity to learn more about your modeling process as well as an opportunity to compare new and existing models. This is why the SigOpt dashboard is created with the flexibility that allows you to create the visualizations that you need to make your modeling decisions and compare different runs and models.



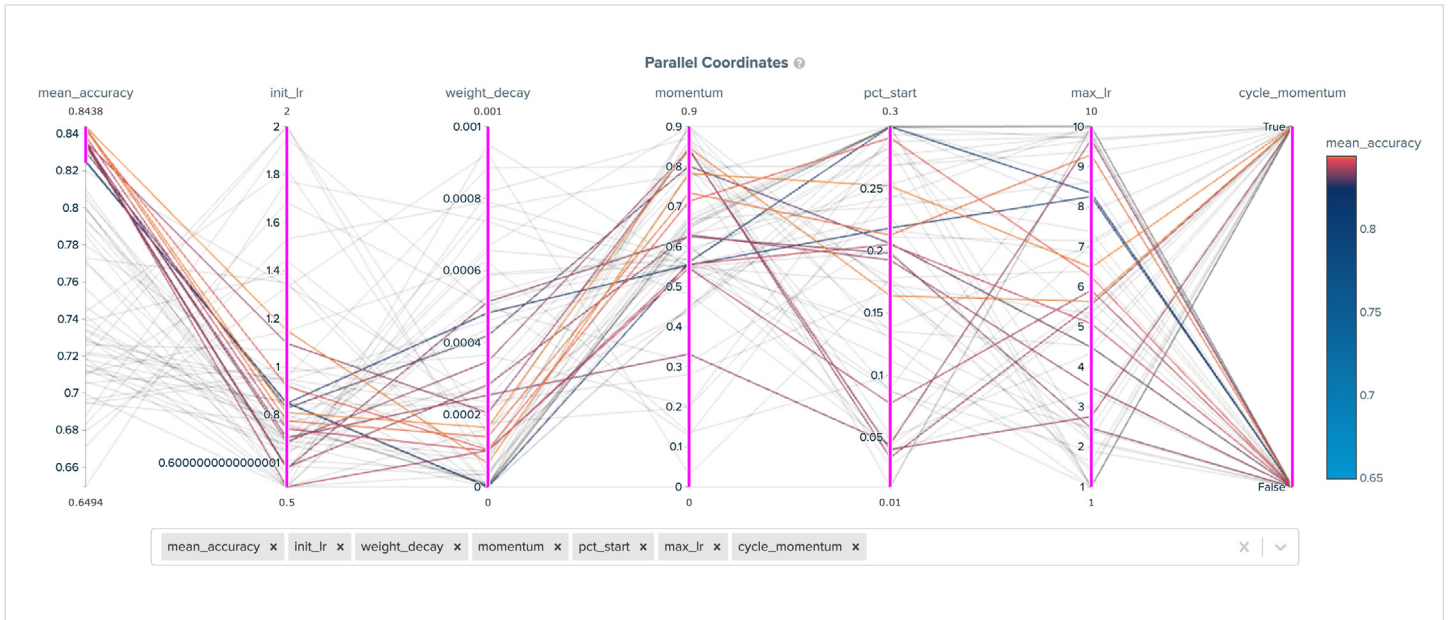
2. Deriving insights from training and tuning jobs

Challenge: Most deep learning (DL) models are black boxes, so it can be difficult to derive what does and does not work when training them. It is possible to develop custom plots and run custom analysis, such as parameter importance or parallel coordinates; however, it is difficult to piece together these plots and analytics into a bigger picture of how a model is performing.

SigOpt solution: SigOpt pairs API-enabled tracking and optimization with a dashboard that populates the full history of runs. It also offers visualization of runs to better understand convergence, parameter importance, and other analyses to get a sense of what is driving performance. It even allows you to

compare performance of metrics across runs to inform model selection and artifact storage that allows users to upload their own data, plots, images, and other inputs into the dashboard. This combination means modelers have the convenience of everything in one place to analyze model performance, and, ultimately, understand their modeling problem in greater depth.

Business results: The Numenta project benefited from the SigOpt Intelligent Experimentation Platform in that it provided access to features unavailable in other similar products. These features allowed the team to design and run novel experiments that were either more efficient or more insightful. The Numenta team relied on these types of insights to iterate their novel ResNet-50 architecture, which, ultimately, enabled them to make the right adjustments to evolve a sparse and performant architecture.



As you execute runs and experiments, SigOpt populates analytics that help you more deeply understand your modeling problem space, such as parallel coordinates and parameter importance.



3. Determining manual vs. automated workflow

Challenge: Manual steps in the workflow often take significant developer time, which is resource intensive. This is particularly true when it comes to optimization, where the Numenta team felt the pain of troubleshooting open source schedulers and optimizers. The engineers ended up spending so much time optimizing that they considered building a full-fledged solution—before they found the SigOpt Intelligent Experimentation Platform. It is possible to manually log runs, set up compute for training runs, implement your own algorithms for hyperparameter optimization and create your own charts, although it is best to outsource all these tasks to existing packages.

SigOpt solution: The SigOpt Intelligent Experimentation Platform automates the painful tasks in the workflow, like logging and hyperparameter optimization, via the SigOpt API or command line, which can easily be folded into any workflow. The Numenta team automatically managed tracking, scheduled jobs, and optimized hyperparameters.

Business results: The SigOpt Intelligent Experimentation Platform helped the Numenta team do a wide variety of experimentation to boost productivity while minimizing the team hours, project time, and compute resources needed. Additionally, the team got better results overall. This conserved resources and gave better guidance on how to adjust future model training for optimal experimentation results.

“The SigOpt Intelligent Experimentation Platform saved significant wall-clock time, team time, and compute resources while also giving the team unique insights on the modeling space. As a result, Numenta developed a state-of-the-art neural network that was

75% sparse
and still achieved
over 77% accuracy.”²

Subutai Ahmad,
VP of Research, Numenta

Proving the sparse approach to neural networks at scale is a viable path to pursue further

The Numenta team quickly discovered that many of the traditional ways to train dense networks do not consistently apply to sparse networks. In fact, they found that much of what works for dense networks does not work for sparse networks, making existing literature minimally insightful for this project. It was necessary to innovate learning relative to sparse network training.

As a result of this clarity, the Numenta team found a set of hyperparameters that were not conventional for dense networks but worked extremely well for sparse networks. This set of hyperparameters will be used as the basis of future studies. Secondly, the network was trained in such a way that it could be quantizable, or run at scale.

The SigOpt Intelligent Experimentation Platform saved the Numenta team significant wall-clock time, team time, and compute resources while also giving the team unique insights on the modeling space. As a result of the insights gained by working with SigOpt, Numenta developed a version of the neural network that was 75 percent sparse and still achieved 77.1 percent with Top-1 accuracy.³

This experiment demonstrated that a sparse approach to neural networks at scale is a viable path. The result, which relied on experimentation enabled by SigOpt, could ultimately lead to far more generalizable and intelligent networks. More immediately, this experiment showed that a more deliberate approach to experimentation, as enabled by SigOpt, can deliver better modeling results.

Learn more

For details on the SigOpt Intelligent Experimentation Platform, visit sigopt.com.



1. Results reported by Numenta during their experimentation.

2. Results reported by Numenta and, for accuracy, available via the SigOpt Dashboard.

3. Results reported by Numenta and, for accuracy, available via the SigOpt Dashboard.

SigOpt and Intel do not control or audit third-party data. You should consult other sources to evaluate accuracy.

SigOpt and Intel® technologies may require enabled hardware, software, or service activation.

No product or component can be absolutely secure.

Your costs and results may vary.

© Intel Corporation. Intel, the Intel logo, and other Intel marks are trademarks of Intel Corporation or its subsidiaries. Other names and brands may be claimed as the property of others.