datamachines

linode

About Data Machines

Data Machines is a data analytics company specializing in designing, building, and using cloud architecture to engineer and share solutions to problems in data analytics, DevOps, machine intelligence, and data science.

datamachines.io

How Data Machines Does Big Data

When Dr. Brian Dennis goes to work each day, he has the privilege of knowing his work impacts diverse challenges such as threat finance, human trafficking, cybersecurity, decision support, healthcare, investing, and litigation.

Data Machines Corp. is a data analytics company specializing in designing, building, and using cloud architecture to engineer and share solutions to problems in data analytics, DevOps, machine intelligence, and data science. The company's systems and code are used daily by over 2,500 researchers, data scientists, and research transition partners to enable advanced R&D projects. Data Machines works with broad research teams composed of government, industry, and academia in white-label transitioning of important technologies.

Dr. Dennis joined Data Machines in 2018 after holding research positions at Invincea Labs and Lockheed Martin. As the company's Director of Data Engineering, he continues his 20+ years of applied data-centric research in the defense industry, working with organizations such as the Defense Advanced Research Projects Agency (DARPA). DARPA, a research and development agency of the United States Department of Defense, is responsible for developing emerging technologies for use by the military.

"Working with DARPA is particularly exciting because we're on the frontline with researchers who are doing things others might not hear about for five to 10 years and that will have an incredible impact in 10 to 15 years," said Dr. Dennis. "Our role is to provide infrastructure for them. We're helping them solve problems; we're being enablers for their work. Day-to-day, it's super stressful and challenging, but when you see the results, it's great."

One DARPA project that Data Machines supports is a program called "Learning with Less Labels." The objective is to find new and advanced machine learning algorithms that use far fewer labeled examples to achieve the same results.

> "Linode not only had great pricing for GPU instances but also was using the exact, hard-to-find NVIDIA cards that we were purchasing for our own data centers."

Dr. Brian Dennis
Director of Data Engineering, Data Machines Corp.

Dr. Dennis offered a simplified explanation of the project: "So let's say you're trying to classify an image. 'Is this a picture of a cat, or is this a picture of a dog?' If you're Google or Facebook, it's easy to get a billion labeled examples and train up these huge models that work pretty well. But if you're not Google or Facebook, it's hard to get a billion examples. In fact, in some domains that the government cares about, a billion examples might not even exist. So the focus of this program is on how advanced researchers can develop those algorithms and show for a fact that, for orders of magnitude, fewer labels they get the same quality for a task like image classification."

"Learning with Less Labels" is multifaceted with aspects focused on object detection, image classification, and machine translation. In a research approach common to DARPA, teams are taking a broad range of approaches: Some are trying to combine multiple different machine learning algorithms; some are diving deep down into technical processes; others are using massive amounts of compute, whereas others are working on getting algorithms to work well on one GPU card. Then there is an independent government evaluation team working to identify successful approaches and help each team adapt based on shared learning.

Data Machines contracts to provide the infrastructure that supports all of these teams worldwide, including researchers in Australia, Canada, Israel, The Netherlands, and throughout the U.S.

That's what led Data Machines to Linode.

"When you say machine learning today, that means GPUs," said Dr. Dennis. "All of the researchers want as much GPU capacity as they can get. We do run some GPUs in our own data center, but we needed more to support the full team around the world. That's when we started looking at public cloud providers to see if there was a way to make resources available to these researchers at a reasonable price.

"We got sticker shock from other larger providers, so I looked at Linode, with whom I'd been a personal customer for nearly a decade," continued Dr. Dennis. "Linode not only had great pricing for GPU instances but also was using the exact, hard-to-find NVIDIA cards that we were purchasing for our own data centers. Plus, Linode offers GPUs in different increments, which is hard to find with other providers. That increased flexibility is helpful when scaling up or bursting from development to testing."

Dr. Dennis admits that part of his responsibility is choosing the best infrastructure solution for each project, so he's constantly on the lookout for the best options and evaluating other providers. "Each project we support is a snowflake, so we have to keep an eye out for individual best solutions, balancing cost, elasticity, flexibility, and security," he said. "But what I've learned, and the advice I would share with others is this: There are other choices besides the big three-letter providers that can be just as valuable to you. Alternative cloud providers like Linode can be flexible and way more responsive. You should get outside the box in terms of contacting other providers.

"And from our experience at Data Machines, I can tell you that Linode brings many assets to the table, including cost, reliability, customer service, and simplicity in pricing and deployment. We love that Linode offers a relatively small but well-focused set of services that are easy to stand up. All the extra stuff is just a drag on our time."

Dr. Dennis added that the reliability and customer service provided by Linode are huge selling points. "Linode's reliability has been great, which is extremely important for the world-class researchers we serve. And Linode's customer service is top-notch and timely. You're working with a real person. If you have a customer with really tight demands, it's great to have good service on your side."

linode

The World's Largest Independent Open Cloud

Our mission is to accelerate innovation by making cloud computing simple, affordable, and accessible to all.

Linode accelerates innovation by making cloud computing simple, accessible, and affordable to all. Founded in 2003, Linode helped pioneer the cloud computing industry and is today the largest independent open cloud provider in the world. Headquartered in Philadelphia's Old City, the company empowers over 800,000 developers, startups, and businesses across its global network of 11 data centers.