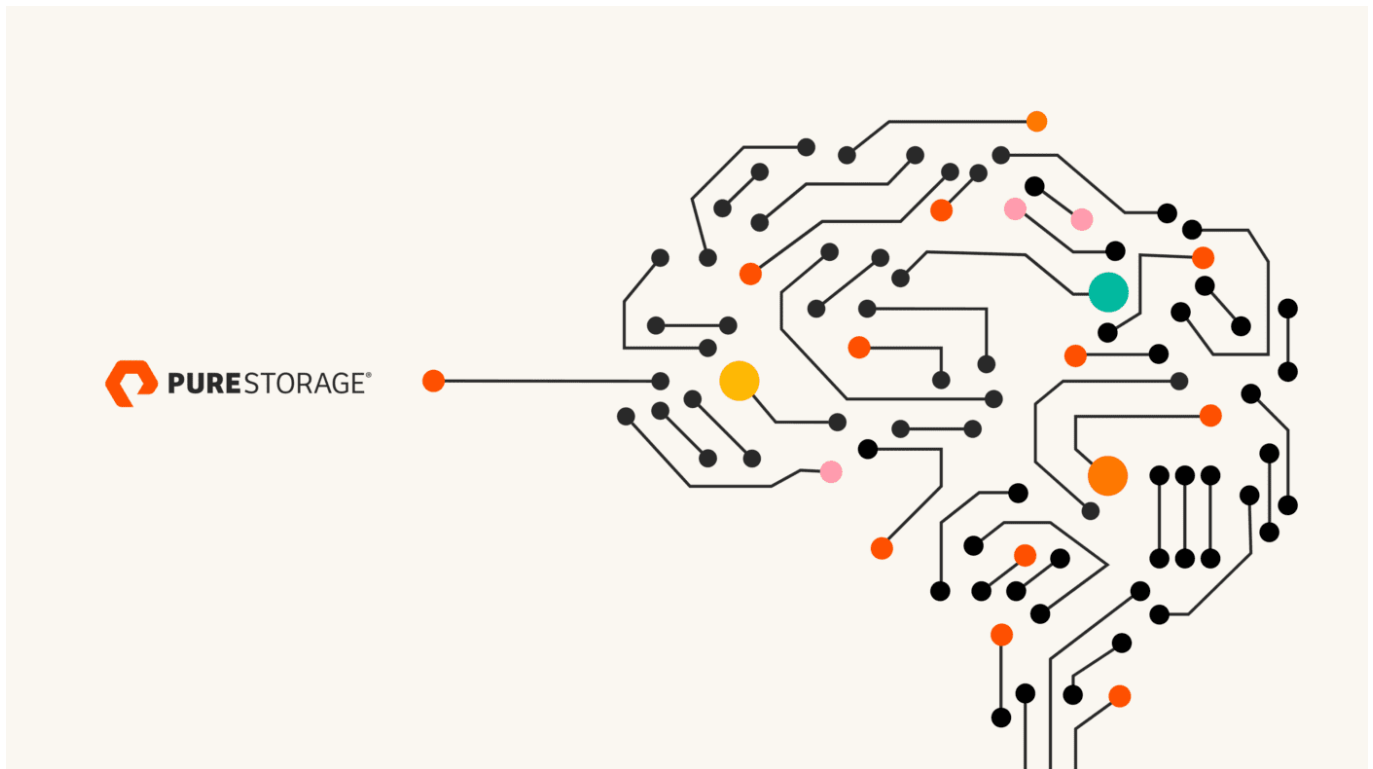


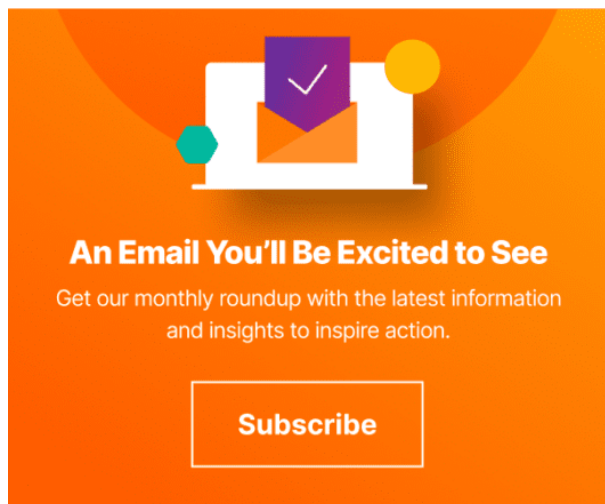
# Unlock Unstructured Data's Value With Machine Learning



Enterprises across every industry are increasingly recognizing the power of machine learning. It can turn any organization's data into valuable insights—insights that have the potential to revolutionize every aspect of the business.

Machine learning (ML) is a type of data analysis based on the concept that systems can use data to learn with little to no human effort needed. ML systems process data to identify patterns, ferret out anomalies, and recognize subtle correlations that people wouldn't notice. It gives organizations deeper insight into how and why some workflows are profitable or efficient and others aren't. It can help prevent fraud, eliminate production bottlenecks, inform sales and marketing professionals of what tactics work with which target audiences, and so much more.

The more organizational data an ML system has to process, the more accurate its output. And running ML workloads can generate its own massive amount of data, which needs to be stored and managed. Based on today's data trends, however, getting enough information to feed into the system isn't the problem.



## The Explosive Growth of Unstructured Data

Every year, the amount of data that businesses collect, store, analyze, and manage increases substantially. Digital transformation has resulted in most organizations using software in the form of automated platforms and applications to track customer accounts, sales, production processes, employee productivity, customer satisfaction and feedback, financials, and so on. All of that data is extremely valuable, especially when it's analyzed together by ML systems to root out the hidden correlations.

Most of that information is unstructured data, which can consist of information from all of those digital platforms and applications, as well as from sensors, telemetry systems, social media accounts, and more. To get maximum value out of ML systems, it's critical that organizations find a way to integrate all of that unstructured data into a unified platform.

Traditionally, that was pretty difficult. Unstructured data doesn't fit neatly into formatted tables. It consists of both files and objects, which were typically stored separately. It was hard to recognize what was relevant to keep and sometimes impossible to find. With a variety of disparate systems collecting and generating separate stores of data, IT admins were hard-pressed to manage or access it efficiently. And for some organizations, the sheer volume of unstructured data makes it impossible for a human workforce to handle it adequately.

Take YouTube, for example. Every single minute, users across the globe upload 500 hours of video content to the platform<sup>1</sup>. The amount of video uploaded in one day on YouTube couldn't be viewed by a single person in their entire lifetime. As a result, YouTube needs a way to filter videos to deliver the best content to its users, with minimal effort on their part.

YouTube relies heavily on machine learning algorithms to filter videos into various categories. The algorithms also flag and remove objectionable and explicit content, and enforce copyright protections on each of the uploads.

If YouTube had to employ humans for these tasks, it would need millions of employees. Machine learning algorithms make it easier to analyze large volumes of unstructured data like videos. YouTube can quickly understand video content to serve it up to the appropriate users. It's also able to optimize the performance of ads shown on its videos to drive a high ROI from them.

# Make Machine Learning Work for You: Choosing the Right Tools

To unlock the value of your unstructured data through machine learning, you first need the right tools—and that includes your storage platform. Traditional storage infrastructure isn't sufficient because it's highly siloed and typically separated using a wide variety of architectures optimized for specific workloads. Unstructured data can vary based on size, file and object count, processing requirement, and file and object protocol. To get the most value out of unstructured data, you should have the ability to analyze it all together. The ideal solution is a storage platform that can consolidate all forms of unstructured data so you can simply and efficiently store, access, manage, and analyze it.

Here's where Pure Storage® can help. With Pure [FlashBlade®](#), you get a unified fast file and object (UFFO) storage platform that serves as an all-in-one, high-performance, scale-out solution for all of your unstructured data. It's designed to handle even large-scale analytics and your biggest machine learning workloads—allowing you to gain a competitive edge by realizing your data's true value.

***To learn more about how Pure Storage can help you get the most out of your ML systems, [contact us](#).***

***Like this article and want to read more? [Sign up](#) for our monthly Perspectives email today. And we promise not to spam you, just inform and inspire you!***

---

<sup>[1]</sup> <https://blog.youtube/press/>

