CLOUD MACHINE LEARNING

Artificial intelligence and machine learning are steadily making their way into enterprise applications in areas such as customer support, fraud detection, and business intelligence. There is every reason to believe that much of it will happen in the cloud.

The top cloud computing platforms are all betting big on democratizing artificial intelligence. Over the past three years, Amazon, Google, and Microsoft have made significant investments in artificial intelligence (AI) and machine learning, from rolling out new services to carrying out major reorganizations that place AI strategically in their organizational structures.

So, if the cloud is the destination for your machine learning projects, how do you know which platform is right for you? In this article, we'll explore the machine learning offerings from Amazon Web Services, Microsoft Azure, and Google Cloud Platform.

WHAT ARE THE BENEFITS OF ML IN TH CLOUD?

- **The cloud's pay-per-use model is good for bursty AI or machine learning workloads.**
- **The cloud makes it easy for enterprises to experiment with machine learning capabilities and scale up as projects go into production and demand increases.**
- **The cloud makes intelligent capabilities accessible without requiring advanced skills in artificial intelligence or data science.**
- **AWS, Microsoft Azure, and Google Cloud Platform offer many machine learning options that don't require deep knowledge of AI, machine learning theory, or a team of data scientists.**

You don't need to use a cloud provider to build a machine learning solution. After all, there are plenty of open source machine learning frameworks, such as TensorFlow, MXNet, and CNTK that companies can run on their own hardware. However, companies building sophisticated machine learning models in-house are likely to run into issues scaling their workloads, because training real-world models typically requires large compute clusters.

The barriers to entry for bringing machine learning capabilities to enterprise applications are high on many fronts. The specialized skills required to build, train, and deploy machine learning models and the computational and special-purpose hardware requirements add up to higher costs for labor, development, and infrastructure.

These are problems that cloud computing can solve and the leading public cloud platforms are on a mission to make it easier for companies to leverage machine learning capabilities to solve business problems without the full tech burden.

There are many good reasons for moving some, or all, of your machine learning projects to the cloud. The cloud's pay-per-use model is good for bursty AI or machine learning workloads, and you can leverage the speed and power of GPUs for training without the hardware investment. The cloud also makes it easy for enterprises to experiment with machine learning capabilities and scale up as projects go into production and demand for those features increases.

Perhaps even more importantly, the cloud makes intelligent capabilities accessible without requiring advanced skills in artificial intelligence or data science skills that are rare and in short supply.

AWS, Microsoft Azure, and Google Cloud Platform offer many options for implementing intelligent features in enterprise applications that don't require deep knowledge of AI or machine learning theory or a team of data scientists.

THE SPECTRUM OF CLOUD MACHINE LEARNING SERVICES

It's helpful to consider each provider's offerings on the spectrum of general-purpose services with high flexibility at one end and special-purpose services with high ease-of-use at the other.

For example, Google Cloud ML Engine is a general-purpose service that requires you to write code using Python and the TensorFlow libraries, while Amazon Rekognition is a specialized image-recognition service that you can run with a single command. So, if you have a typical requirement, such as video analysis, then you should use a specialized service. If your requirement is outside the scope of specialized services, then you'll have to write custom code and run it on a general-purpose service.

It's worth noting that all three of the major cloud providers have also attempted to create general-purpose services that are relatively easy to use. Examples include the Google Prediction API, Amazon Machine Learning, and Azure Machine Learning Studio. They fall somewhere in the middle of the spectrum. At first, it might seem like this type of service would give you the best of both worlds, since you could create custom machine learning applications without having to write complex code. However, the cloud providers discovered that there isn't a big market for simple, general-purpose machine learning. Why? They're not flexible enough to handle most custom requirements and they're more difficult to use than specialized services.

HOW IS HARDWARE IMPACTED BY MACHINE LEARNING WORKLOADS?

- **Machine learning workloads require greater processing power**
- **The amount of processing required could be expensive**
- **GPUs are the processor of choice for many ML workloads because they significantly reduce processing time**
- **Google and other companies are creating hardware that's optimized for machine learning jobs**
- **To help people get started with AI, Amazon offers a camera that can run deep learning models**

Hardware is an important consideration when it comes to machine learning workloads. Training a model to recognize a pattern or understand speech requires major parallel computing resources, which could take days on traditional CPU-based processors. In comparison, powerful graphics processing units (GPUs) are the processor of choice for many AI and machine learning workloads because they significantly reduce processing time.

BENEFITS OF ML IN THE CLOUD: CONCLUSION

Since Azure, Google Cloud, and AWS all provide good general-purpose and specialized machine learning services, you will probably want to choose the platform that you've already chosen for your other cloud services. However, to avoid vendor lock-in when using a general-purpose service, you may want to

use an open-source machine learning framework that is supported by all three vendors. At the moment, the framework with the broadest support is TensorFlow, although the field is changing rapidly, so we expect cross-platform support for more frameworks soon. The main holdout is Google, which previously supported only TensorFlow, but even Google is now introducing support for scikit-learn and XGBoost.